Differentiation of human embryonic stem cells towards neural progenitor cells to study the effect of paracetamol on prenatal brain development

Dissertation submitted for the degree of *Philosophiae Doctor*

A monograph by

Bengt Martin Alfred Falck

Department of Biosciences

Faculty of Mathematics and Natural Sciences

University of Oslo

2020

Contents

Acknowledgments

My gratitude goes to the good people I have been fortunate enough to get to know in my life. I am incredibly happy and fortunate to have a family in blood and close friends and my amazing partner. I would do anything for these people. At this point in my life I am slightly wiser (well I think so), so I´ve stopped looking too hard for more meaning in life, since it was here already, my family and friends. You all know who you are - if nothing else for the electronic spam you get from me at irregular times. I will be a better friend from now on and reach out on a more continuous basis, whether you like it or not!

Special gratitude goes to the people who were there to give me feedback, advice, and support when I was well into advanced covid19/forest-craziness, there truly are some fantastic colleagues (not just science-related) in diverse places. Never stop what you are doing. Goodness can be difficult to find and you are all worth well over your weight in .. diamonds, just to top that old adage with something more valuable than gold. Never hesitate when you are in need of help, I will be there.

P.S. Grandma, I miss you and I wish you were still here for this.

P.S.[2], *Dubbdäck eller ej?*, *Honkland*, *hushållsost eller 128kbs pavarotti på youtube.com fejsbook edition*, *Shroomers united*, and *ÜS/KUWTF* - just keep being awesome!

Experiment, thesis, and general contributions

The two main parts of this thesis

Part I - *Differentiation of human embryonic stem cells to neural progenitor cells for neurotoxicology studies*

Pages … 4 - 110

Part II - *Paracetamol and human neurodevelopment*

Pages … 111 - 145

List of figures and tables

Part I

iv

Supplemental material

List of abbreviations

Abbreviation	Full name
5-HT	Serotonin or 5-hydroxytryptamine
ADHD	Attention deficit hyperactivity disorder
ASD	Autism spectrum disorder
BMP	Bone morphogenic protein
CGE	Caudal ganglionic eminence
CNS	Central nervous system
CTPs	Corticothalamic projection neuron precursors
ddPCR	Digital droplet polymerase chain reaction
EpiSCs	Epiblast derived stem cells
FACS	Fluorescence-activated cell sorting
GE	Ganglionic eminences
GW	Gestational week
hESCs	Human embryonic stem cells
hiPSCs	Human induced pluripotent stem cells
hPSCs	Human pluripotent stem cells
ICC	Immunocytochemistry/immunofluorescence
ID	Intellectual disabilities
IPs	Interneuron progenitors
iPSC	Induced pluripotent stem cells
iSVZ	Inner subventricular zone
LGE	Lateral ganglionic eminence
LINE-1	Long interspersed nuclear element 1
LSX	Small molecule inhibitors LDN-193189, SB431542, and XAV939
mESCs	Mouse embryonic stem cells
MGE	Medial ganglionic eminence
MZ	Marginal zone
NPCs	Neural progenitor cells

oSVZ	Outer subventricular zone
PCs	Principal components
PCPs	Prefrontal cortex precursors
REACH	Registration, Evaluation, Authorization, and Restriction of Chemicals
RGs	Radial glial cells
RPC	Rostral patterning center
scRNA	Single-cell RNA sequencing
snATAC	Single-nucleolus assay for transposase-accessible chromatin sequencing
SVZ	Subventricular zone
TFs	Transcription factors
t-SNE	t-distributed stochastic neighbor embedding
UMAP	Uniform manifold approximation and projection
WB	Western blot

Ph.D. project objectives

The PharmaTox Strategic Research Initiative was established on January 1st, 2015 with support from the Faculty of Mathematics and Natural Sciences at the University of Oslo, Norway. The initiative brings experts of different disciplines such as neuro-, pharma-, and epidemiology together with bioinformaticians and statisticians to work on common problems concerning human neurotoxicity and neurodevelopment (Amundsen et al., 2015; Bjørnstad et al., 2015; Brandlistuen et al., 2013). A year and a half later I was hired to work in a large project involving two core groups and other scientists with the original aim to, as originally stated [sic], ''to translate findings of epigenetic changes upon paracetamol exposure in human embryonic stem cells (hESCs), neural progenitors and neurite outgrowth, with findings in umbilical cord blood of children exposed to paracetamol in fetal life with neurodevelopmental delays'. My main supervisor is Ragnhild Eskeland, now an associate professor at the Institute of Basic Medical Sciences, and in the tradition of PharmaTox, we work in cooperation with Robert Lyle´s group at Ullevål University Hospital and several other scientists with diverse expertise. My contract and Ph.D. candidature was set from September 13, 2016, to September 13, 2020.

PharmaTox is currently working with a selection of common drugs. Many of these can be considered common medications, that might be taken during pregnancy. Broadly these drugs can be classified into three main groups: analgesics, antidepressants, and anticonvulsants. In the analgesic group, the first drug on the to-do list of selected medications is the widely known painkiller paracetamol. Paracetamol has been used for well over 100 years now (Cahn & Hepp, 1886) and part II of this thesis will be a study on paracetamol's effect on neurodifferentiation.

Initial project objectives were as follows

1. To set up and characterize paracetamol exposure on viability, cell renewal, pluripotency, and cell cycle in hESCs.
2. To measure the impact of exposure to paracetamol on hESCs and after neural differentiation.

3. To characterize the effect of paracetamol on gene expression, DNA methylation, selected histone marks, and H2A.Z in undifferentiated and differentiated hESCs and study how this impacts neuronal differentiation.

4. To correlate and translate the findings within PharmaTox.

As projects tend to do, objectives changed to:

Current project objectives are now

1) Set up and create an in-house protocol for neurodifferentiation of human embryonic stem cells to neural progenitor cells. This protocol should be robust enough for neurotoxicology studies.

2) Study the effects of paracetamol on neurodifferentiation utilizing our own protocol.

   a. Methods employed are standard laboratory methods combined with multi-omic single-cell sequencing.

3) Side project *LINE-1 mobile genetic elements*, and similar endogenous viral remnants. The goal is to study the change in element mobilization after exposure to paracetamol and medications used in the treatment of mental illnesses, e.g. schizophrenia and bipolar disorder.

LINE-1 side project

Outside the two major parts that I will describe in my thesis, I have also worked on a mobile genetic element project, long interspersed nuclear element 1 (LINE-1) retrotransposition. Mobile elements such as these can be linked with paracetamol as well as several mental diseases such as schizophrenia and various spectrum disorders. In the LINE-1 study, I have designed a drug-library around known retrotransposition-activating compounds (i.e. cocaine and methamphetamine) and drugs involved in the treatment of many human mental disorders (i.e. antidepressants, stimulants, antipsychotics, and mood stabilizers). This drug library has been trialed in a Master´s student project by experiments performed by the talented student MSc. Guro Cecilie Mustorp. Some of the future prospects in this project could be to use this library with the cells generated with our neurodifferentiation method.

Part I - Differentiation of human embryonic stem cells to neural progenitor cells for neurotoxicology studies

Early human brain development is a process with many unknown factors and mechanisms, especially for developing areas in the brain which we believe have many human-specific functions, such as transitory structures ganglionic eminences and the developing neocortex (Lancaster et al., 2013; Ozair et al., 2013; Suzuki & Vanderhaeghen, 2015). I will also venture to be human-centric enough to agree that the human brain likely is the most complex organ that has evolved among living organisms (on Earth) and it is also evident that our brains are still incapable of understanding most, if not all, of its parts, which also luckily means there is room for many exciting research projects within those subjects.

We started with the belief that our work in the cell-lab with neuronal differentiation of hESCs would take a relatively short period of time until we could start analyzing the first results from treatment. This did not quite work out as we expected, as things sometimes do in science. We abandoned the initial method, as well as a second one. We realized these methods would not work in our neurotoxicology studies as we needed to maintain a certain rigidity and reproducibility during a longer period of time (around 20 days). Once we realized that our current practices would not suffice, we started developing a more customized method for differentiation of hESCs towards NPCs, that would enable us to treat the cells with i.e. paracetamol and reproduce the experiment with comparable outcomes.

We had a breakthrough after substantial developments of the method in late 2018. This allowed us to apply our method successfully in small test runs. Then we did two repeated runs of the method, fully scaled for treatment with paracetamol, and to yield material for several downstream experiments. We were now able to gather and sequence material for both method validation and our investigation of paracetamol. We have since validated our cell-culture method and investigated paracetamol´s effects using several downstream methods; brightfield imaging, immunocytochemistry (ICC), Western blotting (WB), droplet digital polymerase chain reaction

(ddPCR), fluorescence-activated cell sorting (FACS), bulk RNA-sequencing, EPIC-array methylation assays, single-cell RNA sequencing (scRNA-seq), and single-cell assay for transposase accessible chromatin sequencing (snATAC-seq). Starting early 2019 until now, we have generated scRNA-seq, scATAC-seq, and DNA methylation datasets. These datasets alone contain substantial amounts of data and in their being relatively novel as they are obtained with recently developed sequencing technologies, they also provide challenges in terms of how to analyze the data for answering our particular questions.

It is worth mentioning that there is still debate ongoing on the effects of paracetamol, even though, as mentioned, it has been in use during the last 100 years. Because the effects of paracetamol are unknown in the context of brain development, we were curious about what we would observe once the analysis was completed. This thesis will rely heavily on scRNA-seq datasets as these are completed and analysis for these were possible to write before my contract ended. For our later publications, ATAC- and methylation profiling will be integrated, for a multi-omics investigation on the method and the effects of paracetamol. Even though this project was delayed, we now have data enough for several more publications than we initially planned.

Introduction

Medication in pregnancy

The developing human fetus is sensitive. Medication taken during pregnancy will always carry a potential risk as it can affect prenatal development in unexpected ways. This risk needs to be balanced with the benefit of the medication and communicated to mothers in a responsible way that does not cause undue harm. Traditionally toxicologists at pharmaceutical companies, universities-, and other agencies, obtain safety data on risk compounds with animal studies, epidemiological studies, or case reports. These studies have several limitations where an important factor is an interspecies translatability (Ward, 2001). There are obvious ethical issues in including pregnant women in drug trials, and sometimes the risk factors of previously used compounds can be difficult to determine. Indeed, less than 10 % FDA-approved medications

between 1980 and 2000 have satisfactory safety data (Andrade et al., 2004). Moreover, long-term outcomes have received limited attention and the need for faster and more accurate models of developmental toxicity is increasing. We have during modern times been very efficient in creating new compounds, increasing our exposure to these - and combinations thereof - while we sadly have not been as sufficiently accomplished at testing them for safety even after some has been in use for a long time.

The sensitivity of the developing brain

The developing brain and central nervous system are particularly sensitive to exposure from chemicals and pharmaceuticals which ultimately can lead to neurodevelopmental disorders such as schizophrenia, autism, and ADHD (van Thriel et al., 2012). The sensitivity is partly due to interplay between several complex mechanisms such as cell migration, patterning, proliferation, apoptosis, cell differentiation, synaptogenesis and pruning, neurite outgrowth, neurotransmitter turnover, and myelination (Kadereit et al., 2012). This extensive interplay of pathways means that transcriptional programs and the epigenome can be particularly sensitive during certain developmental windows. There exists a body of evidence and examples of prenatal insults, such as compound exposure or maternal stress, that can disrupt the epigenome and in turn, be linked to neurodevelopmental disorder (Bollati & Baccarelli, 2010; Kundakovic & Jaric, 2017; Raciti & Ceccatelli, 2017). There is an ongoing effort to understand more about effects on the transcriptional and epigenomic changes in neurodevelopment in the context of neurodevelopmental toxicity (Raciti & Ceccatelli, 2017). It has been shown that there exist human peculiarities in key events in the development of our brain which is different when compared to other mammals (Kanton et al., 2019). These unique aspects of brain development highlight that it is important to make sure that there is no human-specific toxicity with compounds that are correlated with adverse development in epidemiological studies.

Alternative models for neurotoxicity testing

Recently, hESCs have proven valuable for developmental toxicity studies as an alternative to animal models or immortalized human cell lines (Colleoni et al., 2012; Krug et al., 2013;

Schulpen et al., 2015). These cells offer unique features in their cytogenic stability, proliferation capacity, and potential to differentiate to all three germ lineages. Moreover, in vitro differentiation of hESCs mimic gene expression patterns of early embryonic development (Abranches et al., 2009; Fathi et al., 2011; Hay et al., 2008), and thus can potentially be utilized to identify targets of drug-induced developmental toxicity.

hESCs, and human induced pluripotent stem cells (hiPSCs), continues to show promise for neurotoxicology studies as a complement to animal models. hESC- or hiPSC-based in vitro studies provides the possibility to detect human-specific adverse outcomes in terms of neuronal development and can be tuned for different developmental windows, decrease animal suffering, and are cost-effective and rapid to deploy (Bal-Price et al., 2018). A three-dimensional approach can be powerful for *in vivo* comparison, but the complexity and the multitude of variables carry the risk of decreasing reproducibility, throughput, cost, and further complicates an analysis, whereas the simplicity of a monolayer system can be a strength in terms of exploratory neurodevelopmental toxicology. There is an urgent need to evaluate the safety and impact on human health of thousands of chemicals on the market currently targeted for re-evaluation in accordance with the EU regulation REACH (Registration, Evaluation, Authorization, and Restriction of Chemicals) (Combes et al., 2003). Assessing all those chemicals for human neurotoxicity would result in an enormous need for laboratory animal testing which would be in conflict with the demand for the 3 R´s, Refinement, Reduction, and Replacement in terms of laboratory animals (Guhad, 2005).

Our approach to single-cell sequencing is that we might be able to measure drug-induced changes that are more subtle in hESCs differentiation towards, and along, the neural lineage. Slight changes, such as an affected expression program or a shift in population numbers for a particular cell type, are events that could lead to increased risk of disease and could be of variable severity depending on when in development it would take place - examples being neuronal migration and layering in the brain, or slight shifts of more or less excitatory or inhibitory neurons. Using human cells in early neurodevelopment might also provide a way to

see some human specific effects, which could be difficult, or impossible to discover in animal models.

hESCs biology in comparison with mouse embryonic stem cells

Embryonic stem cells are isolated from the blastocyst's inner cell mass and are pluripotent (multi-potent) - they can develop into any cell type except cells found in the placenta. A cell that is capable of also creating placenta cell-progeny is called totipotent (all-potent), and those can only be cells isolated from the morula stage, before the blastocyst formation. The blastocyst is a structure shared in mammals, which is formed in early development after approximately day 5, post-conception. From the inner cell mass of the blastocyst, the epiblast will form and give rise to three primary germ layers called *endoderm* (*endo* - inside), *mesoderm* (*meso* - middle), and *ectoderm* (*ecto* - outside), in a process known as gastrulation where also the embryos primary body axes are established. The endoderm layer will form the gastrointestinal and respiratory systems. The mesoderm will become skeletal and cardiac muscle, bone, cartilage, connective tissue, lymphatic tissue, and more. The ectoderm will give rise to the nervous system: brain, spinal cord and peripheral nerves, skin, and lining for nostrils, mouth, hair, nails, and anus. hESC-lines commonly in use today are pluripotent and isolated from the blastocyst. Besides pluripotency, other hallmarks of the embryonic stem cells are indefinite self-renewal and rapid growth, and unsurprisingly, cancer tumors often activate the same genes that give these characteristics in stem cells (Clarke & Fuller, 2006).

About a decade ago, induced pluripotent stem cells (iPSCs) were said to soon replace the use of embryonic stem cells (Castelvecchi, 2018). iPSCs are embryonic stem cell-like cells retro-converted from primary cell types (Takahashi & Yamanaka, 2006) and are sometimes said to in essence be the same as hESCs. iPSCs could be a powerful system for the investigation of diseases and in regenerative treatment since they can be converted from cells isolated from patients and donors. However, there is still ongoing concern about the safety of iPSCs whether these are comparable hESCs, e.g. whether they display the same epigenetic geography and

memory, whether they live as long, and whether they could be more tumorigenic. hESCs are still in heavy use and are often used for comparison to iPSCs in studies (Castelvecchi, 2018).

There are differences between mouse embryonic stem cells (mESCs) and hESCs. mESCs acquire a naïve pluripotency state and the conversion of hESCs to naïve state, similar to what the mESCs are in natively, have seen some effort (Duggal et al., 2015). Naïve pluripotency in mESCs is said to be advantageous since a primed state would show bias towards specific lineages. The naivety in *naïve* indicates unbiased pluripotency and a stem cell which has an earlier developmental state akin to the pre-implantation embryo (Xu et al., 2016). In comparison to most available mESC lines, hESCs lines will instead have a cell-state more alike the post-implantation embryo. Naturally, it would be particularly important to be aware of these differences when doing research that requires hESCs biology to represent the pre-implantation embryo. Neurodifferentiation methods that utilize strong induction by combinations of inhibitors and proteins would most likely not be significantly affected whether the hESCs are in a primed state or not.

Human forebrain development

The developing forebrain and cortex are complex structures that are thought to be the most divergent when comparing the human brain to our close relatives in evolutionary terms, the chimpanzees, and other great apes (He et al., 2017; Kanton et al., 2019; Somel et al., 2009; Sousa et al., 2017). The developmental process is poorly understood and has been complicated to study, however, more advanced brain organoid and 'brain on a chip' models together with single-cell omics have provided insights in recent years (Bradley, 2020; Kanton et al., 2019; Rifes et al., 2020). There are cell atlases on mouse brain development by the Allen Institute for Brain Science (A. R. Jones et al., 2009; Sunkin et al., 2013) available online for anyone to explore, and they also host data for human brain development. However, data for the human brain is not as well-annotated, or as detailed, as the available mouse data (Marshall & Mason, 2019; Pressler & Auvin, 2013). Thus, for structures in the human brain that are very divergent from other mammals, it can be difficult to use mouse data for direct comparison with datasets based on

divergent brain regions in humans. Some recent single-cell sequencing-based studies also seem to specifically start with mouse data for annotating corresponding human cell types (Rifes et al., 2020) (check http://linnarssonlab.org/ for recent ongoing efforts on human brain developmental atlas).

The early human embryo is divided into three layers; endoderm, mesoderm, and ectoderm 'Embryological development of forebrain chapter, 2009' (Medina, 2009). Ectoderm is the outermost layer and will eventually become the central nervous system (CNS). A part of the ectodermal layer forms the neural plate which also establishes a developmental axis where gradients of transcription factors (TFs) interplay to orchestrate a sense of direction in the overall structure so that our important brain parts end up and develop where they are supposed to. The neural plate will then fold on itself and form the neural tube. Once it has closed, it fills with cerebrospinal fluid. The anterior, or front part, of the neural tube, expands to form three primary vesicles that create the forebrain, the midbrain, and the hindbrain (Figure 1 *Telencephalon in development* A.). The forebrain eventually separates into a rostral part, the telencephalon and a part behind called the diencephalon. In terms of human development, the forebrain will be visible in the fetus after 5 weeks *in utero* as a small portion towards the front of the fetus.
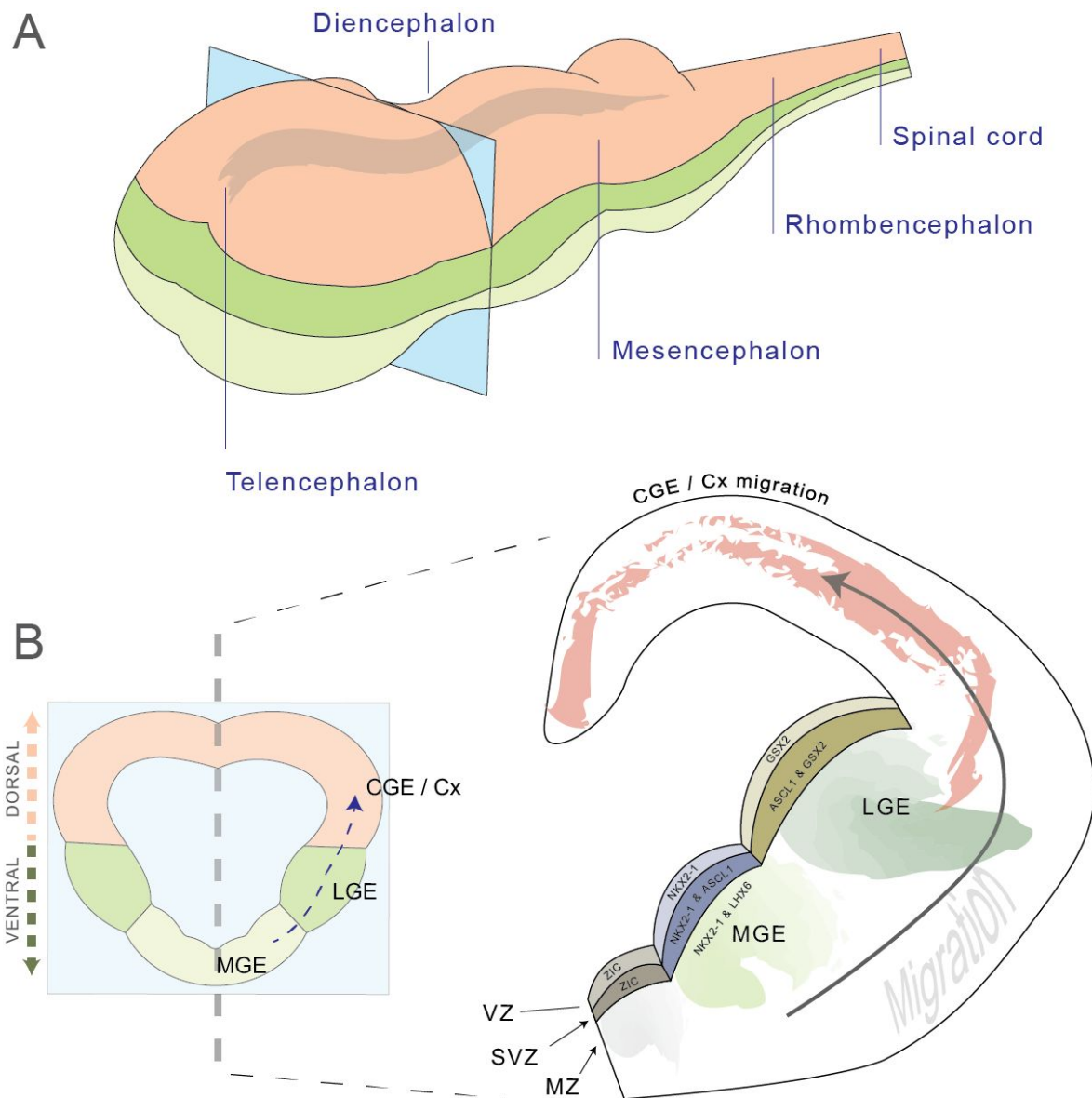
Figure 1 *Telencephalon in development* Developing forebrain, in Latin *procencephalon*, splits into the *telencephalon* and *diencephalon* A) Telencephalon is in the front, anatomical for this direction is the Latin word *rostral*. After telencephalon, we have the second part of the forebrain, the diencephalon. Thereafter, the midbrain, in Latin *mesencephalon*, and hindbrain, in Latin *rhombencephalon*. These parts ultimately end as the spinal cord. B) A cross-section of telencephalon shows the transitory structure known as the *ganglionic eminences* (GE) where migrating neuronal precursors arise and mature as they migrate to the cortex initially from the medial ganglionic eminence (MGE), then through the lateral ganglionic eminence (LGE), to the caudal ganglionic eminence (CGE) to arrive at the developing neocortex. Each hemisphere is mirrored and depicted on the right is a single hemisphere in early development. To help create and guide migrating neurons, accumulated in the MGE are radial glial cells (RGs) that have a similar function to stem cells as they renew and proliferate, but they importantly produce neural progeny that migrate towards the cortex. Radial glial grown perpendicular to the migration path. The migrating cells will have different characteristics depending on where in developmental time they are.

To guide cell migration and axon formation during brain development a transitory structure called the ganglionic eminence (GE) is formed and is present in early neural development (Figure 1 *Telencephalon in development B.*). GE consist of three regions starting at the medial ganglionic eminence (MGE), then lateral ganglionic eminence (LGE), and finally caudal ganglionic eminence (CGE). These regions are sources for the developing brain's first excitatory neurons, GABAergic neurons (neurons which later in development become inhibitory after a phenomenon known as the GABA-shift). GABAergic neurons will mature as they are guided towards the neocortex and other structures. Eventually, the GE will have become the basal ganglia, and neurons will have migrated from the GE to their final destinations.

In vitro **neurodifferentiation**

hESCs can be used effectively in studies of human-specific adverse outcomes in fetal neurodevelopment and can leverage hESCs cost-effectiveness and relative higher throughput in comparison to animal studies (Bal-Price et al., 2018). Although few, there have been attempts at developing *in vitro* stem cell models for use in developmental neurotoxicology (Krug et al., 2013), however, there is room for improvement, with regards to increased protocol robustness, simplicity, efficiency, definition, and reproducibility. In response to these needs, we have developed a protocol for use in neurodevelopmental toxicity testing. This protocol is based on hESCs in differentiation towards neural progenitor cells (NPCs) recapitulating early neurodevelopment.

The protocol uses a monolayer cell culture and differentiation of hESCs to NPCs. A strong emphasis has been placed on keeping the protocol as simple and robust as possible for improving reproducibility, and its ease of use for new users, while keeping the costs of the protocol to a minimum. We have taken additional steps to define our protocol at a single-cell level with the use of single-cell RNA sequencing at key time points in our protocol. Further, we will demonstrate our method by completing the 20-day differentiation scheme repeated times, during treatment with paracetamol (Part II of the thesis).

Part I - Results

A neurodifferentiation protocol

Introduction

SMADs (or Smads) are proteins that are important signal transducers and regulate large cell development and growth programs. By shutting down SMADs with selective inhibition and only allowing a chosen set to be active, you can limit the paths a stem cell can take when it differentiates. We use a mix of small molecule inhibitors that only allow differentiation towards ectoderm, closing down paths towards endoderm and mesoderm (Figure 2 *LSX small molecule inhibitors effectively block endoderm and mesoderm fates*). In 2009, Chambers and colleagues (Chambers et al., 2009) developed a dual SMAD-inhibition protocol for rapidly converting human pluripotent stem cells (hPSCs) into NPCs. By now this protocol is well known and extensively used or iterated upon. Chambers *et al* used Noggin (Valenzuela et al., 1995), a bone morphogenic protein (BMP) inhibitor and Activin/BMP/TGF-β pathway inhibitor SB431542 (Patani et al., 2009) in combination to demonstrate rapidly induced neural differentiation of hPSCs to NPCs.

Small molecule ALK2/3 inhibitor LDN-193189(Chambers et al., 2013; Kreitzer et al., 2013; Maroof et al., 2013; Yu et al., 2008) can be used as an alternative to Noggin, and while LDN-193189 almost retains the effect of Noggin, LDN-193189 was shown to produce a lower percentage of forebrain *FOXG1*-positive cells (Maroof et al., 2013; Yuan et al., 2015). Tankyrase inhibitor XAV939 (S.-M. A. Huang et al., 2009) when used in combination with SB431542 and LDN-193189, was shown to improve the differentiation of forebrain-type NPCs from both hESCs- and iPSCs-lines (S.-M. A. Huang et al., 2009; Maroof et al., 2013).
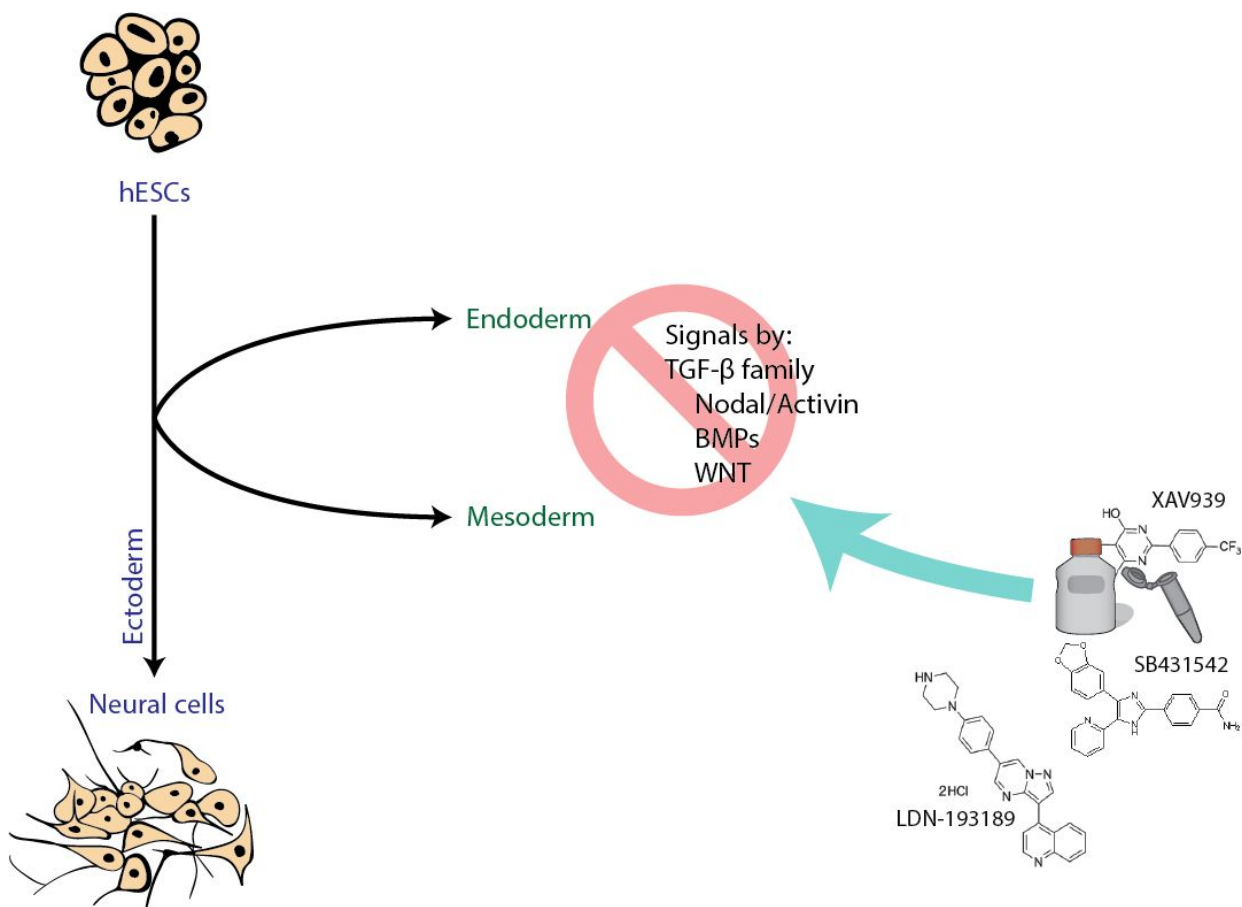
Figure 2 *LSX small molecule inhibitors effectively block endoderm and mesoderm fates* Figure show a simple cartoon on blockage of the endoderm and mesoderm lineages by addition of LDN-193189, SB431542, and XAV939, into the medium for *Part I: Induction* of the presented neurodifferentiation protocol. This inhibitor mix will block any fate except the neuroectoderm fate by inhibiting TGF-β superfamily signaling, e.g. Nodal/Activin, BMP, and Wnt-signaling.

At present, there are many cell differentiation protocols, for both continuous differentiation (differentiate cells through several points in a lineage - e.g. stem cell to neural ectoderm, then to neural tube formation, etc.), and direct reprogramming that can skip steps in order to directly convert e.g. a fibroblast to an astrocyte (Caiazzo et al., 2015). Many differentiation protocols have in common that they focus less on stability as in terms of cell survival, aiming to generate a novel cell-type, but usually together with an aim to have high survival rates. Once cells have been generated, that cell-type can be expanded upon. However, when attempting to do drug treatment over a longer differentiation window, and with the usual requirement to be able to

consistently reproduce results, varying cell survival and/or uncertainty of final cell type, will make it hard to convincingly study drugs that are thought to have subtle effects.

The basis of our work is centered around existing protocols for generating NPCs (Chambers et al., 2009; Kreitzer et al., 2013; Maroof et al., 2013) where we build upon these with changes and improvements that could benefit neurotoxicity testing. For this protocol, we have specifically improved on reproducibility, cell-survival, and simplicity. We also take advantage of single-cell sequencing to profile our cells, thereby adding biological knowledge as well as providing high-resolution sequencing results for cell-types at each time-points. Advantages of our protocol include defined cell seeding amounts, widely used cell substrates and media, minimized use of organic solvents (Pal et al., 2012), tweaked media composition to have daily media changes (where a drug of interest then can be added regularly). We believe our easy-to-use protocol can provide human-specific insights in addition to being relatively cost-efficient, making it a robust addition to the arsenal of neurotoxicity studies. We believe that it can be rapidly established in any laboratory with access to a basic cell culture facility.

Protocol start

The protocol consists of four time points and three major parts, the start (day 0), day 7, then day 13, and day 20 (Figure 3 *Protocol overview)*. Before day 0, HS-360 hESCs were maintained for at least 2-3 passages after thawing them from liquid nitrogen storage, and typically allowed to reach 85-95 % confluency before the cells were moved to a new vessel. A common mistake is to let the cells grow until confluency. The cells should rather be passaged before reaching a stationary phase (Handbook, 2011). LSX inhibitors and growth factors FGF2 and EGF should be prepared as a master stock which then is subdivided into aliquots and frozen. Avoid freeze/thaw cycles by using aliquotes sized for one day's medium change. A 6-well culture plates is used for regular hESCs maintenance where each well in this plate format usually provides around 800 000 to 1 500 000 hESCs (at confluency). The protocol was optimized for a 12-well format which allowed more throughput (up-scaling requires re-optimization, not just surface area conversion). The 12-well format is useful for testing more than one condition and will lower costs slightly.

Reducing to smaller than this format is possible, but might result in too little material for most standard analysis, but the 24-well format is suitable for e.g. ICC, using 13 mm coverslips.

ESCs were stabilized in routine culture for 2-3 passages. The protocol was initiated with healthy and homogeneous hESCs colonies, at about 85-95 % confluency (Figure 4 *Brightfield microscopy of time points in the protocol,* image A.).
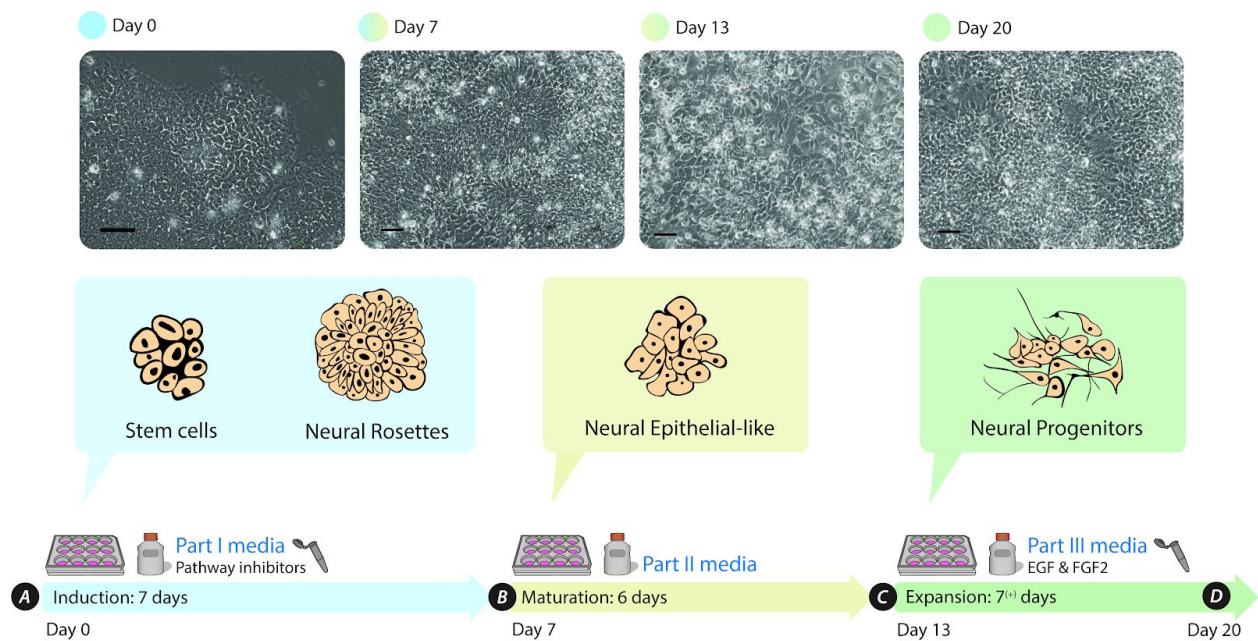
Figure 3 *Protocol overview*. At the start of the protocol period, d0, healthy hESCs were seeded as single cells in a set concentration per area. Medium was changed from hESCs media to neural induction medium together with small molecule inhibitors LDN-193189, SB431542, and XAV939 (abbv. LSX). LSX-mediated rapid neural induction initiates *Part I: Induction*, and at day 7 neural rosettes had appeared, with their distinct radial morphology, and these neural rosette cells were re-seeded as single cells, at the start of *Part II: Maturation*, where LSX inhibitors are no longer present. At day 13, cells had a more rectangular and plain appearance, and they were again single-cell seeded, for the final *Part III: Maturation*, with growth factors EGF and FGF2, added to the medium. At the end of this final part, day 20, we observed NPCs in densely populated wells with occasional gaps in the cell monolayer, traversed by dendrites.
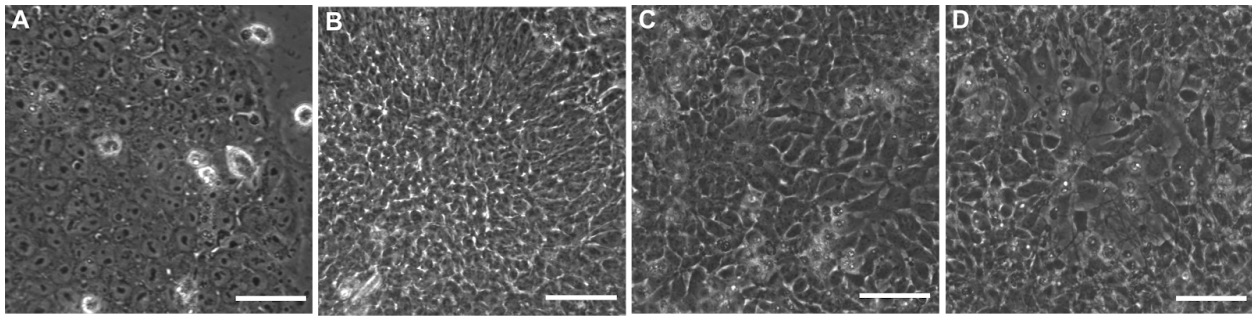
Figure 4 *Brightfield microscopy of time points in the protocol.* One representative image per time point (day 0, day 7, day 13, and day 20) is shown. A. A healthy hESC colony with small round cells that have relatively small cytoplasm. B. Day 7 radial formations, called rosettes, and are sometimes likened to neural tube formation *in vivo*. C. Day 13 cells that have a more rectangular, and plain shape, compared to day 7. D. Day 20 cells in a dense monolayer, with occasional gaps where traversing dendrites were observed. Images were taken at 40x magnification. Scalebar = 50 μm.

hESCs colonies were single-cell suspended using the cell detachment product *Accutase*. This is a solution of proteolytic and collagenolytic enzymes that works well with sensitive cell-lines. hESCs are more sensitive to the mechanical force produced by pipetting manually, therefore, extra time in Accutase is preferable to repeated pipetting. *Single-cells* in this context would be 1-10 cells, as cells tend to clump. The goal was to achieve a counting and a seeding density (cells/area) which was accurate every time an experiment was repeated while avoiding exaggerated mechanical force. More advanced cell counters, such as Countess, provide better cell number approximations. The incubation time required with Accutase will vary between cell-lines, as attachment is something that can be inherently different between different lines. For HS-360, 7 minutes in Accutase, at 37 °C, was used and worked equally well for other parts of the protocol. Temperature fluctuations when working with hESCs, or cells in differentiation, were avoided by using a hotplate, set to 37 °C, in the sterile hood.

Due to the inherent sensitivity of the cells, the pipetting technique might need some tuning. If too much force is used, the most sensitive cells will burst upon leaving the pipette tip, and viability will drop significantly, and surviving cells can lose their differentiation capacity. A gentler touch is required with a maximum of 10-12 resuspensions. A resuspension is the collection of detached cells from the well; a careful ejection against the well-wall, while tilting the plate slightly, so that

when the suspension is ejected at roughly 12' clock in the well, the cells flow downwards and help slightly to detach other cells. This results in a high viability cell suspension, that after a subsequent washing step, is easy to count. Described techniques can be trained on hESCs in advance to using the protocol, and cells can be observed under the cell-lab microscope for how they detach during Accutase treatment. To practice pipetting and measure viability, optimizing detachment and counting will be worth the effort. Again, accurate cell-counting and cell-viability is **crucial** when starting out in each part of the protocol. Going below the recommended cell density will either result in cells dying, or differentiation towards unwanted cell fates.

Day 0 to day 7. Part I: Induction

hESCs were resuspended in hESCs growth medium with added RHO/ROCK pathway inhibitor Y27632, which was added to avoid premature cell-death during the first 24 hours. We have tested several starting cell densities for neural induction that also should work in toxicology settings. An optimal starting cell density was determined to be 60 000 cells per well (12-well culture plate), and hESCs were seeded at this density 24 hours before (day -1) induction media was added (day 0). The 12-well plate is designed for 1ml media per well thus cell suspension was made such that we had 60 000 cells/ml which is convenient before seeding.

Omitting RHO/ROCK pathway inhibitor from the wash medium (or other intermediate medium) is an option, but it is good practice to include inhibitor in intermediate steps in the beginning as it prevents cell death. Cell viability should be between 80 % and 100 % for a successful start of part I. Sometimes the cell count is affected by aggregates in the counting dye, which when loaded on a counting-chip, is mistaken for dead cells by the machine. Keep counting dye, such as Trypan blue, out of sunlight, and centrifuge dye before use (take dye from the top of the vial).

Cells were evenly spread when seeded in wells. A good and consistent (i.e. repeatable) technique, was used to ensure this, moving the plate side to side which then was repeated in the

perpendicular axis, 4-5 repetitions per axis. This was also repeated when cells were placed on the shelf in the incubator.

Even walking around with a plate can cause cells to aggregate towards the well center, or spread unevenly in other ways. Incubators in heavy use should be avoided. Thus a dedicated incubator is recommended for hESCs (also to avoid infection risk when using antibiotic-free hESC-medium), and when using the protocol. Temperature fluctuations is another problem with incubators that are in heavy use. It is worth mentioning that old or unstable incubators sometimes vibrate a lot which causes cells to aggregate in the middle of a well before they have attached.

The day after seeding hESCs, cells growing in a pattern similar to what is shown in 'Figure 5 *Day 1 cell density and colony morphology*', were observed at 20x magnification (cells also had a 'spiky' appearance due to added Rock-inhibitor). If there were no similar formations, and only an sparse amount of cells, cells will not be efficiently induced, and an uneven spread will contribute to inefficient neural conversion.
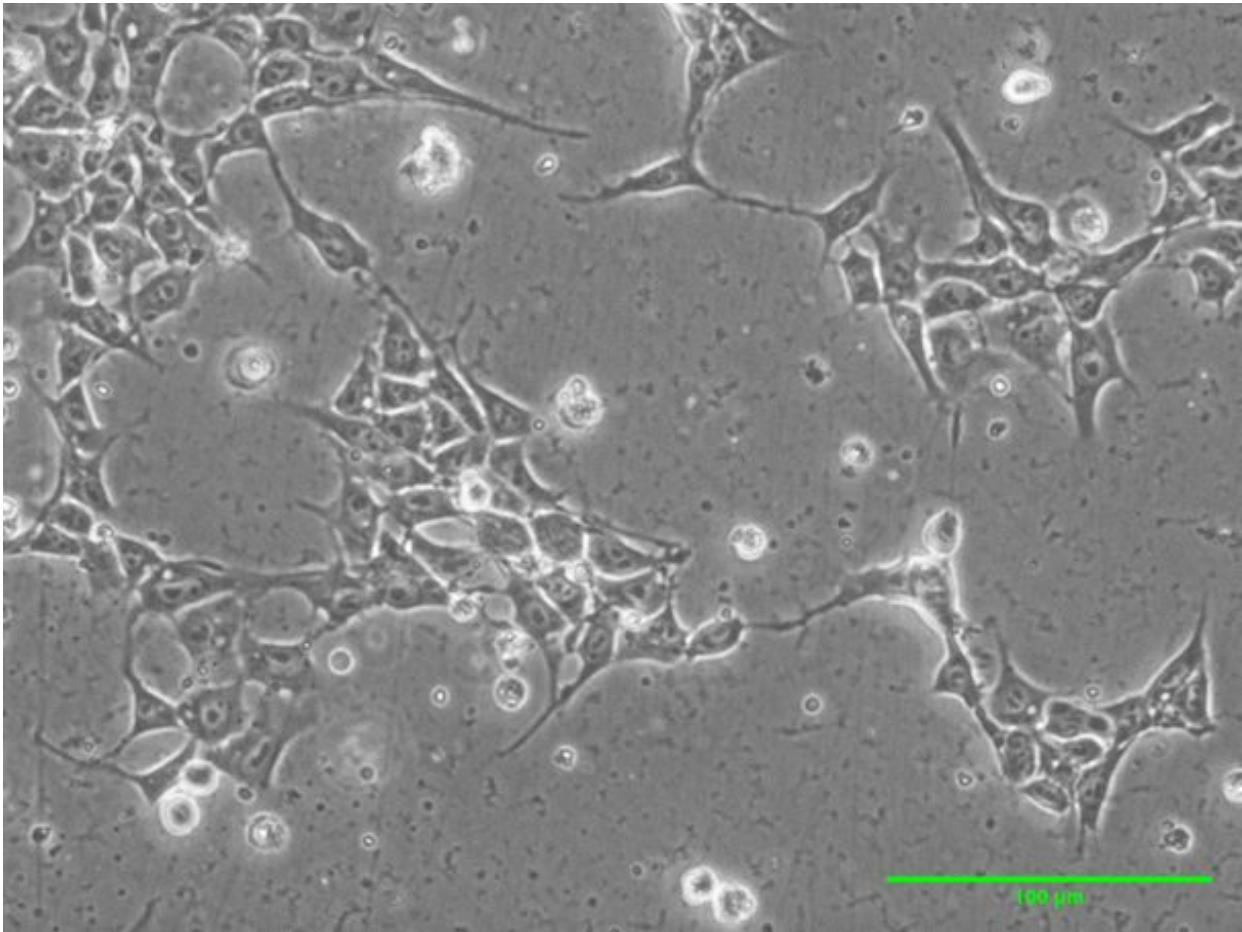
Figure 5 *Day 1 cell density and colony morphology.* At day 1, cells were evenly spread with similar morphology to what was shown in the image above. Cell morphology is affected by the RHO/ROCK pathway inhibitor making the cells somewhat more spiky in appearance. 20x magnification, scalebar = 100 μm.

24 hours after hESCs were seeded, the hESC growth medium with RHO/ROCK pathway inhibitor was removed and replaced with the part I induction medium (Table 1 *Medium composition for Part I: Induction*). If any drug testing was done, the compound to be tested was added to the medium before the addition to plates. In that way the compound was pre-mixed and local effects were avoided, since directly adding a compound creates a sudden locally high concentration of the compound.

Table 1 *Medium composition for Part I: Induction.*

Media	Component	Amount (500 ml)
Part I medium	Advanced DMEM/F12	485 ml
	GlutaMAX Supplement	5 ml (1%)
	Pen Strep	5 ml (1%)
	N2 Supplement	5 ml (1%)
Added fresh	SB431542	10 µM final concentration
	LDN-193189	100 nM final concentration
	XAV939	2 µM final concentration

The part I medium contained LSX inhibitors and a neural-specific supplement that contributed to a more optimal growth environment for neural induction, and the subsequent parts II and III. This supplement was N2, which in part II was combined with another neuro-optimizing supplement called B27, and this combination was used in part III as well, albeit with a modified B27 concentration.

N2 supplement is often seen in neurodifferentiation protocol medium formulations (Sünwoldt et al., 2017), and offers a good neuro-supporting environment, and is added to a basic DMEM-type of medium. N2 is said to accelerate neural commitment while increasing the survival and conversion of post-mitotic neurons.

Columnar orientation and radial patterning (radial patterning is also called rosettes, structurally akin to neural tube formation *in vivo*) were noticeable in the wells starting by day 4 to 5. By day 6 to 7, rosettes were clearly visible and confluency was high. Here, one should not expect a uniform layer, instead, there will be more differential cell structures in the colonies, and there is usually space between larger patches, and larger cells can be seen outside rosette-areas (Figure 6 *Day 4 cells*, and Figure 7 *Day 7 cells*).
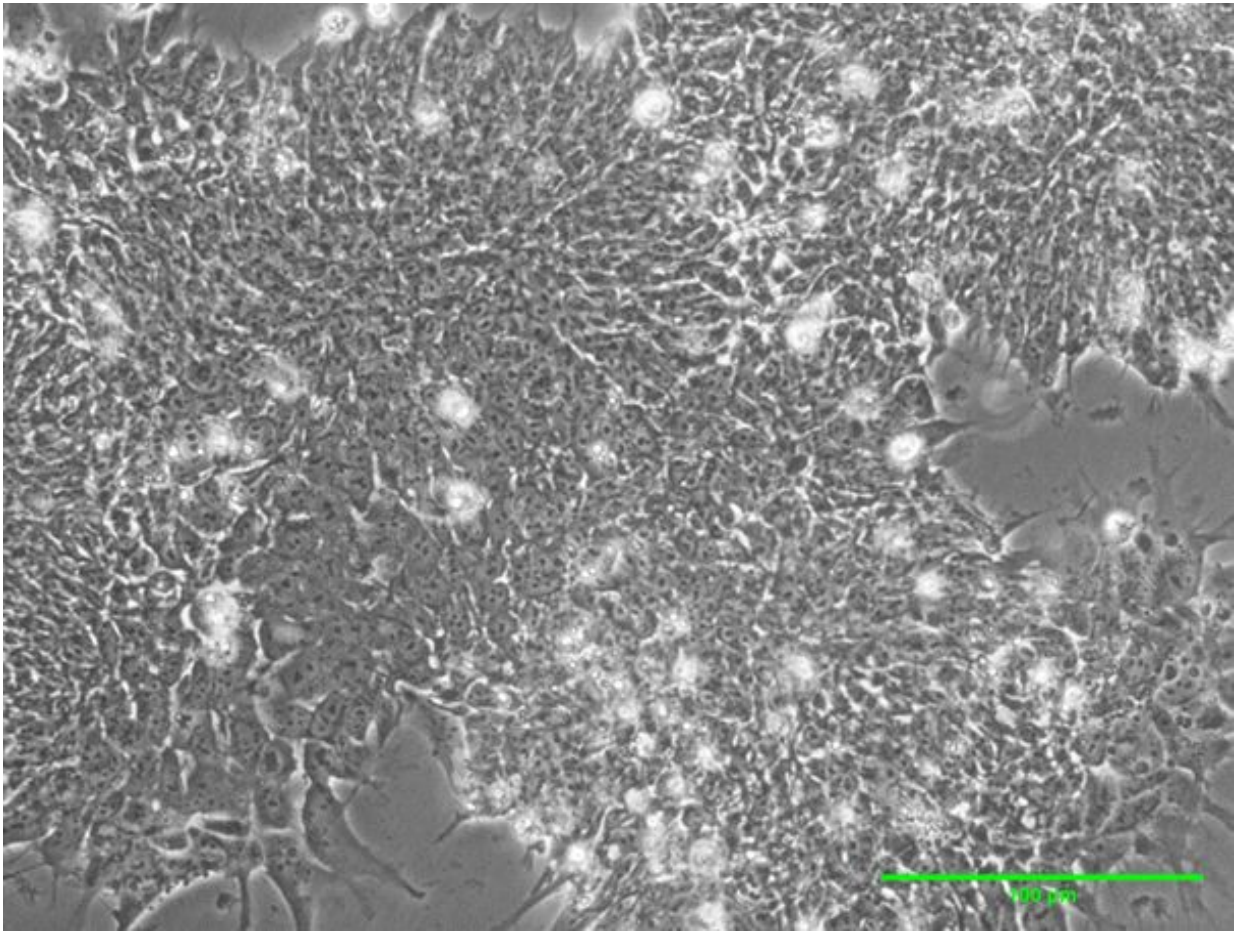
Figure 6 *Day 4* cells. At day 4 radial patterning was visible in the cell monolayer. 20x magnification and scalebar = 100 μm. A uniform layer is not to be expected at this time point. Outside denser rosette-like areas, larger cells were observed. Repeated empty areas between large patches of cells were common.
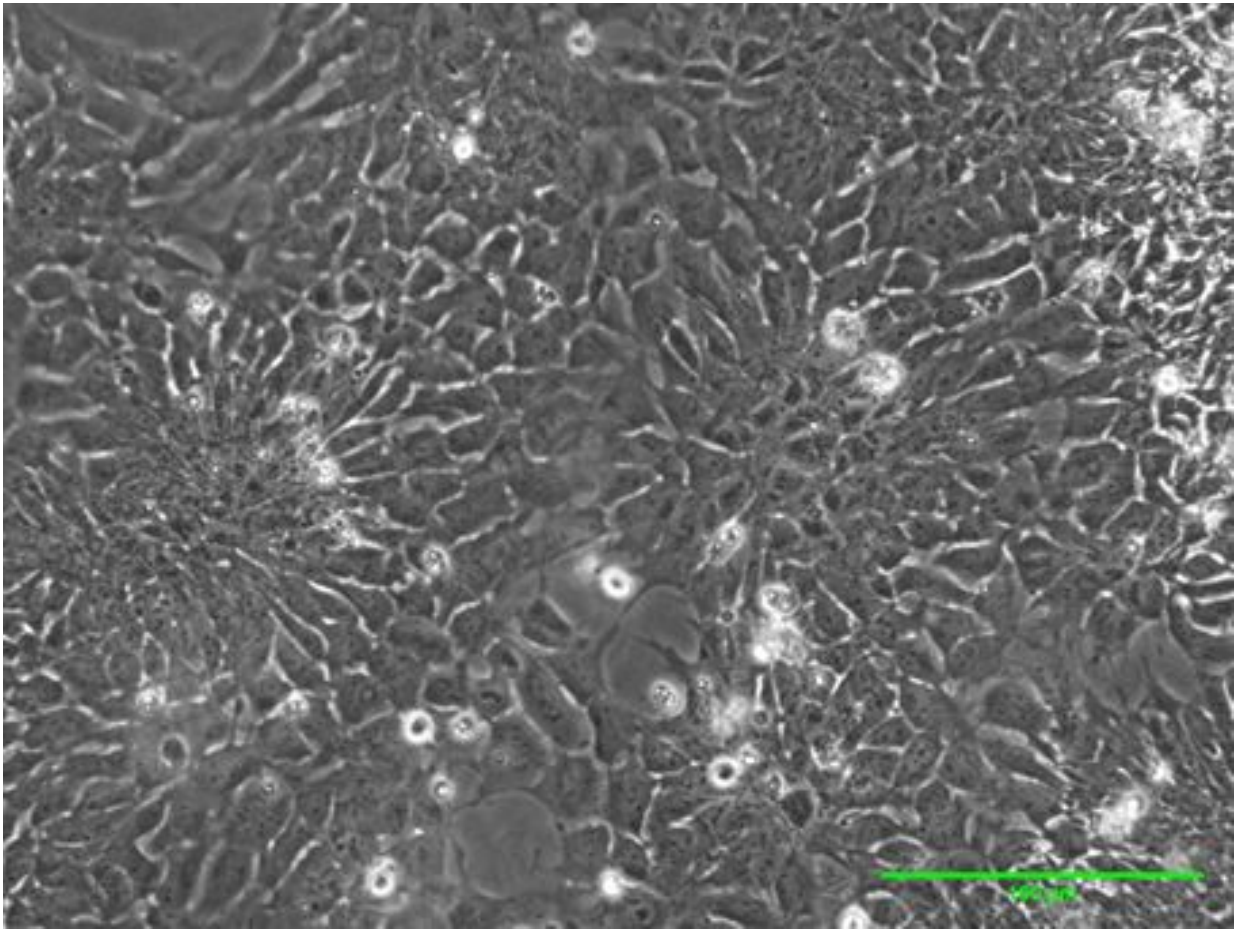
Figure 7 *Day 7* cells. On day 7 distinctive radial patterning (rosettes) were observed in the wells. This was a characteristic dense columnar arrangement of neuroectodermal precursor cells, a rosette formation, and was seen in large patches with some space in between and large cells on the outer rims. 20x magnification and scalebar = 100 μm.

Day 7 to Day 13. Part II: Maturation

Part II: Maturation occurred during a 6 day period after cells had been seeded at high density. 450 000 cells per well were seeded per well, in the 12-well plate format, compared to 60 000 cells per well at the start of the part I. Previously, cells were grown on well surfaces coated with Geltrex. Geltrex was here combined with polyornithine and fibronectin to create a more neural-supportive environment. Polyornithine and fibronectin were added to the wells and incubated for 2-3 hours at 37 °C. The left-over coating solution was then removed, and Geltrex added. At this point, the plate would either be stored at 4 °C (wrapped in parafilm to avoid the

plate drying out), or used directly after an hour of incubation at 37 °C. Plates that were taken out of storage were incubated similarly, before use. Because of the long preparation time, it is recommended to prepare plates the day before. Cells were counted, assayed for viability, and split, by the same techniques as described in part I of the protocol.

B27, another popular neuro-supportive supplement was combined with N2. B27 promotes neuronal cell long-term survival, rate of growth, and attachment (Table 2 *Medium composition for Part II: Maturation*). Reaching the end of part II, closing in on day 13, we observed very confluent cell monolayers (Figure 8 *Day 13 cells*).
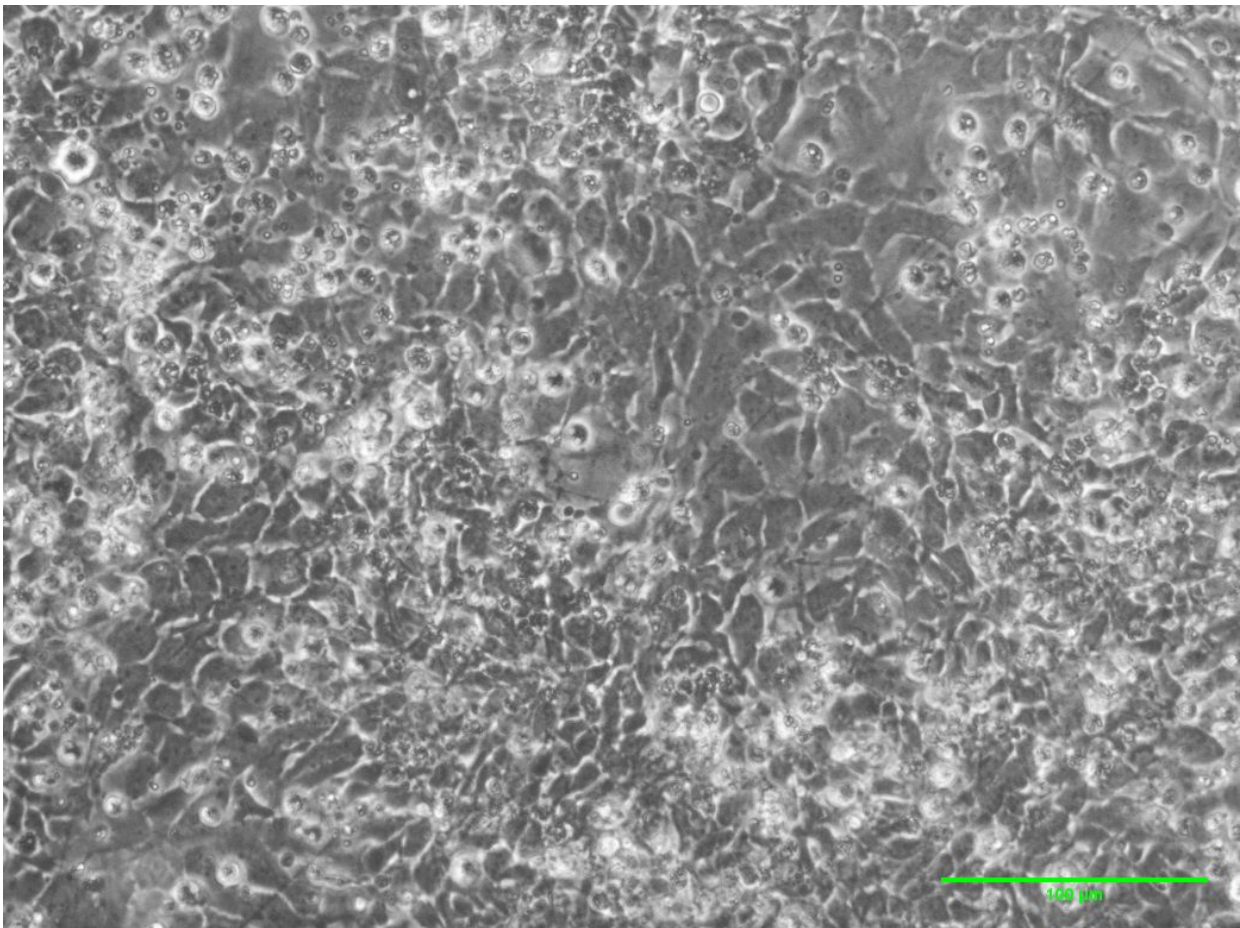
Figure 8 *Day 13 cells*. A typical well area at day 13. Cells were observed to be more rectangular compared to day 7. More debris was also observed, likely due to very confluent wells. 20x magnification. Scalebar = 100 μm.

Table 2 *Medium composition for Part II: Maturation.*

Media	*Component*	*Amount (500 ml)*
Part II medium	Advanced DMEM/F12	480 ml
	GlutaMAX Supplement	5 ml (1%)
	Pen Strep	5 ml (1%)
	N2 Supplement	5 ml (1%)
	B27 Supplement	5 ml (1%)

Day 13 to Day 20. Part III: Expansion

The final phase, part III: expansion, started on day 13. This phase lasted for 7 days but it is likely that the cells at day 20 can be expanded to generate mature neural progeny, or for expanding NPCs. For part III, cells are processed using the same procedures as described for parts I and II, and the cell density is the same as for part II, 450 000 cells per well.

Usually, there were enough cells per well at the end of part II to enable two new wells, using a 1:2 passage ratio, in part III. At day 20 we observed an intricate dendritic-like mesh between gaps in the thick cell layer (Figure 9 *Day 20 NPCs*). Growth factors FGF2 and EGF were introduced and added daily during part III (Table 3 *Medium composition for Part III: Expansion*). Day 20 cells have been maintained for 4-5 subsequent passages. We have also frozen, thawn, and expanded cells at all protocol timepoints. For further passages after day 20 we recommend a density of 450 000 cells per well to be maintained for the first 4-5 passages, then to adjust the cell number if needed.
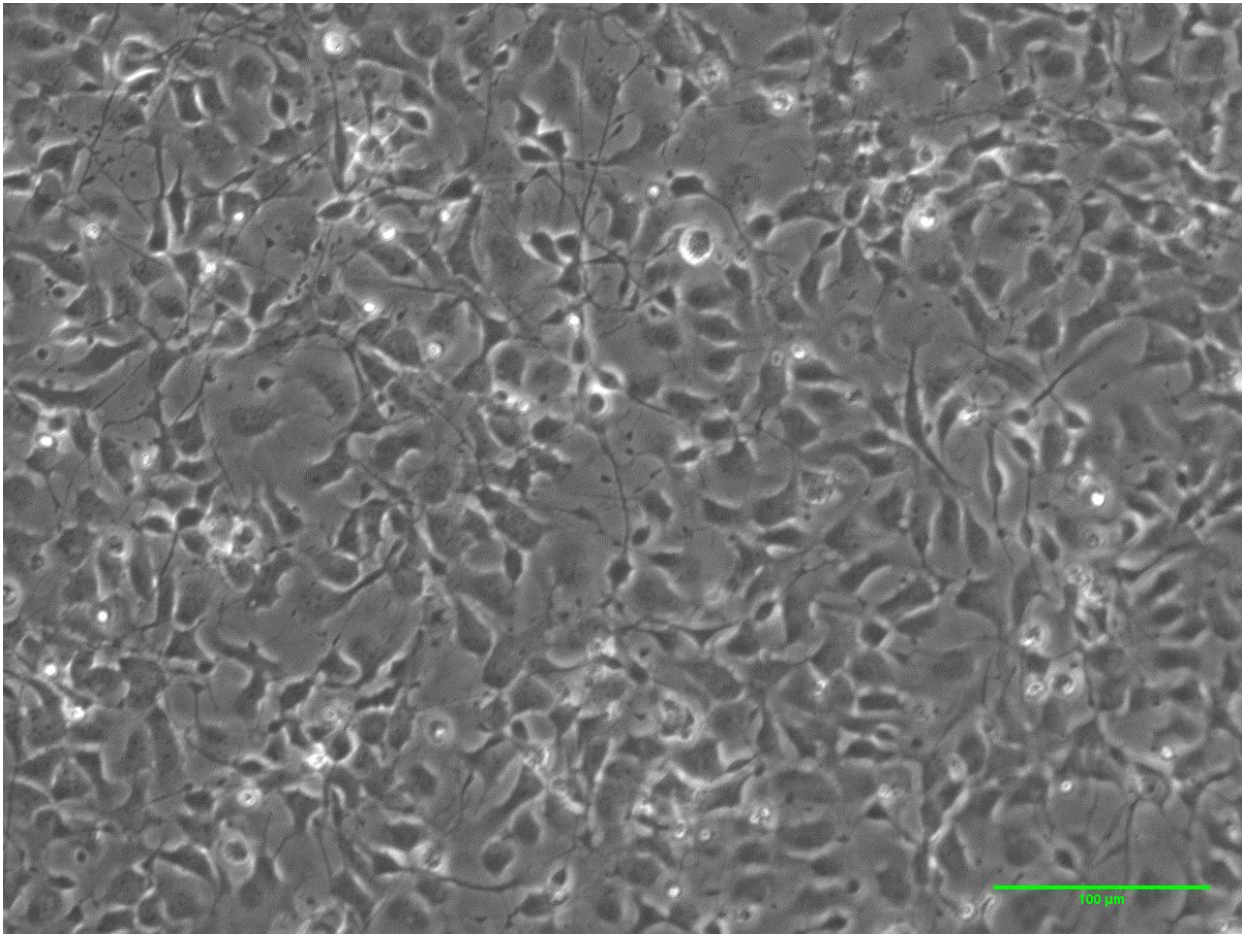
Figure 9 *Day 20 NPCs*. A homogenous culture with cells sharing similar morphology was seen at day 20. The image shown of day 20 cells was chosen to show one area that was slightly less populated to reveal the morphology and cell outgrowths. 20x magnification. Scalebar = 100 μm.

Table 3 *Medium composition for Part III: Expansion*

Media	Component	Amount (500 ml)
Part III medium	Advanced DMEM/F12	487.5 ml
	GlutaMAX Supplement	5 ml (1%)
	Pen Strep	5 ml (1%)
	N2 Supplement	5 ml (1%)
	B27 Supplement	2.5 ml (0.5%)
Added fresh	Human bFGF2	10 ng/ml final concentration
	Human EGF	10 ng/ml final concentration

ICC/IF on selected markers

To visualize how our cells were progressing we performed ICC with antibodies targetting common markers of pluripotency and neurodifferentiation. The markers investigated were not particularly difficult to stain (Table 11 in methods contains the list of primary and secondary antibodies used) and when looking for presence or non-presence of markers, a simple microscope capable of fluorescence will do.

Pluripotency panel

To investigate pluripotency, well-known transcription factors (TFs) and filament markers were used. Our pluripotency panel (Figure 10 *Pluripotency panel*) consisted of TFs *SOX2*, *OCT4*, and the intermediate filament protein Nestin. *SOX2* is a pluripotency factor that was present during the whole protocol and plays an important role in differentiation towards NPCs (Shuchen Zhang et al., 2019). *OCT4* forms a heterodimer with *SOX2* and these proteins work together to activate pluripotency genes in hESCs, but *OCT4* is stem cell-specific, where the levels drop rapidly after day 0. Thus, *OCT4* is not detected at day 7. Nestin is a filament protein protein used as a marker to follow neural fate decisions, however, Nestin is also present in hESCs, although at lower levels, where the filaments are organized slightly differently when compared to cells at day 20. As our pluripotency panel shows, *OCT4* is present on day 0 with a uniform expression in all cells, but is not detected on day 7, and remains absent at all subsequent time points. *SOX2* has a variable expression as is expected, but is present at all time points as an important pluripotency factor. Nestin is present from day 0 to day 20, but seems to vary in filament organization, depending on timepoint.
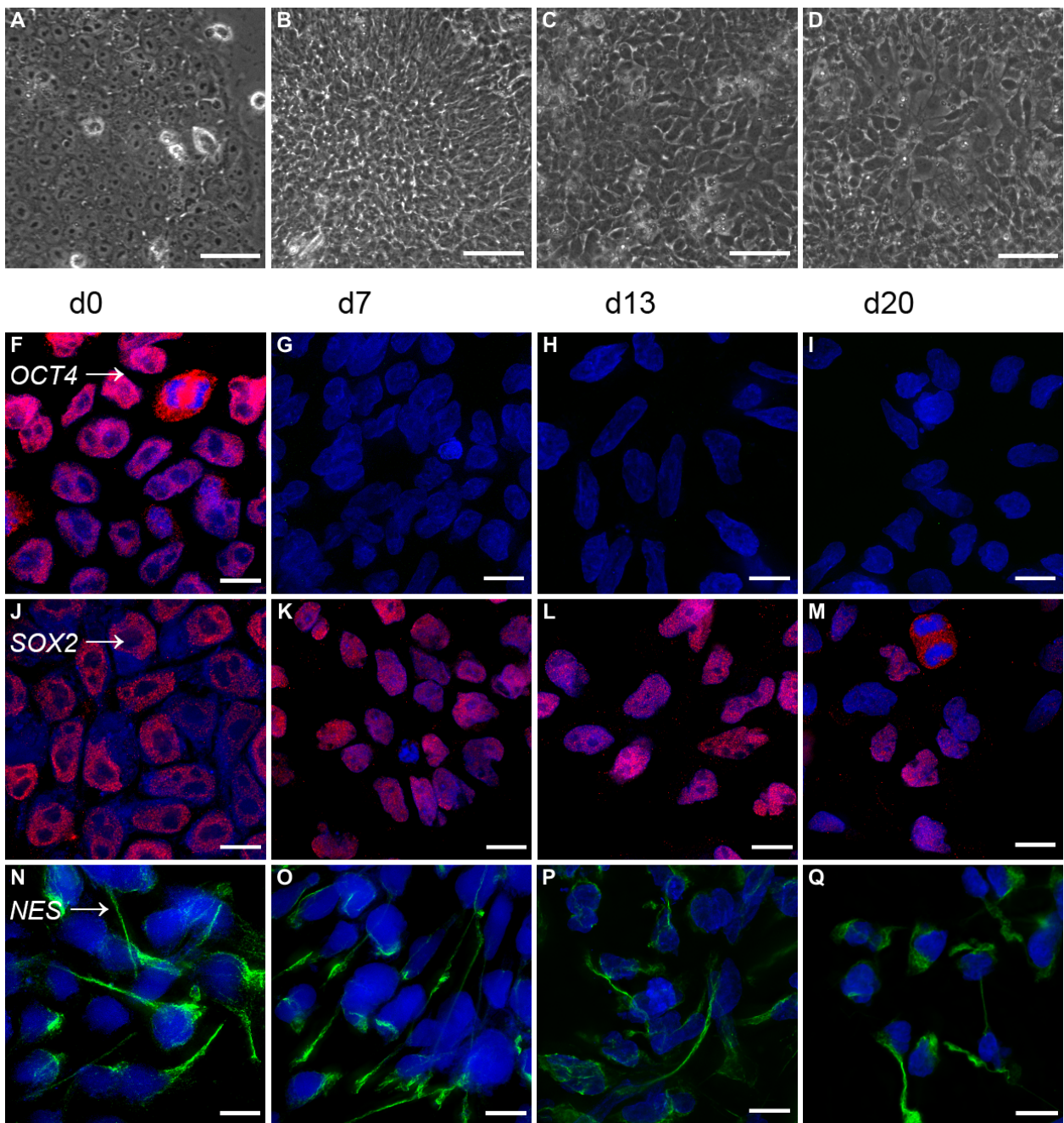
Figure 10 *Pluripotency panel*. Based on pluripotency factors *OCT4* and *SOX2*, and filament protein Nestin. The top-row, A to D, shows brightfield images of live cells at day 0, 7, 13, and 20 (from left to right). Subsequent rows are immunofluorescence images of fixed cells at days 0, 7, 13, and 20 (from left to right). F to I shows pluripotency factor *OCT4*/Pou5f1. J to M shows pluripotency- and proliferation factor *SOX2*. N to Q shows filament protein Nestin. Brightfield images; 40x magnification, scalebar = 100μm. Fluorescent images; 100x magnification. Scalebar = 10 μm.

Neurodifferentiation panel with TFs *PAX6*, *OTX2*, and filament protein βIII-tubulin

The neurodifferentiation panel consisted of TFs *PAX6* and *OTX2* and the filament and microtubule element βIII-tubulin. *PAX6* is a key neural TF in early neurodevelopment (Shuchen Zhang et al., 2019). *OTX2* is a head organizer TF and patterning factor, and is involved in regional patterning of the forebrain and midbrain which makes it critical in gastrulation and early neural development and has been shown to be an important factor in the development of the medial ganglionic eminence and septum (transitory structure in early brain development) (Hoch, Lindtner, et al., 2015). βIII-tubulin, also known as Tuj-1, is almost exclusively found in the neuronal lineage and correlates with early neurogenesis, but has been observed in some cancers and in the testis (Cicchillitti et al., 2008; Person et al., 2017). As shown in our neurodifferentiation panel (Figure 11 *Neurodifferentiation panel*), starting at day 7, we had an expression of βIII-tubulin at all three depicted timepoints (A to C), which we expected for early neurogenesis. The βIII-tubulin antibody was sadly not used on day 0 cells (due to limited supply of cell material, to cover several other experiments at that point), but a report suggested that βIII-tubulin might be expressed in stem cells normally, thus what could be of interest here would be an expression that increased at later stages of neurodifferentiation (Foudah et al., 2014; Garza-Manero et al., 2019). Nonetheless, βIII-tubulin should be antibody-stained at day 0 for completion and comparison, in the next ICC batch together with *OTX2* and *PAX6*, to show their status at day 0. Forebrain-midbrain patterning factor *OTX2* was also visible (D to F), but somewhat differently expressed on day 20 where it seemed to be present in both cells cytoplasm and nucelus. *PAX6* was measured the strongest at day 7, and decreased to be relatively low at day 13, and detected only in a few cells at day 20.
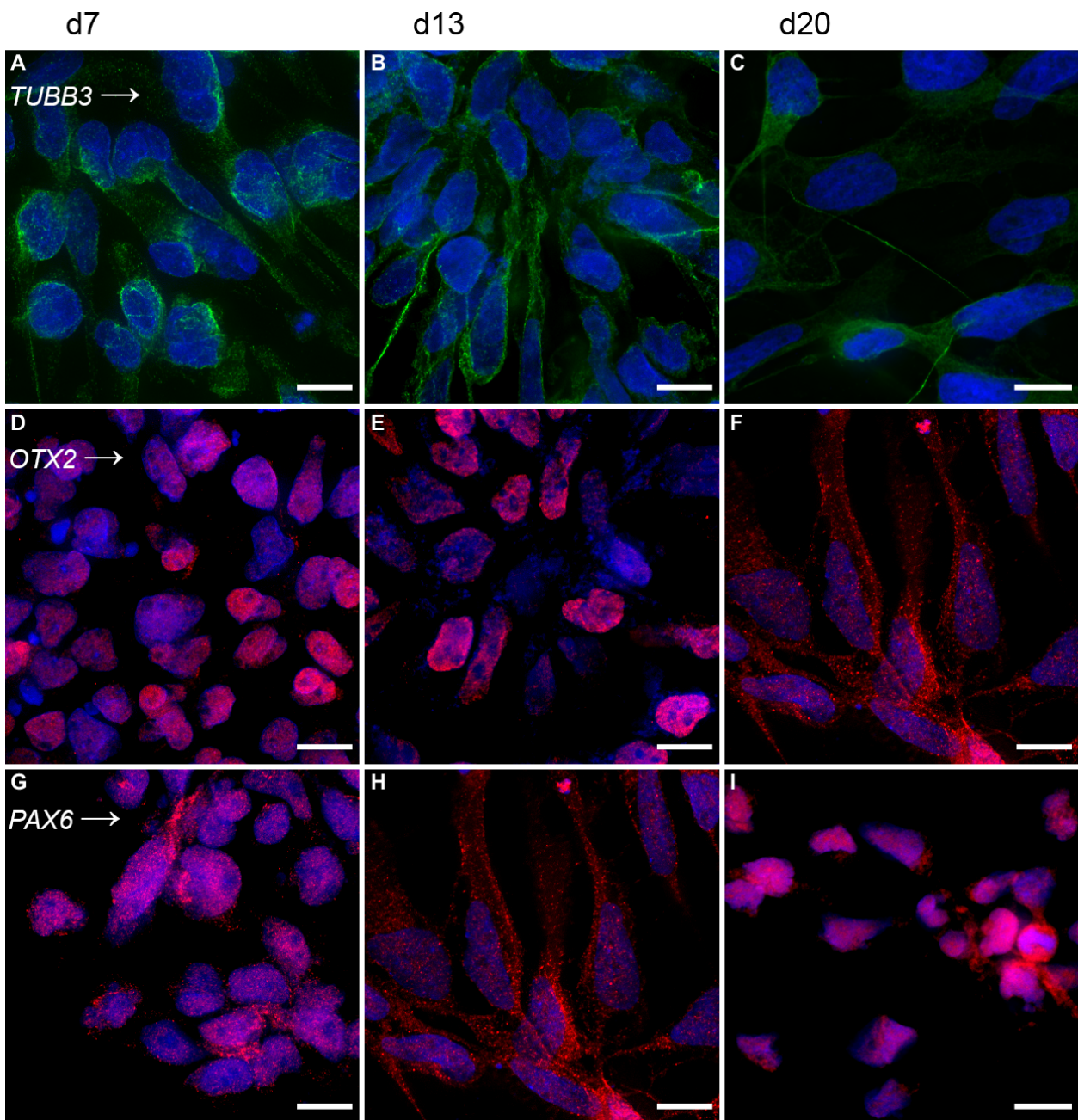
Figure 11 *Neurodifferentiation panel*. Based on neuro-markers βIII-tubulin (*TUBB3*), *OTX2,* and *PAX6*. Starting from the left, images show fixed cells at day 7, 13, and 20. Top-row (A to C) represents the microtubule element βIII-tubulin, the second-row (D to F) forebrain-midbrain patterning factor *OTX2*, and the third-row (G to I), neurogenesis factor *PAX6*. 100x magnification. Scalebar = 10 μm.

ddPCR of markers expressed in cells harvested at day 0, day 7, day 13, and day 20

Droplet digital PCR was run on samples from day 0, 7, 13, and 20. Markers targetted were *OCT4*, *NANOG*, *NES*, *VIM*, *PAX6*, *OTX2*, *FOXG1*, *TUBB3*, and *MAP2*. The day 0 samples included 5 replicates, day 7 included 6 replicates, day 13 included 4 replicates, and day 20 included 4 replicates. *OCT4*, *NES*, and *OTX2* have been described previously (*ICC/IF on selected markers*). ddPCR analysis expands upon the marker list used for ICC/IF. *NANOG* is a classical stem cell marker, similarly to *OCT4,* but with a more variable expression (Y. Liu et al., 2006). *VIM*, vimentin, is an intermediate filament protein involved in a wide spectrum of cellular functions, and is present in many cells where it assembles into filament structures across the cytoplasm (Danielsson et al., 2018). *FOXG1*, Forkhead box G1, has a critical role in forebrain development (Hettige & Ernst, 2019) and *MAP2* is a common marker for neurons and neuronal differentiation (Soltani et al., 2005).

As observed for day 0 samples (Figure 12 *Day 0 ddPCR of selected markers*), there was a low level of *OTX2* expression. *OTX2* has been shown to be required for ESC transition into epiblast derived stem cells (EpiSCs). The EpiSC-state has been described as a more true stem cell state for hECSs as these cells have been suggested to be more alike the primed state of the postimplantation epiblast than mESCs, which more resemble the naïve ground state of the preimplantation epiblast (Acampora et al., 2013; Xue et al., 2011). Therefore the measured presence of *OTX2* on day 0 in the hESCs was not that surprising. We also observed a strong expression of *OCT4*. We observed expression of *NANOG*, *NES*, *VIM*, and *TUBB3,* but not expression of the neuronal markers *PAX6*, *FOXG1*, and *MAP2*.
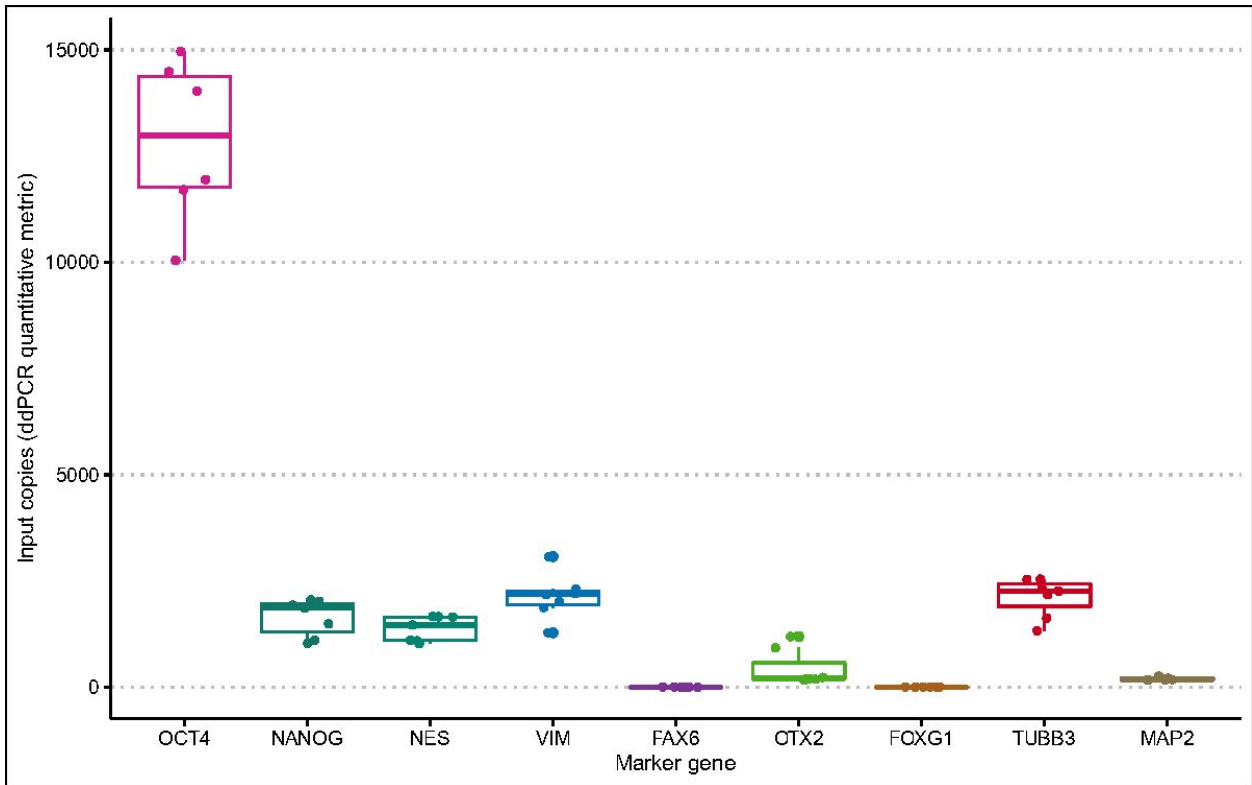
Figure 12 *Day 0 ddPCR of selected markers*. At day 0 stem cell marker *OCT4* expression was high. Further, expression of *NANOG*, *NES*, *VIM*, and *TUBB3* was detected, but not expression of *PAX6*, *FOXG1*, or *MAP2*. A slight expression of *OTX2* was seen, which might be surprising, but *OTX2* has been shown to be required in the epiblast (EpiSCs) transition of embryonic cells. hESCs have been suggested to be more like EpiSCs, than like preimplantation naïve, as in mouse ESCs. *Input copies* on the y-axis are the quantitative metric for ddPCR.

A week into neural induction on day 7 (Figure 13 *Day 7 ddPCR of selected markers*) we observed the expected loss of expression for stem cell markers *OCT4* and *NANOG*. We measured a higher expression of filament genes *NES* and *VIM*. Expression of important neuronal markers *PAX6*, *FOXG1*, and *MAP2*, was detected. We also observed *OTX2* expression, rising significantly compared to on day 0.
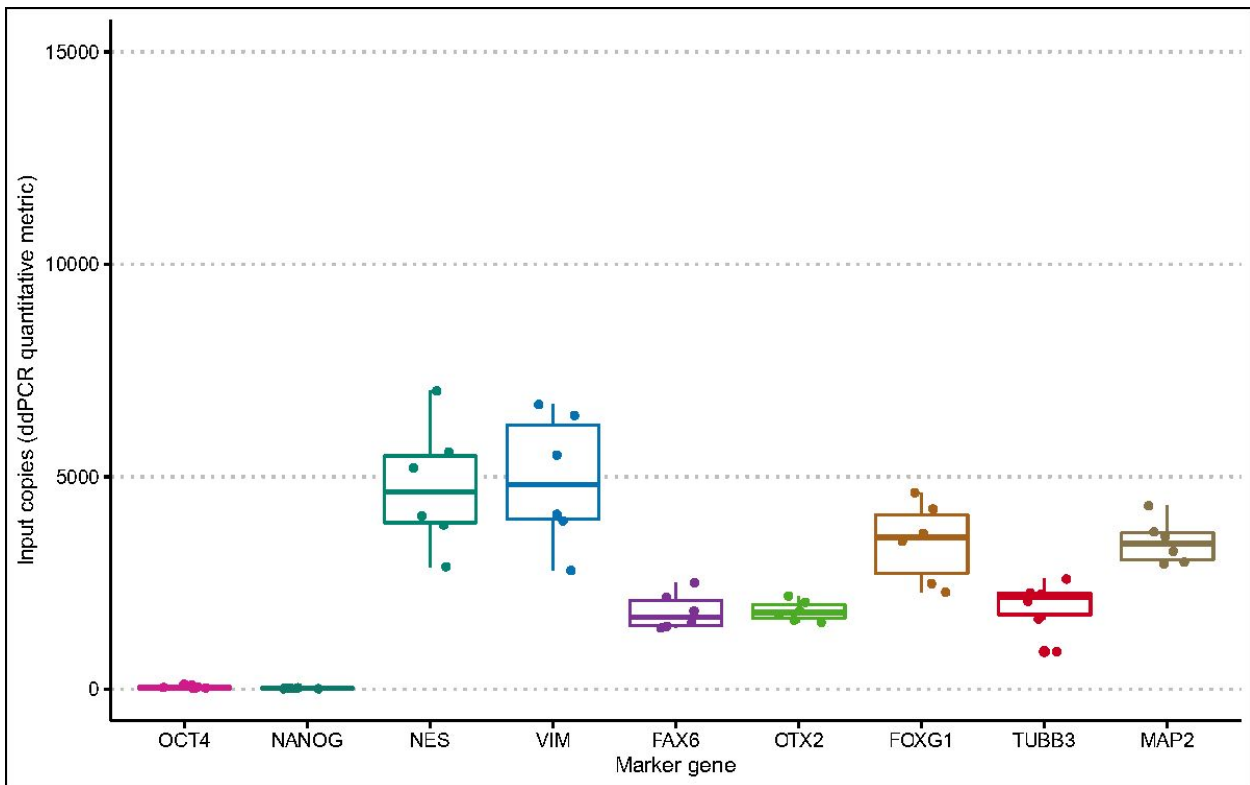
Figure 13 Day 7 *ddPCR of selected markers*. At day 7, we observed complete loss of expression of the stem cell markers *OCT4* and *NANOG*. Filament markers *NES* and *VIM* had increased in expression and the expression of *PAX6*, *FOXG1*, and *MAP2* neuronal markers was evident. *OTX2* had increased significantly in expression compared to on day 0. *Input copies* on the y-axis are the quantitative metric for ddPCR.

After neural induction, at the end of part II: maturation part, *PAX6* expression dropped significantly, compared to on day 7, which could be an indication of the development of particular structures in the forebrain, as *PAX6* is required in the patterning of the forebrain, and thus is present at much lower levels for certain regions (Figure 14 *Day 13 ddPCR of selected markers*) (Holm et al., 2007; L. Jones et al., 2002; Parish et al., 2016; Quintana-Urzainqui et al., 2018). *VIM* and *FOXG1* expression were approximately doubled in comparison to on day 7, indicating a strong forebrain-fate commitment. *TUBB3* was expressed at a significantly higher level compared to on day 7. *MAP2* seemed to be expressed at a similar level as on day 7.
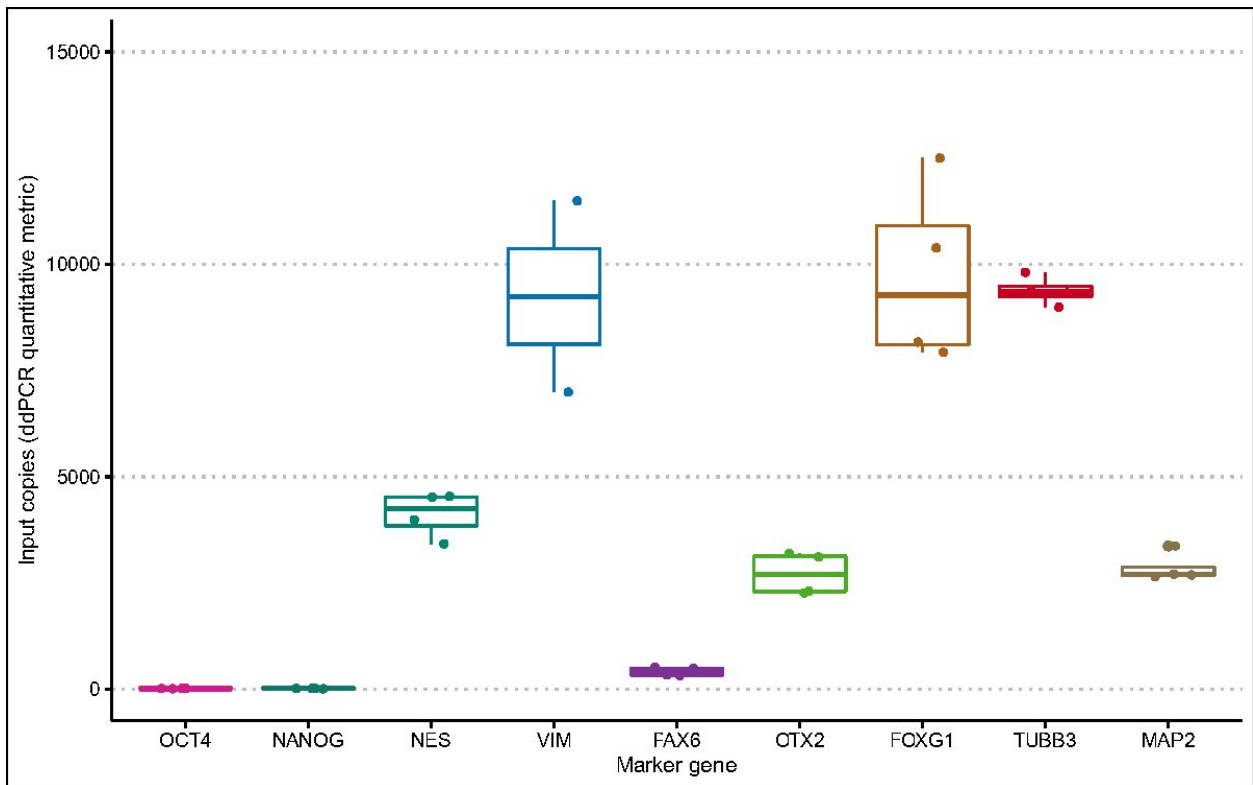
Figure 14 Day 13 *ddPCR of selected markers*. *PAX6* expression was significantly reduced at day 13, compared to on day 7. *VIM* expression approximately doubled, while *FOXG1* expression had more than doubled since on day 7, showing forebrain-fate commitment. *TUBB3* increased significantly in expression, compared to on day 7. *MAP2* seemed to maintain the same level of expression as on day 7. *Input copies* on the y-axis are the quantitative metric for ddPCR.

At the end of part III: expansion (Figure 15 Day 20 *ddPCR of selected markers*), filament markers *VIM* and *TUBB3* increased in expression compared to on day 13. The mature neuronal marker *MAP2* increased markedly in expression when compared to on day 13. A slight increase in the expression of forebrain-midbrain patterning factor *OTX2* was observed, while *NES* and *FOXG1* seemed to be expressed at comparable levels to on day 13.
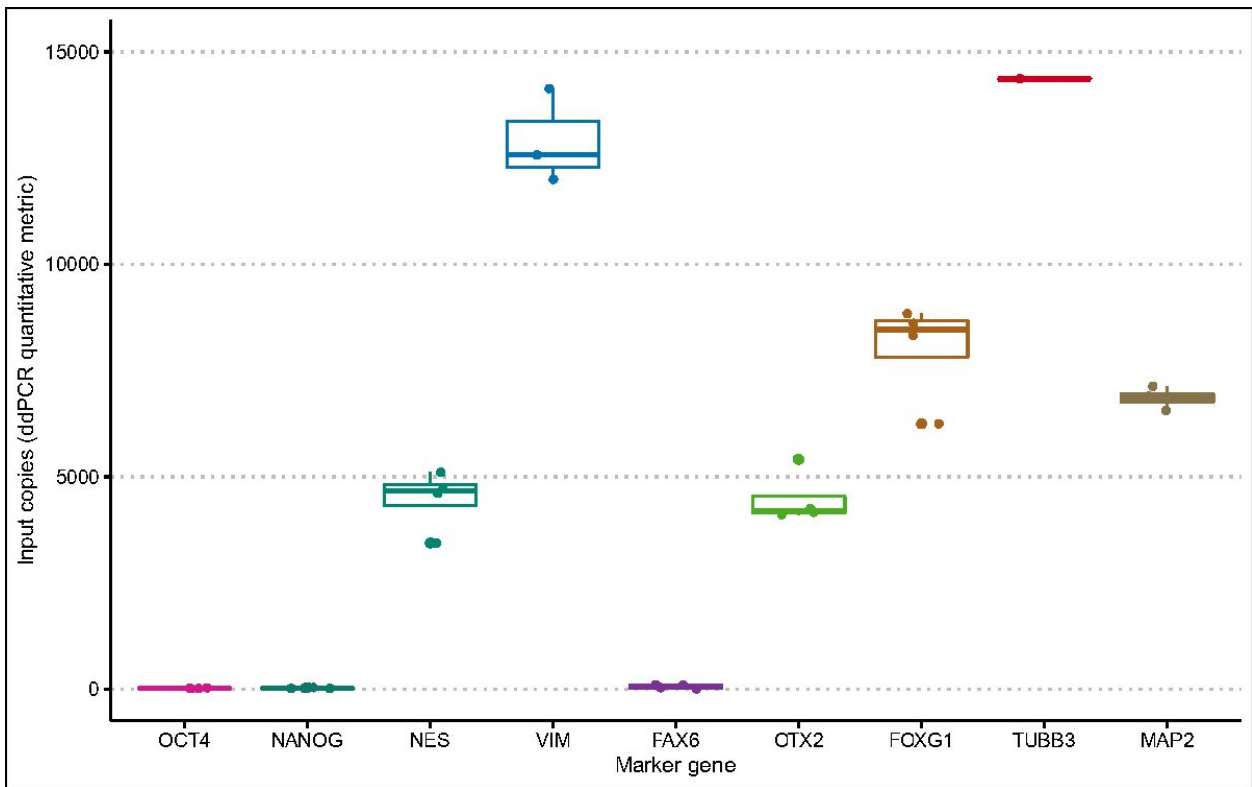
Figure 15 Day 20 *ddPCR of selected markers*. *VIM*, *TUBB3*, and *MAP2* increased markedly in expression compared to on day 13. *OTX2* increased slightly in expression, *NES* and *FOXG1* seemed to be expressed at comparable levels to on day 13. *Input copies* on the y-axis are the quantitative metric for ddPCR.
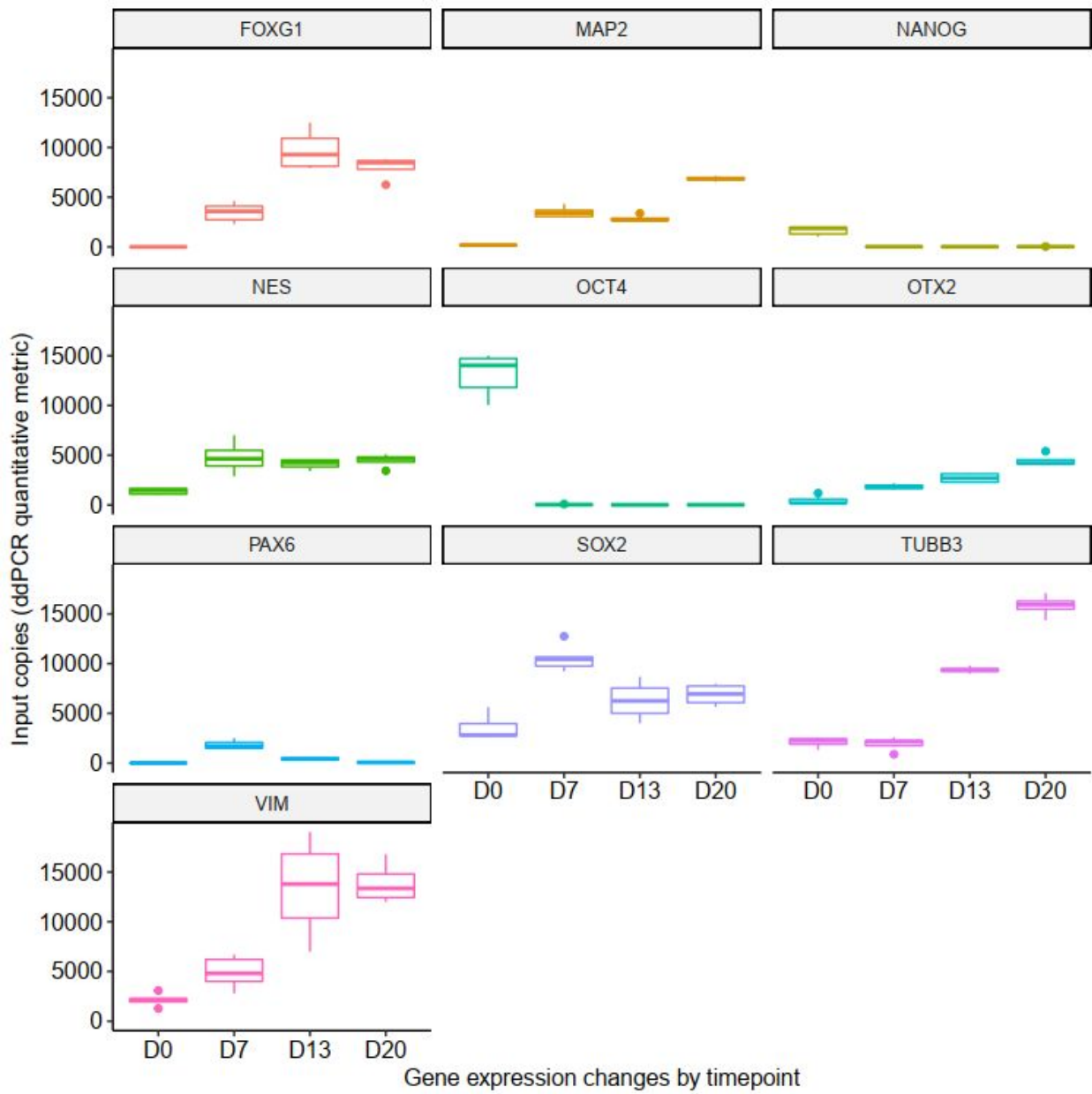
Figure 11-15B Changes in gene expression, all timepoints. Summary figure for Figure 11 to 15 plotting each marker in its own facet with marker gene expression in order of day 0 (D0), day 7 (D7), day 13 (D13), and day 20 (D20).

Single-cell whole-genome RNA sequencing

Brief introduction on figure types

For single-cell plots, individual cells are often plotted as a dot. Hundreds, or thousands of cells, are therefor plotted at the same time in order to convey a figure. Cells are placed on the plot based on underlying data in order of how much similar/dissimilar they are to each other. However this data is composed of a large number of parameters and those are often placed in order of contribution to the observed variation, in components called principal components (PCs). There are various techniques used to reduce multi-dimensional data to create a more readable 2D-representation while trying to preserve the local structure. Commonly used techniques in a dimensional reduction for single-cell data are t-distributed stochastic neighbor embedding (t-SNE) (Kobak & Berens, 2019; Maaten & Hinton, 2008) and uniform manifold approximation and projection (UMAP) (Becht et al., 2018; McInnes et al., 2018). In the thesis, UMAP is primarily used and it also seems to have become a more popular choice in recent single-cell sequencing studies (Becht et al., 2018; Luecken & Theis, 2019).

UMAP plots with expression levels of individual cells

Many popular single-cell analysis packages, like Seurat (Butler et al., 2018), offer ways to plot expression of a gene in cells while having all the cells represented as a UMAP or tSNE reduced dimension plot. Sometimes the order in which points are plotted on these plots can affect how the expression is displayed as these points on these plots can overlay each other in unexpected ways. By using hexagonal cell representation of single-cell data the cells can be binned into tiny hexagons in which the mean expression is plotted thus avoiding this potential problem. This is done with an R-tool called Schex (Freytag & Lister, 2020). Therefor Schex was used when possible. Plotting methods, representation, metrics, etc. does not seem to have a clear standard, as of yet, in single-cell studies.

Determining clusters without prior cell-type knowledge, or annotation reference, using R-tool scClustViz

An effective tool to visualize multiple single-cell clustering results is scClustViz (Innes & Bader, 2018) from Bader Lab at the University of Toronto. scClustViz is an interactive tool that will help visually aid in selecting an appropriate cluster amount by showing differentially expressed genes between clusters and a silhouette-type of a plot, where a good cluster has a positive silhouette value. The tool is particularly useful when characterizing heterogeneity since different cell-types or cell-states are expected to have differentially expressed genes and thus and this tool elegantly displays this metric in a way to facilitates picking a particular cluster amount. While there is a multitude of software solutions to do more complex analysis for all of the single parts that this tool provides a solution too, but the fact that the ties together this complexity into one interactive tool makes it very useful and a good way to explore and investigate clusters. Several built-in functions allow a researcher to compare clusters, do simple annotation, explore the varying numbers of clusters, visualize data, metadata, and quality control metrics on reduced dimension plots such as UMAP and t-SNE. Once more knowledge on the underlying biology is gained, as in this case: what cell-types he have and wherein forebrain development we are, the number of clusters can be increased or reduced, to fit a biological picture better.

Day 0: hESCs

1647 cells were used in the analysis of day 0 after quality control and filtering (SI Day 0 Figure I and SI Day 0 Figure II). The number of clusters for day 0 was initially set to 3 after an inspection in scClustViz (Figure 16 *Day 0 clustering and differentiation prediction* A.).

Movement of differentiation

CytoTRACE differentiation prediction indicated several starting points of differentiation, outer areas of cluster 1 and 0, the more differentiated area was predicted as largely cluster 2 (Figure 16 *Day 0 clustering and differentiation prediction* B.). Based on how the prediction plot for day 0 compares to subsequent time points (Figure 19 *Day 7 clustering and differentiation prediction*,

Figure 23. *Day 13 clustering, and differentiation prediction*, and Figure 26 *Day 20 differentiation prediction and end-of-differentiation markers*), and the integration of run A and run B (SI Day 0 Figure II C.), the day 0 timepoint is much more homogeneous where prediction and separation of clusters are resolved upon differences between cell-states in a similar cell-type, hESC, rather than several cell-types.

Predicted start-of-differentiation-markers are *JPT1*, *TOP2A*, *UBE2C*, *TPX2*, *KIF23*, *PIF1*, *AURKA*, *CKS2*, *BUB3*, and *CENPF*. Information on *JPT1* and hESCs are hard to find but GeneCards (genecards.org) describes apoptosis modulation and signaling as a related pathway and further mentions *JPT1* playing a role in cell adhesion and regulation of cell-cycle. The topoisomerase enzyme variant *TOP2A* is expressed uniquely in human pluripotent stem cells and has been reported to be downregulated in differentiation (Ben-David et al., 2015). *UBE2C* has been shown to decrease in differentiation towards NPCs and is associated with cell-cycle control and proliferation (Fatima et al., 2020; Re et al., 2014). *TPX2* is implicated in cell-cycle and expressed in proliferating cells in G1/S transition (Kufer et al., 2002) and is seen dysregulated in tumors to promote proliferation, metastasis, and tumorigenicity (Lei et al., 2016; F. Li et al., 2019; Yan et al., 2013). *KIF23* is involved in mitosis and cell-cycle and has been shown downregulated upon hematopoietic differentiation of hESCs (P. Li et al., 2018). *PIF1* is a helicase that prevents genome instability and is involved in telomere maintenance (Paeschke et al., 2013). The depletion of *AURKA* (aurora kinase A) has been shown to result in differentiation and affected self-renewal (D.-F. Lee et al., 2012). *CKS2* overexpression promotes cell proliferation and is associated with poor breast cancer prognosis and is essential during early embryogenesis (N. Huang et al., 2019; Martinsson-Ahlzén et al., 2008), *BUB3* is part of the SAC-machinery and has a spindle-assembly checkpoint function and helps with kinetochore-microtubule attachment (Toledo et al., 2014). The centrosomal protein *CENPF* is another cell-cycle associated gene involved in the segregation of chromosomes during cell division (Rattner et al., 1993; Jiaxu Wang et al., 2017).

The predicted end-of-differentiation markers were *DDIT4*, *CRABP2*, *USP9X*, *UGP2*, *BNIP3*, *LITAF*, *AC106864.1*, *PODXL*, *APELA*, and *IRX2*. *DDIT4* (DNA damage-inducible transcript 4) encodes a mTORC1 (mammalian target of rapamycin complex 1) inhibitor which regulates cell growth, survival, and proliferation and can respond to change in energy levels and stress (Baser et al., 2019; Sofer et al., 2005). *CRABP2* has been identified as a possible marker of primed or naive human embryonic stem cells (Xu et al., 2016). *USP9X* is a stem-marker and has been shown essential for TGFβ-mediated embryonic development (Dupont et al., 2009). *UGP2* is a metabolic gene, according to GeneCards, and encodes for UDP-glucose pyrophosphorylase 2. *BNIP3* is involved in autophagy and apoptosis. *BNIP3* has been shown upregulated in hESCs growing in hypoxic conditions, but autophagy played the major role here whereas apoptosis did not (Abaci et al., 2010). Autophagy has been shown to work in concert with pluripotency proteins *OCT4*, *SOX2*, and *NANOG* to maintain pluripotency homeostasis, inhibition of autophagy affected pluripotency negatively even though there were high levels of pluripotency proteins present (Cho et al., 2014). *LITAF* (Lipopolysaccharide-induced TNF factor) plays a role in protein trafficking and lysosomal degradation (S. M. Lee et al., 2012). LncRNA *AC106864.1* is interesting as there are no publications found but it seems important and shows up in the top genes for cytoTRACE. In the same genomic location, there is also pseudogene *WRBP1* (tryptophan-rich basic protein pseudogene 1) which has only one reference: *The Sequence of the Human Genome*. Pseudogenes are often attributed to having no functional relevance although some have shown to be regulatory RNAs e.g. siRNAs derived from pseudogenes (Chan et al., 2013; Roberts & Morris, 2013). *PODXL*, podocalyxin-like protein, is a surface protein part of the CD34-family and was recently been discussed as a feature of pluripotency in combination with particular markers (Kang et al., 2016). *APELA*, or *ELABELA*, is a growth factor that promotes the self-renewal of hESCs (Ho et al., 2015). Finally, downregulating or inhibiting *IRX2* has been reported to hinder hESC differentiation to NPCs (Cohen et al., 2000).
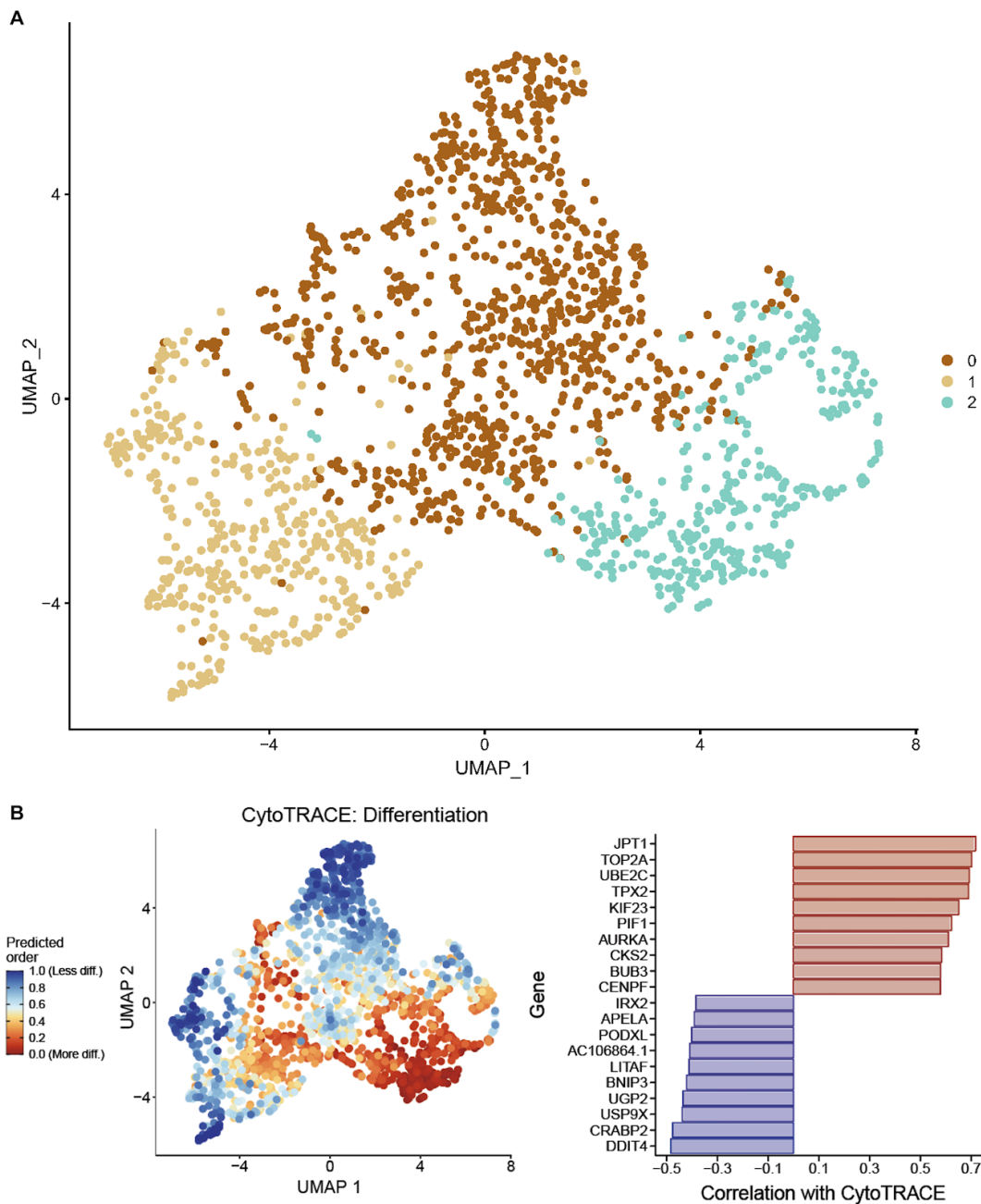
Figure 16 *Day 0 clustering and differentiation prediction.* Results on day 0 for Seurat/scClustViz clustering solution and cytoTRACE differentiation prediction. A. Day 0 hESCs were resolved as 3 clusters based on silhouette metrics in scClustViz, then plotted on a UMAP via Seurat. B. CytoTRACE predicts areas in clusters 1 and 0 to be the start of differentiation, while cluster 2 is the predicted end of differentiation. Predicted start-of-differentiation-markers are genes *JPT1, TOP2A, UBE2C, TPX2, KIF23, PIF1, AURKA, CKS2, BUB3*, and *CENPF*. Predicted end-of-differentiation markers are genes *DDIT4, CRABP2, USP9X, UGP2, BNIP3, LITAF, AC106864.1, PODXL, APELA*, and *IRX2*.

Investigation of markers to determine cell-types at day 0

Automated single-cell cluster annotation at all time points was done with the R-tool scCatch (Shao et al., 2020) which uses a panel of 353 cell types and 686 subtypes, using 20 792 reference marker genes and 2 097 mouse and human references. The day 0 scCatch results (**Table 4** *Day 0 scCatch cell-type predictions*.) predicts cluster 0 as an astrocyte-type cluster but with a relatively low prediction score. Cluster 1 and 2 is more accurately predicted as type embryonic stem cell with a high relative prediction score. Knowing that reference database for all automatic cluster-annotation has a specific repertoire of cell-types, it is likely that the low score, and a name such as an astrocyte that is less likely to be present at day 0, indicates the cluster is not accurately annotated, and likely due to the automated tool not having good enough profiles for many of the cells in this protocol. Although the tool might fare better at day 20 for some more mature progeny. For cluster 0 we see marker *EZR* which is a migratory marker and links the plasma membrane to actin (Sotthibundhu et al., 2016). *SOX2*, NANOG, and *OCT4* have been shown to bind at the promoter regions of ID1 and different ID-genes (*ID2* and *ID4* are also markers in the scCatch table of cell-type markers) seem to be active during differentiation, perhaps suggesting this cluster to be more 'ready', while the other clusters are active cell-cycling and renewal (Stahlberg et al., 2009). *ZFAND5* has been shown overexpressed in blastomeres (Galán et al., 2010). *ALCAM* in hESCs has been implicated in cardiomyocyte differentiation indicating a possible mesoderm-prone fated cluster (Ghazizadeh et al., 2018; Rust et al., 2009).

Table 4 *Day 0 scCatch cell-type predictions* scCatch uses the most variable genes per cluster as input and predicts most likely corresponding cell types based on its internal database. Results are presented as a table where column 1 *Cluster* contains a cluster-ID/number, column 2 *Cluster markers* are the variable genes per cluster it received as input (which can be hundreds of markers), column 3 *Cell type* which is the predicted cell type, column 4 *Pred. Score* which is a prediction score from 0 to 1, column 5 *Cell type markers* which are the predicted relevant markers that scCatch picks out from the large lists in column 2, and finally column 6 *PMID* which are references to papers linked to the markers found by scCatch. **For day 0**, scCatch predicts cluster 0 as *astrocytes*, then - with a higher prediction score - clusters 1 and 2 as *Embryonic Stem Cells*.

Cluster	Cluster markers	Cell type	Pred. Score	Cell type markers	PMID
0	RPS29, MT-ND2, RPS20, RPS27, RPL27A, RPL23, MT-ND5, APOE, HNRNPAB, RPS11, EIF4G2, PPP1R14B, DHCR24, ZFAS1, SPATS2L, AC106864.1, SRRM2,	Astrocyte	0.8	APOE, CLU, EDNRB, EZR, ID1, ID4, LRRC8A, ZFAND5, ZFP36L2, AIF1L, ALCAM, FABP5, HMGCS1, ID2, ITPR2, LITAF, SCD, TTYH1	27806376, 29539641
1	FOXD3-AS1, SFRP2, TPM1, DNMT3B, RAB17, EDNRB, SRRM2, RPS11, TERF1, VSNL1, RPL27A,	Embryonic Stem Cell	0.95	SOX2, PODXL	28436947, 24571984, 24298345, 20976554, 20874104, 20602182
2	IRX2, CRABP2, SOX21, ID4, USP9X, C5orf38, APOE, LITAF, EEF1A1, CKB, ID2, FST, TMSB4X, CLU, IRX1,	Embryonic Stem Cell	0.95	SOX2, PODXL, LIN28A	28436947, 24571984, 24298345, 20976554, 20874104, 20602182

Day 0 cell-states

In the dot-plots of day 0 (Figure 18 *Day 0 dot-plots*), we see *IRX2*, *ID4*, and *SOX21* as markers for cluster 2 at day 0 while cluster 0 and 1 are closely related indicating that scCatch was wrong with its *astrocyte* label. There is not much difference in general between clusters at day 0 and these clusters can all be annotated as hESCs (Figure 17 *Day 0 cell-state annotation*) but it would be interesting to follow up the small differences in the hESCs, such as *IRX2*, *ID4*, and *SOX21* specificity for cluster 2.

43
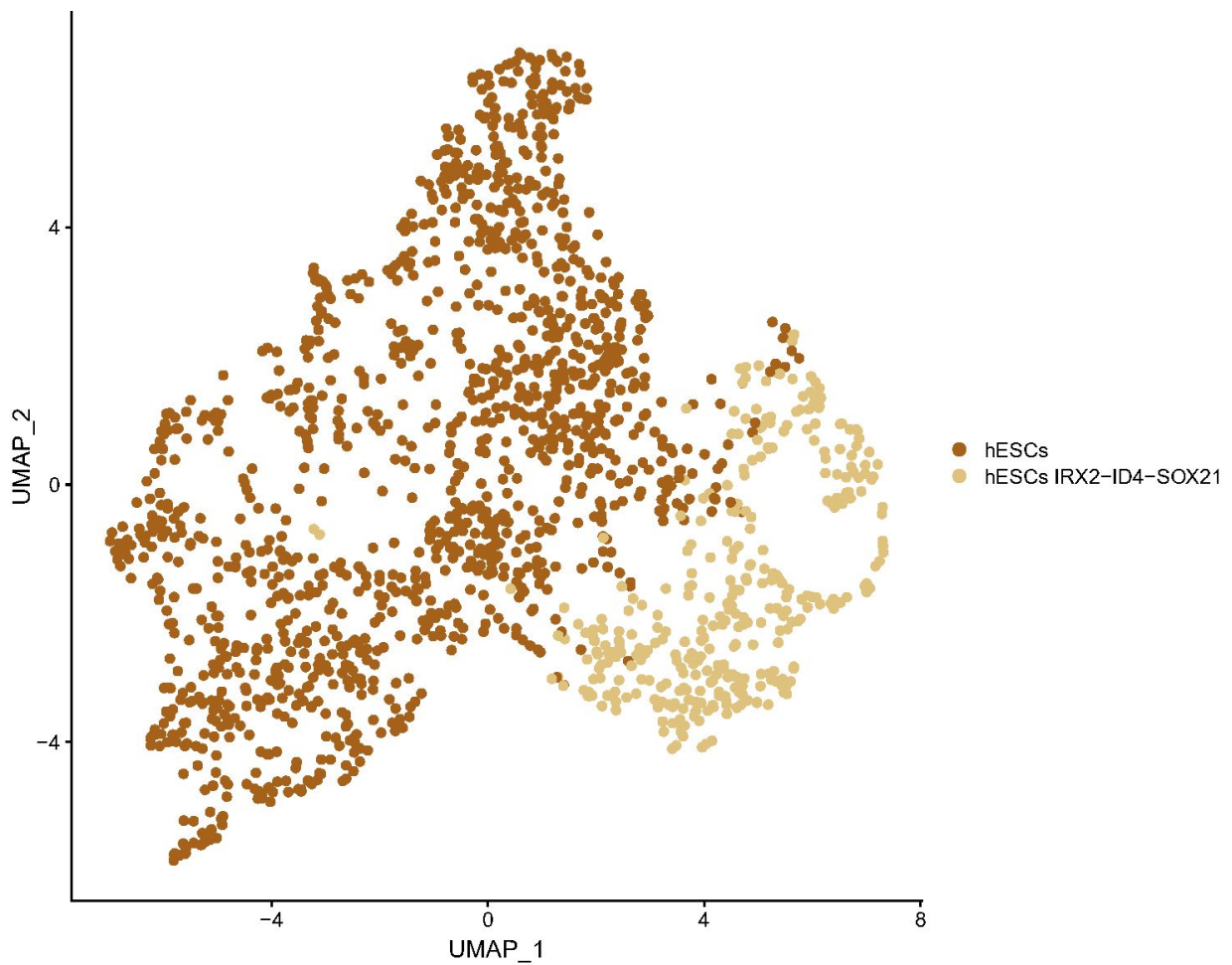
Figure 17 *Day 0 cell-state annotation* Results from differentiation predictions, automatic cell-type predictions, and manual investigation of markers, indicated a relatively homogenous population of hESCs, where one population had higher expression of *IRX2*, *ID4*, and *SOX21*. Annotation for day 0 is more close to *cell-states*, rather than cell-type. Annotation is presented on the UMAP above.
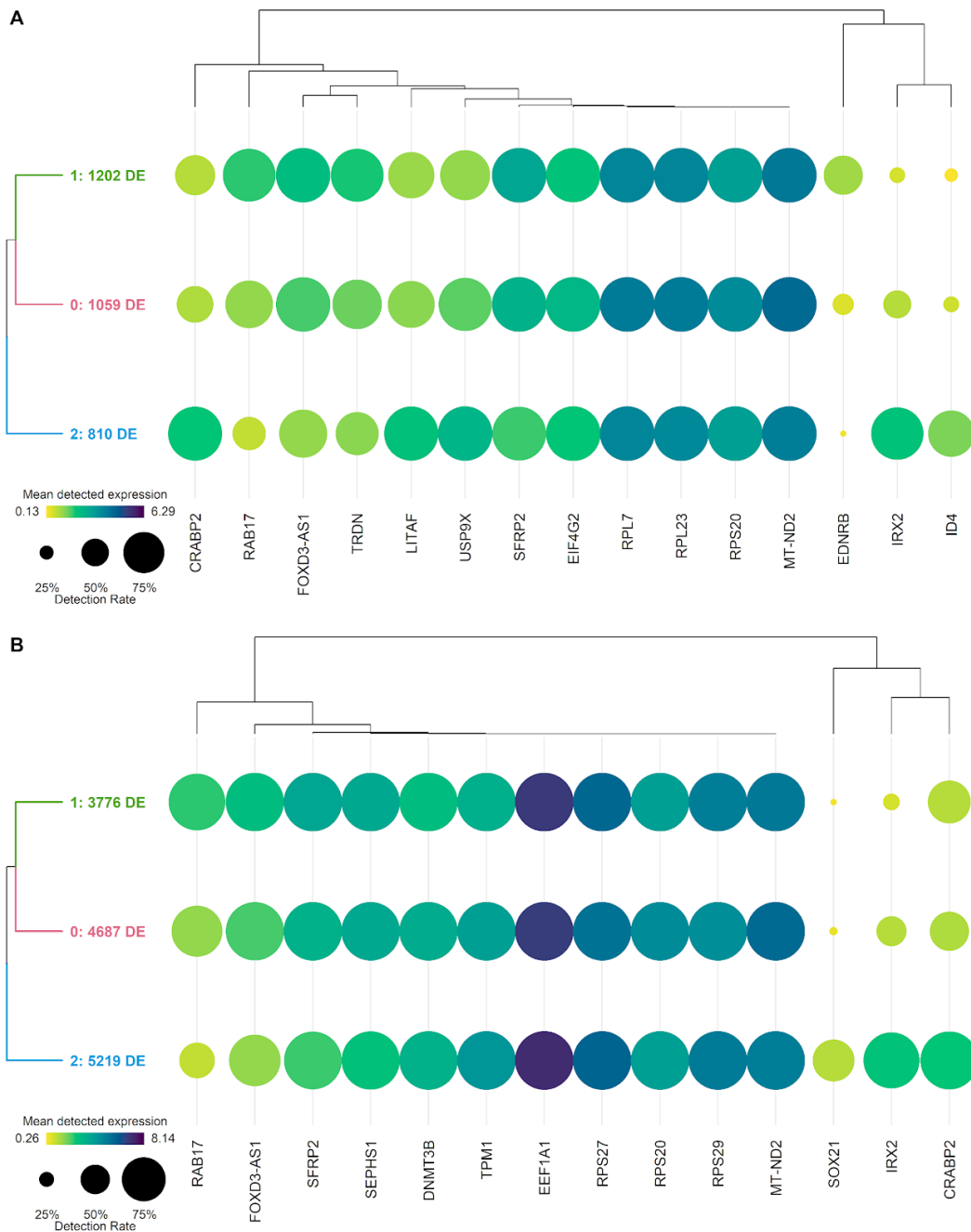
Figure 18 *Day 0 dot-plots.* Visualizing cluster differential expression of marker genes using pairwise tests between a cluster and other clusters. A. Cluster marker genes, where *IRX2* and *ID4* are markers for cluster 2, while cluster 1 and 0 are almost identical. B. Genes more unique to a cluster compared to rest, *SOX21* appears a third marker for cluster 2 in this dot-plot.

45

Day 7: End of Part I: Induction

2264 cells were used in the analysis of day 7 after quality control and filtering (SI Day 7 Figure I and SI Day 7 Figure II). At day 7, cells have been under the influence of LSX-inhibitors for a week, rapidly converting to cells in the neuroectoderm lineage. The number of clusters for day 7 was resolved to be 3 (Figure 19 *Day 7 clustering and differentiation prediction* A.).

Movement of differentiation

CytoTRACE differentiation predicts cluster 2 to be the start-of-differentiation which then moves in direction of cluster 0, which is the end-of-differentiation (Figure 19 *Day 7 clustering and differentiation prediction* B.). In contrast to day 0, batch integration of experiment run A and run B show a much better integration (SI Day 7 Figure I, C.). There is more heterogeneity in the cell population at day 7, in comparison to day 0, and R-tool and scRNA-seq analysis package Seurat, which uses distinct anchors/markers to map the experimental runs onto each other, is integrating the runs well.

Predicted start-of-differentiation-markers for day 7, starting at cluster 2, are *HMGB2*, *NUSAP1*, *TOP2A*, *UBE2C*, *TUBB4B*, *KPNA2*, *UBE2S*, *DLGAP5*, *CKS1B*, and *TPX2*. Not much information in this context was found on specifically *TUBB4B* but it is likely specialized tubulin expressed in these cell-types/tissue. Shared with day 0 is *TOP2A*, *UBE2C*, and *TPX2*. *HMGB2* has been shown important for the promotion of differentiation to the neuroectodermal lineage while being important for proliferation and is also expressed in hESCs (Bagherpoor et al., 2017), indicating *HMGB2* to be a possible stem/progenitor/cycling cell marker at this timepoint. *HMGB2* has been shown to be a negative regulator of telomerase activity thus is likely involved in the homeostatic regulation of the renewing populations (Kučírek et al., 2019). *NUSAP1* and *KPNA2* have been shown expressed in relation to the renewal of cells, in rosettes (day 7) and NPCs (Burke et al., 2020). *NUSAP1* has also previously been used as a marker for proliferating progenitors together with *MKI67*, *TOP2A*, *TPX2*, *CENPF*, *HMGB2*, *HIST1H4C* and *UBE2C* (Collin et al., 2019; Teotia et al., 2020). *UBE2S* has been shown to promote self-renewal and

pluripotency in mouse ES cells and seems to have a relation with *SOX2* in neural progenitor cell differentiation (Cui et al., 2018; J. Wang et al., 2016).

Predicted end-of-differentiation-markers for day 7 are *MSH6*, *PIK3R3*, *MXD4*, *MAPK10*, *LINC01551*, *MCM4*, *MCM5*, *MCM6*, *UNG*, and *CALHM5*. There are no genes in common with day 0 end-of-differentiation which makes sense since we are a week into neuroectodermal induction. GeneCards reports *MSH6* as a member of DNA mismatch repair MutS family of proteins and *MSH6* together with similar proteins have been shown expressed in blastomas (Belloni et al., 1999; Stark et al., 2015). *PIK3R3* seems to have less information about its role at this stage, but GeneCards reports that the gene is expressed in neural rosettes. *MXDX* family proteins have been suggested to be specialized Myc-antagonists with a role in neurodevelopment (Hooker & Hurlin, 2006). *MAPK10* is a member of the Toll-like receptor pathway which is involved in CNS development and NPC differentiation (Fathi et al., 2011; Rolls et al., 2007). Another LncRNA, *LINC01551*, shows up at day 7 as a differentiation-marker. Interestingly, GeneCards report this LncRNA gene to target two of 11 reported enhancer sites for *FOXG1*, and this LncRNA has only 4 publications whereas one is on hepatocellular carcinoma (Gao et al., 2019), and another one is, also potentially interestingly, on genetic variations of glutamate in brains of multiple sclerosis patients (Baranzini et al., 2010). Its role at this stage seems unknown and could be interesting to follow up on. Several *MCM* proteins are in the marker list, *MCM4*, *MCM5*, and *MCM6*, and they are components of the pre-replicative complex essential in initiation of DNA replication (Kearsey et al., 1996) but they seem to have some role in the neural lineage as one study showed a link between human cytomegalovirus infection and viability of NPCs when *MCM* genes were dysregulated (D'Aiuto et al., 2012). Further, *MCM4* has been implicated as a *SOX2* target (Zhou et al., 2016) in hESC to NPC differentiation. Not much information is found regarding *UNG* and *CALHM5* gene has only 7 publications but seems to be a pore-forming subunit in a voltage-gated ion channel (UniProt).
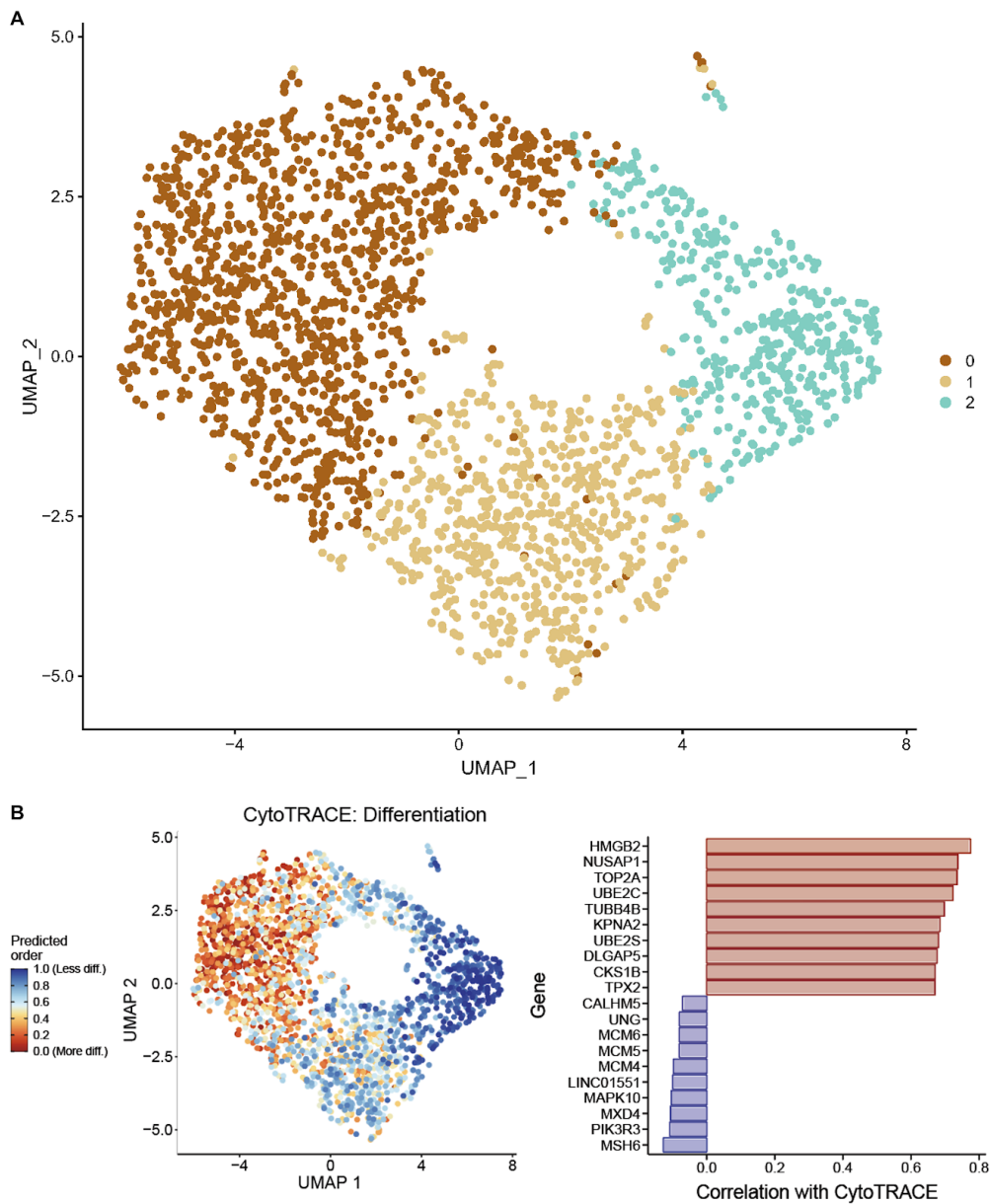
Figure 19 *Day 7 clustering and differentiation prediction.* Results on day 7 for Seurat/scClustViz clustering solution and cytoTRACE differentiation prediction. A. Day 7 was resolved as 3 clusters in scClustViz, then plotted on a UMAP via Seurat. B. CytoTRACE predicts cluster 2 to be the start cell-state of differentiation at day 7, with an intermittent stage as cluster 1, then cluster 0 as end-of-differentiation. Predicted start-of-differentiation markers are genes *HMGB2*, *NUSAP1*, *TOP2A*, *UBE2C*, *TUBB4B*, *KPNA2*, *UBE2S*, *DLGAP5*, *CKS1B*, and *TPX2*. Markers for end-of-differentiation, with a relatively weaker correlation, are *MSH6*, *PIK3R3*, *MXD4*, *MAPK10*, *LINC01551*, *MCM4*, *MCM5*, *MCM6*, *UNG*, and *CALHM5*.

Investigation of markers at day 7

scCatch cell-type predictions for day 7 show low prediction scores where all three clusters are predicted to be astrocytes (Table 5 *Day 7 scCatch cell-type predictions*). Cluster 1 and 0 are closely related to each other when inspecting the dendrogram in dot-plots (Figure 20 *Day 7 dot-plots*). Cluster 2 is the starting point for differentiation in cytoTRACE, cluster 2's identity is likely a renewing progenitor population. For cluster 0, cell-type markers according to scCatch are *ANLN*, *MTHFD2*, *HES4*, *HES5*, *PSAT1*, and *SLC3A2*. Cluster 1 markers are *MTHFD2*, *HES4*, *HES5*, and *SOX4*. Shared scCatch cell-markers between these are *HES4* and *HES5*. Notch-signaling control how fast progenitors commit to maturation and the Notch-pathway has been shown to be active in rosette cultures (Abranches et al., 2009), such as our day 7 cells, and for NPCs (Basak & Taylor, 2007). *HES5* is an important target in the Notch-pathway (Ohtsuka et al., 1999) and works as an effector in Notch-signaling. *HES5* will maintain progenitor state and repress proneural genes, while cells that escape to progress and mature will become *HES5* negative and take on proneural genes such as *MAP2*, *NEUROG2*, *ATOH1*, etc. (Manning et al., 2019). *ANLN* has been implicated at the rosette stage to have a function in the neural rosettes actin cytoskeleton (Shang et al., 2018). There is less information about *MTHFD2* but according to GeneCards its expressed at neural tube formation which is roughly equivalent to rosette-stage *in vitro*, similarly *SOX4* is important in neural tube formation and is another rosette-indicator (Shang et al., 2018).

By the markers reported by scCatch, clusters 0 and 1 can reasonably be annotated as maturing neural rosette cells, while cluster 2 can be annotated as progenitors, a pool of cells that resupply each other and/or push new cells towards neural maturation. Cluster 2 marker *FOS* is another Toll-like receptor pathway (TLR) member similar to *MAPK10* (TLR member *MAPK10* was predicted as a start-of-differentiation marker on day 7). However, *FOS* protein can be phosphorylated by MAPK kinases, and depending on where we are in time and tissue the outcome can be both activation or repression of target genes (Hurd et al., 2002). *JUN*, or *c-JUN*, has been shown to, with *c-FOS*, to be required for neuronal differentiation in PC12 cells (Eriksson et al., 2007) and both have been shown to interact with BMP-signaling (Peng et al.,

2002), and LDN-193189 was a potent inhibitor of the BMP pathway as described in the introduction of the method. Markers thus far mean cluster 2 will be annotated as *Day 7 Progenitors*, clusters 1 and 2 will be fused together and annotated as *Day 7 Maturing Rosette Cells*.

Table 5 *Day 7 scCatch cell-type predictions*. **For day 7**, scCatch predicts clusters 0, 1, and 2 to be *astrocytes,* with a relatively low overall prediction score. For cluster 0, scCatch determines cell-type markers to be: *ANLN*, *MTHFD2*, *HES4*, *HES5*, *PSAT1*, and *SLC3A2*. Cluster 1 cell-type markers are *MTHFD2*, *HES4*, *HES5*, and *SOX4*. Cluster 2 cell-type markers are *ANLN*, *FOS*, *JUN*, *CLDN10*, *PHGDH*, *PLIN3*, and *SLC3A2*.

Cluster	Cluster markers	Cell type	Pred. Score	Cell type markers	PMID
0	MKI67, TOP2A, NUSAP1, UBE2C, HIST1H4C, CDK1, ASPM, TPX2, NDC80, CENPF, KIF23, H2AFX, DLGAP5, NCAPG,	Astrocyte	0.76	ANLN, MTHFD2, HES4, HES5, PSAT1, SLC3A2	27806376, 29539641
1	HIST1H4C, HIST1H1B, HIST1H1D, HIST1H1E, HIST1H1C, PCLAF, RAD51AP1, HIST1H3B, TYMS, MKI67, MYBL2,	Astrocyte	0.73	MTHFD2, HES4, HES5, SOX4	27806376, 29539641
2	PLK1, AURKA, CCNB1, CDC20, HMMR, CENPE, UBE2C, CENPA, SGO2, TPX2, ASPM, TUBA1C, TOP2A, CENPF, PIF1,	Astrocyte	0.77	ANLN, FOS, JUN, CLDN10, PHGDH, PLIN3, SLC3A2	27806376, 29539641
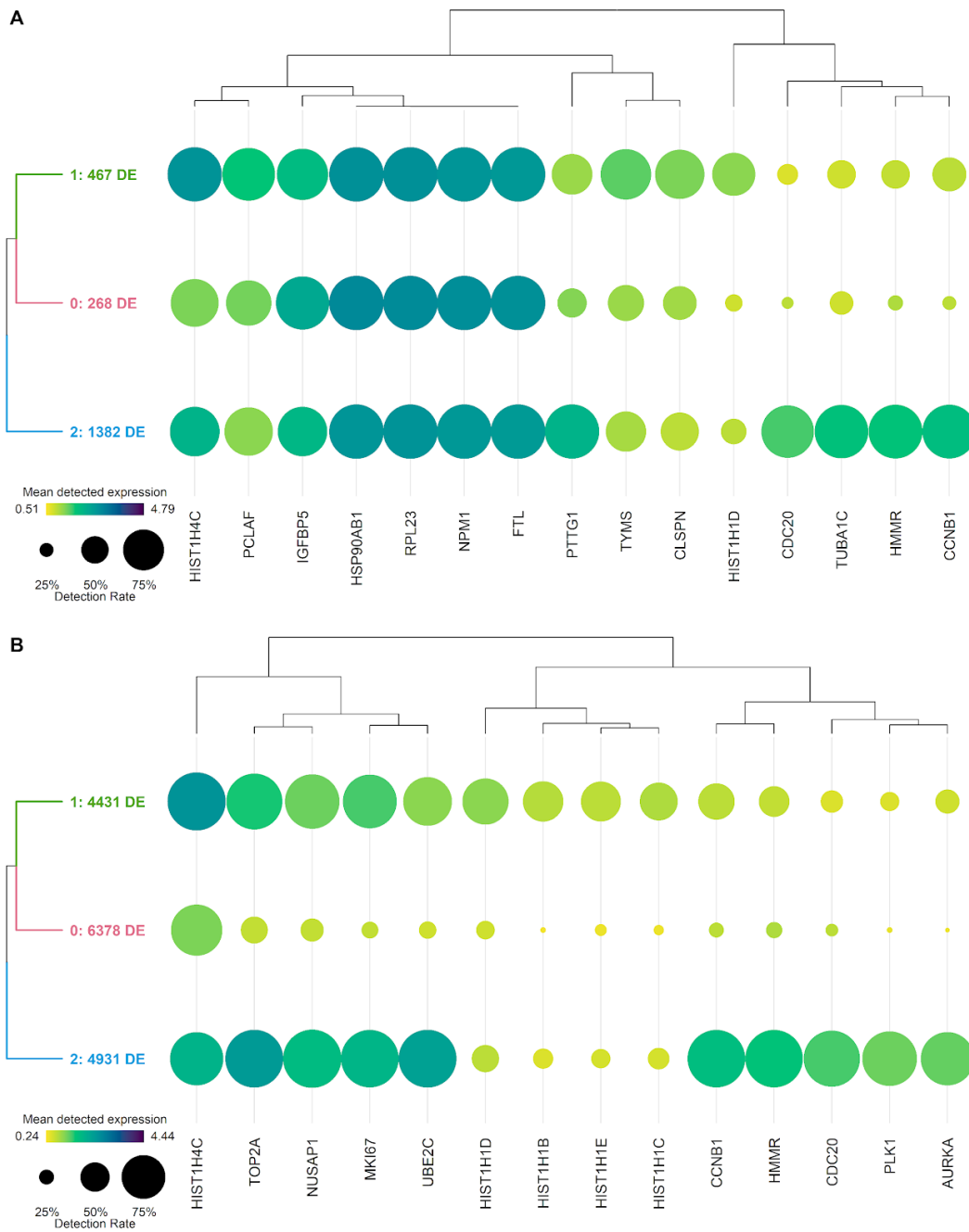
Figure 20 *Day 7 dot-plots* Visualizing cluster differential expression of marker genes using pairwise tests between one cluster and all other clusters. Here, the dot-plots are showing a greater heterogeneity on day 7 when compared to day 0. A. Marker genes after pairwise tests are *CDC20, TUBA1C, HMMR*, and *CCNB1* for cluster 2 (suggested to be progenitors). Cluster 1 decreases in expression of these genes and cluster 0 have further lost expression of these genes. B. Genes more unique to a cluster compared to rest. The genes for cluster 2 are *CCNB1, HMMR, CDC20, PLK1*, and *AURKA*. Those genes then decrease in expression gradually, via cluster 1, then cluster 0 (and those two are suggested to be rosette clusters).

51

Progenitor cells are reported to be found within rosette structures

The weak correlation for genes predicted to drive differentiation in cytoTRACE suggests that it is not very evident in the data where cells are heading. Thus defining a path of differentiation, and trying to define cell-types might not be as relevant as later time points. Likely we have more different cell-states of rosette/neuroectoderm, rather than cell-types, similar to day 0. However, a proliferating cluster should be possible to identify. Validation of progenitor cells, which are reported to be within rosettes (Conti & Cattaneo, 2010), and are detected by *NES* and *PAX6* expression. However, *NES* has a very uniform expression pattern, and so does *PAX6*. *PAX6* is possibly, but very slightly, expressed higher in cluster 2 compared to the other clusters (Figure 21 *Day 7 progenitor-specific markers* A. and B.).
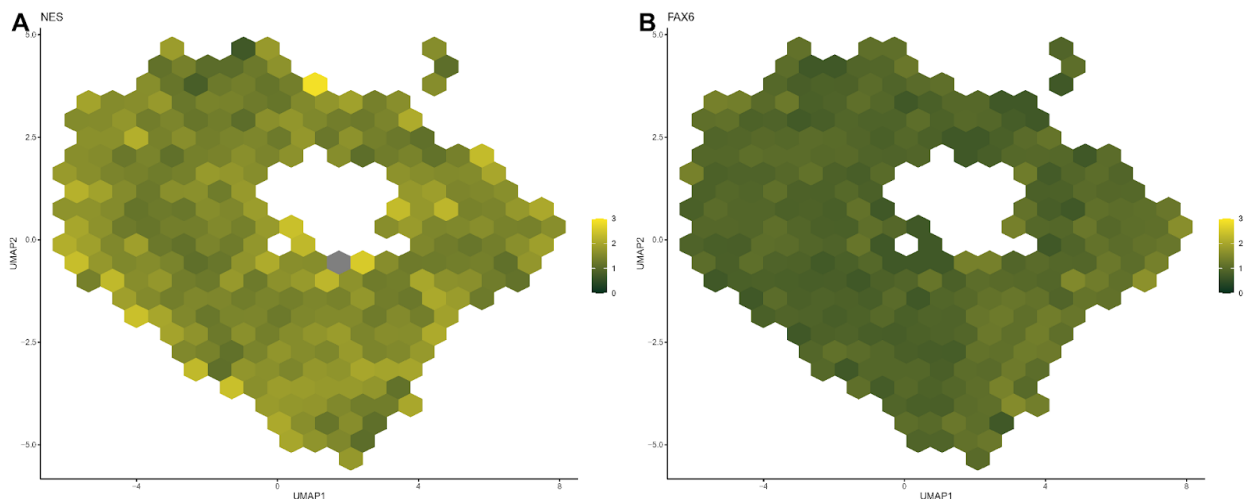
Figure 21 *Day 7 progenitor-specific markers*. It was reported in the literature that progenitor markers at the rosette-stage might be found expressed in the rosette interior, two of these were *NES* and *PAX6*. A. Expression of progenitor marker *NES*. B. Expression of progenitor marker *PAX6*. Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 4.

Day 7 cell-types

Finally, Day 7 clusters 0 and 1 will be annotated as *Rosette I*, and Rosette *II*, while cluster 2 will be annotated as *Progenitors/rosette* (Figure 22 *Day 7 Cell-types*). Annotation of cluster 2 tries to reflect that identifying progenitors at day 7 was less straightforward.
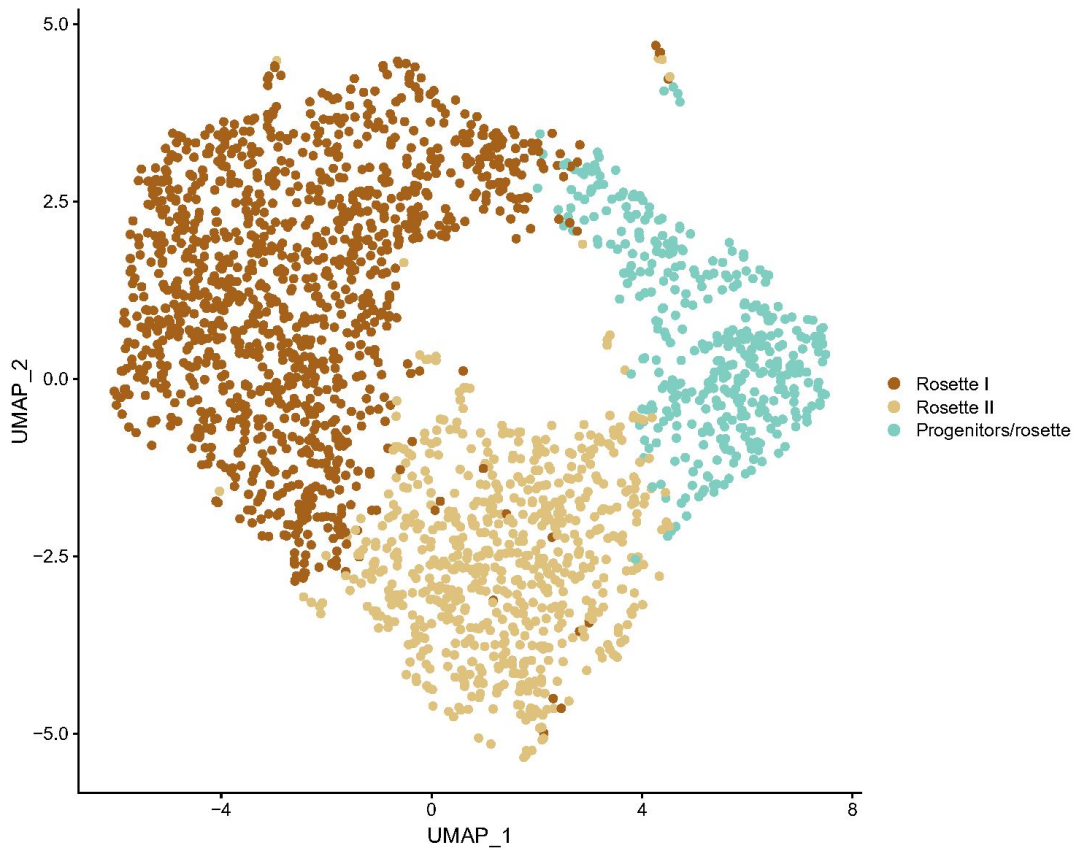
Figure 22 *Day 7 cell-types*. Results from differentiation predictions, automatic cell-type predictions, and manual investigation of markers, suggest that cluster 0 and 1 can be annotated as *Rosette I* and *Rosette II* respectively, while cluster 2 are named *Progenitors/rosette*. Annotation is presented on the day 7 UMAP.

Day 13: End of Part II: Maturation

By day 13, cells have been maturing in B27 and N2 supplemented media for a week, and this maturation might be visible as increased heterogeneity. 1958 cells are included at the day 13 time point following quality control and filtering (SI Day 13 Figure I, and SI Day 13 Figure II). Initial clusters were resolved to 8 clusters which suggest a large increase in heterogeneity (Figure 23 *Day 13 clustering, and differentiation prediction* A.)

Movement of differentiation

CytoTRACE prediction starts in cluster 2 then differentiation moves towards and through clusters 3, 5, and 0. The end of differentiation seems to be at two locations, cluster 6 (left-most cluster) and then via cluster 4 ending in cluster 7 (bottom cluster) (Figure 23 *Day 13 clustering, and differentiation prediction* B.). Predicted start-of-differentiation markers are genes *HMGB2*, *NUSAP1*, *CENPF*, *TOP2A*, *TUBB4B*, *MKI67*, *TPX2*, *SMC4*, *H2AFX*, and *ASPM*.

Start-of-differentiation markers shared with days 0 and 7, are *TOP2A*, and *CENPF*. Markers shared with day 7 are *HMGB2*, *NUSAP1*, and *TPX2*. *TUBB4B* shows up on day 13, similar to day 7, and as mentioned on day 7, TUBB4B did not have much information available. But different tubulins have been reported to be specialized for different tissue and developmental programs (Breuss et al., 2017), so *TUBB4B* is likely specific for the neural lineage. *MKI67*, similarly to *NUSAP1,* is a marker for proliferating progenitors. Then we have *SMC4* which has been indicated as a cell division/cell-cycle/cell proliferation marker (Bhattacharyya et al., 2009) but is otherwise described by UniProt as a component of the condensin complex that is part of converting interphase chromatin into condensed chromosomes. *H2AFX*, or *H2AX*, is a histone variant for histone H2, an alternative name is *histone variant H2A.X*, and this histone is, according to GeneCards, involved in Wnt-signaling and androgen receptor (AR) transcription via *PKN1*. As is commonly known, DNA wraps around histone octamers, H2A, H2B, H3, or H4, to form a more organized structure called chromatin were exchanging one of these and/or making modifications are a form of epigenetic changes/regulation of e.g. gene expression since it involves changes outside of the DNA code itself. Histone variant *H2AX* is further implied in

DNA damage response (N. Ayoub et al., 2008) and chromatin remodeling via Ring Finger Protein 8 (encoded by *RNF8*) *(Luijsterburg et al., 2012)*. *H2AX* seems to be functionally related to apoptosis and proliferation in neuronal precursors in the subventricular zone and in granule cell precursors (later cerebellum), and limiting *H2AX* restricts the proliferation of adult neural stem cells (Barral et al., 2014; Fernando et al., 2011). *H2AX* also seems to be much more phosphorylated upon Zika-infection in neural progenitor cells (Devhare et al., 2017). Finally, *ASPM* is another marker for day 13 start-of-differentiation which is a neuroblast-division associated gene and a decisive factor for increased cerebral cortical size (Bond et al., 2002; Villanea et al., 2012). These markers suggest cytoTRACE is predicting the progenitor area at day 13 accurately. Many new and interesting start-of-differentiation markers on day 13 might indicate an increase in heterogeneity for the progenitor cell pool at this stage.

Predicted markers for end-of-differentiation at day 13 are genes *TAGLN3*, *DCX*, *STMN2*, *ELAVL4*, *SOX4*, *SYT1*, *DLL3*, *MIAT*, *MLLT11*, and *MXD4*. *TAGLN3* (Transgelin 3) has been shown to be a marker for early post-mitotic neurons in the chick brain (Ratié et al., 2014; Ware et al., 2016). *DCX*, or doublecortin, is a neurogenesis marker encoding a microtubule-associated protein essential for brain development and a common marker for immature and migrating neurons (Ayanlaja et al., 2017; Couillard-Despres et al., 2005; Seki et al., 2019; Walker et al., 2007). *STMN2* (Stathmin-2) is a microtubule regulator associated with neuronal growth and necessary for axonal outgrowth and regeneration (Klim et al., 2019). *ELAVL4* encodes an RNA-binding protein that takes part in precise control of mRNA abundance and processing, a function important for neurons (Bronicki & Jasmin, 2013). *ELAVL4* has distinct mRNA isoforms which were recently shown to depend on their alternative 5' UTRs in early neurons and radial glial progenitors (Popovitchenko et al., 2020). The same study showed that an upstream RNA-binding protein called *CELF1* regulates *ELAVL4* to decide the development of glutaminergic neurons where dysfunctional *ELAVL4* can impair the development of functional glutamatergic neurons. Further in our day 13 end-of-differentiation marker list is *SOX4* which was a marker for day 7 as well, important in neural tube formation and a rosette-indicator but is also a marker for intermediate progenitor cells in their maintenance and formation (C. Chen et

al., 2015). *SYT1* (Synaptotagmin-1) is an important membrane protein in synaptic vesicles critical in the release of neurotransmitters and triggers this release by acting as a Ca^{2+}-sensitive sensor (Courtney et al., 2019; H.-K. Lee et al., 2010). *DLL3* (Delta-like 3) is a Delta family member and Notch-ligand involved in the commitment of neural progenitor cells to differentiation in the neural tube and developing nervous system while its expression also depends on TFs Neurog2 and Ascl1 (Henke et al., 2009). *MIAT* is a lncRNA expressed exclusively in a subset of postmitotic neurons and in differentiation neural progenitors (Roberts et al., 2014). This lncRNA was shown, in mice, to respond to retinoic acid (Mohamed et al., 2010) (common reagent for neural induction), retinoic acid is included in its precursor form, Vitamin A, in the B27 supplement used for part II and III in our protocol. *MLLT11* (Mixed-lineage leukemia; translocated to chromosome 11) has been shown involved in central- and peripheral nervous system development and is expressed, among other areas, in the basal region of the forebrain (Yamada et al., 2014). Finally, and also shared with day 7, we have *MXD4*, a member of the *MXDX* family suggested to be specialized Myc-antagonists with a role in neurodevelopment (Hooker & Hurlin, 2006).

Observations so far suggest that cluster 2 can be annotated as progenitors while clusters 6 and 7 are two relative end-stage clusters, and those two might represent two distinct lineages on day 13.
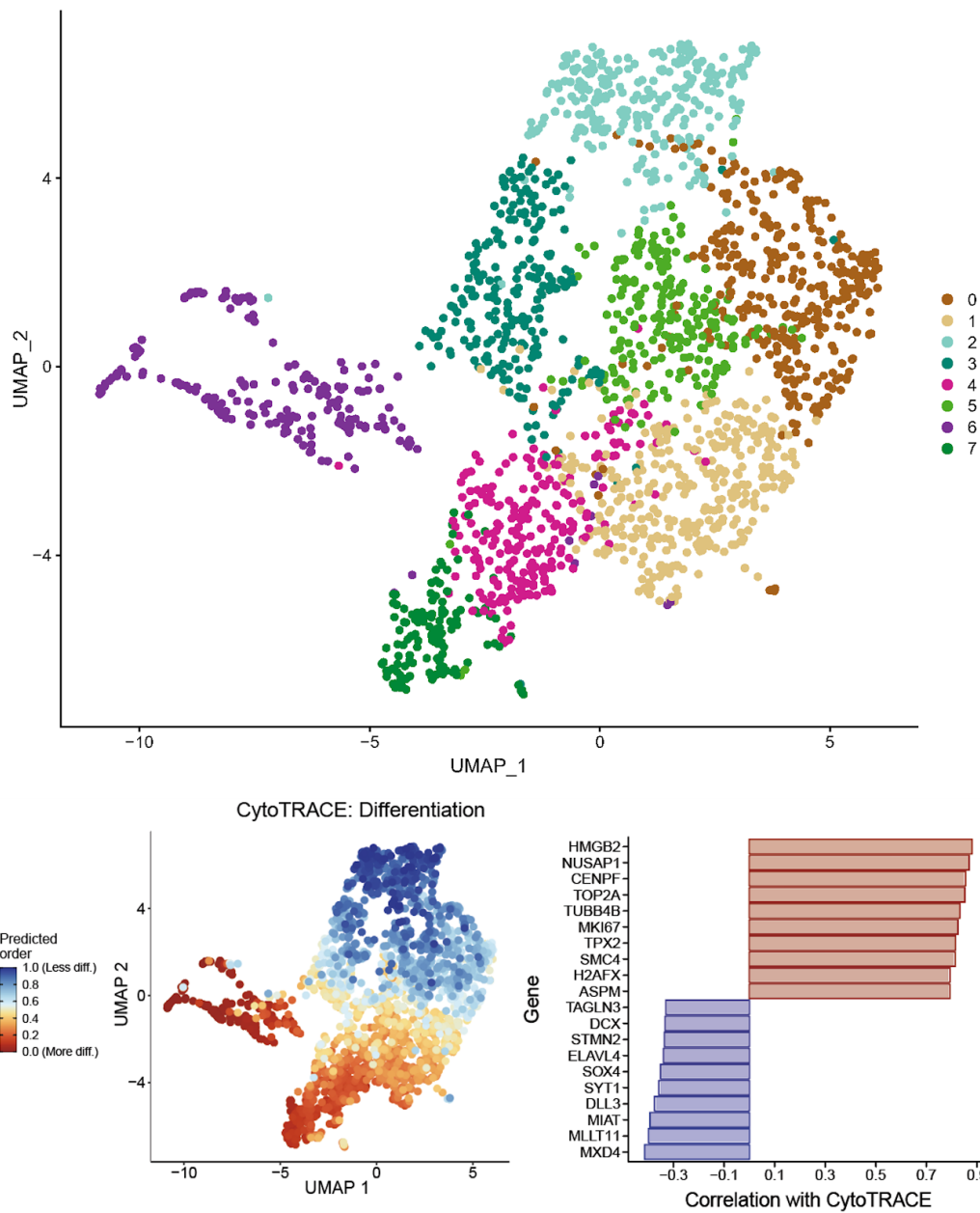
Figure 23 *Day 13 clustering, and differentiation prediction.* Results on day 0 for Seurat/scClustViz clustering solution and cytoTRACE differentiation prediction. A. Day 13 was mapped with 8 clusters based on silhouette metrics in scClustViz, then plotted on a UMAP via Seurat B. Differentiation is predicted by cytoTRACE to start with cluster 2, then move through clusters 3, 5, and cluster 0. Differentiation end seems to be at two locations, clusters 6 and 7. Predicted start-of-differentiation markers are genes *HMGB2*, *NUSAP1*, *CENPF*, *TOP2A*, *TUBB4B*, *MKI67*, *TPX2*, *SMC4*, *H2AFX*, and *ASPM*. Markers for end-of-differentiation are *TAGLN3*, *DCX*, *STMN2*, *ELAVL4*, *SOX4*, *SYT1*, *DLL3*, *MIAT*, *MLLT11*, and *MXD4*.

Investigation of markers to determine cell-types at day 13

scCatch is more successful at day 13 with the cell-type predictions on the 8 clusters, and the seemingly better matches to cell-types in its reference database are evident in the cell-type marker lists (Table 6 *Day 13 scCatch cell-type prediction*). Cluster 0, 2, and 3 are annotated as NPCs, and cluster 1 as Astrocytes. These four clusters have a long list of markers and will need some investigation. Cluster 5 and 7 are annotated as hESCs by scCatch but the prediction is based solely on *SOX2* expression, and we know from before to expect *SOX2* expression throughout the protocol. These two clusters need further investigation. Cluster 7 was also predicted by cytoTRACE to be one of the most differentiated clusters at this time point, which makes it contra-indicatory that this cluster is hESCs. Cluster 6 was named by scCatch as neural stem cells (NSCs). NSCs at this stage and onwards are very similar to NPCs or radial glial cells (RG) and they seem to be used interchangeably depending on the study, and usually denote progenitors of some type. NSC is sometimes used to mean a stem cell in the neural lineage, but of a mature sort, and could thus be an end state at day 13. Nonetheless, this cluster needs investigation and has fewer marker genes, only 7, in comparison to the first 5 clusters which had 15, or more, markers each.

Table 6 *Day 13 scCatch cell-type prediction*. **For day 13**, there are multiple genes predicted as cell-type markers by scCatch, whereas the first 5 markers for each cluster is; for cluster 0 *NPCs: NES, DCX, DPYSL3, ATAD2, AURKB*, for cluster 1 *Astrocytes*: *SOX9, ANLN, APOE, ID4*, and *MAFB*, for cluster 2 *NPCs*: *DCX, ANP32E, APOLD1, ASPM*, and *AURKB*, for cluster 3 *NPCs*: *DCX, ANP32E, AURKB, BIRC5*, and *BUB1*. Then for cluster 4 *Astrocytes*: *ANLN, C1orf61, ID4, LHFPL6*, and *ATP1A2*, for cluster 5 *hESCs*: *SOX2*, for cluster 6 *NSCs*: *PROM1, NES, ADGRG1, DCX*, and *NEUROD1*, and finally for cluster 7 *hESCs*: *SOX2*.

Cluster	Cluster markers	Cell type	Pred. Score	Cell type markers	PMID
0	HIST1H4C, HIST1H1B, HIST1H1E, HIST1H1D, HIST1H2AH, HIST1H1C, HIST1H3B, TYMS, PCLAF, HIST2H2AC,	Neural Progenitor Cell	0.86	NES, DCX, DPYSL3, ATAD2, AURKB, C21orf58, CDCA4, CDCA5, CDK1, CENPK, CENPM, CENPU, DHFR, DUT, FBXO5, FEN1, GAP43, HELLS, HJURP, HMGB2, ID4, PCLAF, KIF11, KIF15, KIF22, KIF23, MAD2L1, MELK, MIS18BP1,	25843934, 21275797, 29539641
1	FABP7, UBE2S, DLGAP5, TPX2, MKI67, UBE2C, CCNB1, HES5, GTSE1, NUSAP1, HMMR, TOP2A, MGST1, KPNA2, NDC80,	Astrocyte	0.85	SOX9, ANLN, APOE, ID4, MAFB, RGMA, ETV5, FABP7, GJA1, GNG11, HES1, HES4, HES5, ID2, ITM2C, MLLT11, PLIN3, PTN, PTPRZ1, SOX4, SPRY1, SYT1, TAGLN3, VEPH1	28336567, 27806376, 29539641
2	PIF1, CENPA, AURKA, PLK1, CDCA3, UBE2C, CENPE, ASPM, TOP2A, SGO2, KPNA2, PSRC1, CDK1, CKAP2L,	Neural Progenitor Cell	0.81	DCX, ANP32E, APOLD1, ASPM, AURKB, BIRC5, BUB1, BUB3, C21orf58, CCNA2, CCNB1, CCNB2, CDC20, CDCA3, CDK1, CDKN3, CENPF, CENPN, CENPW, CKAP2, CKAP2L, CKS1B, CKS2, CLSPN, DEPDC1B, DLGAP5,	21275797, 29539641
3	PTTG1, CCNB2, CDC20, DYNLL1, BIRC5, UBE2S, HMGB2, ARL6IP1, HMMR, TUBA1C, NUCKS1, CCNB1, DLGAP5,	Neural Progenitor Cell	0.81	DCX, ANP32E, AURKB, BIRC5, BUB1, C21orf58, CCNB1, CCNB2, CDC20, CDK1, CDKN3, CENPF, CKS1B, DLGAP5, DTYMK, DUT, PIMREG, FBXO5, GAP43, HJURP, HMGB2, HMGB3, HMGN2, PCLAF, KIF23, MAD2L1, NUF2,	21275797, 29539641
4	NUSAP1, TOP2A, SMC4, IGFBP5, TPX2, SMOC1, ASPM, TUBB4B, HMGB2, MKI67, H2AFX, NKX2-1, MAD2L1, CDK1,	Astrocyte	0.8	ANLN, C1orf61, ID4, LHFPL6, ATP1A2, ELAVL4, FABP5, FABP7, HES1, HES4, HES5, ID2, NSG1, PDPN, PTN, SPRY1, TAGLN3, VEPH1	27806376, 29539641
5	PCLAF, CLSPN, TYMS, RRM2, LINC01551, TK1, CDC6, FEN1, DUT, FABP7, ATAD2, GINS2, CDC45, CHAF1A,	Embryonic Stem Cell	0.91	SOX2	28436947, 24571984, 24298345,
6	INSM1, SRRM4, ELAVL4, ONECUT1, DLL3, STMN2, NEUROG1, ISL1, ST18, TAGLN3, NEUROD1, ELAVL3, NHLH1,	Neural Stem Cell	0.91	PROM1, NES, ADGRG1, DCX, NEUROD1, NCAM1, DPYSL3	28795331, 25082219, 25843934,
7	DLX6-AS1, PPP2R2B, DLX5, DLX6, MEST, COL1A2, SELENOP, RBP1, GNAS, VIM, MAFB, MXD4, RALYL, SFRP1, SEMA3E,	Embryonic Stem Cell	0.92	SOX2	28436947, 24571984, 24298345,

Cluster analysis to determine cell-types

Due to the increased heterogeneity, dot-plots for day 13 represent many clusters and gene combinations (Figure 24 *Day 13 dot-plots*). Therefor analysis will follow a different flow than what it did at time points day 0 and day 7. The analysis here will be by clusters-groups, as in closer related clusters will be grouped, rather than to analyze single clusters in a long row which might be confusing.

Cluster 2 NPCs and cluster 0 NPCs are grouped together with cluster 5 hESCs one step above. Cluster 2 was predicted as a start-of-differentiation area in cytoTRACE and the marker genes produced here were suggesting this area to be likely progenitor which are in charge of renewing the cells and keeping a steady pool of progenitors that can self-renew as well as go on and continue differentiating. Cluster 2 as in Figure 24 A., has a strong expression of *PIF1*, *PLK1*, *CDCA3*, *CENPA*, and *AURKA*. *PIF1* was also seen as a top start-of-differentiation maker at day 0, and prevents genome instability and is involved in telomere maintenance. *AURKA* was important for cell renewal and seen on day 0 and day 7. Cluster 0 has a stronger expression of histone variants; *HIST1H2AH*, *HIST1H1B*, *HIST1H1E*, and *HIST1H1D,* compared to cluster 2. Cluster 0 is weaker in *PIF1*, *PLK1*, *CDCA3*, and *AURKA* expression.

Histone linker variants of H1 (*HIST1H2AH*, *HIST1H1B*, *HIST1H1E*, and *HIST1H1D*) are important in embryonic development and have specific functions (Pan & Fan, 2016) and have been further implicated in olfactory neuron maturation (J.-Y. Li et al., 2012). Most of these linker variants (H1B, H1E, H1D) are reported to be transcribed in S phase of cell cycle thus aside from linker variants having a particular function in the day 13 cells, this might imply that cluster 0 is different to cluster 2 and also different in that it has more cells in S-phase. Indeed upon inspection of cell-cycle assignments, cluster 2 has almost entirely G2M-phased cells while half of cluster 0 is predominately in S-phase (SI Day 13 Figure I, D.) Thus part of the explanation for the split of related clusters 2, 0, and 5 could be what cell-cycle these progenitors are in. Indeed, these histone linker proteins are expressed in cluster 5 as well, but at a lower level compared to cluster 0, and at a higher level compared to cluster 2. Cluster 5 has more mixed cells

from S, G1, and G2M than cluster 2 and 0 (SI Day 13 Figure I, D.). In conclusion for this group, these three clusters seem to be of progenitor type, where clusters 2 and 0 largely separates according to G2M and S-phase, and cluster 5 is different due to its mix of cell-cycle cells, and/or possibly somewhat more committed since the orientation of this cluster is more towards the more differentiated clusters.

Then we have clusters 3, 1, and 4. As seen in Figure 24 A., markers that are expressed stronger in cluster 3 *NPCs* compared to cluster 1 *Astrocytes* are *TOP2A*, *NUSAP1*, *BIRC5*, *CLSPN*, *CCNB2*, *MKI67*, *TPX2*, *DLGAP5*, *RRM2*, and *CDC20*. Cluster 3 has a relatively low expression of *PLK1*, *CDCA3*, *CENPA,* and *AURKA,* whereas cluster 1 has no expression of these markers. In Figure 24 B., *UBE2C* and *ASPM* are expressed stronger in cluster 3 compared to cluster 1. *UBE2C* was mentioned earlier as a cell-cycle marker and *ASPM* as a neuroblast marker. *SOX9* and *PCDH19* seem slightly higher in cluster 1, compared to 3, but not at any particular high degree. *SOX9* is reported to be important for the maintenance and induction of neural stem cells (Vong et al., 2015), and *PCDH19* has been reported to be important for neuronal cell adhesion (Pederick et al., 2018).

In Figure 24 A., marker genes *IGFBP5* and *DLX5* are expressed stronger in cluster 4 compared to clusters 1 and 3. While markers that were expressed strongly in cluster 3 compared to cluster 1 are less expressed for this cluster as well, in comparison to cluster 1. In Figure 24 B. we see one more marker for cluster 4, *CRABP1*. *CRABP1* is responsive to retinoic acid and has been shown to protect cells from excess RA (J. Y. Won et al., 2004), and its increased expression might indicate more commitment towards differentiation for cluster 4 and this also explains why this cluster is placed closer to cluster 6 and 7 (UMAP) and not directly grouped with cluster 3 and 1 (dendrogram, UMAP). *CRABP1* and retinoic acid have also been shown, through ERK signaling, to slow down cell cycle progression, and prepare a cell for differentiation (Park et al., 2019). To sum up this group, cluster 3, 1, and 4 are rather similar, as the dendrogram and similarities in gene expression show, and these clusters are situated between predicted start- and end of differentiation (Figure 23 *Day 13 clustering, and differentiation prediction*). Cluster 4 seems to

be the more mature cluster, out of the three, which will be taken into account when annotating. E.g. cluster 4 would be a more committed cell-state that respond to retinoic acid. Cluster 1 and 3 could be annotated as progenitors, but that has slowed down, to get ready to commit or re-enter the progenitor-state.

Cluster 6 *NSCs* and cluster 7 *hESCs* are predicted as the most differentiated areas by cytoTRACE on day 13. There are indeed some distinct markers for these two clusters. Starting with cluster 6 *NSCs,* in Figure 24 A., markers are *DLL3*, *ONECUT1*, *ELAVL4*, *INSM1*, and *SRRM4*. Figure 24 B., add a stronger *MAP6* expression, compared to the rest, and a very strong expression of *TAGLN3*. As mentioned in the cytoTRACE results, *DLL3* is a Delta family member and Notch-ligand involved in the commitment of neural progenitor cells to differentiation and is related to *ASCL1*. *ONECUT1* has been reported to be a downstream *PAX6* target and a glutaminergic neuron marker (Kee et al., 2017; Klimova et al., 2015). Furthermore, *ONECUT1* was shown correlated to *ONECUT2*, *NKX2-1* (relevant for day 20), and RFX4 in ventral telencephalon neurons (Trevino et al., 2020). *ELAVL4* encodes an RNA-binding protein needed for precise control of mRNA abundance and processing particularly in neurons where upstream protein where *CELF1* regulates *ELAVL4,* to decide the development of glutaminergic neurons. *INSM1* seems to be ubiquitously expressed in the developing nervous system but is sometimes mentioned in the transition of olfactory progenitors to progeny (Rosenbaum et al., 2011). *SRRM4* depletion has been shown to impair neurite outgrowth and later formation of cortical layers in the developing forebrain (Raj & Blencowe, 2015). Cluster 6 is very likely a mature, relative for the timepoint, neural cluster in comparison to the rest, except 7. Cluster 6 could be annotated as e.g. *ventral forebrain fated cells*.

Cluster 7 *hESCs* is the last cluster to analyze for day 13. It was the second cluster predicted as an end-of-differentiation cluster and was named h*ESCs* by scCatch based on just *SOX2* as a marker. In Figure 24 A., markers specific for cluster 7 are *DLX6*, *DLX6-AS1*, *PPP2R2B*, and *DLX5*. Importantly, what Cluster 7 lacks expression of, in comparison to cluster 6, are genes *DLL3*, *ONECUT1*, *ELAVL4*, *INSM1*, and *SSRM4*. Figure 24 B., does not add any more markers.

DLX6-AS1 (*DLX6* antisense RNA 1) is a lncRNA that is developmentally regulated and a strong marker for developing forebrain (Kohtz & Fishell, 2004). *DLX6-AS1* has been shown to increase in progression of neurodifferentiation and this lncRNA was shown to be 6123-fold change higher in interneurons compared to other cell-types and is a good interneuron marker, or interneuron precursor marker (S. J. Liu et al., 2016). *DLX5* and *DLX6* are two homeobox TFs that are commonly used as markers for developing GABAergic interneurons (de Lombares et al., 2019; Paina et al., 2011). *PPP2R2B,* Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B beta isoform, encode (as the name implies) a serine/threonine-protein phosphatase enzyme. GeneCards report that *PPP2R2B* is implicated in the negative control of cell division and cell growth. In the context of neurodifferentiation, *PPP2R2B* has been investigated related to the disease *spinocerebellar ataxia* (E. O'Hearn et al., 2012; Sato et al., 2010; Wardle et al., 2009), and has been shown enriched as an anterior/posterior gene in microarray (Kudo et al., 2007), and has been shown to disrupt neuron morphology when overexpressed (E. E. O'Hearn et al., 2015). Everything reported in the literature so far, suggest that cluster 7 could be annotated as *GABA-interneuron precursors*.

Figure 24 *Day 13 dot-plots*. Visualizing cluster differential expression of marker genes using pairwise tests between one cluster and all other clusters. A. and B. gene markers are analyzed in the main text to their complexity on day 13. Starting from the top on the left side of the dot-plot A. and dot-plot B., dendrograms are the same, and they show cluster 3 *NPCs* and 1 *Astrocytes* as related, then cluster 4 *Astrocytes* one step above in relation to those two clusters. Cluster 2 *NPCs* and cluster 0 *NPCs* are grouped together, with cluster 5 *hESCs* one step above those two. Cluster 6 *NSCs* and 7 *hESCs* are separated at the parent branch. A. Cluster markers. B. Genes more unique to a cluster compared to rest.

Day 13 cell-types

Finally, the day 13 clusters are annotated as follows (Figure X *Day 13 cell-types*). Cluster 2 as *Progenitors I: G2M*, cluster 0 as *Progenitors II: S-phase*, and cluster 5 as *Progenitor III: Mixed phase*. Then cluster 1 as *Sensitized progenitors I*, cluster 3 as *Sensitized progenitors II*, and cluster 4 as *Committed progenitors, RA responded*. And then cluster 6 as *Ventral forebrain fated cells* and cluster 7 as *GABAergic interneuron precursors*.
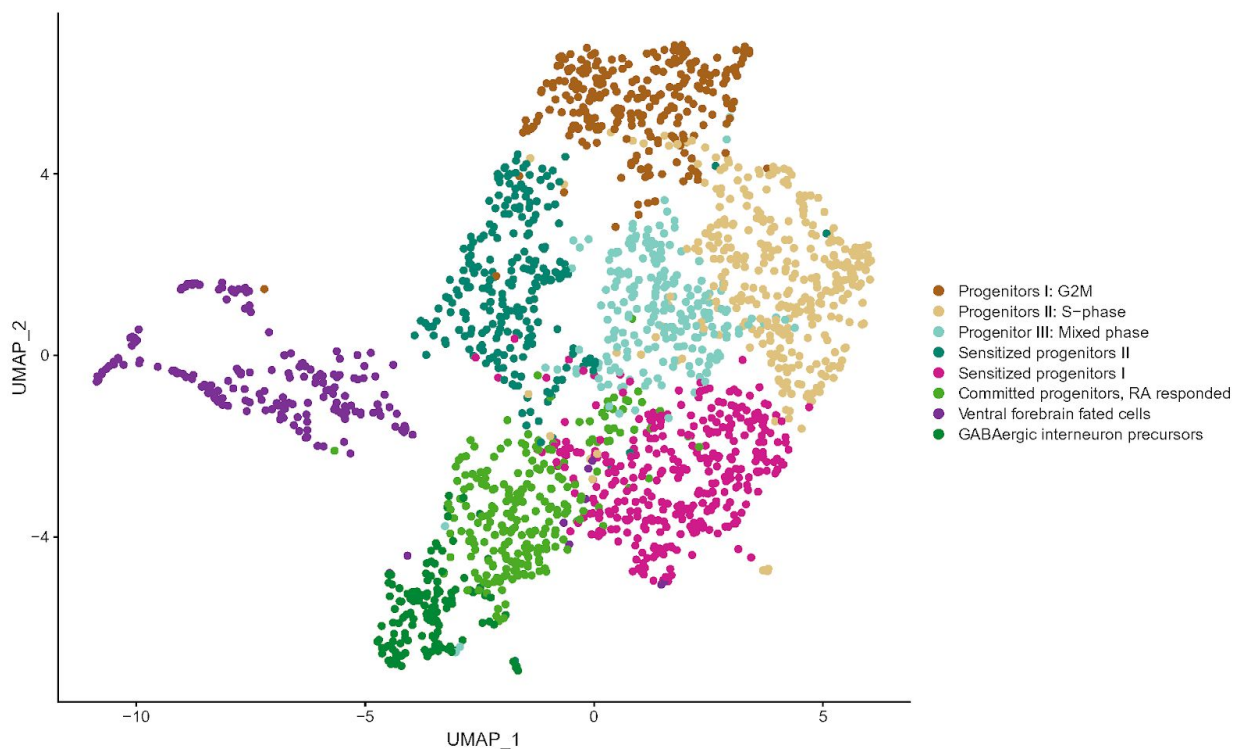
Figure 25 *Day 13 cell-types* The day 13 clusters were annotated as follows: cluster 2 as *Progenitors I: G2M*, cluster 0 as *Progenitors II: S-phase*, and cluster 5 as *Progenitor III: Mixed phase*. Then cluster 1 as *Sensitized progenitors I*, cluster 3 as *Sensitized progenitors II*, and cluster 4 as *Committed progenitors, RA responded*. And then cluster 6 as *Ventral forebrain fated cells* and cluster 7 as *GABAergic interneuron precursors*.

Day 20: End of Part III: Expansion

2838 cells were used in the analysis of day 20 after quality control and filtering (SI Day 20 Figure I, SI Day 20 Figure II). Day 20 is the final day in the protocol and is expected to display the highest heterogeneity. As indicated at day 13, ventral forebrain fate is likely at day 20 with GABAergic interneuron maturation, and possibly some layer-formation occurring. Remembering that this is just a monolayer of cells, we expect a mix of cells that is just reminiscent, and not as advanced in structure, of the emergence of the ganglionic eminences, in particular the medial ganglionic eminence which is one of the major migratory start sites of GABAergic interneurons. Again, in our case, we should expect much less of a clear-cut structure since we have a monolayer and not a more complex model such as an organoid.

Movement of differentiation

CytoTRACE prediction for day 20 (Figure 26. *Day 20 Differentiation prediction and end-of-differentiation markers* A.) show differentiation to proceed from the bottom-left, up, then right along a bend, and ending at a sharp tip. Then a disconnected cluster of cells is shown in the bottom-right where half is undifferentiated cells and the other half is differentiated cells. This cluster could be an effect of poor quality or might be a different lineage that will not fit the trajectory, it will be investigated further. Top 4 of end-of-differentiation markers predicted are *MLLT11*, *STMN2*, *DCX*, and *SOX4*, and they are plotted by mean expression organized in hexagons by Schex (Figure 26. *Day 20 Differentiation prediction and end-of-differentiation markers* B.). Mature neuronal marker *MAP2* is a common marker in many studies (Soltani et al., 2005) and has an average expression all-over, and interestingly expression is particularly strong at the bend on the end (Figure 26. *Day 20 Differentiation prediction and end-of-differentiation markers* C.). As seen in the differentiation prediction marker list (Figure 26. *Day 20 Differentiation prediction and end-of-differentiation markers* D.), predicted start-of-differentiation genes are *HMGN2*, *HMGB2*, *NUCKS1*, *HMGB1*, *SMC4*, *CENPF*, *NUSAP1*, *DEK*, and *CKS1B*. Predicted end-of-differentiation makers are *MLLT11*, *STMN2*, *DCX*, *SOX4*, *MIAT*, *TERF2IP*, *INA*, *SYT1*, *TAGLN3*, and *BASP1*.
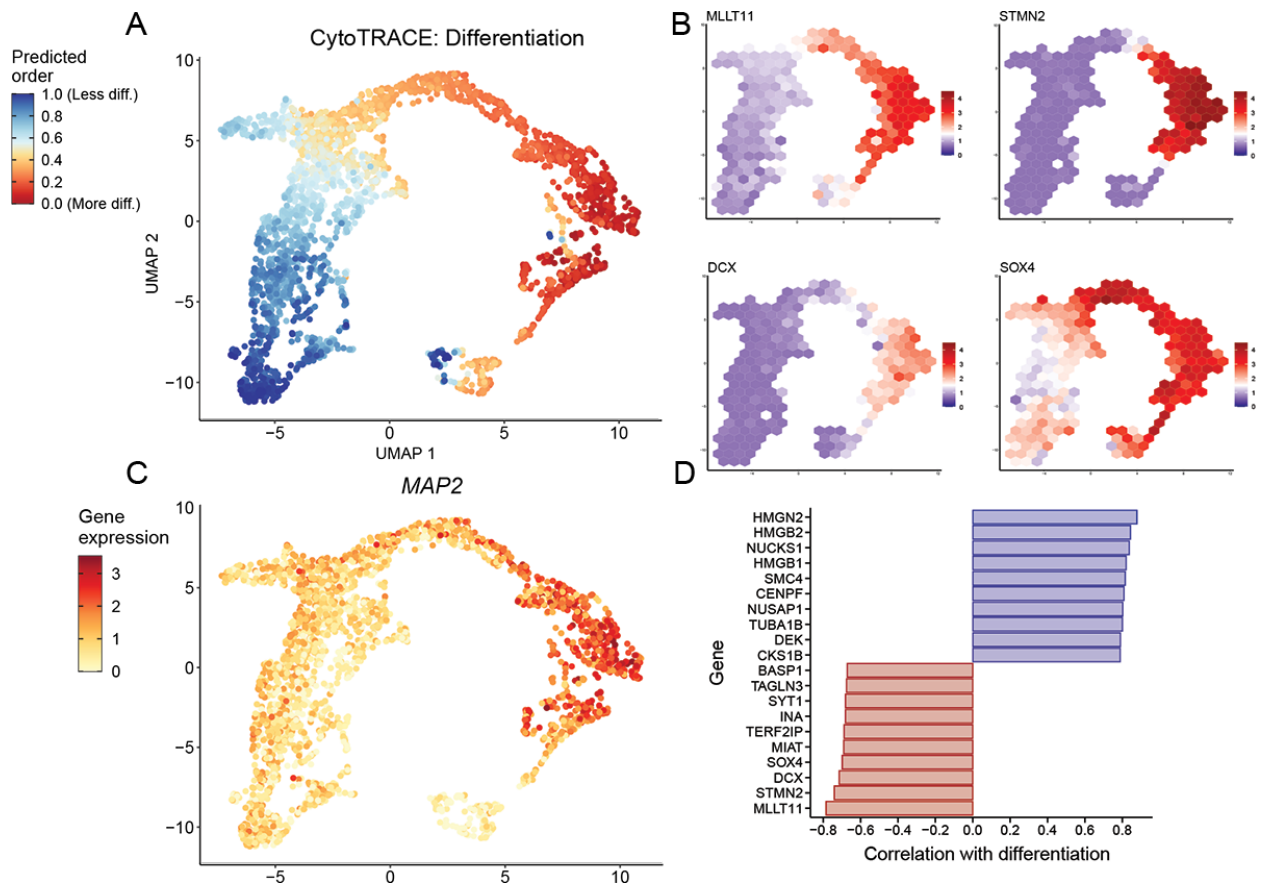
Figure 26 *Day 20 differentiation prediction and end-of-differentiation markers* A. Differentiation is predicted to move from bottom left, top, then right, ending in a tail. Notice the last cluster of cells at the bottom-right, cut off slightly more from the rest, which has some undifferentiated cells and some differentiated cells. B. Selected end-of-differentiation markers *MLLT11*, *STMN2*, *DCX*, and *SOX4* with mean expression and arranged as hexagons C. Mature neuronal marker *MAP2*, run through the cytoTRACE tool which shows *MAP2* expression increasing at the end of differentiation. D. Predicted start-of-differentiation markers are *HMGN2*, *HMGB2*, *NUCKS1*, *HMGB1*, *SMC4*, *CENPF*, *NUSAP1*, *DEK*, and *CKS1B*. Predicted end-of-differentiation makers are *MLLT11*, *STMN2*, *DCX*, *SOX4*, *MIAT*, *TERF2IP*, *INA*, *SYT1*, *TAGLN3*, and *BASP1*.

Shared markers at the start-of-differentiation are *HMGB2* (day 7) and *NUSAP1* (day 7 and day 13), *SMC4* (day 13), and *CENPF* (day 0 and day 13).

New start-of-differentiation markers at day 20 are *HMGB1*, *HMGN2*, *NUCKS1*, *CKS1B*, and *DEK*. *HMGB1* (High mobility group box-1) has been shown highly expressed in early brain development (Zhao et al., 2011) and *HMGB1* is important for survival and proliferation of NPCs (L. Wang et al., 2014). Similarly to *HMGB1*, *HMGN2* (High mobility group nucleosomal binding domain 2) is highly expressed in early brain development and in NPCs and is important for pluripotency (Garza-Manero et al., 2019). *NUCKS1* is important for cell-cycle progression (Ostvold et al., 1985), mentioned as a GWAS candidate for Parkinson's disease, and seems important for cell proliferation (S. Singh et al., 2019). *CKS1B* (CDC28 protein kinase regulatory subunit 1b) is active in dividing NPCs, and in the presence of FGF2, helps maintain cell proliferation (Darr, 2009). *DEK* is a chromatin remodeling gene that is important for proliferation and Wnt signaling The new genes in the list of top start-of-differentiation markers at day 20, together with the shared ones from the other time points, suggest strongly that this area is the progenitor zone which later could be annotated as *NPCs*.

Shared end-of-differentiation markers at day 20 with other time-points are *MLLT11* (share with day 13), *STMN2* (shared with day 13), *DCX* (shared with day 13), *SOX4* (shared with day 13), *MIAT* (shared with day 13), *SYT1* (shared with day 13), and *TAGLN3* (shared with day 13).

New end-of-differentiation markers at day 20 are *TERF2IP*, *INA*, and *BASP1*. The *TERF2IP* (TERF2 Interacting Protein), also known as RAP1, assists projection neurons, which have become multipolar, to move non-radially in migration to the cortical plate (nonpolar projection neurons move radially) (Jossin & Cooper, 2011). *INA* (Internexin Neuronal Intermediate Filament Protein), also called Alpha-Internexin, is a neurofilament part of the axo-skeleton and is suggested to help with intracellular transport to dendrites and axons (Kaplan et al., 1990; Schult et al., 2015). *BASP1* (Brain acid soluble protein 1) has been shown to promote axonal regeneration and neurite outgrowth (Shin et al., 2013) and have a role in synaptic function

(Behan et al., 2009). These new markers suggest a progressed maturation stage in comparison to other time points and suggest the emergence of projection neurons, migration of neurons, development of neurites, and synaptic function.

PAGA pseudotime

Similar to the prediction of how differentiation can be predicted moving through a population of cells, an analysis of pseudotime, with partition-based graph abstraction (PAGA), can be used to determine a trajectory of which time, or *pseudotime* rather, moves through a start towards an end (Wolf et al., 2019). This is also a good way to explore heterogeneity in cell populations, and can similarly to cytoTRACE help in the interpretation of complex differentiation processes. By using scanpy, a popular tool in python (Wolf et al., 2018), PAGA-analysis was performed on day 20 cells using the UMAP dimensional reduction to guide PAGA's build of pseudotime trajectory (Figure 28 *Day 20 PAGA solution and pseudotime*). In the PAGA-analysis cluster numbers (labels) are not the same as in Seurat or scClustViz (Figure 28 *Day 20 PAGA solution and pseudotime* A.), however, this label-number is irrelevant and does not inform on any order or direction, more importantly, the number of clusters is the same and the trajectory path is similar to cytoTRACE differentiation prediction. Behind the colored PAGA-graph is the grey silhouette of the underlying kNN graph, upon which the high-dimensional gene expression data is represented (connecting each cell in what appears almost to look like an intricate spider-web). Here, clusters obtained with the same Louvain algorithm that underlies Seurat and scClustViz clustering as well. Here, the clusters are colored nodes in the PAGA-graph, manually placed directly in front of the grey web. Below the PAGA-graph (Figure 28 *Day 20 PAGA solution and pseudotime* B.) is the pseudotime trajectory where the start of *pseudotime* is the same as the predicted start of differentiation by cytoTRACE. We then follow pseudotime along the same general path but observe some more details of how PAGA connects clusters along the way. Before displaying the PAGA-figure, here is the UMAP showing clustering done by Seurat and scClustViz (Figure 27 *Day 20 clusters*), which is the same UMAP data used to guide PAGA and cytoTRACE (however without transferring the same label-numbers which unfortunately makes the figures a bit less pedagogic).
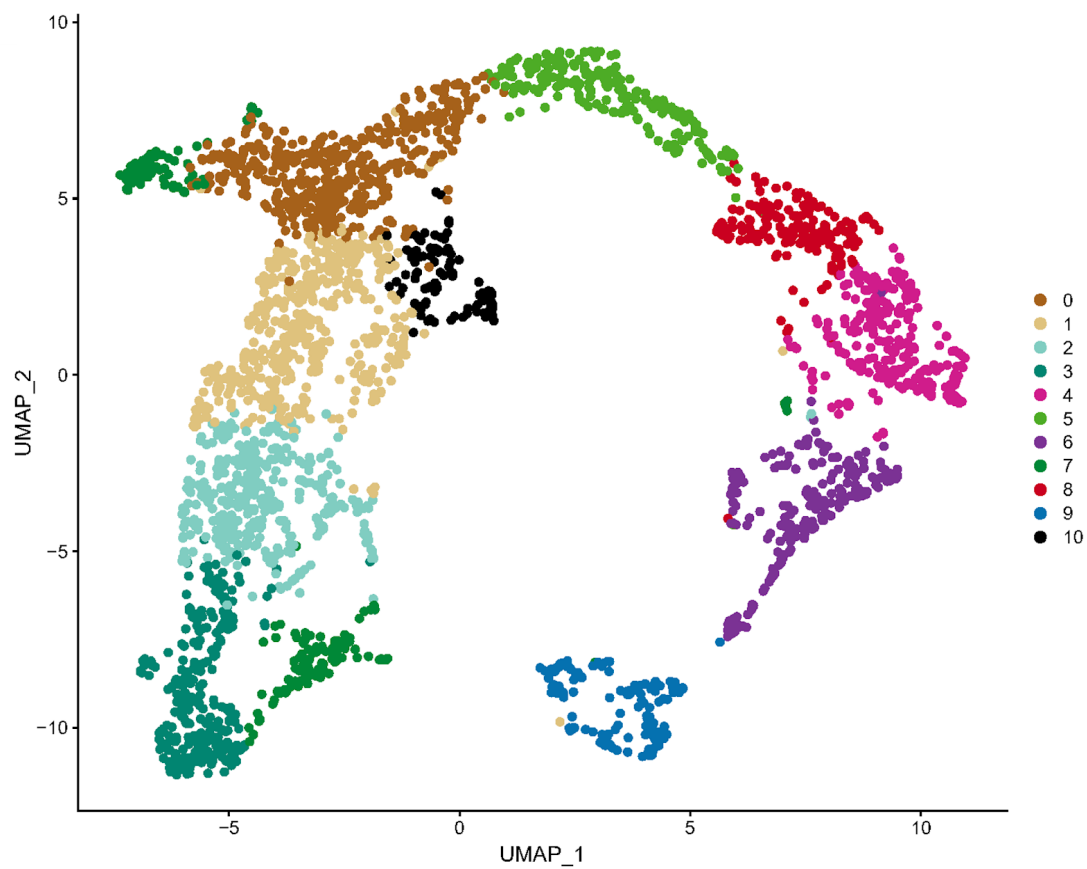
Figure 27 *Day 20 clusters.* scClustViz clustering solution resolved 11 clusters at day 20, presented here on a UMAP via Seurat.
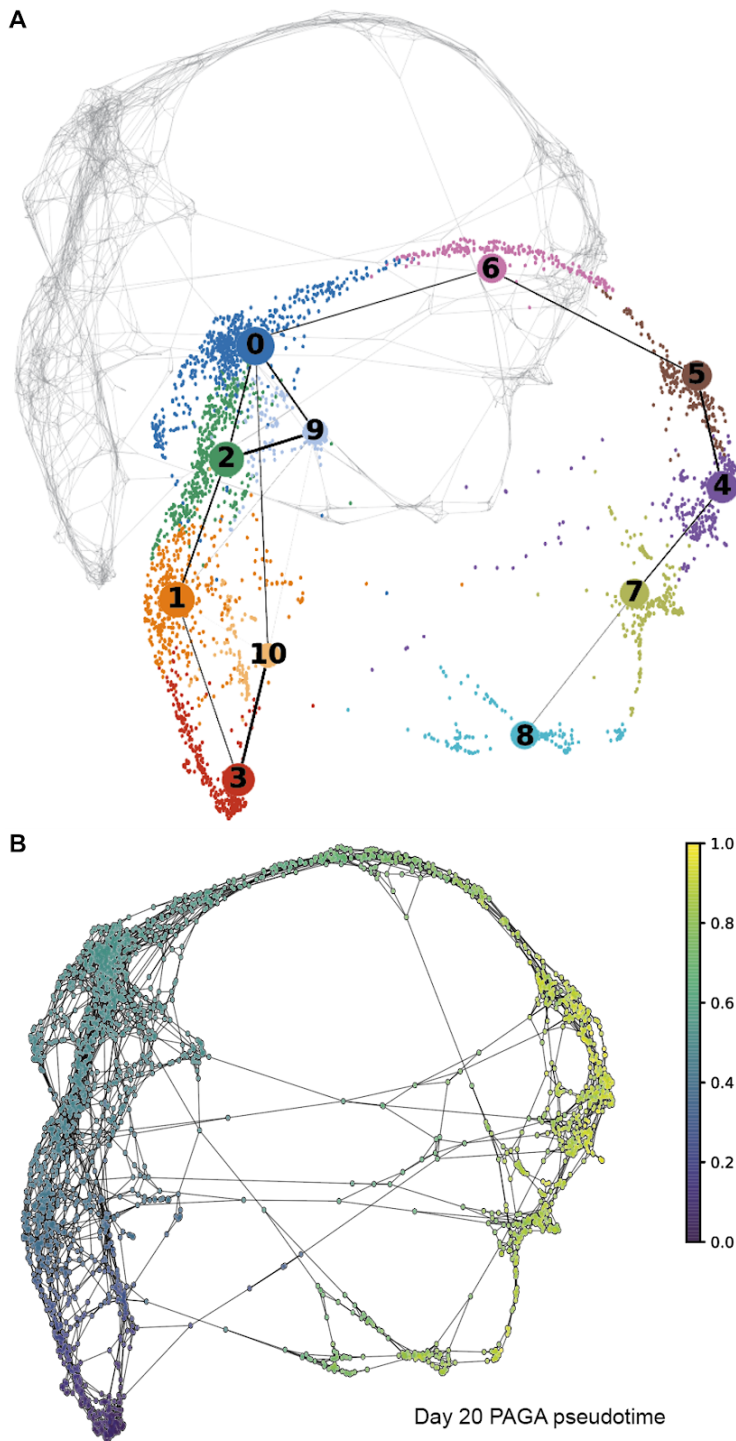
Figure 28 *Day 20 PAGA-graph and pseudotime* A. PAGA solution for day 20. Colored circles with a number represent clusters, as they were determined by the Louvain algorithm (J. Zeng & Yu, 2018). Behind the colored PAGA-graph is a grey silhouette that represents the underlying kNN graph, which the high dimensional data is ordered on. B. The result of PAGA pseudotime prediction is very similar to differentiation prediction, in that predicted *pseudotime* starts in the bottom-left part of the map then it moves in the same general trajectory. Here, additional details on cluster relationships are shown by using PAGA, in comparison to cytoTRACE where we just see the color gradient.

Investigation of markers to determine cell-types at day 20

The marker investigation at day 20 is expected to more complex than previous time points. scCatch at day 20 predictably reveals more heterogeneity in the cell population in comparison to previous time points (Table 7 *Day 20 scCatch cell-type prediction*). Cluster 0, 1, 2, and 3 and 10 are predicted as *Neural Stem Cells*. However, Cluster 2, 3, and 10 predictions are based on just 2-3 genes per cluster. Clusters 4 and 6 are predicted to be *Neurons* and among the markers here we recognize *MAP2*, *SYP*, and *RELN*, as well as several *DLXX* types of genes for cluster 6, suggesting this is a GABAergic interneuron cluster. Cluster 5 and 8 are predicted to be *Astrocytes* based on a long list of markers so these will need some investigation. Cluster 7 is the only cluster predicted to be *NPCs* which, considering markers *ASPM*, *AURKB*, *BIRC5*, *CCNA2*, *CDC20*, and *CENPF*, as used earlier to annotate progenitors, seems to be an accurate prediction. Neighboring clusters are likely also within the NPCs category, but this will need some investigation. Finally, cluster 9 is called a *Microglial cell* cluster and corresponds to the bottom-right cluster that was split in half in differentiation prediction on day 20 (Figure 26 *Day 20 differentiation prediction and end-of-differentiation markers*).

Table 7 *Day 20 scCatch cell-type prediction.* **For day 20**, there are multiple marker genes predicted as cell-type-specific markers by scCatch. These markers will be discussed in the main text.

Cluster	Cluster markers	Cell type	Pred. Score	Cell type markers	PMID
0	TUBB, VIM, ID4, MALAT1, C1orf61, CALM2, UBE2T, LINC01551, NKX2-1, AL139246.5, TUBA1A, UBE2S, SOX9, SHISA2, TUBB4B, TFDP2, HNRNPA1, NPM1,	Neural Stem Cell	0.94	SOX2, SOX9, ADGRG1, DCX, NCAM1, DPYSL3, HMGB2	28795331, 16797497, 19525879, 21275797, 28771884
1	MCM6, MCM5, PCNA, MCM3, MCM4, GINS2, CDCA7, STMN1, CDC6, UHRF1, H3F3B, CLSPN, UNG, HELLS, NASP, DTL, DUT, UBC, WDR76,	Neural Stem Cell	0.87	NES, SOX9, DCX, NCAM1, DPYSL3, HMGB2	28795331, 16797497, 25843934, 21275797, 28771884
2	HIST1H1B, HIST1H4C, PCLAF, TYMS, RRM2, MYBL2, RAD51AP1, TK1, CDCA5, RANBP1, CENPM, HIST1H3B, CDK1, HIST1H1D, ESCO2,	Neural Stem Cell	0.83	NES, HMGB2	28795331, 25843934, 28771884
3	PIF1, CENPA, PLK1, AURKA, DEPDC1, CDCA3, KIF14, CDCA2, FAM83D, TTK, KIF20A, CKAP2L, CDCA8, SGO2, NDC80, BUB1, DLGAP5, PSRC1,	Neural Stem Cell	0.93	NES, SOX2, HMGB2	28795331, 16797497, 25843934, 21275797, 28771884
4	SLC17A6, SNCG, NEFL, SHISA6, TTC9B, NEFM, ONECUT3, MAPT, NTRK1, SCN3B, STMN4, L1CAM, STMN2, C1QL1, GAP43, INA, KLHL35, RTN1,	Neuron	0.88	APBB1, MAP2, SLC17A6, ENO2, SYP, PPP2R2B, DPYSL5, NFASC, RELN	7867517, 24142904, 17651422, 27587997, 29539641, 26060301
5	NEUROG1, RGS16, GADD45G, TLX3, MFNG, DLL3, SRRM4, CDKN1C, DLL1, HES6, INSM1, CKB, ZBTB18, KIF19, OS9, TAGLN3, FEZF2, KLHDC8A,	Astrocyte	0.85	ANLN, BTG2, CDC42EP4, DKK3, GLUL, IER2, L1CAM, PEA15, PON2, SCG2, SCG3, SSTR2, ADD3, AIF1L, BDH2, BTBD17, DDR1, DHCR7,	27806376, 29539641, 26060301
6	DLX2, DLX6-AS1, DLX1, ARX, DLX5, GAD2, SLC32A1, VWC2, GAD1, GRIA2, SPOCK1, TAC3, DLX6, MIR7-3HG, GNAS, KIF5C, MAP1B, RBP1,	Neuron	0.86	MAP2, ENO2, SYP, ARX, PPP2R2B, DLX1, DLX2, DLX5, DLX6, NFASC, SST	24142904, 27587997, 29539641, 26060301
7	CDC20, CCNB2, PTTG1, BIRC5, HMGB2, TPX2, NUCKS1, CCNB1, CENPF, HMMR, HMGN2, ARL6IP1, H2AFZ, DLGAP5, CENPE, HMGB3, KIF4A,	Neural Progenitor Cell	0.86	ADGRG1, DCX, NCAM1, DPYSL3, ANP32E, ASPM, AURKB, BIRC5, BUB1, CCNA2, CCNB1, CCNB2, CDC20, CDCA3, CDKN3, CENPF, CENPU,	19525879, 21275797, 29539641
8	CHRNA3, LHX9, GNG8, DCC, POU3F1, NHLH1, KLHL35, NTRK1, NEUROD1, ST18, ELAVL4, DACH1, POU2F2, INSM1, STMN2, TPPP3, ISL1,	Astrocyte	0.86	CPE, FTH1, IER2, JUND, KIF21A, L1CAM, RND3, SCG2, SCG3, SPON1, SSTR2, ADGRG1, BTBD17, DNER, ELAVL4, GNG5, HLA-C, HMGCS1,	27806376, 29539641, 26060301
9	DLX5, CLIC5, DLX6, FGF8, DLX6-AS1, DIRAS3, ARHGAP36, PNOC, TMEM132B, RIPOR2, GRHL3, IGSF1, SELENOP, AL109807.1, GDF10, EDN3,	Microglial Cell	0.86	APLP2, APOE, B2M, C20orf27, CD164, CD63, CD83, CMTM6, CMTM7, CSRP2, CTSC, CTSD, CYBA, DAB2, FBLN1, FXYD6, GLUL, GNB4, GNG5,	29539641, 26060301, 29262845
10	FGF17, WNT7B, FGF8, TRH, SMS, PEG10, PRSS23, COL2A1, SMOC1, EFEMP1, FBLN1, MAF, TTYH1, UNC13C, L1TD1, BOC, SFTA3, PCDH19, RPL21,	Neural Stem Cell	0.83	SOX1, NCAM1, HMGB2	16797497, 21275797, 28771884

Cluster analysis to determine cell-types

For dot-plot heatmaps (Figure 29 *Day 20 dot-plots*), analysis at day 20 will be done group-wise, using the clusters that are grouped together in the dendrogram, visible in the leftmost part of the figure, similar to what was done for day 13.

RGs - Radial glial, the cycling/renewing cells

Cluster 3 was predicted as start-of-differentiation and start-of-pseudotime in PAGA analysis. Inspecting Figure 29 A. and B., cluster 3 has stronger expression of *PIF1*, *CENPA*, *PLK1*, *AURKA, CDC20, CCNB2*, and *DEPDC1*, compared to cluster 2. Cluster 2 is also after cluster 3, in terms of position in differentiation movement and pseudotime. And after cluster 2 on relative position on the time-axis, comes cluster 1, which has have lower levels of *CCNB2*, *BIRC5*,

73

RRM2, *CDC20*, *PIF1*, *CENPA*, *PLK1*, *AURKA*, and *DEPDC1,* compared to the first clusters 3 and 2. As almost all of these markers are mentioned in relation to progenitors that are rapidly renewing and cycling, and these markers are decreasing as we move forward in differentiation and pseudotime, it suggests that these first clusters are progenitors, where some are slowing down in cell-cycle, to become ready for the commitment to differentiation. At this point in time progenitors can also be called RGs (radial glial cells), which would be a general name for the progenitor's pool at this timepoint. Minichromosome maintenance complex genes, e.g. *MCMC2*, are useful markers for detecting slowly cycling putative progenitors (Maslov et al., 2004), and in cluster 2 wee see *MCM3* expressed higher compared to starting cluster 3, while *TOP2A*, *CENPF,* and *GMGB2* is decreasing in expression. Cluster 2 seems thus to be slowing down. In cluster 1, *MCMCX* genes increase in expression, e.g. *MCMC2*, *MCM5*, and *MCM6*, while *NUSAP1*, *TOP2A*, *UBE2C*, *HMGB2*, and *MKI67* drop expression level. Again suggesting slow down of cell-cycle and slow down in renewal processes. These initial three clusters could thus be annotated as *RG*, *RG II: slowing down*, *RG: ready to commit*. Alternatively, the first two clusters can be merged and named as *RGs*, while the third one is called *RG II: stopped to commit*.
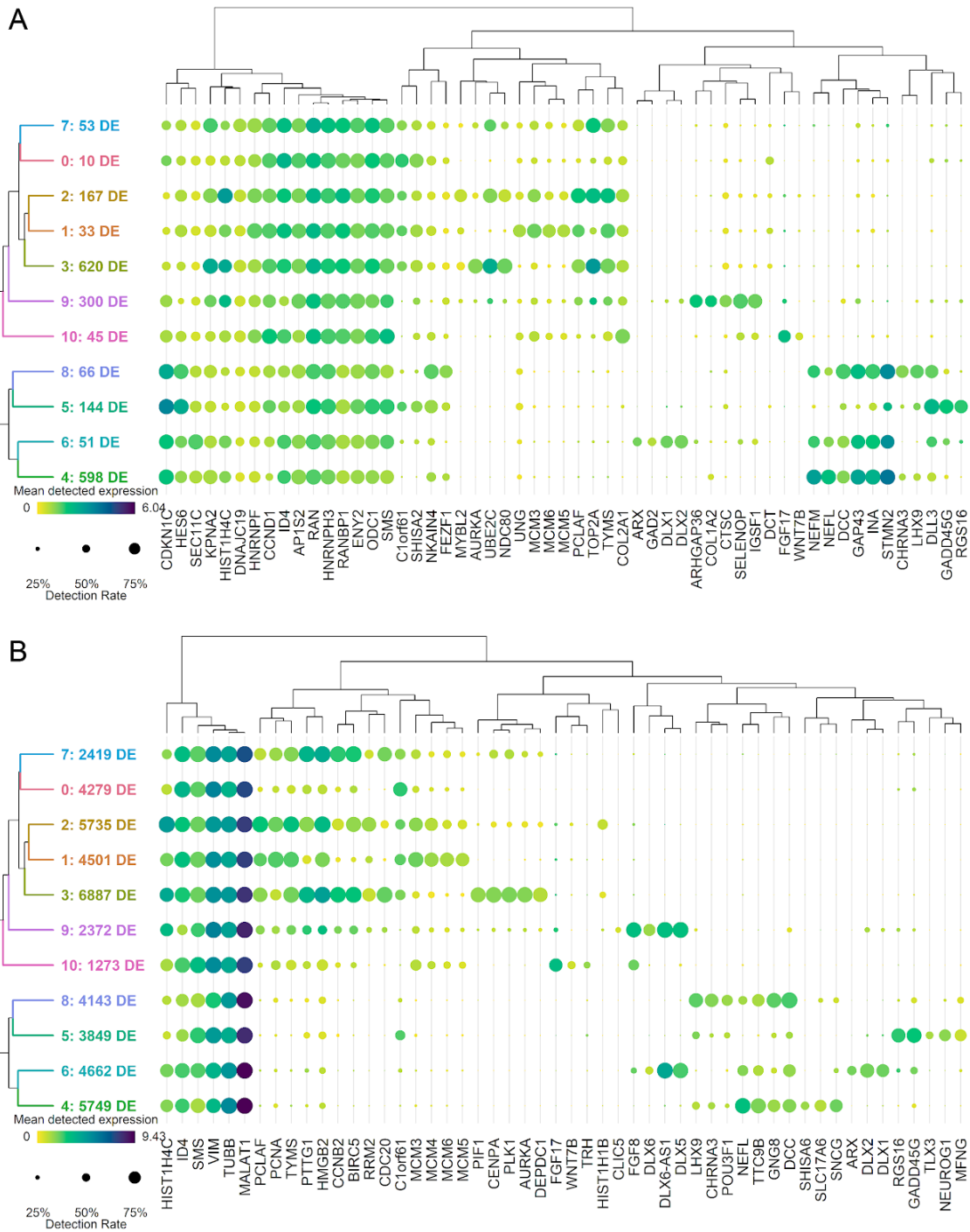
Figure 29 *Day 20 dot-plots.* Visualizing cluster differential expression of marker genes using pairwise tests between one cluster and all other clusters. A. and B. gene markers are analyzed in the main text to their complexity on day 20. Starting from the top on the left side of the dot-plot A. and dot-plot B., dendrograms are the same, and they with progenitor-like clusters 2 *NSCs*, cluster 1 *NSCs*, and cluster 3 *NSCs* on the parent branch. We then have cluster 7 *NPCs* and cluster 0 *NSCs* grouped. Next, we have cluster 9 *Microglial cells* which will be regarded as an independent cluster for now. Then we have cluster 10 *NSCs* and this cluster is the black-colored side-cluster, just before the top next to cluster 0 (Figure 27 *Day 20 clusters*).

HES5 is a good marker for Notch-signaling at this timepoint (Basak & Taylor, 2007; Ohtsuka et al., 1999; Ziller et al., 2014) (Figure 30 *HES5 - a marker for Notch-signalling*), and Notch-signaling fades within the middle of cluster 0, just after *committed* RG in cluster 1.
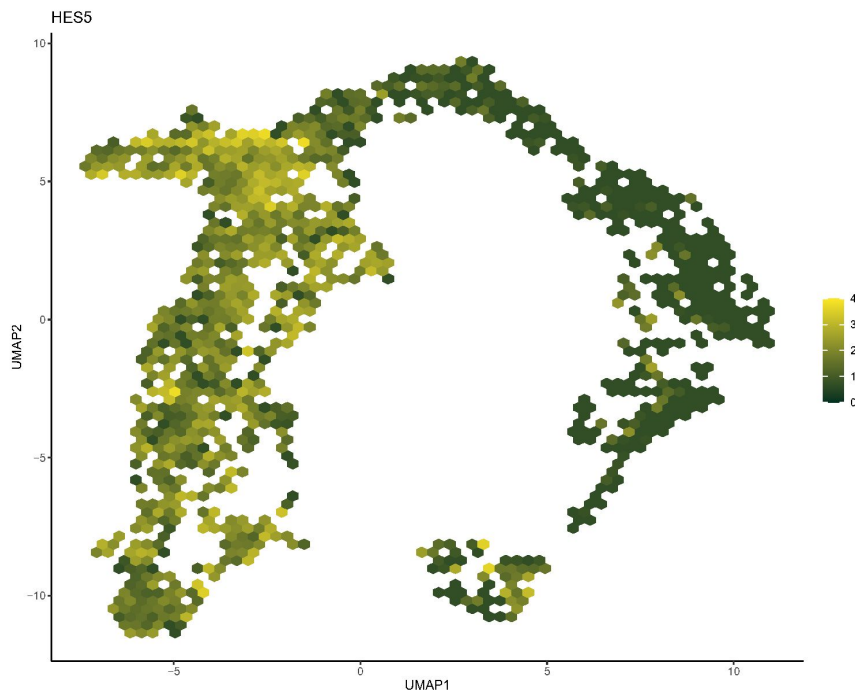
Figure 30 *HES5 - a marker for Notch-signalling*. Expression of *HES5* is higher in progenitor-like clusters, overall expression pattern of *HES5* then fades out around cluster 0. Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 4.

Where Notch-signaling fade, intermediate progenitors (IPs) accumulate, this is the commitment of RG to the neural fate (Mizutani et al., 2007; Nelson et al., 2013; Yoon et al., 2008). At this point, cells can either re-enter cell-cycle or proceed to become neural progeny. As cluster 0 is situated right at this boundary, with flanking cluster 7 (half of cluster 7, more on that later) and cluster 10, and then cluster 5 ahead, straight on the differentiation path and pseudotime trajectory, it suggest that cluster 0 could be annotated as *Intermediate progenitors, IPs*.

<u>Layer V precursors</u>

After the IPs, comes cluster 5. As it is next in line on differentiation prediction and pseudotime it could be expected that it has mature markers indicating neural progeny formed from the IPs. Indeed, proneural marker *ASCL1* is expressed strongly in this cluster (E. J. Kim et al., 2007, 2008) and also at the end of the trajectory, at the very tip (Figure 31 *Proneural marker ACSL1*).
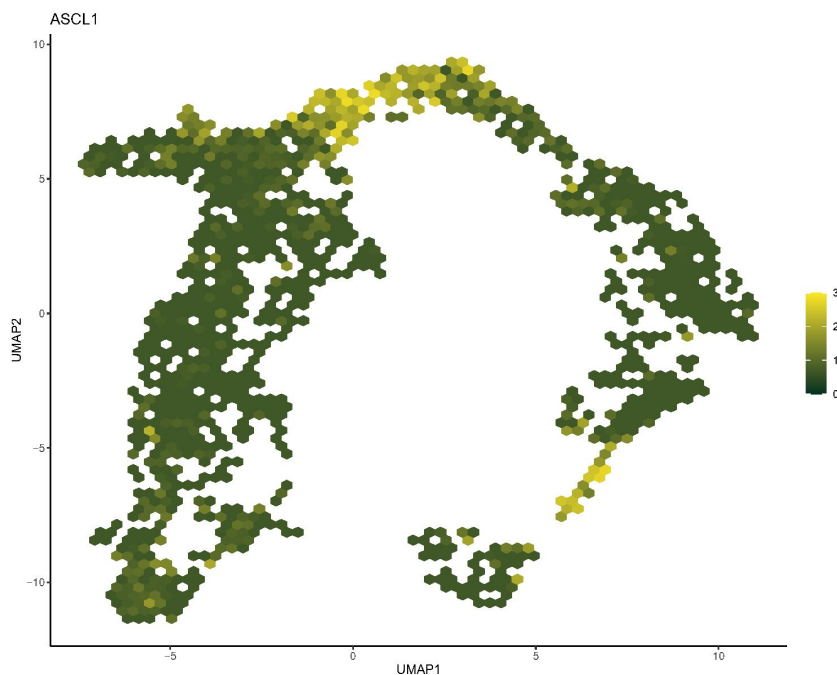
Figure 31 *Proneural marker ACSL1*. Expression of *ASCL1* seen at the end of cluster 0, into cluster 5. Expression can also be noticed at the end of cluster 6. Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 3.

Interestingly, *FEZF2* (Figure 32 *Layer V marker FEZF2*), which is a marker for neuronal subtype differentiation, patterning of forebrain and olfactory, and deep layer V neurons, are expressed strongly here together with *ASCL1* (Castelo-Branco et al., 2003; Hirabayashi et al., 2004; Roth et al., 2010; Shimizu et al., 2010; Tantirigama et al., 2014). Early born projection neurons first occupy the deepest layers, VI, V, and IV, as neurons grow in an "in and out" sequence, and then later-born neurons occupy more superficial layers, III, II, and I (Götz & Huttner, 2005)**.** The spinal cord, midbrain, and hindbrain will have main projections coming from layer V, while layer VI neurons projects towards the thalamus. Layer II and III, together

with a portion of neurons from layer V, will form intracortical projections (O'Leary & Sahara, 2008). Neurons that populate deep layers are Notch-signaling dependent and would appear as this signaling fades, as reported for neuroepithelium and early-RG (Edri et al., 2015). Cluster 5 might then be annotated as *Layer V precursors*, but more evidence for this layer might be needed since *FEZF2* also control forebrain and olfactory patterning, and is not just a marker for layer V - where we are in relative *in vivo* development would be a good clue but we must remember we are still in a monolayer culture so we can also expect a mix of cells that are not directly comparable.
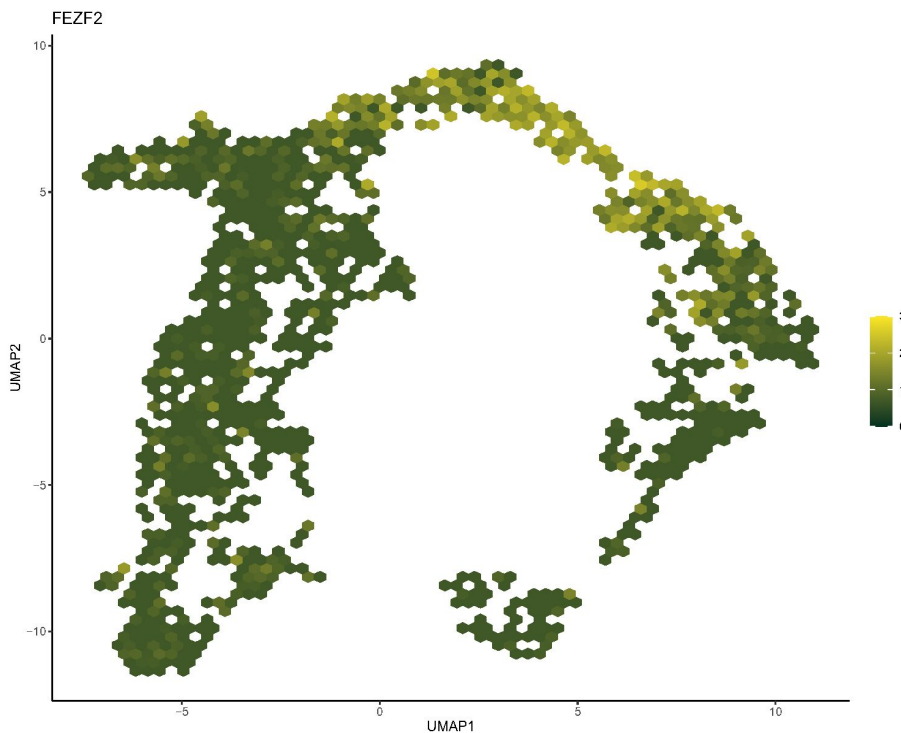
Figure 32 *Layer V marker FEZF2.* Expression measured primarily in cluster 5 where expression fades out when moving into cluster 8. Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 3.

Layer IV/5-HT precursors

For Cluster 8, *FEZF2* expression is seen here as well but decreasing towards the right flank. The boundaries on the sides of previous cluster 5 *layer V precursors* are interesting since we have the end of Notch-signaling on the left flank and *FEZF2* expression fading on the right flank, into

cluster 5, suggesting that we observe patterning. *FEZF2* inhibition would cause postmitotic neurons to go into apoptosis, and *FEZF2* promotes Wnt/β-Catenin signaling in the forebrain by repression of *LHX9* and *LHX2*, which are negative regulators of Wnt signaling (Siwei Zhang et al., 2014). Interestingly, *LHX9* expression is seen increased just on the right flank of cluster 5 where *FEZF2* expression is dropping in strength (Figure 33 *Wnt repressors LHX9 and LHX2)*. *LHX2* expression is strong across the RG clusters that also have strong Notch-signaling then it decreases in strength as Notch-signaling decreases and is much lower at the left flank of cluster 5.
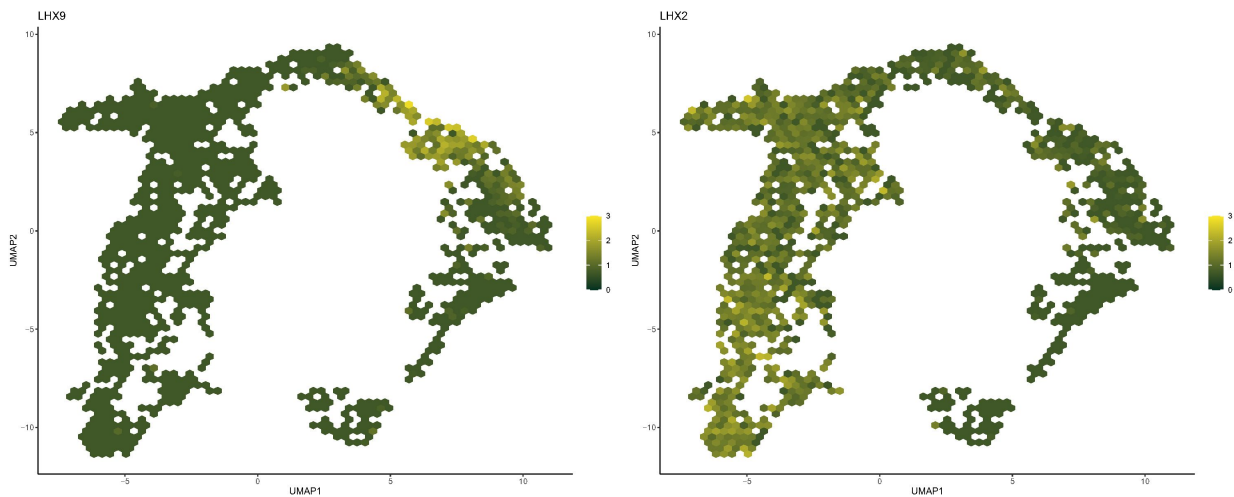
Figure 33 *Wnt repressors LHX9 and LHX2.* Expression of LHX9 (left) seen primarily in cluster at end of cluster 5, and clearly seen in cluster 8, when it after this cluster fades out. LHX2 (right) has an expression in the same areas as *HES5*, and fades out to lower levels around cluster 0. Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 3.

Wnt/β-Catenin signaling play important roles in patterning in the forebrain and a low presence of this signal is required for telencephalon emergence whereas the diencephalon would see a relatively higher signal, together with *FGF* and *SHH* expression (Wilson & Houart, 2004). *SHH*, *FGF1*, or *FGF2*, are not detected at any substantial level (Figure 34 *SHH and FGF2*).
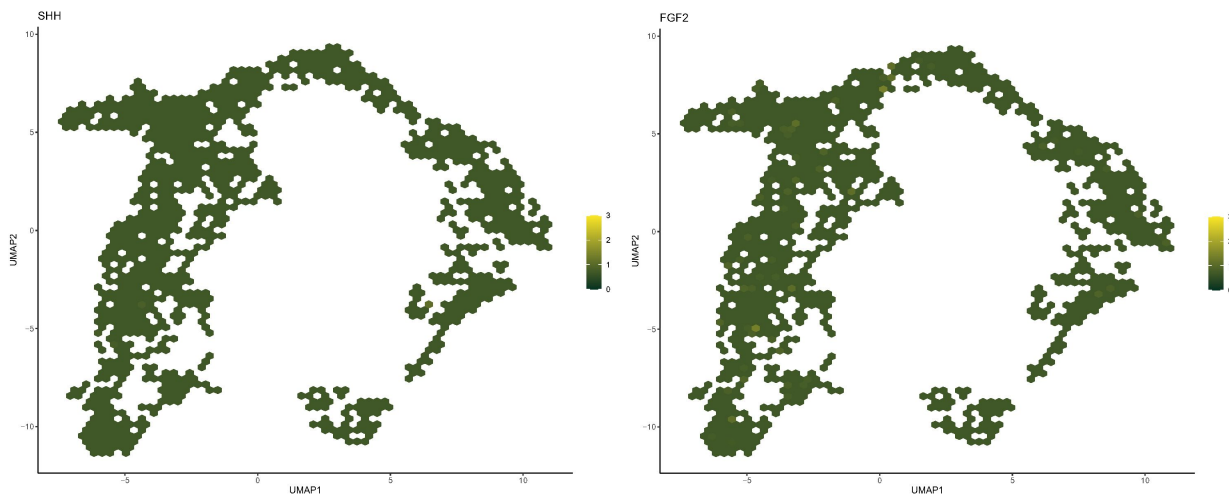
Figure 34 *SHH and FGF2.* These two markers are not detected at any significant levels. Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 4.

An interesting feature of cluster 8 is a high expression of *GAP-43* (Figure 35 *Volcano plot comparing expressed genes between clusters 5 and 8*, Figure 36 *GAP-43 expression on day 20*) when compared to cluster 8. *GAP43* is very important in the development of forebrain serotonergic innervation, and *GAP43*/5-HT is particularly dense in layer IV (Donovan et al., 2002). *In vivo* serotonergic neurons (5-HT) develop very early and as is commonly thought, abnormalities with 5-HT signaling could contribute to causing depression, schizophrenia, depression, and other mental diseases. For now, cluster 8 will be annotated as *layer IV/5-HT precursors*.
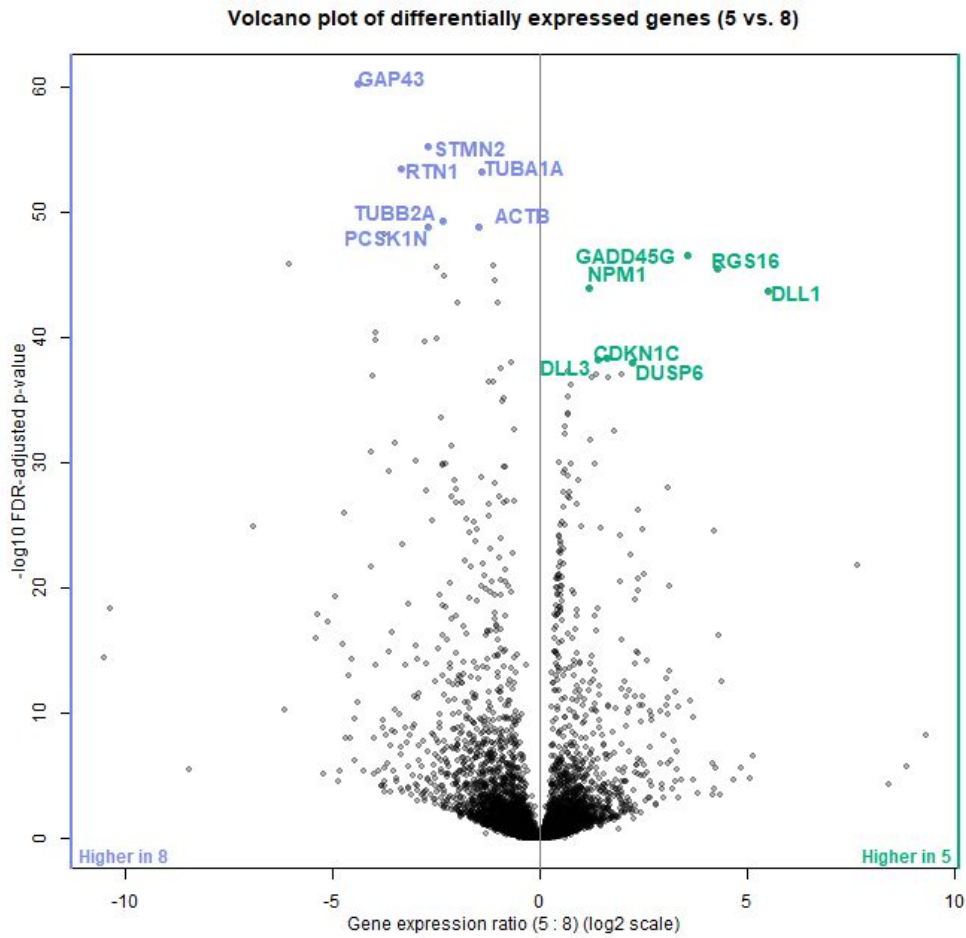
Figure 35 *Volcano plot comparing expressed genes between clusters 5 and 8*. Genes with higher relative expression in cluster 8 are *GAP43, STMN2, RTN1, TUBA1A, TUBB2A, ACTB, and PCSK1N*. Genes expressed higher in cluster 5 are *GADD45G, RGS16, NPM1, DLL1, CDKN1C, DLL3,* and *DUSP6*. Axes are presented by scClustViz, -log10 FDR-adjusted p-value on y-axis, and gene expression ratio in log2 scale on the x-axis. FDR = false discovery rate.
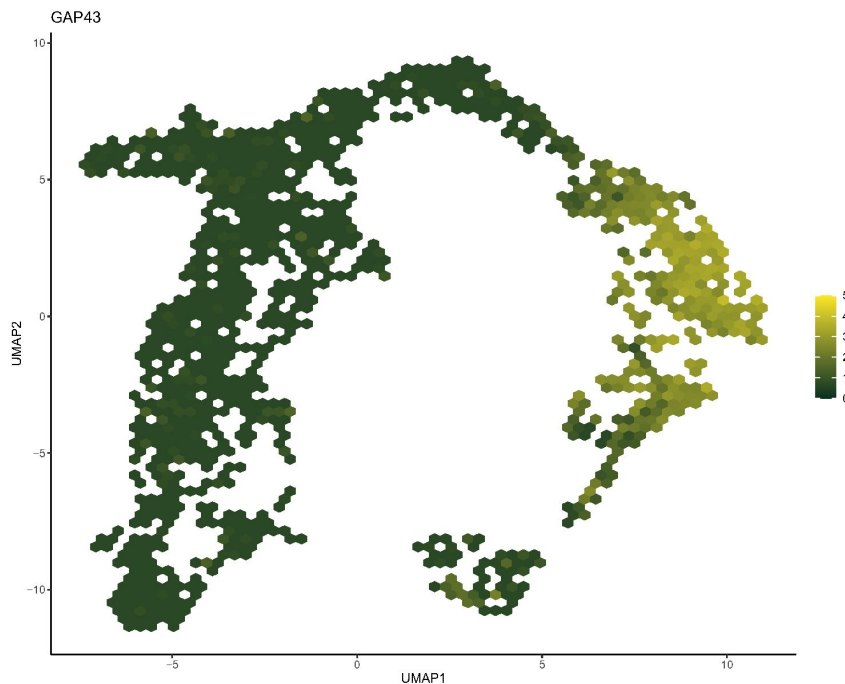
Figure 36 *GAP-43 expression on day 20.* Expression of *GAP43* increases rapidly halfway into cluster 8, is strong for cluster 4, then fades into cluster 6. Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 5.

5-HT/Chol/Glut precursors

Cluster 4 markers from the dot-plots (Figure 29 *Day 20 dot-plots* B.) show a bit higher prevalence of *SHISA6*, *SLC17A6*, and *SNCG*, then *NEFL*, *TTC9B*, and *GNG8*, for cluster 4 in comparison to subsequent cluster 6 which show markers *DLX6-AS1* and *DLX5* suddenly high. According to GeneCards, *SHISA6* is involved in the maintenance of synaptic transmission and regulates AMPA-type (a glutaminergic ion channel, there is also kinic acid- and N-methyl-D-aspartic acid- (NMDA) channels) glutamate receptor (AMPAR) keeping the glutamate receptor activated in presence of glutamate. For *SCL17A6*, UniProt writes that *SLC17A6* assists in glutamate uptake in glutamate transmembrane transporter activity. *SNCG* (γ-synuclein) is important for signal-transduction (Liang et al., 2015) such as Akt and mTOR but also seems to regulate serotonin transporter function (Falck & Hillarp, 1959; Wersinger & Sidhu, 2009). *NEFL*, Neurofilament light, is an important component of vertebrate axons (Yum et al., 2009). *TTC9B* has just 8 publications in GeneCards and most are related to postpartum depression (Kaminsky & Payne, 2014; Osborne et al., 2016). The gene encodes for a protein

82

called Tetratricopeptide Repeat Domain 9B and seems very unexplored. GTEx analysis shows that it seems to have a very brain-specific expression where levels are the highest in the cerebellum and frontal cortex (Figure 37 *GTEx gene expression results for TTC9B*). The specificity to these particular areas of the cell populations, roughly overlapping *GAP-43*/5-HT expression area (Figure 36 *GAP-43 expression on day 20*, Figure 38 *TTC9B expression on day 20*), and the limited knowledge in literature about this protein, would make it an interesting protein to investigate in future experiments. As for annotation, cluster 4 could be named *5-HT/Chol/Glut precursors*, it also seems that in an area such as this (relative to where we might be in neural development) it is common with a mix of glutaminergic, GABAergic (seen clearly for next cluster, cluster 8), and cholinergic fibers and glutaminergic input seems to have an excitatory effect on cholinergic neurons (Hur et al., 2009).
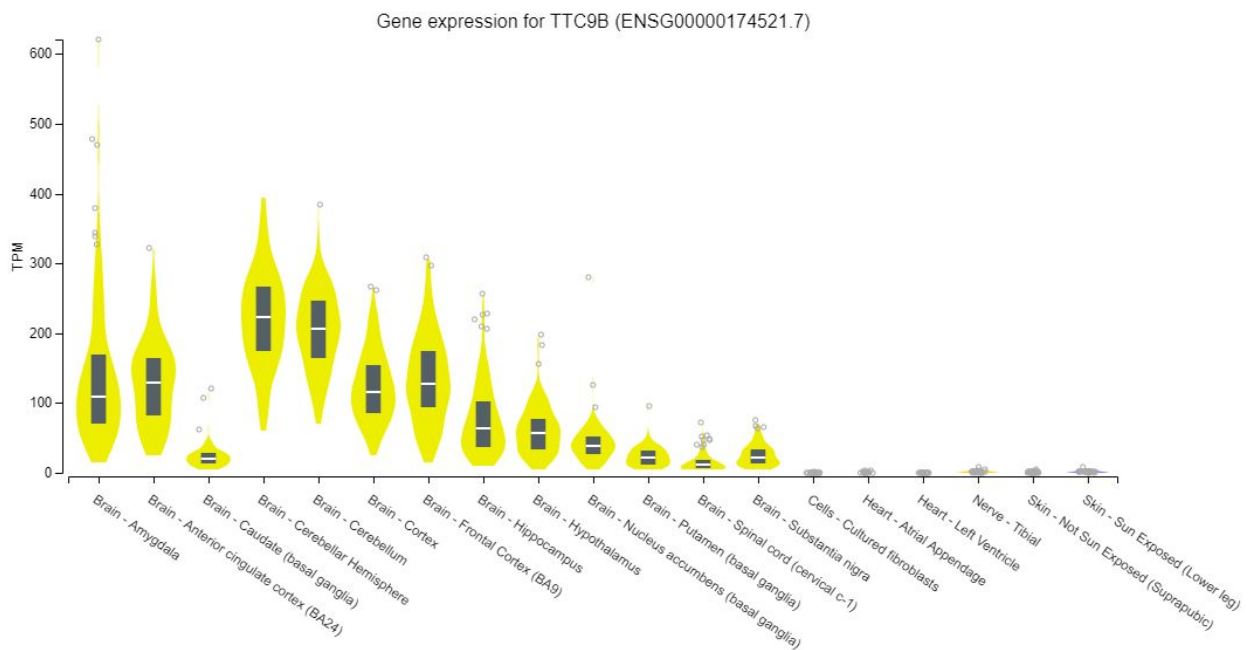
Figure 37 *GTEx gene expression results for TTC9B* (ENSG00000174521.7). TTC9B has a brain-specific expression, where levels are the highest in the cerebellum and frontal cortex.
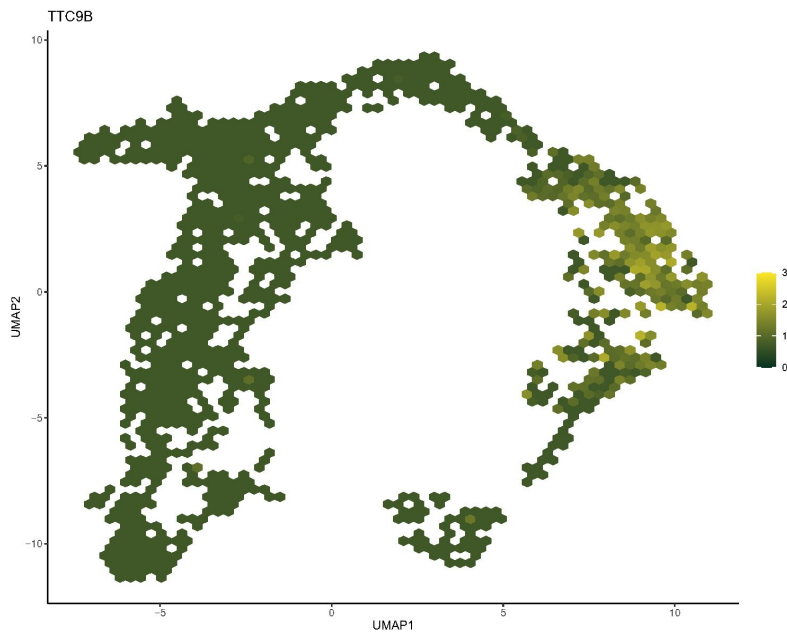
Figure 38 *TTC9B expression on day 20.* TTC9B is expressed primarily in cluster 4. Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 3.

For cluster 6, there is an expression of *DLX2*, *DLX1*, *DLX6-AS1*, *DLX5*, *ARX*, and *GAD2* in cluster 6 (Figure 39 *GABAergic markers DLX2, DLX1, DLX6-AS1, DLX5, ARX, and GAD2*). These markers are strong indicators of GABAergic interneurons. Additionally, there is strong neural maturation *DCX* marker expression and expression of *PCP4* (Purkinje cell protein). A fitting annotation for cluster 6 would then be *GABAergic interneurons*, however, to attribute full neuronal functionality to these, other experiments will have to be made, such as patch-clamp.
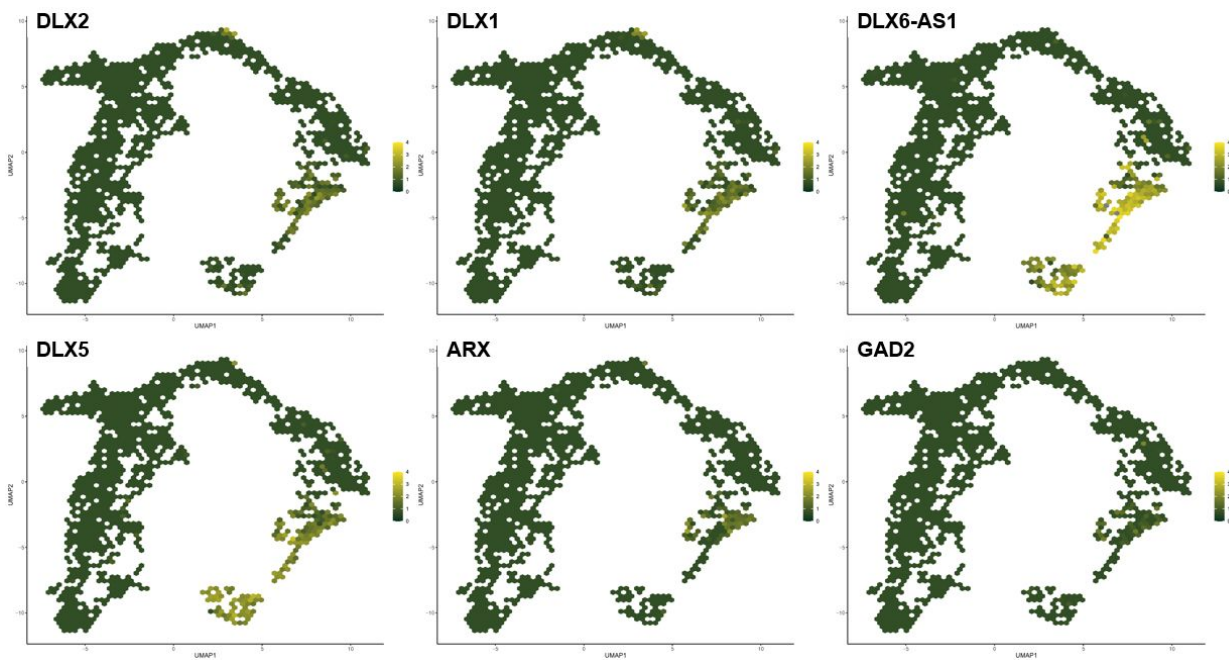
Figure 39 *GABAergic markers DLX2, DLX1, DLX6-AS1, DLX5, ARX, and GAD2*. Markers *DLX2, DLX1, and ARX* are expressed in cluster 6. *DLX6-AS1* is expressed strongly in cluster 6, and also into cluster 9. *DLX6* increase in expression once in cluster 6, then get stronger and ends at cluster 9. *GAD2* has a slight expression in cluster 6. Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 4.

Interneurons fated for developing thalamus, corticothalamic precursors (CTPs)

Cluster 9 is the disconnected island at the bottom of the d20 UMAP (Figure 27 *Seurat and scClustViz clustering day 20*). Markers for cluster 9 are *ARHGAP36*, *COL1A2*, *CTSC*, *SELENOP*, *IGSF1* and share *DLX5* and *DLX6-AS1* with cluster 6 *GABAergic interneurons*, but has a relatively stronger expression of *FGF8* (Figure 29 *Day 20 dot-plots visualizing cluster differential expression of marker genes using pairwise tests between one cluster and all other clusters*). Comparing differentially expressed genes between the GABAergic interneuron cluster 6 and this cluster 9 (Figure 40 *Volcano plot of differentially expressed genes between cluster 9 and 4*) show that *SPARC*, *ID3*, *VIM*, *FSTL1*, *NPM3*, and *GSTP1* is expressed higher when compared to cluster 6.
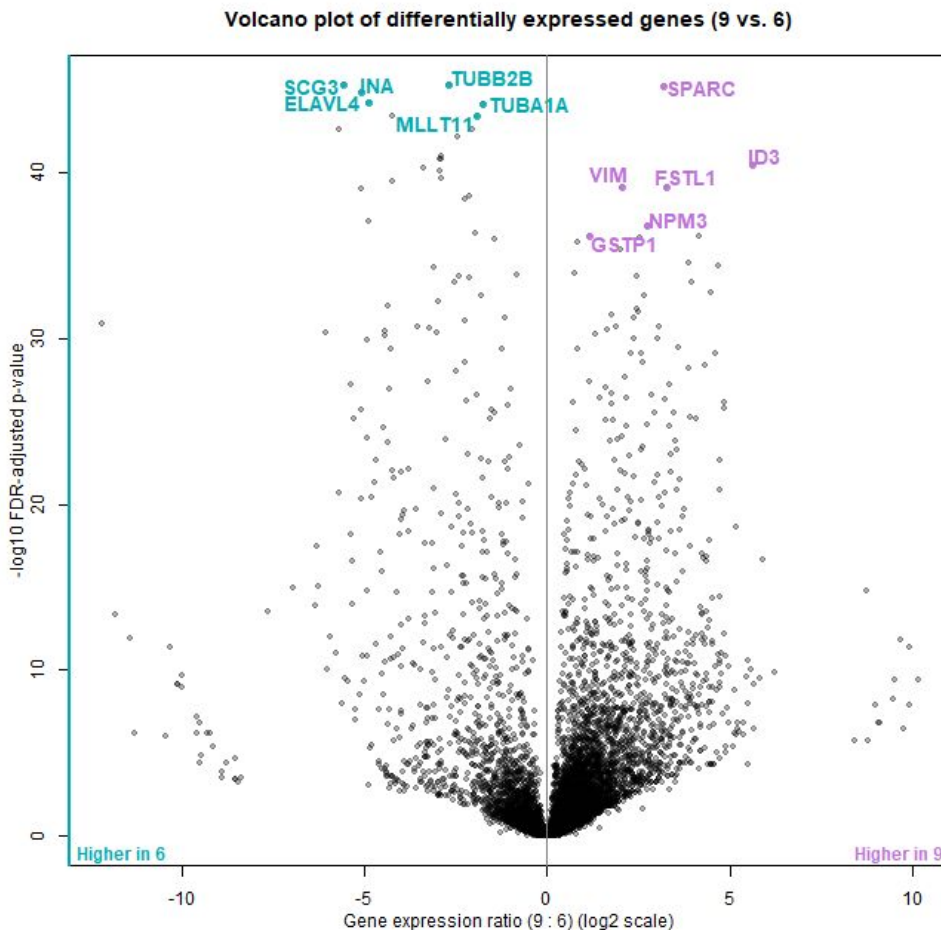
Figure 40 *Volcano plot of differentially expressed genes between cluster 9 and 4* Genes expressed relatively higher in cluster 9 are *SPARC*, *ID3*, *VIM*, *FSTL1*, *NPM3*, and *GSTP1*. Genes expressed higher in cluster 6 are *TUBB2B*, *INA, SCG3, ELAV4, TUBA1A,* and *MLLT11*. Axes are presented by scClustViz, -log10 FDR-adjusted p-value on y-axis, and gene expression ratio in log2 scale on the x-axis. FDR = false discovery rate.

According to GeneCards, *SPARC* is involved in bone collagen calcification but more relevant, it is involved in extracellular matrix (ECM) formation. The other genes are generic, like *VIM* which is a common progenitor/cycling cell marker at this stage, and both *SPARC* and *VIM* are expressed in the Notch-signaling areas so they are not very specific to cluster 9, they are just specific when comparing to cluster 6. Instead, we inspect cluster 9 vs all the other cells to see if there are better clues (Figure 41 *Differential gene expression between cluster 9 and all other cells*).
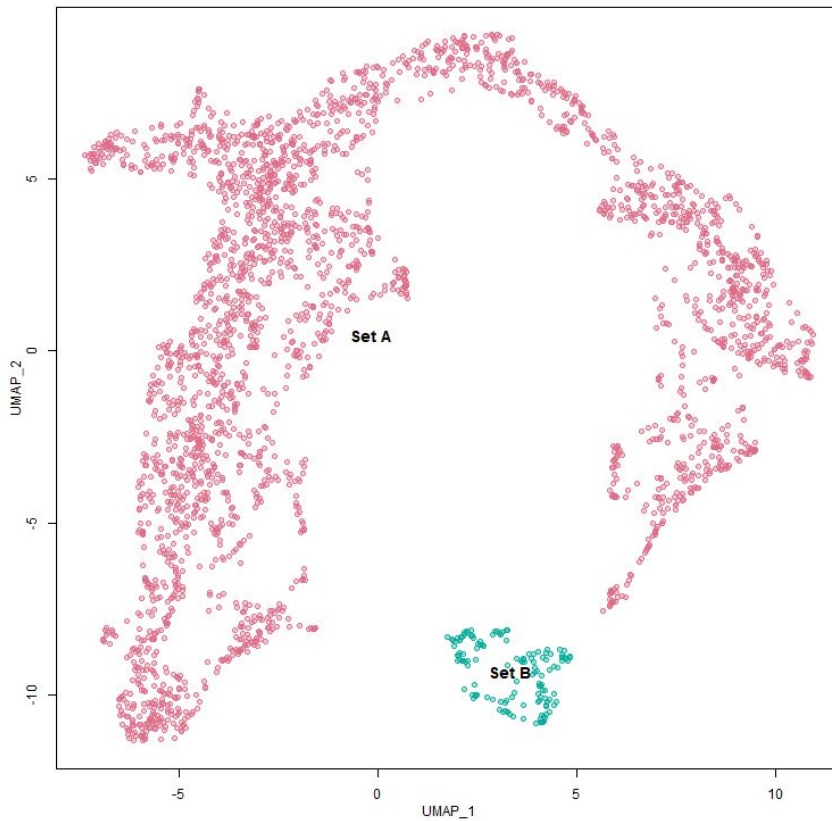
Figure 41 *Differential gene expression between cluster 9 and all other cells* Figure show which cells were selected and included in Set A (red), and in Set B (green), for gene expression comparison.

Cluster 9 has a higher expression of genes (top-10) *DLX5*, *CLIC5*, *DLX6*, *FGF8*, *DIRAS3*, *ARHGAP36*, *DLX6-AS1*, *PNOC*, *TMEM132B*, and *RIPOR2* (Figure 42 *Volcano plot of differentially expressed genes between cluster 9 and all other cells)*. Of these, the DLXX genes are perhaps less interesting as they were expressed high in cluster 6 *GABAergic interneurons* (Figure 39 *GABAergic markers*), but in comparison to the interneurons, cluster 9 does not show expression of other GABAergic markers *DLX2*, *DLX1*, *ARX*, and *GAD2* that cluster 6 had.
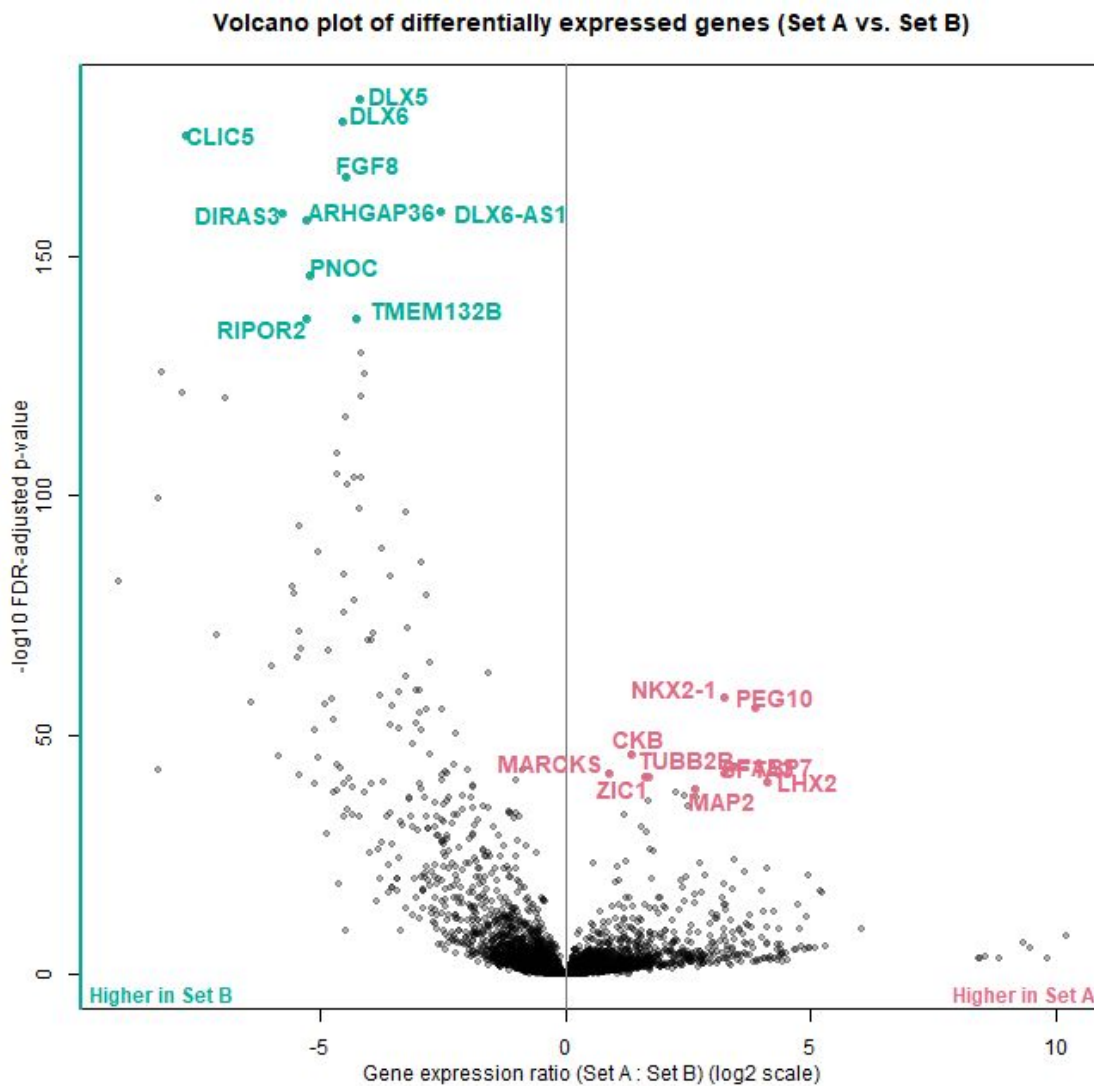
Figure 42 *Volcano plot of differentially expressed genes between cluster 9 and all other cells* Genes expressed relatively higher in cluster 9 are *DLX5*, *CLIC5*, *DLX6*, FGF8, *DIRAS3*, *ARHGAP36*, *DLX6-AS1*, *PNOC*, *TMEM132B*, and *RIPOR2*. Genes expressed relatively higher in all other cells at day 20 are *NKX2-1, PEG10, CKB, TUBB2B, FABP7, MARCKS, LHX2, ZIC1, and MAP2*. Axes are presented by scClustViz, -log10 FDR-adjusted p-value on y-axis, and gene expression ratio in log2 scale on the x-axis. FDR = false discovery rate.

FGF8 has a strong expression in cluster 9. And also at the end of cluster 6, and at cluster 10 (which will be talked about next). *FGF3* and *FGF8* have been implicated in ear development at the developing forebrain stage (Domínguez-Frutos et al., 2009; Kelley et al., 2006; Ladher et al., 2005; Theil et al., 2008). Perhaps clusters 9 and 10 are geared towards sensory development. A weak expression of *FGF3* is noticed also where *FGF8* is expressed (Figure 43 *FGF3 and FGF8 expression in cluster 9 and closely related areas*) (Peyre et al., 2015). *FGF8* is expressed in the rostral forebrain as the telencephalon develops, however, *FGF8* controls multiple patterning sites, and *SHH* is required to maintain its expression here (Storm et al., 2006) and we do not have *SHH* expression as shown earlier, although the protein itself could be present (Figure 34 *SHH and FGF2*).
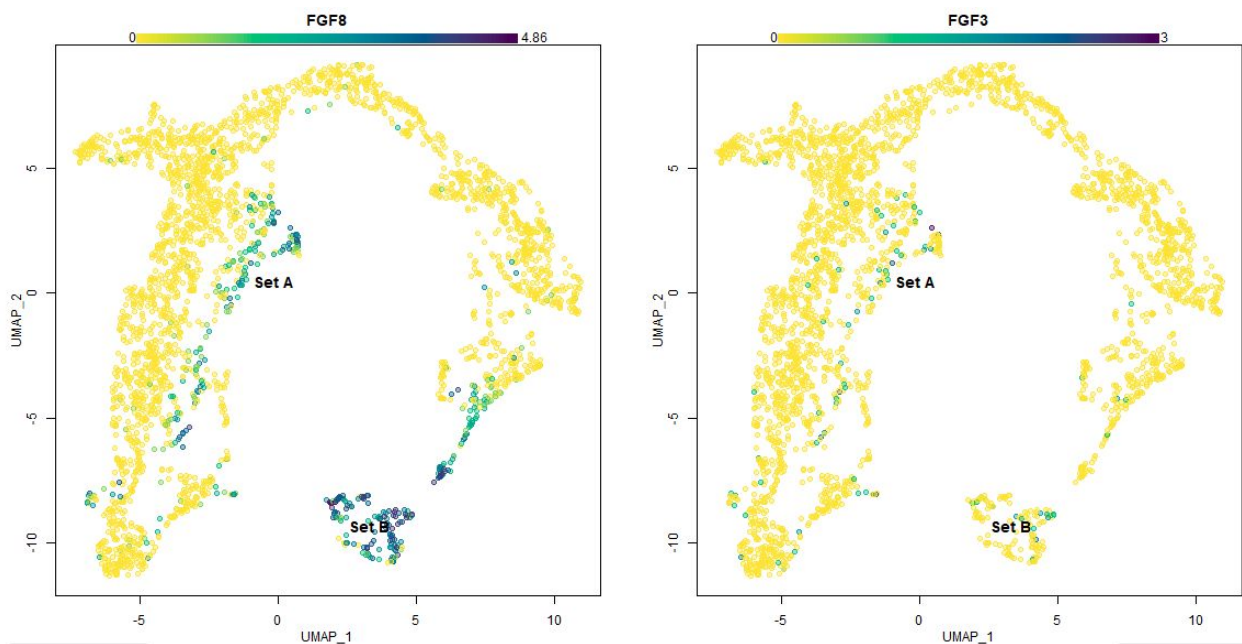
Figure 43 *FGF3 and FGF8 expression in cluster 9 and closely related areas FGF8* is strong in cluster 9 and seem to have an expression at the end of the tip at cluster 6 and in cluster 10, on the opposite diagonal side of the UMAP. *FGF3* is seen lowly expressed in the same areas. Expression levels are 0 to 4.05 for *FGF8* and 0 to 3 for *FGF3*, these levels, and legends, are set internally to scClustViz, and could not be adjusted.

Another feature of cluster 9 is the gene *PNOC*, which according to GeneCards encodes a protein that is processed to several products that include nocistatin, nociceptin, and orphanin. Nociceptin is a neuropeptide that can bind to the nociceptin receptor to increase sensitivity to pain, while nocistatin is said to inhibit nociceptin (Heinricher, 2005). *PNOC*, with FGF signaling, seems to be implied in ear development (Anwar et al., 2017) and has otherwise been used as an interneuron-related marker (Ouwenga et al., 2018), which could suggest these are sensory-fated interneurons and explain the retention of *DLX5*, 6, and *DLX6-AS1* expression. *DIRAS3* (GTP-binding protein Di-Ras3) did not seem to have been researched much in this context but when checking the expression profile in GTEx it seems expressed in the hypothalamus and basal ganglia (Figure 44 *GTEx expression for DIRAS3*).
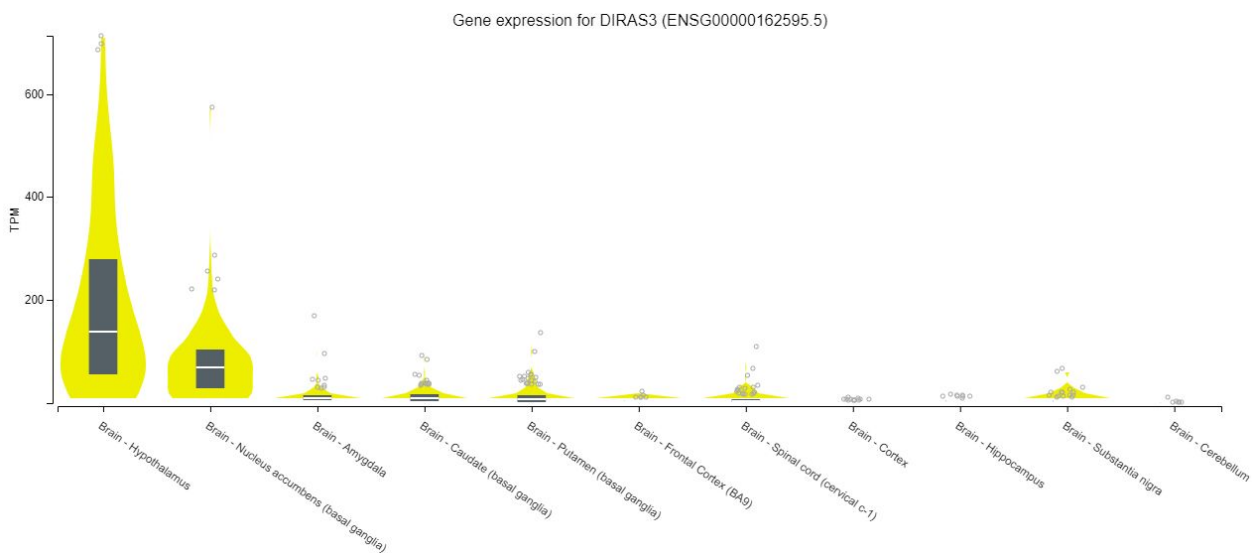
Figure 44 *GTEx expression for DIRAS3* High expression of *DIRAS3* is seen in hypothalamus and basal ganglia.

ARHGAP36 (Rho GTPase Activating Protein 36), has, as the name implies, activity related to GTPases (GeneCards). GeneCards list only 10 publications for *ARHGAP36*, however, so there is not much relevant information immediately available. One study suggests *ARHGAP36* to active Hedgehog and antagonize PKA (Ma et al., 2019). *ARHGAP36*, together with *VIP*, is an interneuron marker (Sugino et al., 2019), in the mouse. *TMEM132B* (Transmembrane Protein 132B), has been implicated in intracranial aneurysm (Farlow et al., 2015) and GWAS on nicotine dependence (Rose et al., 2010), and GWAS on bipolar disorder (Winham et al., 2014). A more

relevant study has shown it expressed in corticothalamic/subplate projection neurons early in development (Molyneaux et al., 2015). *DIRAS3* was shown above to be expressed in high levels in the thalamus, and since *PNOC* was also suggested as an interneuron marker, and that we have several - but not all - of the general interneuron markers, this could indicate that cluster 9 cells are interneurons fated for developing thalamus, i.e. corticothalamic projection neuron precursors, from now on abbreviated as CTPs. Finally, the gene *RIPOR2* (RHO Family Interacting Cell Polarization Regulator), is according to GeneCards sensory-related, and the protein encoded is part of hair cell stereocilia important for hearing. There is not much information on *RIPOR2* but some of the top GO-terms on GeneCards are relevant, such as cell adhesion, negative regulation of cell adhesion, for e.g. interneuron migration. *RIPOR2*, or *FAM65B* as it is called on GTEx, also has a general expression in many brain areas, including the hypothalamus (Figure 45 *GTEx expression for RIPOR2/FAM65B*). For now, cluster 9 will be annotated as interneurons fated for developing thalamus, *CTPs*, but more investigation down the line will be useful as there are many sensory indications in the literature as well.
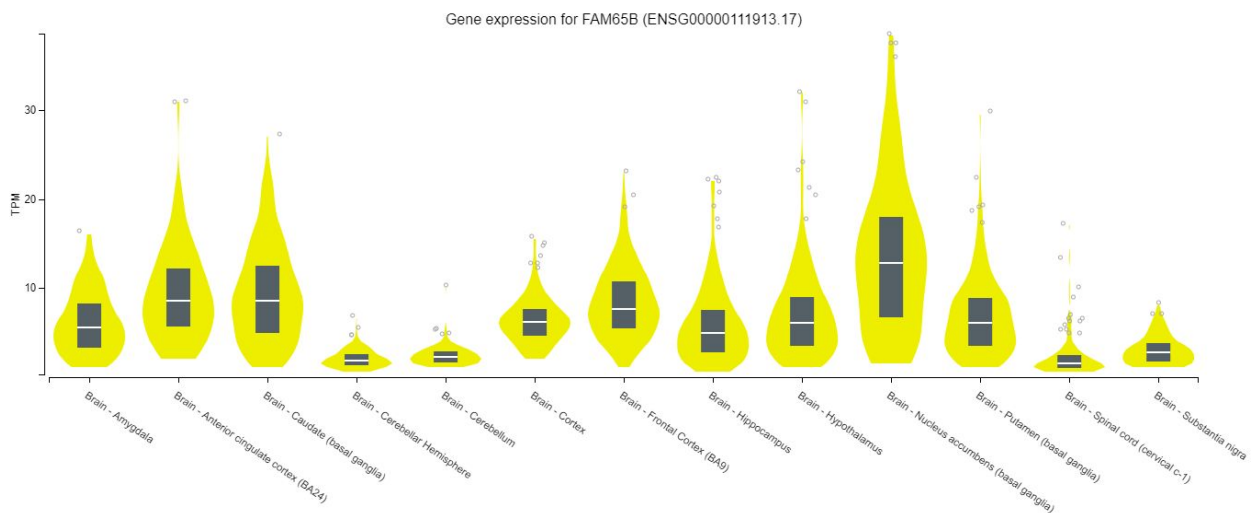
Figure 45 *GTEx expression for RIPOR2/FAM65B* Expression of *RIPOR2* can be seen in several areas in the brain, but less in the cerebellar hemisphere, cerebellum, spinal cord, and the substantia nigra.

Prefrontal cortex precursors

Cluster, 10, seems to only have one relevant marker, *FGF17* (Figure 29 *Day 20 dot-plots visualizing cluster differential expression of marker genes using pairwise tests between one cluster and all other clusters.*). Comparing cluster 10 to all other cells (Figure 46 *Volcano plot comparing expression in cluster 10 vs all other cells*), show that *FGF17* is indeed a very clear marker of this particular area.
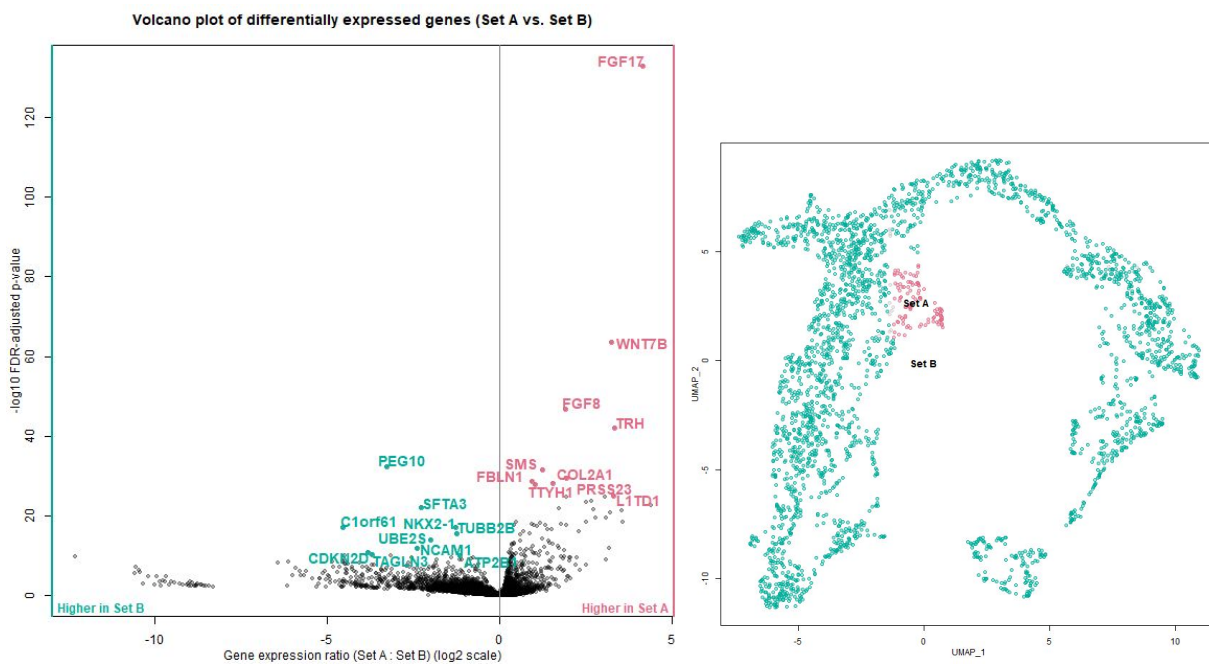
Figure 46 *Volcano plot comparing expression in cluster 10 vs all other cells* Left, in particularly *FGF17* seems to be the most distinct marker for set A. Right, a UMAP showing Set A (red), and Set B (all other cells, in turquoise). Left, volcano plot axes as presented by scClustViz, -log10 FDR-adjusted p-value on y-axis, and gene expression ratio in log2 scale on the x-axis. FDR = false discovery rate.

FGF17 (and *FGF8*) is secreted in something called the rostral patterning center (RPC) as the telencephalon grows and becomes more patterned (Hoch, Clarke, et al., 2015). *FGF17* seems to be required in specific roles in rostral-caudal cortical patterning and dorsomedial prefrontal cortex (DPC) development (Hoch, Clarke, et al., 2015) (mouse study). The very high expression of *FGF17* in cluster 10 could mark the cells in this cluster as having developed a certain developmental potential for a cell-type or tissue not directly on the developmental trajectory predicted for the other clusters. The DPC is an interesting area as it has been shown to be

92

implicated in long-term memory and to play roles in high-level functions such as a sense of self, morality judgments, empathy, altruism, and decision making (Ferrari et al., 2016; Isoda & Noritake, 2013). A dysfunctional DPC is also implicated in social difficulties found in ASD (Ajram et al., 2017). The DPC is also implied in connection to ADHD symptoms (Bayard et al., 2018). *TRH* (Thyrotropin-releasing hormone) is also seen as expressed higher in cluster 10 and has been shown detected in forebrain neurons projecting to cortical regions (Heuer et al., 1998). Another gene expressed in cluster 10 is *SMS*. *SMS* encodes the spermine enzyme which has been shown downregulated in the cerebral cortex of suicide completers (G. G. Chen et al., 2010). Polyamines, such as produced involving *SMS*, have been shown to influence GABA-receptor function and they can affect transmission through NMDA- and AMPA-receptors (G. G. Chen et al., 2010). Therefor changed levels of polyamines have been implicated in many disorders such as mood disorders, anxiety, schizophrenia, and delayed development of speech (G. G. Chen et al., 2010; Kesler et al., 2009; Larcher et al., 2020). For now, cluster 10 will be annotated as *prefrontal cortex precursors, PCPs*.

<u>A split technical-, or low-quality cluster</u>

When inspecting cluster 7, it seems to be split between opposite sides on the UMAP which is strange, since no other cluster show this behavior where it is split in two. This could be an indication for a cluster with low-quality cells (Figure 47 *Day 20 Cluster 7, split into two parts*). There is also a possibility of more trajectories that diverge which the UMAP cannot accurately show in reduced dimensions. This could be investigated in the future with more advanced trajectory methods or using several branching solutions for PAGA. Cluster 7 has no distinct markers separating it from the other clusters (Figure 29 *Day 20 dot-plots visualizing cluster differential expression of marker genes using pairwise tests between one cluster and all other clusters*). Furthermore, when selecting cluster 7 and comparing differentially expressed genes in a volcano plot - vs all other cells (Figure 48 *Volcano plot showing genes expressed more in cluster 7 vs all other cells*), there are few markers that are useful for annotating this cluster. It would anyway be a problem as it is split into two almost opposite areas in the UMAP. Therefore, for now, this cluster will remain unnamed and will be set aside for a more thorough investigation at a later date.
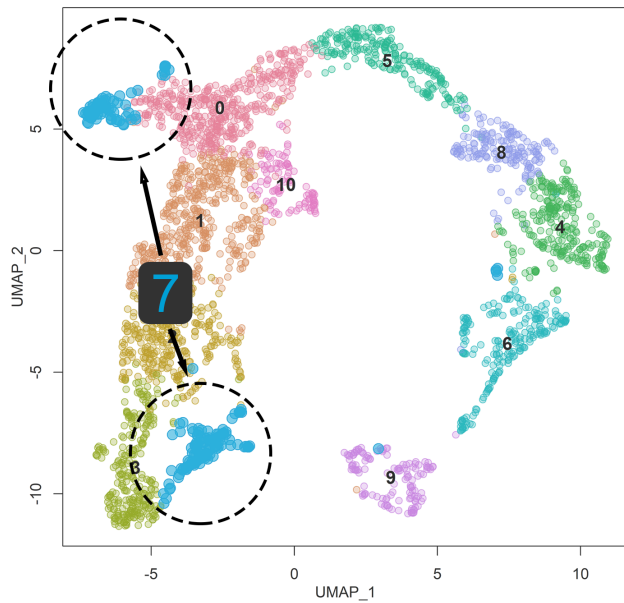
Figure 47 *Day 20 Cluster 7, split into two parts* Cluster 7 is marked in blue. scClustViz has an internal color-scheme with unfortunately built-into the package code itself.
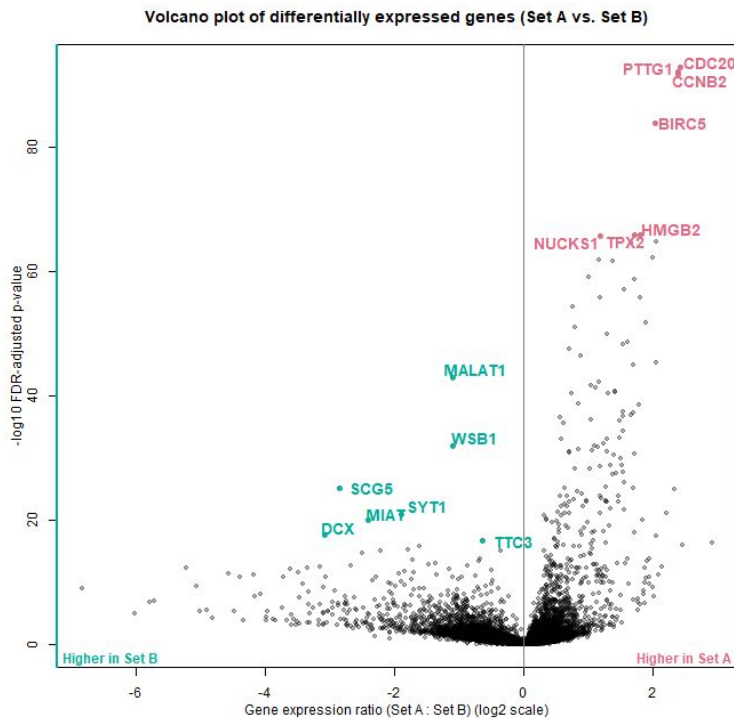
Figure 48 *Volcano plot showing genes expressed more in cluster 7 vs all other cells.* Genes shown to be higher in expression in cluster 7 are *CDC20*, PTTG1, *CCNB2*, *BIRC5*, *HMGB2*, *TPX2*, and *NUCKS1*. For rest of the cells at day 20, genes expressed higher are *MALAT1, WSB1, SCG5, SYT1, MIAT,* and *DCX.* Axes are presented by scClustViz, -log10 FDR-adjusted p-value on y-axis, and gene expression ratio in log2 scale on the x-axis. FDR = false discovery rate.

Day 20 cell-types

Finally, we can annotate the cell-types for day 20. (Figure 49 *Day 20 cell-types*). Clusters 3 and 2 were merged and annotated as *RGs*, and cluster 1 as *RGs: stopped to commit*. Cluster 8 was annotated as *PCPs* (*prefrontal cortex precursors*), cluster 0 as *Intermediate progenitors*, and cluster 5 as *Layer V precursors*. Cluster 8 was annotated as *Layer IV/5-HT precursors*, cluster 4 as *5-HT/Chol/Glut precursors*, cluster 6 as *GABAergic interneurons*, and cluster 9 were annotated as *CTPs* (*Corticothalamic projection neuron precursors)*.
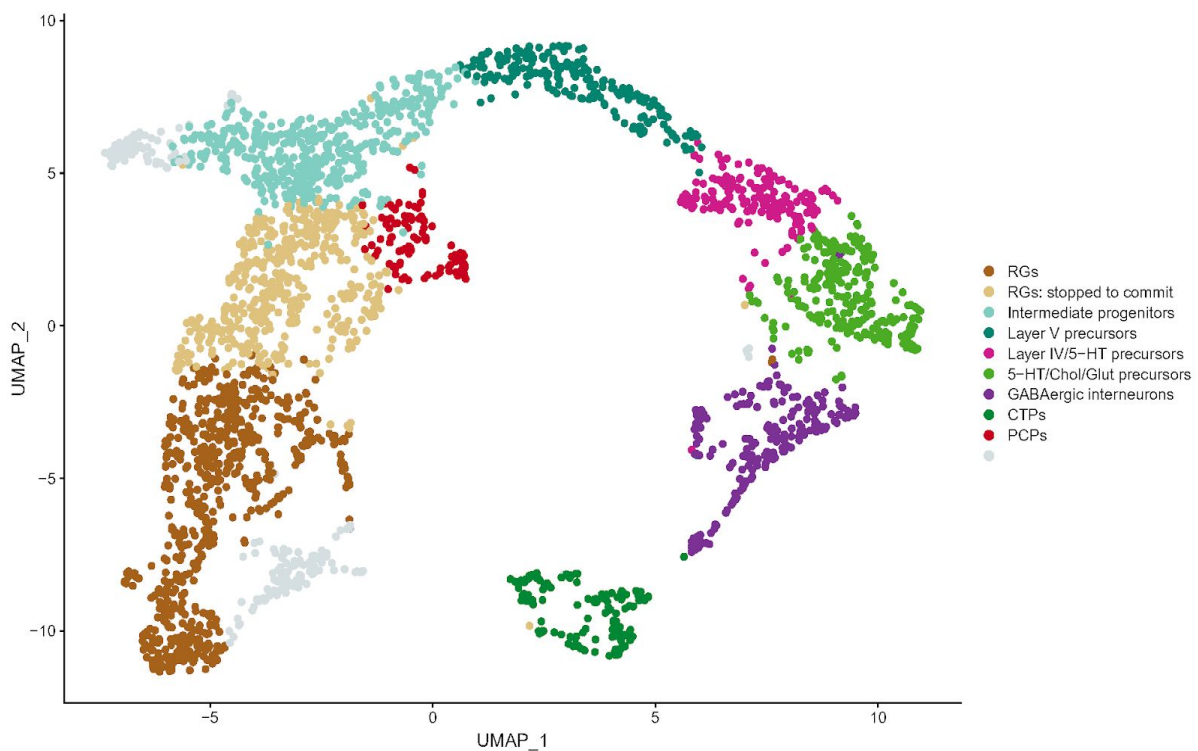
Figure 49 *Day 20 cell-types*. RGs = Radial glial cells, 5-HT/Chol/Glut = Serotonergic/Cholinergic/Glutaminergic, CTPs = Corticothalamic projection neuron precursors, PCPs = Prefrontal cortex precursors. Grey color indicates a split cluster, with cells that could not be annotated, and is a possible low-quality cluster, but will be investigated further. Annotation presented on a UMAP via Seurat.

All time-points trajectory solutions; PAGA-graph and cytoTRACE

Finally, to round off the results section of part I of this thesis, an all time-points trajectory is presented using PAGA-graph and cytoTRACE differentiation prediction so as to wrap up what has been shown visually at previous chapters. 8708 cells from day 0, day 7, day 13, and day 20 was included for this PAGA- and cytoTRACE analysis.

Movement of differentiation from day 0 to day 20

CytoTRACE differentiation prediction where day 0, 7, 13, and 20 are included, 8708 cells in total (Figure 50 D*ifferentiation prediction all timepoints*). Impressively, cytoTRACE predicts differentiation to move accurately from the hESCs via day 7 and day 13, to the endpoint day 20. Day 7 and day 13 appear more merged together when all four timepoints are included which maybe suggests that between day 7 and day 13, either many starting and ending markers are shared, or differences in start to end in terms of how much the cells change are less than between day 0 and 7, and day 13 and day 20.
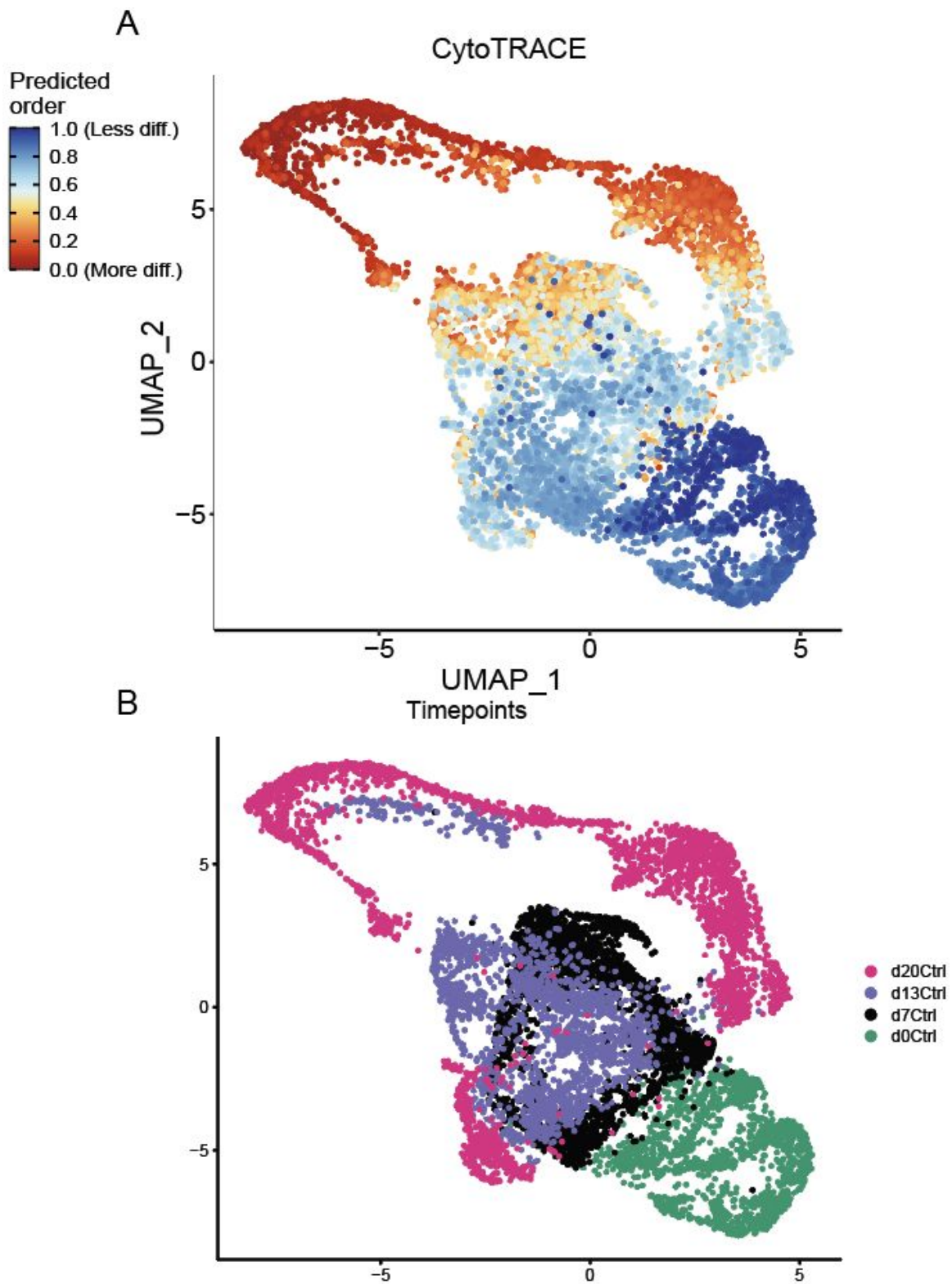
Figure 50 *Differentiation prediction all timepoints*. 8708 cells are included from day 0, day 7, day 13, and day 20 are included. A. Prediction of differentiation moving from day 0 hESCS, through day 7 and day 13, to end at day 20. B. All four timepoints colored on a UMAP for comparison.

PAGA-graph pseudotime prediction, day 0 to day 20

PAGA-graph and pseudotime prediction, using 8708 cells from all days, show one interesting difference to differentiation prediction (Figure 51 *PAGA-graph with clusters based on the Louvain-algorithm and pseudotime prediction, all time points*), when cytoTRACE predict differentiation day 7 and day 13 was not separated very well. But PAGA manages to a greeter degree to show pseudotime moving from day 0, then through day 7, then through day 13, and finally ending at day 20. Even though the change in color (which indicates how pseudotime progress) is slight, it is visible in PAGA, but not by using cytoTRACE.
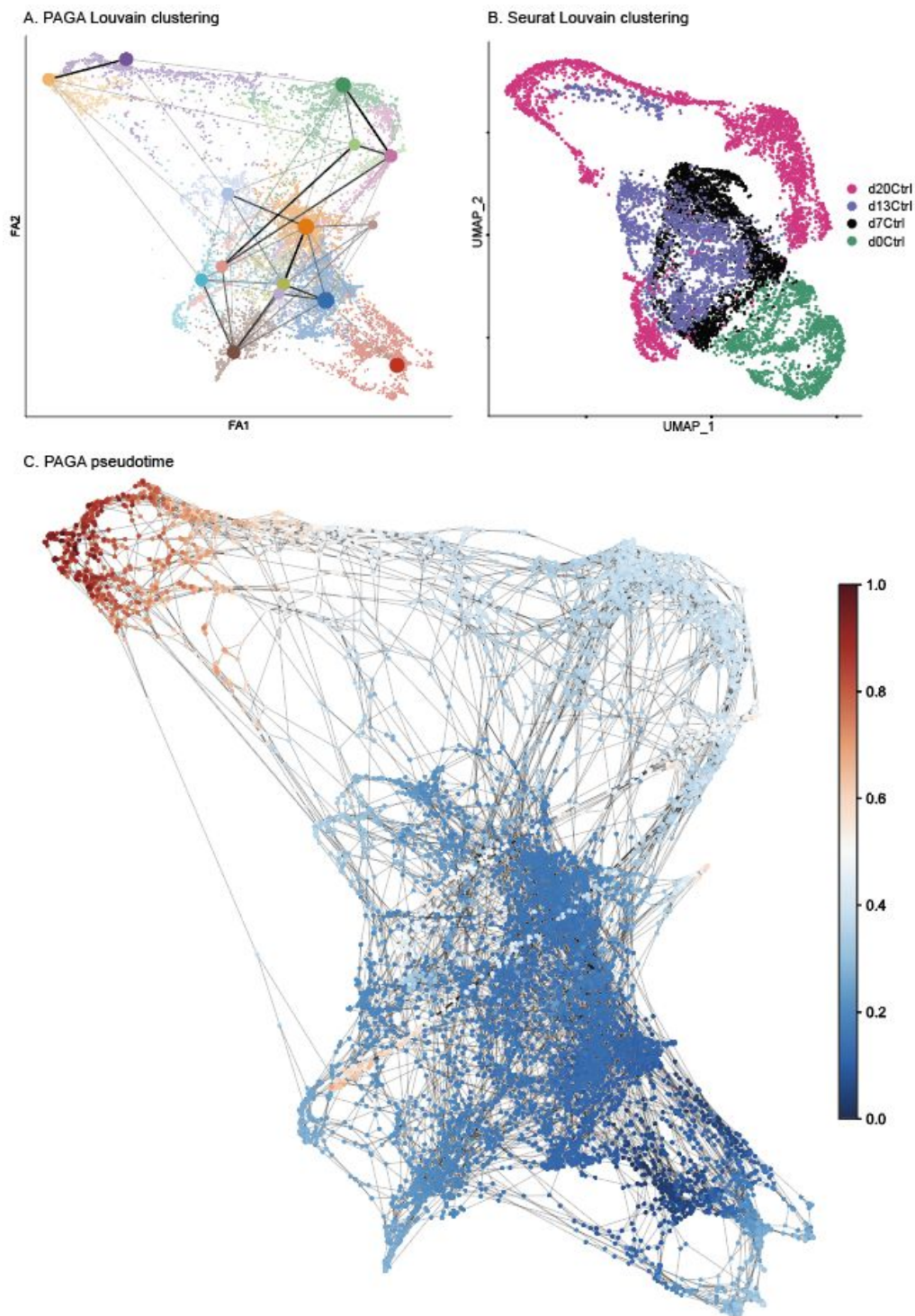
Figure 51 *PAGA-graph with clusters based on the Louvain-algorithm and pseudotime prediction, all timepoints*. A. Louvain-algorithm resolution was set to a number which resolved 15 clusters, those clusters were then used in the PAGA-graph. Colored circles, with a number, represents clusters calculated by the Louvain algorithm. B. For comparison, the Seurat UMAP is shown. C) PAGA pseudotime predicts time to begin (dark-blue) where the day 0 hESCs are (bottom-left). As a slight difference to cytoTRACE, PAGA seems to be better at predicting trajectory, pseudotime to can be seen moving through day 7, then day 13, and ending on day 20.

100

Part I - Discussion

Where the hESC-based toxicology field stands as of now

The last decade has seen an increase in developmental disabilities such as ADHD, ASD, and intellectual disabilities (ID), diagnosed in children aged 3-17 years in the USA (Zablotsky et al., 2019). Evident is that children from poorer communities and rural areas are more at risk. A problem in the USA is the possibility of overdiagnosis due to pharmaceutical companies having aggressive advertisement and physician outreach, with the opioid crisis being a terrible display of the more extreme end of this (Vadivelu et al., 2018). Alarmingly, another study from the previous decade, 1997-2008, also showed an increase in developmental disabilities. ADHD does not seem to be a more common diagnosis for males, since recent research has shown that females are underdiagnosed, and women who are diagnosed with anxiety or depression could benefit from additional screening for ADHD. Together with a growing body of evidence suggesting gender-biased diagnosis for ADHD, there are also other gender biases in medicine ("Male GPs Are Less Likely to Assess Cardiovascular Risk in Female Patients," 2016; Martin et al., 2018; Stafford, 2009), and in the psychiatry profession (S. McCarthy, 2016; Richmond, 2019). It is then important to mention that the HS-360 hESC line used in this study, is male, but the H9 line is female, thus it could be included as a comparison in future single-cell sequencing. It would be good to take some of these factors into account, depending on what drug is investigated using this protocol, although using only one cell-line per gender would not prove or disprove sex-based differences as more than one representative cell line is needed for significance.

In terms of toxicology methods utilizing pluripotent stem cells, there are still tens of thousands of chemicals registered for use, especially in areas where testing is less required such as in agriculture (Bal-Price et al., 2018; Luz & Tokar, 2018). One embryotoxicity test that does not require the sacrifice of pregnant test animals, is the mEST-test (mouse embryonic stem cell test) (Genschow et al., 2002), which has been frequently used.. It is, nevertheless, a mouse-based embryotoxicity test. For neurodevelopmental testing, a rudimentary test needs to become a requirement for chemical safety in the industry. It's not just the pharmaceutical field that should have this requirement as the risk of exposure is also high for agricultural workers, factory

workers in the technical industry, etc (Fritsche et al., 2018; J. Li et al., 2019). In terms of what exists for human-specific testing using human embryonic, or induced, stem cells, there is a rosette-induction test where morphology, viability, and bulk gene expression is assessed (Shinde et al., 2015; Waldmann et al., 2014). NPC proliferation tests exist but many of the suggest protocols use differentiation based on mouse embryonic stem cells while the human NPC testing seems to be based on established cell lines, and not *during* differentiation, (Behl et al., 2015; Radio et al., 2015; Sohn et al., 2017), or just one specific time-frame, such as a week of rosette development (B. Huang et al., 2017).

Tests intended for radial glial are not really established to a satisfactory degree when considering the diverse and critical roles they play during brain development and are mostly seen discussed as a possibility in some cells that are showing glial markers (Masjosthusmann et al., 2018; Uzquiano et al., 2018). There are tests for radial glial migration where the distance glial cells migrate are measured (usually manually in e.g. ImageJ) (Bal-Price et al., 2018). In terms of longer differentiation and neurogenesis, there was one study that investigated valproic acid (VPA) (Ehashi et al., 2014), which is a strong teratogen, similar to thalidomide - the culprit (together with the human error of prescribing it) behind a period of time where children were born with missing limbs (Khalil et al., 2020; Meganathan et al., 2012). VPA is most commonly seen in epilepsy treatment thus sometimes posing a problem considering risk vs benefit in removing it during pregnancy as seizures for the mother can be a greater risk than the potential teratogenicity for the developing fetus. The VPA-study examined 22-genes for which expression is altered by teratogens, by qRT-PCR, which suggest that one would need some prior knowledge of what genes to examine and some knowledge of disease effect. Another study used NPCs as a starting point for differentiation to investigate the effects of polyamidoamine dendrimers (Y. Zeng et al., 2016).

There seems to be a need for human-specific testing that covers a longer differentiation period while also providing a possibility for exposing cells with a drug, throughout the duration of the protocol. A robust and specific protocol that enables the use of hESCs in the aforementioned analysis also seems like a useful feature. Moreover, our single-cell analysis shows that there is much complexity in cell culture on day 20, more than we expected. Naturally, before the last 5

years or so, and the advent of single-cell sequencing omics, researchers were not really equipped to fully investigate cell-populations at this level. The cell diversity seen in this protocol is hopefully then useful in the discovery of subtle effects during some of the critical periods we have in human neurodevelopment.

hESCs and neural differentiation in neurotoxicology

Our protocol required a consistently set cell density (cells per a given area) at all points where we used hESCs cells to start an experiment or when cells were transferred during an experiment for a new phase. It might seem straightforward, when protocols describe the use of certain passage ratios, e.g. to passage cells 1:3 for a certain time, then to change to 1:10 for x number of passages, or you see a % of cells being used (also very subjective without a cell count). This is very normal, as the aim here is to remove some tedium with always counting - and removing tedium is something a hard-working human loves - but such shortcuts give rise to many uncertainties when using cell-lines that behave differently, or when different users end up with different results due to such relatively subjective cell passage guidelines. An optimized density will ensure efficient differentiation and cell-survival with the added bonus of having a reproducible protocol when repeating a drug toxicity study. We mainly used the male hESC line HS-360 (Kurtz et al., n.d.; Ström et al., 2010), which in our hands had satisfactory stability in routine culture conditions, and we would recommend it for use with our protocol. We have tested the protocol with hESC line H9, and it worked well, but there can be inherent hESC line differences. Thus, depending on which ones the user decides upon it is good practice to make some comparison experiments before performing all experiments with a single cell-line.

We would also like to repeat paracetamol-treatment using H9 for comparison to a different hESC line (paracetamol-study shown in Part II). The initial state of hESCs is of vital importance in experiments such as these. The efficiency of most, if not all, differentiation schemes depends on a healthy homogenous population of hESCs. It will save the user a lot of additional work taking the time to learn how to optimize hESC culture conditions and how to prune away colonies that show signs of differentiation as early as possible. As constant pruning and selection of cells is

tedious and is also a factor that can bring variance into the experiments when different people are involved, it is good advice to test more than one hESC line and select the one which is more type (counting on that both lines display the same results at the end of the experiments). For those who have worked with hESCs for an extended period, it is a familiar problem that these cell-lines are very sensitive, so that even in routine culturing conditions, the cells can spontaneously start differentiating to other cell-types, stop growing and detach, or become stressed and rapidly die. Different hESC cell lines also seem to have varying stability, such as in maintaining stem-cell characteristics while in routine culture, which is usually preferable. Curiously, and inconveniently, when hESCs spontaneously differentiate, they have been shown to have an inherent preferred lineage to which they commit, e.g. ectoderm, mesoderm, or endoderm (or combinations). During this thesis work hESC lines H9 and HS-360 this tendency was observed, as H9 cells tended to spontaneously differentiate into a neural fate, while the hESC line HS-360 tended to take a mesodermic direction. This phenomenon, with variable differentiation commitment, has been documented in several studies (Pal et al., 2009; Sarkar & M., 2011). Another point that could be made, is that it is often said that induced pluripotent stem cells, iPCSs (adult primary cells reprogrammed into stem cells), are the same (or equal) to hESCs, but since there are inherent differences, also between hESC-lines it seems reasonable to have the benefit of the doubt and investigate this prior to critical experiments.

Forebrain/GABAergic marker *NKX2-1*, *PAX6*, and *FOXG1*

We see a strong expression of typical ventral forebrain marker *NKX2-1*, and a relatively lower *FOXG1* expression in the single-cell data, where both markers appear before day 20 (Figure 52 *NKX2-1 expression at day 13 and day 20* and Figure 53 *FOXG1 expression at day 7, day 13, and day 20*). Why *PAX6* expression was lost already at day 7 to day 13 seemed alarming at first, however, loss of *PAX6* indicates a ventral forebrain identity (Figure 14 *Day 13 ddPCR of selected markers*, and Figure 15 *Day 20 ddPCR of selected markers*). We also see many markers of inhibitory neuron function in the analysis of day 13 and day 20 clusters, suggesting a more ventral fate (Micali et al., 2020). We never added Sonic hedgehog, SHH, which some protocols suggest to more strongly induce ventral forebrain fate. We see a strong expression of NKX2.1,

which is a strong ventral forebrain inducer. The LHX6 marker will most likely be expressed, as these interneurons mature past 20 days as the major cortical interneurons that originate from NKX2.1 cells in the developing telencephalon will maintain expression into adulthood (Hu et al., 2017).
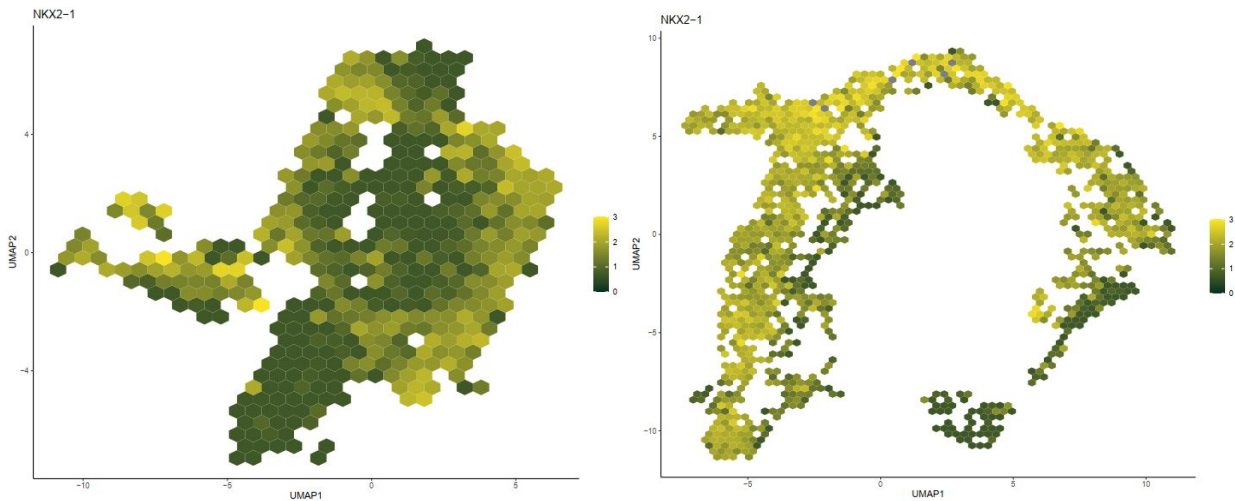
Figure 52 *NKX2-1 expression at day 13 and day 20. NKX2-1* is a ventral forebrain marker and its expression is detected first at day 13 then is strongly expressed at day 20 for all clusters, except the more mature GABAergic interneurons, corticothalamic projection neuron precursors (CPNs), and prefrontal cortex precursors (PCPs). Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 3.
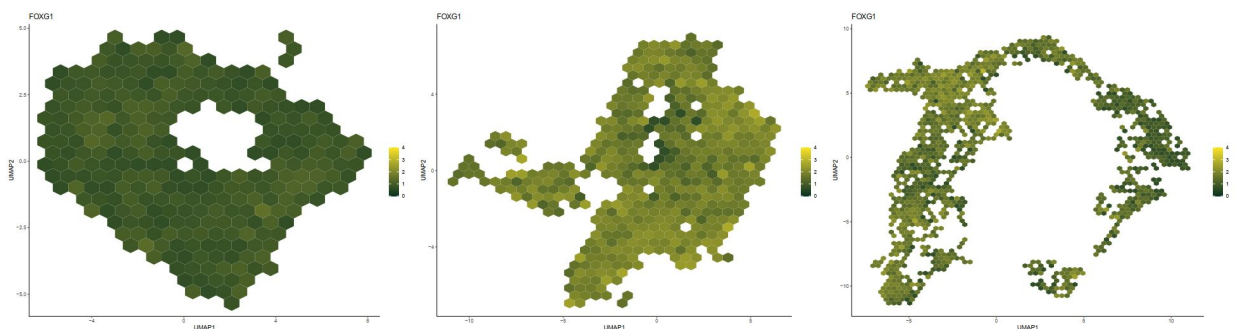
Figure 53 *FOXG1 expression at day 7, day 13, and day 20. FOXG1* is a ventral forebrain marker and its expression is detected first at day 7 (low), then has a uniform expression by day 13, and at day 20, strong *FOXG1* expression is observed, with particularly high expression in areas that also exhibited Notch-signaling (Figure 30 *HES5 - a marker for Notch-signalling*). Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 4.

The DKK1 mimetic XAV939 (X) inhibits BMP and WNT signaling, thus promoting anterior fates (Glinka et al., 1998). By acting as a WNT pathway inhibitor, DKK1 will significantly increase *NKX2-1* expression while decreasing *PAX6* and GLI3 expression (X.-J. Li et al., 2009). As described in the method, XAV939 is added at part I: induction. The addition of XAV939 is a likely contributing factor to the strong ventral forebrain identity that we see appearing in our cells by day 13. *PAX6* is expressed in day 7 and is required for early neurogenesis. *PAX6* seems to be suppressed after day 7 and is detected at very low relative levels at day 13 and day 20. The strong ventral induction of our cells could also be due to possible inherent cell-line biases, and this is currently being tested by more comparisons to H9 hESCs. There are recent studies showing cell-lines and organoids from different donors having biases towards different fates (Kanton et al., 2019; Micali et al., 2020). In experiments not shown in this thesis, we have seen small differences between hESC lines H9 and HS-360 where e.g. *PAX6* was expressed at different levels between the two cell lines.

Now in the end, where are we?

The day 20 timepoint might be described as an in vitro ventricular zone-like/early-MGE. Cortical layer formation starts around GW7 (human gestational week 7) and GW8 where the marginal zone (MZ) is separated from subplate and we see the first types of neurons appearing (Budday et al., 2015). The subventricular zone (SVZ) forms around GW7 when RGs migrate towards more basal locations (Dehay et al., 2015). Then the inner and outer subventricular zones (iSVZ and oSVZ) appear between GW11 and GW13 (Hansen et al., 2010). Our monolayer cell culture limits the level of complexity that can be achieved when compared to using organoids, e.g. a ganglionic eminence with distinct MGE, LGE, and CGE areas is not likely to be observed in the monolayer.

Cell types at the end stage of the protocol

A central goal with the development of the method was to characterize the cells at the endpoint at day 20. A thorough characterization would give increased analytical power and also generate valuable information for further use of the protocol in neurotoxicology studies. It was a

complicated matter to accurately annotate cell-types at this stage of human forebrain development and this annotation will likely see more development before we submit our work to a journal, and likely after submission, in revision. This difficulty of annotation, and tendency to go back and re-iterate and change data analysis, and expectation that things will change, is expected for an area which is rapidly seeing new discoveries every day.

A certain type of gene-set was observed in the clusters of cells that were predicted to be start-of-differentiation at the various time points. These gene-sets expressed many progenitor- or stem cell genes, that from the literature have been described to have functions such as cell renewal, pluripotency, rapid cell-cycle progression, asymmetrical division, and capability to re-enter the progenitor state if a commitment was not made. At day 20, the progenitors were manually annotated as radial glial cells, RGs (Figure 49 *Day 20 cell-types*). At this point they are not just stem cells, but will also function as guides in the migration of their progeny. Intermediate progenitors are a middle step between cycling progenitors and fully committed progeny cells. These could be identified in an area where Notch-signaling was fading, where these cells can either re-enter the cell-cycle or proceed to neural progeny (Figure 30 *HES5 - a marker for Notch-signalling*). After the IPs, we saw an expression of mature neural markers. *FEZF2*, a marker for neuronal subtype differentiation, patterning of the forebrain, and deep layer V neuron precursors, was highly expressed in this area (Figure 32 *Layer V marker FEZF2*). Early born projection neurons first occupy the deepest layers, VI, V, and IV, as neurons grow in an "in and out" sequence. Neurons that populate deep layers are Notch-signaling dependent and therefore it makes sense that these appear here as also Notch-signaling faded among the IPs.

We were observing interesting patterning effects after the IPs. *FEZF2* expression and Notch-signaling had clear boundaries. As stated in the literature, Wnt/β-Catenin signaling in the forebrain is repressed by *LHX9,* and *LHX2* and interestingly *LHX9* expression increased just on the right flank of the layer IV/5-HT precursors while Notch-signaling had a clear boundary at the IPs (Figure 33 *Wnt repressors LHX9 and LHX2*, and Figure 30 *HES5 - a marker for Notch-signalling*). High expression of *GAP43* was seen in this cluster, which is important for the development of forebrain serotonergic innervation, which is particularly dense in layer IV

(Figure 36 *GAP-43 expression on day 20*). In vivo serotonergic neurons (5-HT) develop early and, as is often discussed, dysregulated 5-HT signaling is suggested to significantly contribute to depression, schizophrenia, and other mental diseases. For cluster 4 we saw a diverse expression of genes involved in AMPA-, NMDA-type ion channels, glutamate uptake, signal transduction, and serotonin transporter regulation. Cluster 6 had many distinct markers, suggesting GABAergic activity and function (Figure 39 *GABAergic markers DLX2, DLX1, DLX6-AS1, DLX5, ARX, and GAD2*), which led to that cluster 6 could be annotated as GABAergic interneurons, but with the caveat that we cannot conclude that they are fully functional without measuring signal transduction with e.g. experiments such as patch-clamping. After the GABAergic interneurons, markers for cluster subsequent cluster 9 were more unclear in that their descriptions in literature, and on GeneCards, was harder to put together into a context for a particular cell type. Some markers were shared with the previous GABAergic interneurons (which was expected in terms of that there would be patterning with gradients of markers) and the *in silico* cluster solution will not always overlap perfectly with underlying *biological* diversity). However, cluster 9 did have some distinct features (Figure 42 *Volcano plot of differentially expressed genes between cluster 9 and all other cells*), indicating it interneuron precursors possibly fated for developing thalamus, i.e. corticothalamic projection neuron precursors, which lead to the cluster being abbreviated as CTPs.

The next cluster, cluster 10, seemed to only display one relevant marker, *FGF17*, and *FGF17* together with *FGF8* (a moderate expression of *FGF8* was observed in this cluster) are secreted at an area that was called rostral patterning center, as the telencephalon grows and becomes more patterned. *FGF17* seemed especially important for dorsomedial prefrontal cortex development and precursors for this area seemed like a possibility at this time point. There were also genes involved in polyamine mechanisms which showed a connection with higher cognitive functions and mental disease, genes implicated in cortical projections, thus the cluster was finally named prefrontal cortex precursors (PCPs).

Finally, a split cluster 7 was observed, where half of the cells were on the top of the day 20 UMAP, while the other half was at the bottom. For this thesis, this cluster was set aside after quite some time of analysis (this analysis is not shown to the full extent in the thesis). Cluster 7 could be of low quality, but there are also other possibilities, such that there are more trajectories to account for at this time point, where more advanced and CPU-heavy trajectory analysis could be of help downstream of this thesis.

The established in vitro protocol has thus taken cells on a developmental path, starting as embryonic stem cells, to neural rosettes, and then via maturation, finally to radial glial cells, where radial glial progeny ranged from intermediate progenitors, deep layer neuronal precursors, corticothalamic precursors, prefrontal cortex precursors to GABAergic neurons.

Part I - Conclusions

A protocol for hESC neurodifferentiation to forebrain-like neuronal cells, that enabled simultaneous drug testing, was created and tested successfully. Analysis results from the results of this protocol, repeated twice, were presented.

For annotation, multiple criteria need to be weighted. This can be related to morphology, function, or genetic features. Automatic cell-annotation software and cluster solutions will assist in determining cell-types, but for the developing human brain, we are still at a point where it is necessary to perform manual work to have any accuracy in determining a cell-type. Furthermore, after going through recent single-cell studies in human and mouse, other researchers also demonstrate many challenges and uncertainties when trying to accurately name their cells. In our case, we did not know exactly what cells, or neurodevelopmental stages, we were expecting, except perhaps NPCs at day 20.

In terms of neurodevelopmental toxicology testing, to us, the cell populations we see at the end of our protocol have clear potential to be used in the analysis of drug effects to neurodevelopment. Interneurons have been reported to be very diverse and are intensively studied, and can potentially be involved as factors for diverse adverse outcomes when their development is disturbed by a xenogenic compound (Josh Huang & Paul, 2018; Mayer et al., 2018; van Heusden et al., 2019). One must also take into account that there are also distinct differences in human brain development compared to other mammals, and our closest primate cousins, such as a much-expanded transient subplate zone, a subpial granular zone, additional origins of migrating neurons and interneurons, and additional layers in the cortex (Miller et al., 2014).

Part II - Paracetamol and human neurodevelopment

N-acetyl-para-aminophenol

Paracetamol, or acetaminophen, is a widely used nonprescription analgesic for the treatment of pain and fever. It is commonly in use as an analgesic during pregnancy and a report published in 2005 showed that more than half of the women included in the study had taken paracetamol at some point during pregnancy (Werler et al., 2005). The reported use of paracetamol is not alarming, as it is recommended before other painkillers and considered safe to take during all trimesters when there is a need. Due to the unresolved discussion on how paracetamol works, its also not surprising that currently proposed mechanisms of action for paracetamol are complex, with possible tissue- and developmental-specific effects (Jóźwiak-Bebenista & Nowak, 2014). Paracetamol's name arrives from its chemical name, N-acetyl-para-aminophenol, where para is followed by *cet*, and *amol* indicates the aminophenol ring.

The primary mechanism behind the analgesic effect of paracetamol is surprisingly still under some debate even after well over a hundred years of wide-spread use (S. S. Ayoub & Flower, 2019). Paracetamol is a member of the family of aniline derivatives (Figure 54 *Schematic of paracetamol's parent molecules*) where acetanilide is the original compound discovered by two doctors in the town of Strasbourg in 1886 (Cahn & Hepp, 1886). They were treating a patient for parasitic worms and mistakenly treated the patient using the wrong compound. By serendipity, and luckily for the patient, they used aniline which although not very effective against parasites quite substantially lowered the fever. Aniline was not used for very long however, as it was shown to have several toxic effects, one of the worst being the blood disease methemoglobinemia that would lead to seizures and heart arrhythmia. Following this, there was a push towards finding less toxic variants, and contributing to high activity in this process were not only the substantial antipyretic effects but also the cheap cost and relative ease of synthesizing the aniline variants. Of all the aniline derivatives that were discovered in this process, paracetamol and phenacetin were deemed the best market candidates (Jóźwiak-Bebenista & Nowak, 2014).

In 1948, it was proven by Bernard Brodie and Julius Axelrod that the main analgesic metabolite in the derivatives on the market was actually paracetamol (Brodie & Axelrod, 1948). In 1955, paracetamol was introduced to the US market by McNeil Laboratories under the name 'Tylenol Children's Elixir' and a year later, paracetamol was sold in the modern form of the 500mg tablet in the UK under the trade name Panadol (Jóźwiak-Bebenista & Nowak, 2014). Mixes and variants of anilines have been on the market for quite some time, with recurring research into different derivatives. However, the last derivative, phenacetin, was withdrawn from the US market in 1983, while the last country in Europe to allow it was Poland (withdrawn in 2004). Phenacetin was withdrawn due to its carcinogenic effects. The anilines have a characteristic phenol ring which is a carcinogenic suspect, as the ring can function as an intercalating agent, inserting itself between nucleotides in a DNA strand.
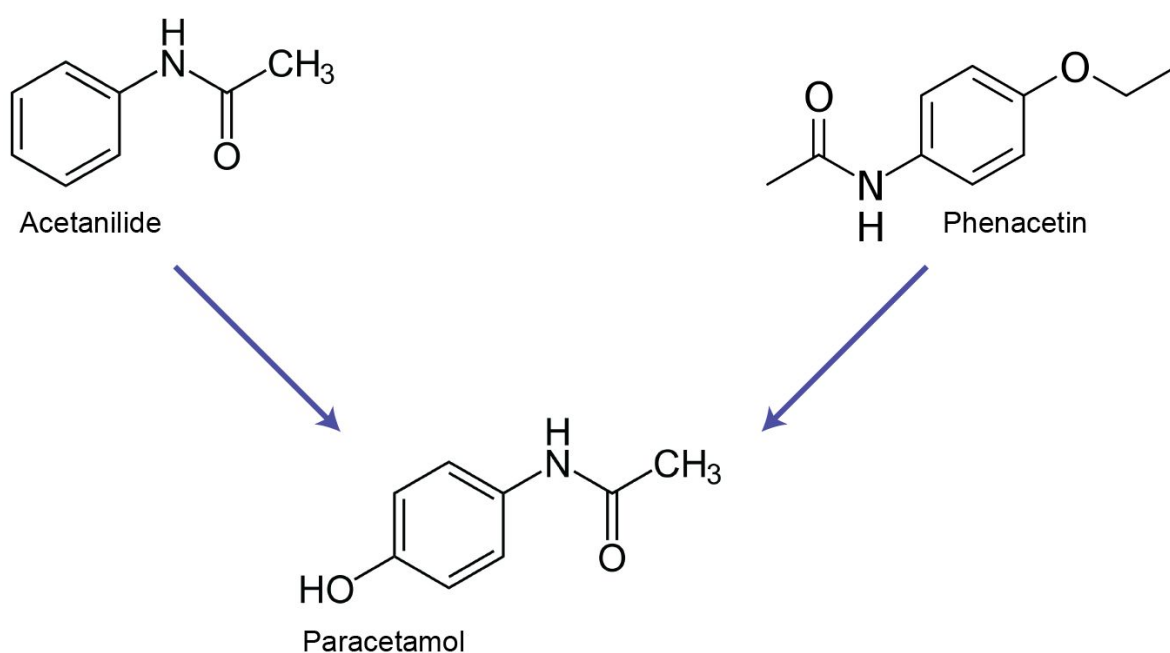
Figure 54 *Schematic of paracetamol's parent molecules* Paracetamol was first discovered as its sister molecule acetanilide which does not have the hydroxyl-group of paracetamol. Other similar molecules have been on the market, but the last one, Phenacetin, was withdrawn from the US market in 1983, and the European market in 2004 (last out was Poland), due to its carcinogenic effects.

Mechanisms of paracetamol

Paracetamol does not have the anti-inflammatory effects of non-steroidal anti-inflammatory drugs (NSAIDs), e.g. ibuprofen, but does have good analgesic and antipyretic qualities. The reason why paracetamol sometimes has been considered a member of the NSAID family is that paracetamol suppresses prostaglandin production, as all NSAIDs do. Paracetamol has a very narrow therapeutic window and is infamous for its liver toxicity, where a dose slightly higher than the therapeutic dosage can result in toxic effects. This is less common for most over-the-counter drugs, and when giving children paracetamol, the dose must be adjusted, as paracetamol metabolism changes with age and the difference between adults and children are quite large. What is interesting is not only a variable capacity of metabolism but also different activity of alternate metabolic pathways (Figure 55 *Paracetamol metabolism*). At birth, detoxification by sulfation is dominant (Adjei et al., 2008; Jóźwiak-Bebenista & Nowak, 2014) whereas the glucuronidation pathway becomes more effective later in life. Oxidation of paracetamol in infants via CYP2E1 is miniscule, similar to glucuronidation. This detoxification pathway requires some years of maturation.

Prostaglandin synthesis from arachidonic acid is catalyzed by COX-enzymes which exist in more than one isoform. The third COX isoform, COX-3, is subject to ongoing discussions regarding its function (Jóźwiak-Bebenista & Nowak, 2014) as it was shown to be sensitive to paracetamol. However, this investigation was carried out with the canine variant of COX-3, and it was later shown that the canine COX-3 is a special splice variant, while human COX-3 was not sensitive to paracetamol (Chandrasekharan et al., 2002; Kis et al., 2005). It has long been thought that paracetamol's analgesic and antipyretic effects were mediated via the compound acting as a COX1/2 inhibitor and also affecting in some way the central nervous system (CNS). The CNS effect of paracetamol was first shown in work published in the 1970s where it was shown that paracetamol reduced prostaglandin synthesis, and ten times more efficiently in the brain compared to the spleen (Flower & Vane, 1972), There have been several studies disproving COX-1/2 inhibition as the main mechanism of paracetamol, and while paracetamol is an effective inhibitor of the COX-2 enzyme it is highly dependent on a local redox environment

(change in oxidation states, e.g. in biology: metabolism of glucose when it is oxidized to CO^2, and oxygen is reduced to H_2O) (Hinz et al., 2008). Other work has suggested that the analgesic effect is caused by paracetamol acting on the 5-HT serotonergic receptor, as was demonstrated in rats when induced pain of the median nerve was blocked by paracetamol (Pickering et al., 2008). Studies subsequent to the proposed COX1/2, COX-3, and 5-HT serotonergic effects, have suggested that paracetamol is a pro-drug and that it is the metabolites of paracetamol that exert the effects, not paracetamol itself.

Paracetamol does have some quite interesting metabolites. In the brain and spinal cord of mice, it was shown that paracetamol is deacetylated to p-aminophenol which reacts with arachidonic acid in a reaction mediated by the enzyme fatty acid amide hydrolase (FAAH) to form N-arachidonoylphenolamine (AM404) (Figure 55 *Paracetamol metabolism*). AM404 is a strong activator of vanilloid receptor subtype 1, later named as transient receptor potential vanilloid 1 (TRPV1), which is a ligand for endocannabinoid receptor type 1 (CB1). The endocannabinoid receptors, and their endogenous ligands - the endocannabinoids -, are involved in fetal, natal, and post-natal development and have been shown to be involved in neural development and in neurodegenerative disease (Basavarajappa et al., 2009). While AM404 will activate TRPV1 it will also act as an endogenous cannabinoid reuptake inhibitor (Jóźwiak-Bebenista & Nowak, 2014). By blocking AM404 activity ahead of paracetamol administration to rates, researchers showed complete removal of the analgesic effects of paracetamol (Bertolini et al., 2006; Jóźwiak-Bebenista & Nowak, 2014). Another proposed mechanism of paracetamol's pain relief is by inhibiting nitrogen oxide formation via the L-arginine/NO pathway (activated by substance P and NMDA receptors) since NO is an important molecule in the spinal cord pain signal transmission (Bujalska, 2004). The complexity surrounding the mechanisms of paracetamol is a good example where multiple approaches to drug toxicity testing are reasonable to sure its safety for use, i.e. in pregnancy. Paracetamol has been shown to to cross the placental barrier freely (Koehn et al., 2019). With that paracetamol has differential fetal metabolization during maturation, there is an additional question of how the effect of paracetamol could change due to developmental timing.
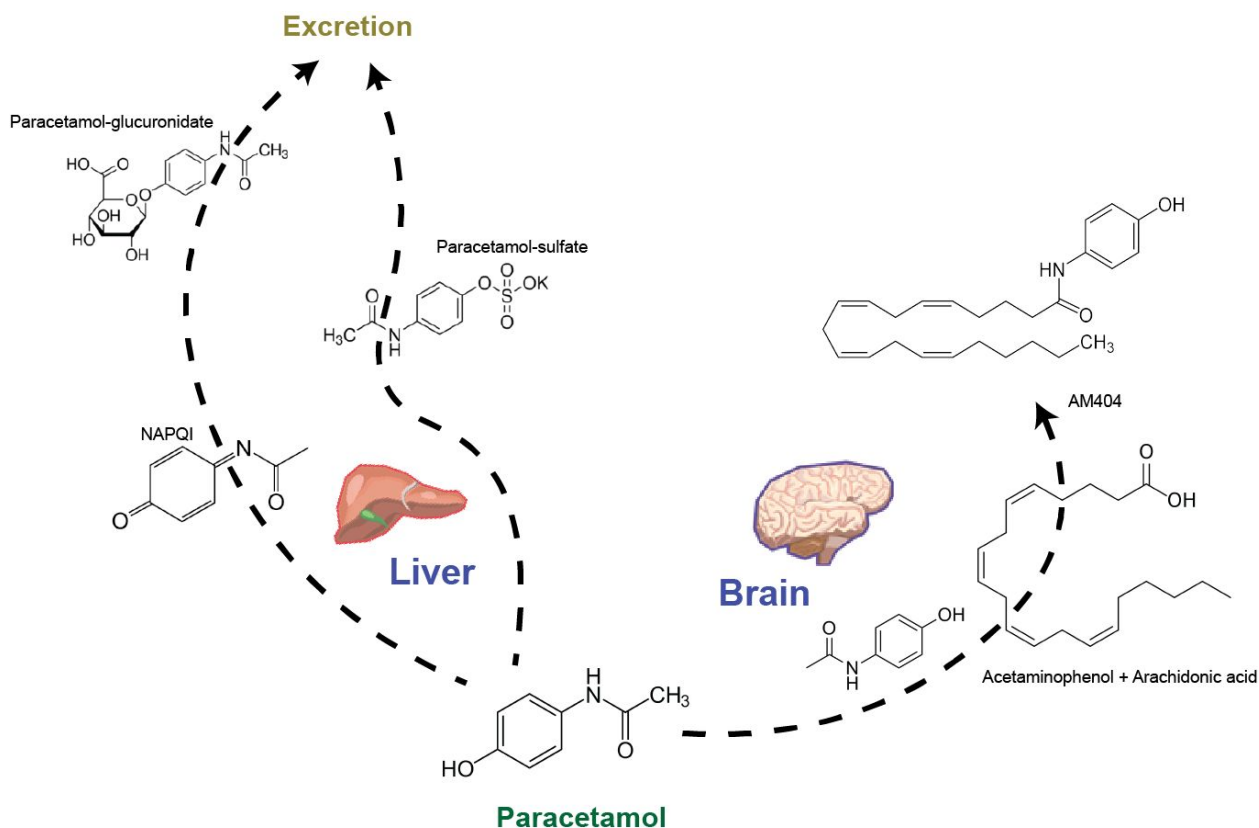
Figure 55 *Paracetamol metabolism* Paracetamol in the brain will be metabolized to acetaminophenol which together with arachidonic acid will form AM404. Paracetamol metabolism in the liver can create a very reactive metabolite NAPQI (Guo et al., 2004). NAPQI is conjugated to glutathione and then rapidly excreted in the urine unless glutathione is depleted with resulting severe liver toxicity. Paracetamol is otherwise excreted as either the glucoronidate- or sulfate-conjugates. The schematic is simple in design and there are more minor metabolites of paracetamol than shown.

Paracetamol and cohort studies

A cohort study is a study on a group that shares a common characteristic i.e. a large group of women in pregnancy and/or after pregnancy. Several cohort studies starting from 2013 have shown an association between the use of paracetamol in pregnancy and an increased risk of adverse neurological outcomes, e.g. language delay and ADHD (Avella-Garcia et al., 2016; Bornehag et al., 2018; Brandlistuen et al., 2013; Liew et al., 2014; Stergiakouli et al., 2016; Thompson et al., 2014). Within PharmaTox, previous work has indicated a possible effect of paracetamol on development, when used during long periods in pregnancy (Brandlistuen et al.,

2013; Gervin et al., 2017; Ystrom et al., 2017). Moreover, recent work from one of the groups in PharmaTox found an association between methylation changes in cord blood, long-term paracetamol exposure (> 20 days) during pregnancy, and an ADHD diagnosis in the offspring (Gervin et al., 2017), by using data from the Norwegian *Mother and Child* cohort study (MOBA) (Magnus et al., 2006).

Studying the effects of paracetamol using our neurotoxicology protocol

Cells were either untreated, the *CTR* group, treated with 100 µM paracetamol, the *P100* group, or treated with 200 µM paracetamol, the *P200* group. Both dosages, P100, and P200 are at relevant therapeutic levels from plasma levels (Graham et al., 2013). Paracetamol has a partition coefficient of 3.2 (between octanol and water) which means that paracetamol will likely diffuse passively through cell membranes. Furthermore, paracetamol's binding to other proteins in plasma is minimal, and it has a large volume of distribution (50 L) after intravenous injection, indicating that paracetamol reaches throughout the body without allocating to tissues (Graham et al., 2013). The P200 dosage is similar to the measured peak concentration in plasma after an intravenously injected therapeutic dose (1 g) of paracetamol, 200 µM, where the lowest concentration measured was 130 µM, and the P100 dosage, 100 µM, is thus just below the lower level. A paracetamol regiment for an adult is recommended to be up to 1 g paracetamol, at a maximum frequency of four times per day (Graham et al., 2013).

Paracetamol has been reported to enter the developing brain and cerebrospinal fluid at higher levels in the fetal rat compared to the adult rat (Koehn et al., 2019), where chronic paracetamol treatment increased the transfer even more (Koehn et al., 2019). One study reported that paracetamol crossed the placental barrier freely (Bremer et al., 2017). While this indicates that the drug crosses at 100% from mother to offspring, a second study using ex-vivo human placenta as a perfusion model reported 44-48% transfer of paracetamol, then for metabolites: 38-40% of paracetamol sulfate and 31-36% of paracetamol glucuronide were transferred. Both studies suggested that paracetamol crossed the placental barrier rapidly and that metabolites produced by the crossed as well, albeit at somewhat lower levels. The recommended dosage of paracetamol is

more than half for a young child compared to the dosage recommended for an adult. This is a potential problem, as the dosage that would maternally be transferred over the placental barrier from an adult dose, would be too high, and could cause risks for the developing fetus.

We changed media with fresh paracetamol added (100 µM group was called P100 and 200 µM group was called P200), once every 24 hours. This regimen was kept for the remainder of the protocol, day 0 to day 20. Most likely, it would never be recommended to take paracetamol for this long (20 days) without seeing a doctor. However, since paracetamol is considered safe to use during pregnancy and the reported use is high, situations, where this can occur, are likely. Interestingly, a sub-therapeutic dose of paracetamol has been reported as neuroprotective (Blough & Wu, 2011; Ghanem et al., 2016; D. S. Maharaj et al., 2004; H. Maharaj et al., 2006; Nazıroğlu et al., 2009; Saliba et al., 2019), which means the half-dose of paracetamol might behave differently than just exhibiting relative lower adverse effects than 200µM.

For our study on paracetamol's effect *in vitro* neurodevelopment of hESCs to NPCs, we had many questions after exploring the literature on the mechanisms of paracetamol. The literature was quite clear on the hepatotoxic effects of paracetamol, but much less clear on the potential neurotoxic effects that might occur during neurodevelopment. Although there are many metabolites of paracetamol, and combinations thereof, we left that complication on the side to initially treat with the pro-drug alone. It does raise questions for when, and what degree - if any, our cells start metabolizing paracetamol, and this would be interesting to answer, e.g. in a collaboration with a mass-spectrometry group where we would supply control and treatment medium for analysis. For these first experiments, we were hoping to measure some of the subtle, long-term effects, that a drug like paracetamol could cause, and that the results would contribute to showing why it is important that we put more national, and international effort, into creating better human neurotoxicity testing, for the many hundreds, if not thousands, of compounds that remain to be evaluated for projects such as REACH.

Part II - Results

Initial hESCs experiments and determination of the *in vitro* dose of paracetamol

As mentioned in the introduction, paracetamol is still under debate regarding many of its biological mechanisms. However, there is also concern about how to consider the many metabolites paracetamol can be converted into, of which several might be relevant to study. Another issue is to what extent paracetamol will be degraded in the cell medium. No reliable data on the half-life in medium exist or on how fast paracetamol and all its metabolites can be transferred across natural barriers, such as the blood-brain and the placental barriers. Paracetamol has also been reported to have negligible binding to carrier proteins and to accumulate in tissues, thus in vitro concentrations may be calculated based on the in vivo concentration in plasma, which was shown to be in the 130-200 µM range after a 1 g dose (2x 500 mg standard paracetamol tablets) (Graham et al., 2013).

Another issue of importance was at what stage during pregnancy exposure to paracetamol would take place. The unborn child will have a changing expression of enzymes and other systems during fetal development. This means that it would be of high interest to know when the offspring itself develops the capacity to metabolize paracetamol, and whether it happens in the fetal liver or brain, or whether metabolites are transferred primarily from the mother. Nonetheless, it was decided to start with the addition of the parent molecule, paracetamol, to the cell medium and carry out treatment experiments, but we knew there were a number of factors involved that could be important to investigate.

About 2 years before the in-house protocol was established, we tested several different concentrations of paracetamol in hESC cultures. hESCs were first grown for 24 hours, to allow the cells to stabilize, attach, and be in the log-phase after 24 hours. After 24 hours, paracetamol was added (with another dose the day after) in increasing concentrations: 0.1, 10, 50, 100, 150, 200, 500, and 1000 µM. After 24 hours of treatment, viability óf hESCs was measured by luminescence using two commercial kits: CellTiter from Promega measuring ATP, and GSH-Glo

from Fisher Scientific measuring GSH (oxidative stress) (Figure 56 *GSH-Glo and CellTiter hESCs viability after paracetamol treatment*). The hESCs seemed to tolerate all the paracetamol dosages well. However, the duration of the tox-tests was short in comparison to our current 20 days long neurodifferentiation protocol, but at the start of the project we did not know what dosages we were going to use, or for how long we were going to treat our cells. We expected to have paracetamol present in the culture for longer periods than 48 hours, so we were not certain how the same concentrations would be tolerated during a longer incubation period.
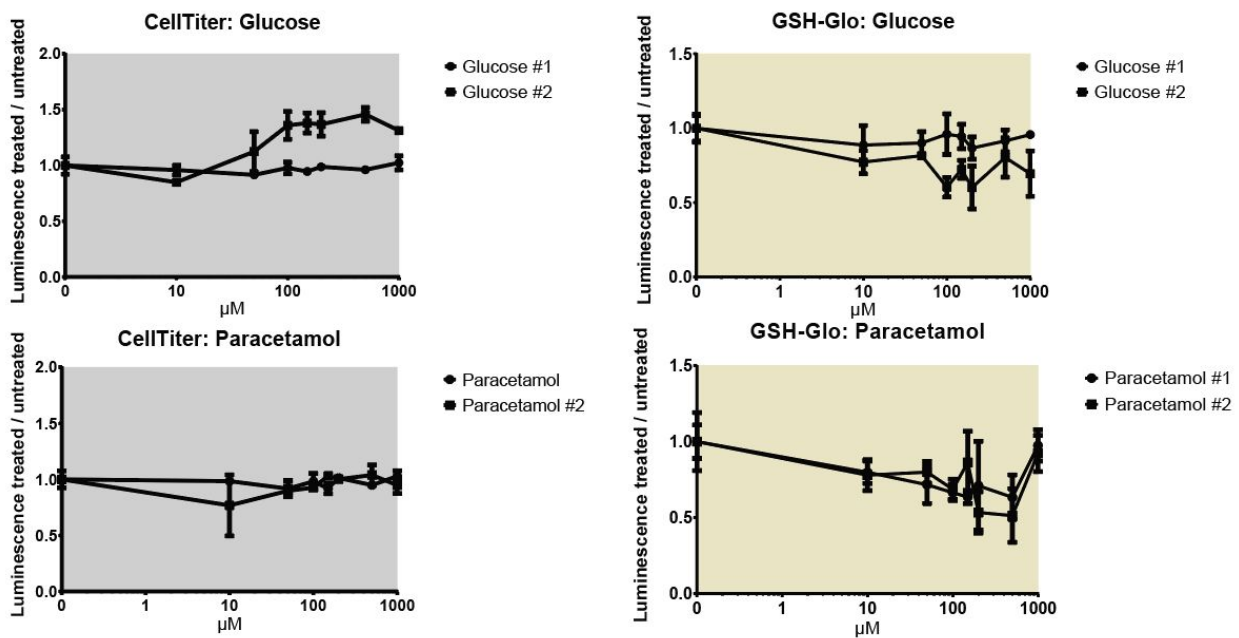
Figure 56 *GSH-Glo and CellTiter hESC viability after paracetamol treatment*. Approximately 12 000 hESCs were seeded, per well, in the 96-well plate format. After 24 hours of cell growth, paracetamol, alternatively control molecule glucose, was pre-mixed into medium, which then replaced the old medium in 96-well plate. After 24 hours of exposure, decreased cell viability was measured by luminescence. hESCs seemed to tolerate all paracetamol dosages well. On the left, for CellTiter, the cell viability, and level of luminescence measured, are relative to levels of cellular ATP, which decreased when the cell viability dropped. On the right, for GSH-Glo, cell viability was measured as oxidative stress, where a lower GSH pool meant less luminescence measured.

Proliferation of cells seemed to increase when they were exposed to paracetamol during protocol runs

For the proliferation discussion and single-sequencing results and discussion, the 100 μM paracetamol treatment group, and cells within, will hereafter be known as P100. Similarly, 200 μM will be known as P200, and control as CTR.

Anecdotally, we observed an increased cell proliferation in our cultures upon paracetamol treatment during the neurodifferentiation protocol. Part II of the protocol is a maturation phase starting at day 7 and lasting to day 13, where at day 13 cells are transferred to part III for the last phase of the protocol. Peculiarly, at day 13 the cell-counts for the transition to part III, yielded a higher number of cells in P100 and P200 groups, compared to CTRL, which could be easily observed as more dense cultures in the culture wells. In addition, the treated cells were morphologically different from the untreated cells. In retrospect, this would have been an interesting observation to address immediately with proliferation assays. Figure 57 shows brightfield images of CTR and P200 cells after 20 days (at the end of the protocol), where there is a noticeable difference in cell density and morphology. In CTR wells, some gaps in the cell monolayer and an intricate dendritic network were observed, whereas this was rarely seen for P200, which displayed a denser cell monolayer, with cells looking morphologically different from the control cells during the third part of the protocol. The increase in proliferation is just an observation at this point as we have not done the proper experiments to prove any change in proliferation. However, as there are further indications of increased proliferation in P100, and P200, showing up in the scRNA-seq data, the proliferation increase due to paracetamol exposure is highly interesting.
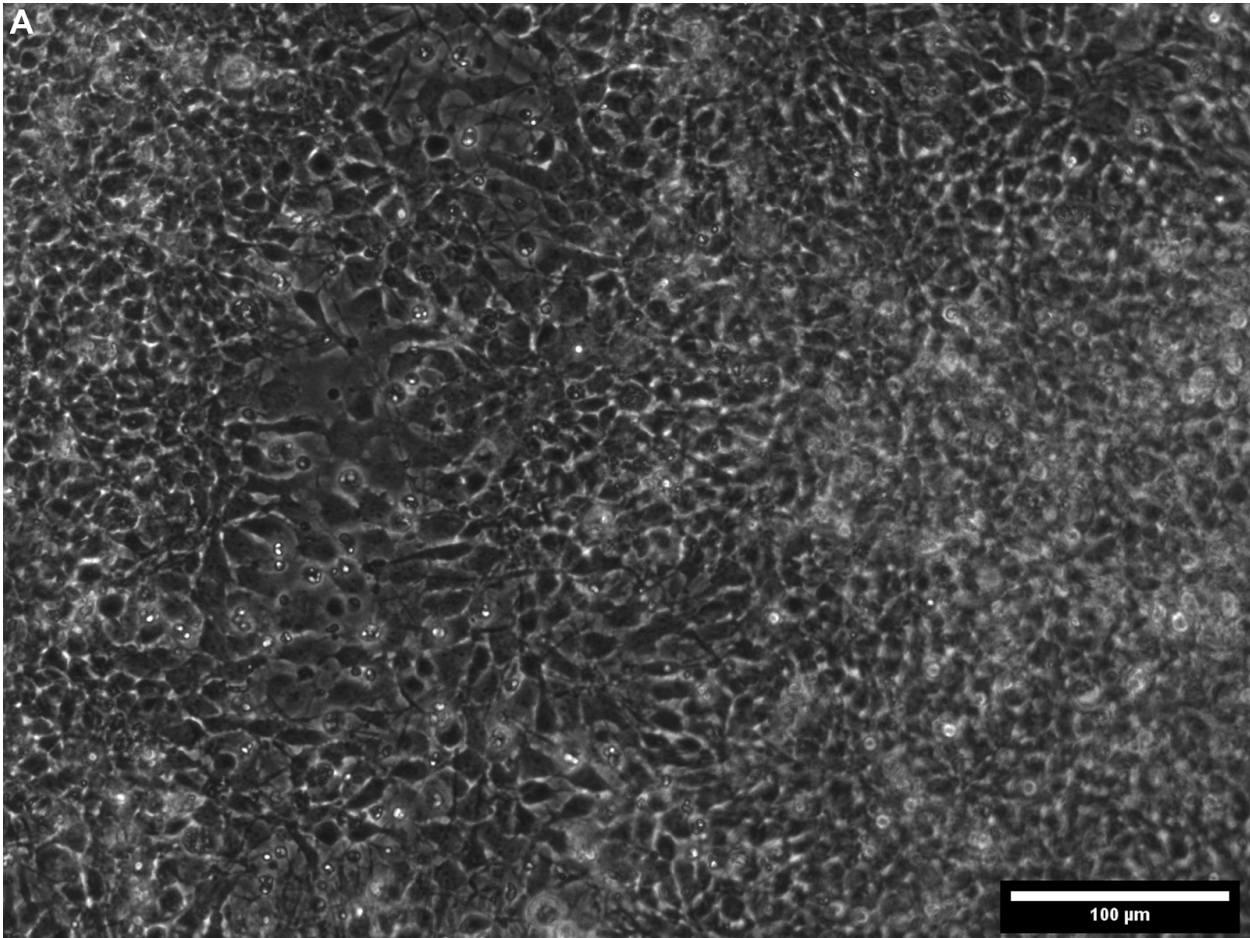
Figure 57 *Brightfield image at day 20 of CTR cells.* The same number of cells (450k) were seeded in CTR and P100/P200, at the start in each part of the neurodifferentiation protocol. But there seemed to be a difference in cell numbers at the end of part II and III in the paracetamol treatment groups. **A.** Image of CTR cells. In more open areas, between cells in CTR, a network can be observed. Magnification is 20x. Scalebar = 100 μm.

Note: Figure 57 is split into Figure 57 and Figure 58 so that could be made as large as possible for the A4 format. Hopefully, once viewed in a PDF, one can zoom in and compare them better.
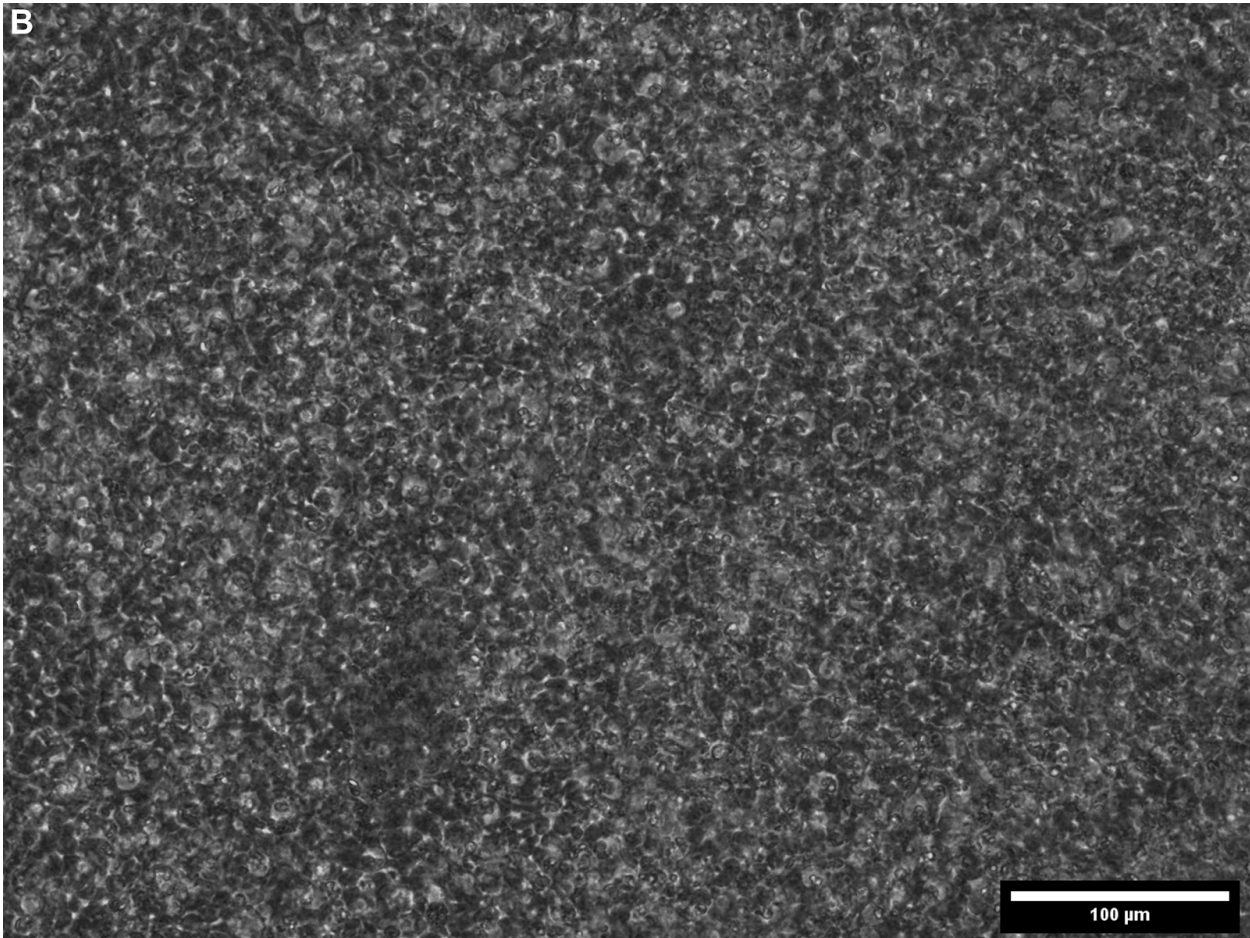
Figure 58 *Brightfield image at day 20 of P200 cells.* The same number of cells (450k) were seeded in CTR and P100/P200, at the start in each part of the neurodifferentiation protocol. But there seemed to be a difference in cell numbers at the end of part II and III in the paracetamol treatment groups. **B.** In P200 wells on day 20, cells grew much denser. Here, in comparison to previous Figure 57 *Brightfield image at day 20 of CTR cells*, there are no gaps in the monolayer, there is a thicker sheet of debris floating above the monolayer, and when images such as these were shown to colleagues in the lab, they would appreciate cell amounts to be the double compared to CTR even though cells were initially seeded at the same numbers. Magnification is 20x. Scalebar = 100 µm.

Single-cell sequencing of cells treated with paracetamol

For this thesis, control datasets used in part I was integrated with treatment data using Seurat (from the Satija Lab) (Butler et al., 2018; Hafemeister & Satija, 2019; Hart, 2019; *Satija Lab*, n.d.). Cell types in the integrated dataset were predicted by having Seurat assigning them these by using cell types established in part I as reference, with the new dataset as a query.

Changes in cell-population distributions in treatment with paracetamol

When comparing amounts of cells, per cell-type, in CTR *versus* paracetamol treatment, some effects were observed that confirmed what was previously seen in cell-culture brightfield microscopy and cell-counting (Figure 57 *Brightfield image at day 20 of CTR cells*, and Figure 58 *Brightfield image at day 20 of P200 cells*). Cell numbers were seen changing in particular cell type groups with paracetamol treatment (Table 8 *Day 20 Predicted cell-types in the treatment population*). In particular, radial glial cells, RGs, which is the relative proliferative and regenerate cell at this stage, increased from approximately 22 % to 32 % when comparing P100 and P200 to CTRL. For P200, some mature progeny cells decrease in numbers, in particular intermediate progenitors, whereas prefrontal cortex precursors (PCPs), displayed a relatively large increase. That PCPs increase in abundance is interesting, since the PCP cluster its located on the side of the intermediate progenitor cells (IPs), close to RGs (Figure 60 *Day 20 P200 and CTR dataset integration*), while IPs are on the direct trajectory to mature progeny, as predicted by cytoTRACE and PAGA pseudotime analysis, possibly suggesting downregulation of that particular trajectory in P200 treatment.

Cluster 7, non-annotated cells from part I

Another noticeable change in the integrated dataset is that "cluster 7", which was un-annotated and split into two parts in thesis part I, has more local structure to it now, likely due to the increased number of cells included in the dataset, as this gives clustering analysis more statistical power (Figure 59 *Day 20 P100 and CTR dataset integration*) (Figure 60 *Day 20 P200 and CTR dataset integration*). This integrated data might, therefore, be helpful when trying to annotate these cells, considering that they are in the region of cells affected by paracetamol treatment, it is relevant to investigate further.

Table 8 *Day 20 Predicted cell-types in the treatment population* Annotated control data from part I was used as a reference to predict cell-types in the treated cell populations. Columns 1-3 show the number of cells according to cell-type per CTR, P100, or P200. Columns 4-6 show the corresponding percentages each cell-type represent out of the group total. RGs = Radial glial cells (stem cell/progenitors), IPs = Intermediate progenitors, 5-HT = Serotonergic, PC = Precursors, IN = Interneurons, CTP = Corticothalamic precursors (CTPs), PCP = Prefrontal cortex precursors. Data is from single-cell sequencing of two repeated protocol runs.

Cell-type	CTR	P100	P200	CTR, %t	P100, %t	P200, %t
RGs	628	933	723	22.13%	33.14%	32.52%
RGs: stopped	451	456	312	15.89%	16.20%	14.04%
Intermediate progenitors	459	358	72	16.17%	12.72%	3.24%
Layer V PCs	222	196	180	7.82%	6.96%	8.10%
Layer IV/5-HT PCs	178	159	79	6.27%	5.65%	3.55%
5-HT/Chol/Glut PCs	264	192	161	9.30%	6.82%	7.24%
GABAergic INs	214	72	55	7.54%	2.56%	2.47%
CTPs	134	116	110	4.72%	4.12%	4.95%
PCPs	102	98	409	3.59%	3.48%	18.40%
Not annotated *	186	235	122	6.55%	8.35%	5.49%
Total	2838	2815	2223			
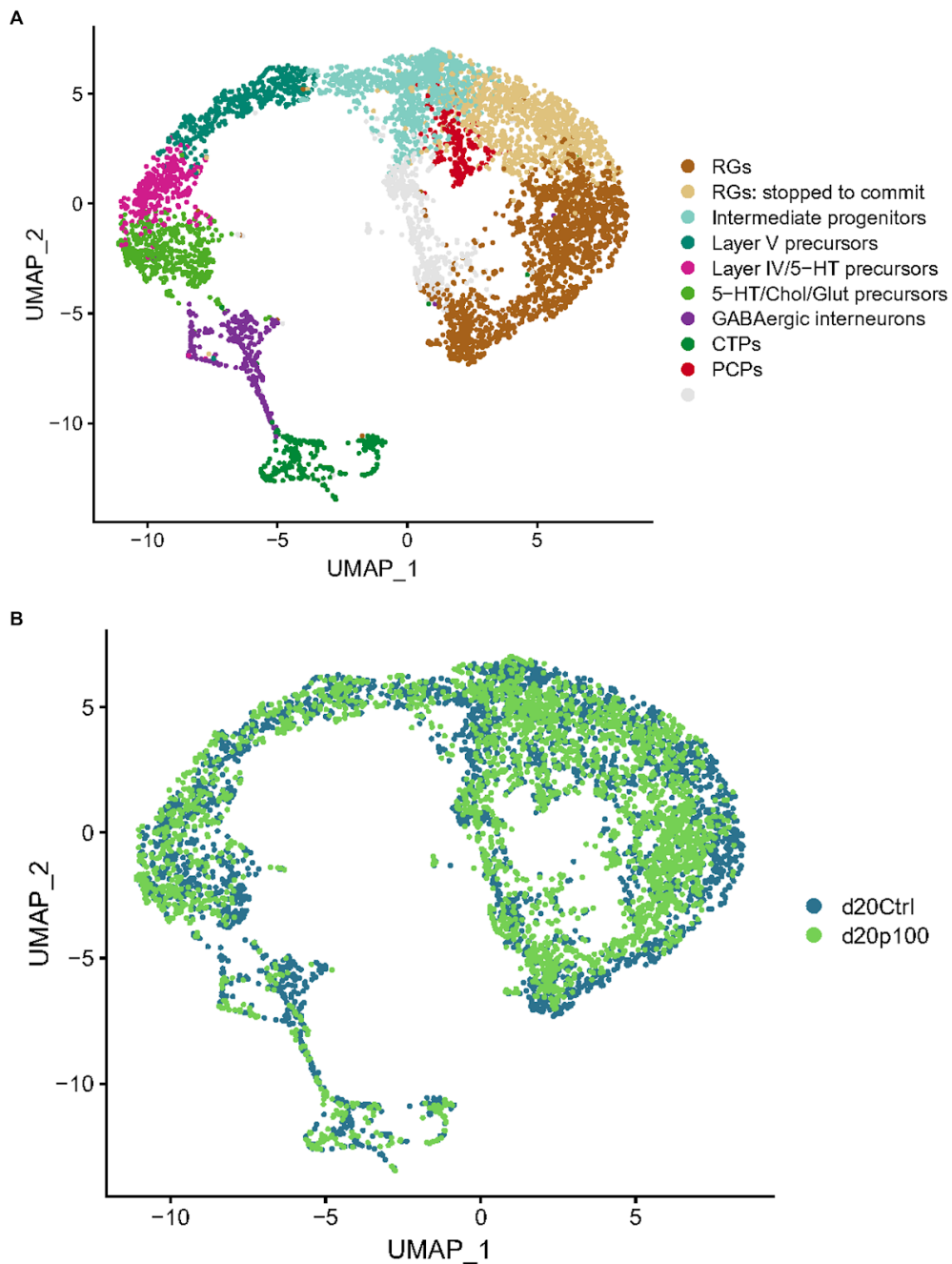
Figure 59 *P100 and CTR dataset integration* **5653 cells** were included in the dataset. A. Cell-types in the integrated dataset were predicted by using the annotated ctrl dataset from part I as reference. B. A UMAP displaying how CTR and P100 cells at day 20 are placed by Seurat integration.
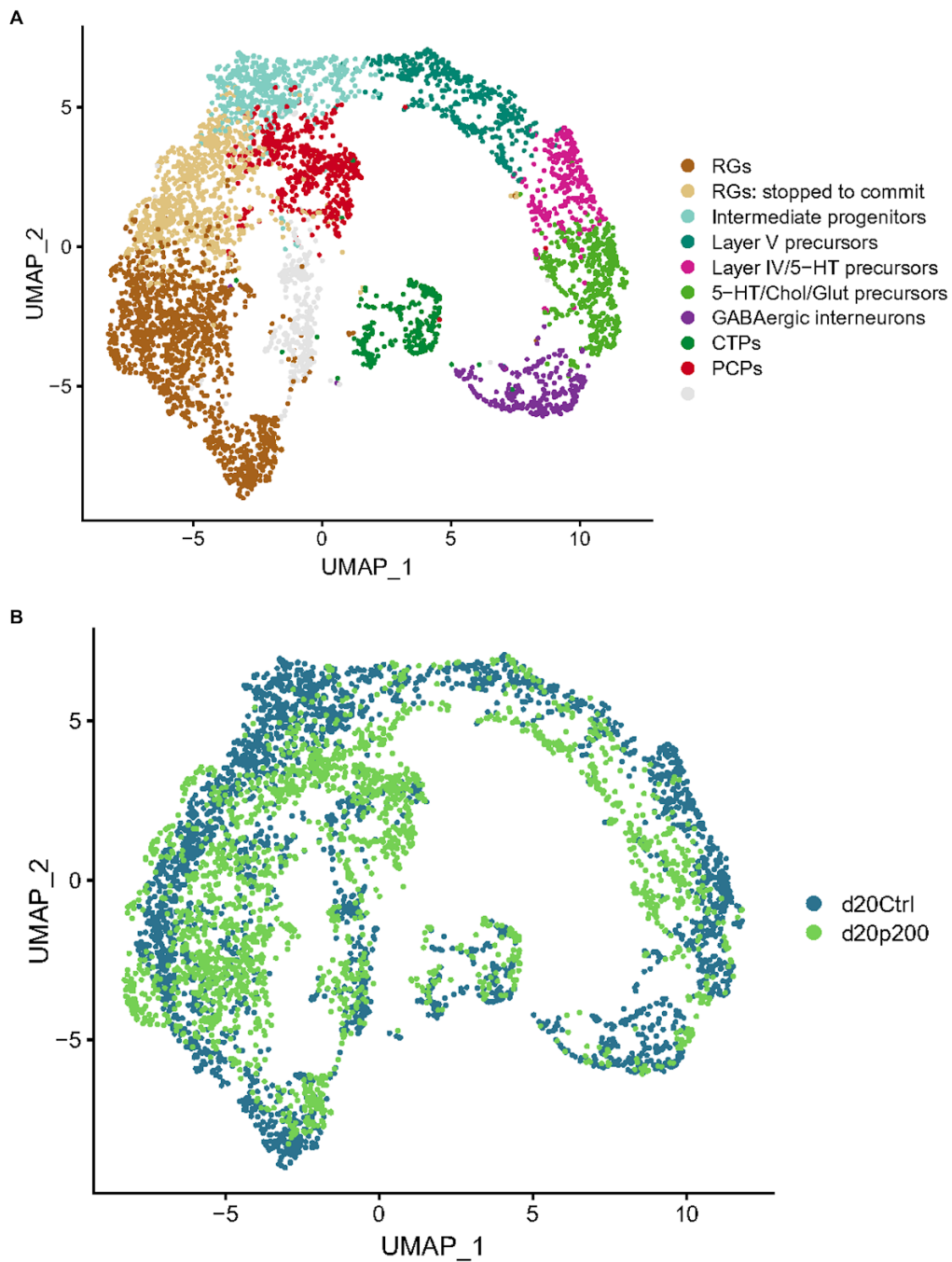
Figure 60 *P200 and CTR dataset integration* **5061 cells** were included in the dataset. A. Cell-types in the integrated dataset is predicted by using the annotated ctrl dataset from part I as reference. B. A UMAP displaying how CTR and P200 at day 20 are placed by Seurat integration. When comparing CTR (bright green) to P200 (blue) more P200 cells are observed at clusters with RGs and PCPs.

Selecting the P200 PCPs and inspecting these cells alone on a UMAP, illustrates how much more P200 cells appear as RGs and PCPs, compared to CTR, at day 20 (Figure 61 *Subsetting P200 and CTR for RGs and PCPs*).
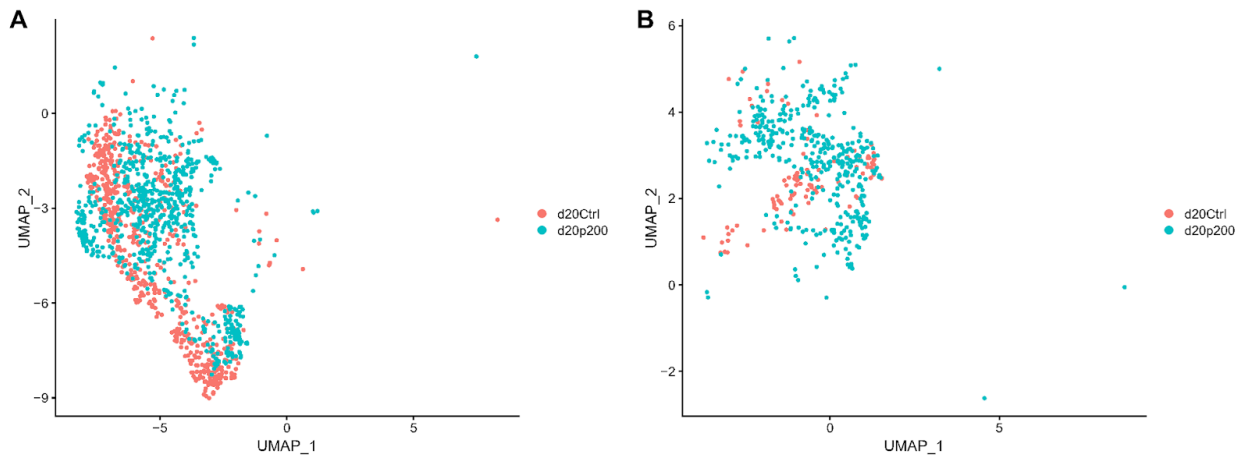
Figure 61 *Subsetting P200 and CTR for RGs and PCPs* Here, RGs (left) and PCPs (right) are isolated from the larger UMAP (Figure 59 *P200 and CTR dataset integration*) and presented alone on a new UMAP to more clearly demonstrate differences in cell numbers between CTRL and P200. A. The UMAP displays only RGs, and demonstrate a much higher number of P200 cells (turquoise) in comparison to CTR cells (red). B. The UMAP displays only PCPs, and similar to the RGs, a higher amount of P200 cells (turquoise) are visible compared to CTR (red).

PAX6 abnormalities and expression changes for several important genes

Experiments previously carried out by scientists in the PharmaTox environment have shown that there were *PAX6* changes in a chick brain model system after treatment with glucocorticoids (Austdal et al., 2016). Indeed, when analyzing the data obtained in this study, *PAX6* expression is suddenly detectable for P200 in RGs and PCPs (prefrontal cortex progenitors), whereas in CTR, *PAX6* was previously not detected at a significant level on day 20 (Figure 62 *Changes in PAX6 and NKX2-1 expression in P200 compared to CTR*). Large changes with *NKX2-1* expression were also noticeable.
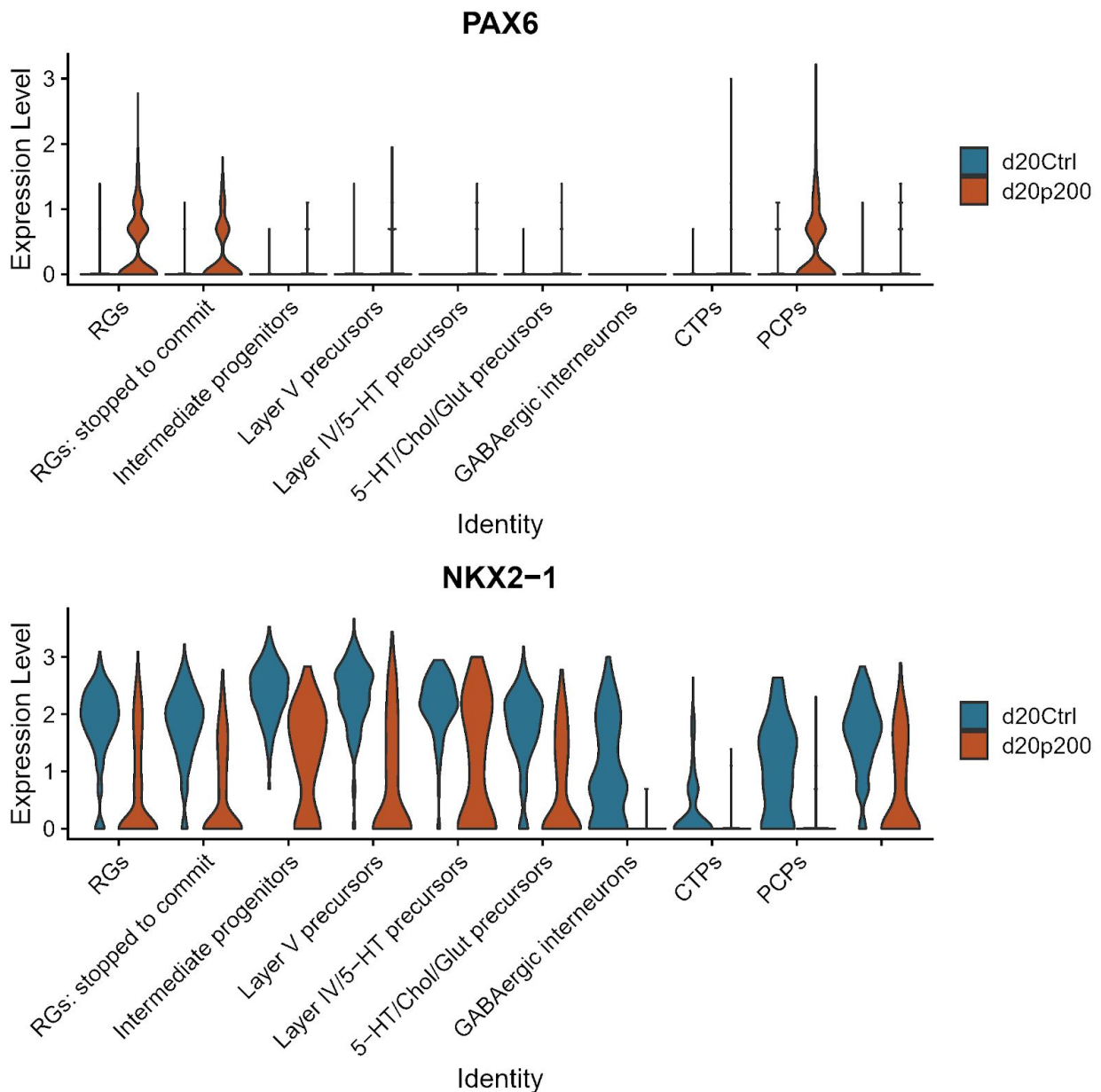
Figure 62 *Changes in PAX6 and NKX2-1 expression in P200 compared to CTR.* Top row, while *PAX6* had disappeared already by day 13 in CTRL, *PAX6* now appears with altered expression, at day 20 in P200 cell types; RGs, RGs: stopped to commit, and PCPs (prefrontal cortex progenitors). Bottom row, the forebrain marker *NKX2-1* is also significantly affected is expressed lower in P200 compared to CTRL, across all clusters except Layer IV5-HT precursors.

When comparing changes in gene expression within the integrated data (CTR *vs*. treatment), P200 stands out on in terms of how much changes in gene expression are induced for particular markers. As observed for the two RG clusters (Figure 63 *RGs and RGs: commit, P100, and P200 vs. CTR*), P200 induced larger changes in comparison to the half concentration. In particular, a large decrease in the expression of *C1orf61*, *SFTA3*, *NKX2-1*, *MGST1*, *FABP7*, *MGST1*, and *MBIP* was observed, when compared to CTRL. Concomitantly, there was a small increase in the expression of *WNT7B*, *PAX6, COL2A1, FTL, GSTP1*, and *ACTG1*. The increase in expression of *PAX6* is not that striking at first glance, but since *PAX6* is an important neurodevelopmental marker and patterning factor, for which its disappearance in thesis Part I suggested to be a ventral telencephalon fate, observing it in the 200 µM group is therefore interesting.
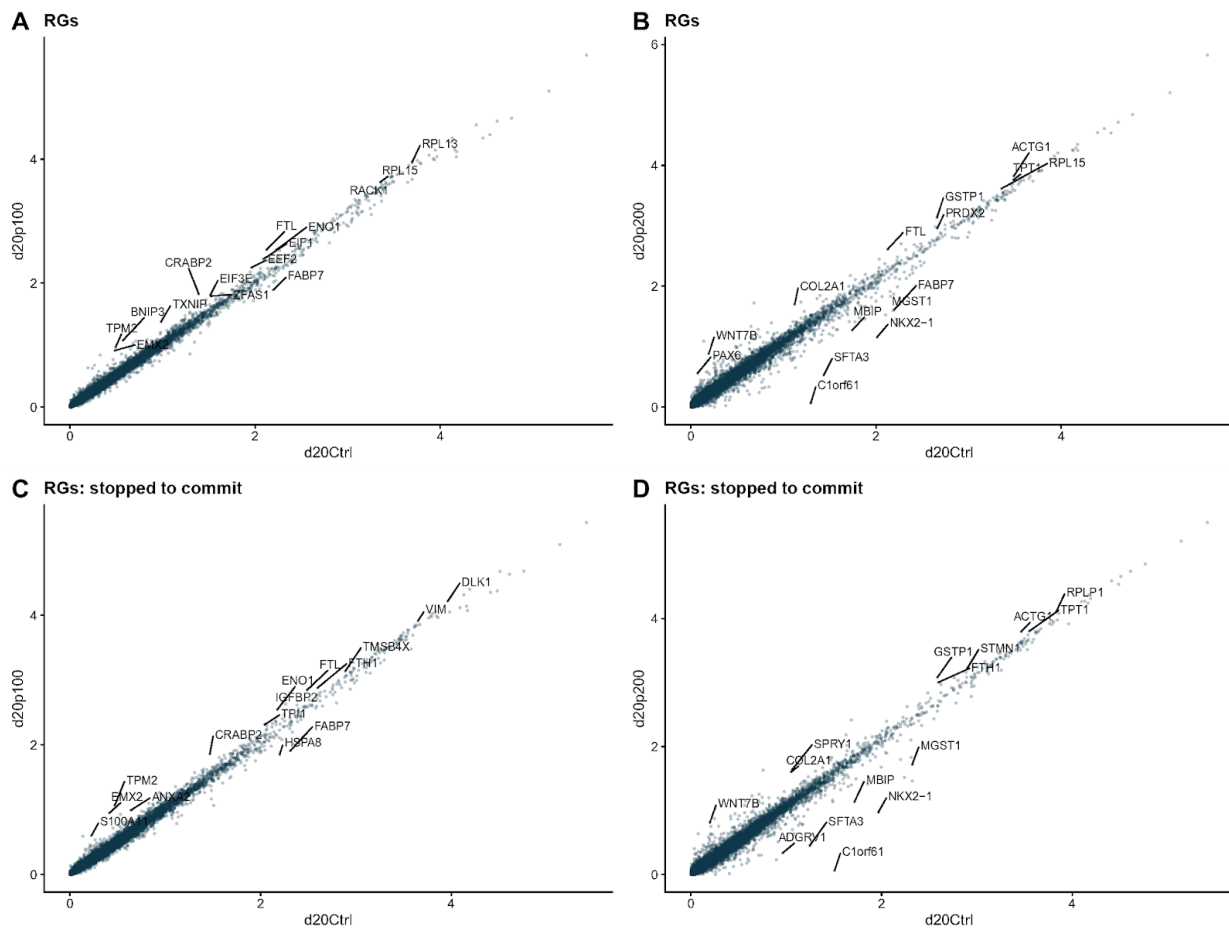
Figure 63 *RGs and RGs: commit, P100 and P200 vs. CTR.* Scatterplots depicting changes in gene expression for P100, and P200 compared to CTR, including the two RG cell types: *RGs* and *RGs: stopped to commit*. A. and C. While P100 usually show smaller changes in terms of fold change, B. and D. P200 induce larger changes, in particular to *C1orf61*, *SFTA3*, *NKX2-1*, *MGST1*, *FABP7*, *MGST1*, and *MBIP* which are downregulated, and then a smaller increase for expression in *WNT7B*, *PAX6*, *COL2A1*, *FTL*, *GSTP1,* and *ACTG1*.

In the PCP population, we observe that gene expression in the P100 treatment group after 20 days is not affected to the same extent as in the P200 treatment group (Figure 64 *PCPs, P100 and P200 vs. CTR*), similarly to what is presented in figure 63. For the P200 cells, *NKX2-1, MBIP*, and *MGST1* gene expression decreased - similarly to what was observed for RGs (Figure 63 *RGs and RGs: commit, P100 and P200 vs. CTR*), and also *HES5* and *SMS. EMX2* and *CTSC* are the two genes that increase their expression the most in the P200 group. The other cell-type groups display similar patterns (SII Figure I *Genes changed for IPs and Layer V precursors after paracetamol treatment*, SII Figure II *Genes changed for Layer IV/5-HT and GABAergic interneurons after paracetamol treatment*, and SII Figure III *Genes changed for 5-HT/Chol/Glut PCs and CTPs after paracetamol treatment*.).
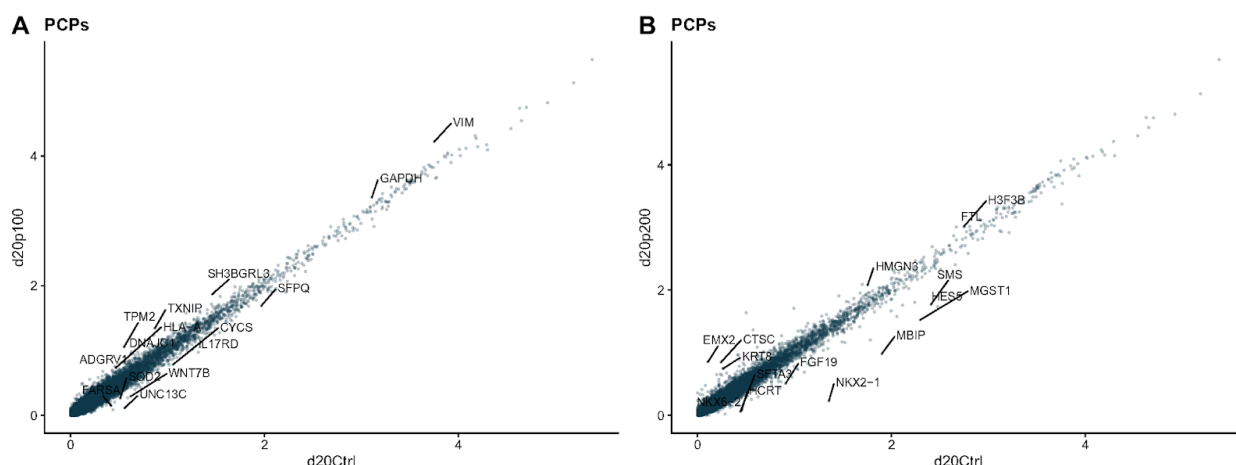
Figure 64 *PCPs, P100 and P200 vs. CTR.* Scatterplots depicting changes in gene expression for P100, and P200 compared to CTR. A. P100 is not affected as much as P200, *VIM* stands out slightly. B. 200 µM again seems to downregulate certain genes, *NKX2-1, MBIP*, and *MGST1. HES5* and *SMS* is also decreased within this cell-type group.

The observation that some common genes were reduced or increased in expression, triggered the analysis of expression of the genes that were observed to be altered the most in the P100 and P200 compared to CTR. These genes were *SFTA3, NKX2-1, HES5, MGST1, MBIP, COL2A1, CXXC5, SPRY1, EMX2, ZIC2, C1orf61, NAV1, PAX6, WNT7B, CRABP1*, and *LHX2*.

In the P100 group, the below dot-plot does not show striking changes in gene expression. The two genes that appear to change significantly are *EMX2* and *C1orf61*. *EMX2* expression increased and *C1orf61* expression decreased in P100 cells when compared to CTR (Figure 65 *Dot-plot on changed genes after P200 treatment, P100 vs. CTR*).
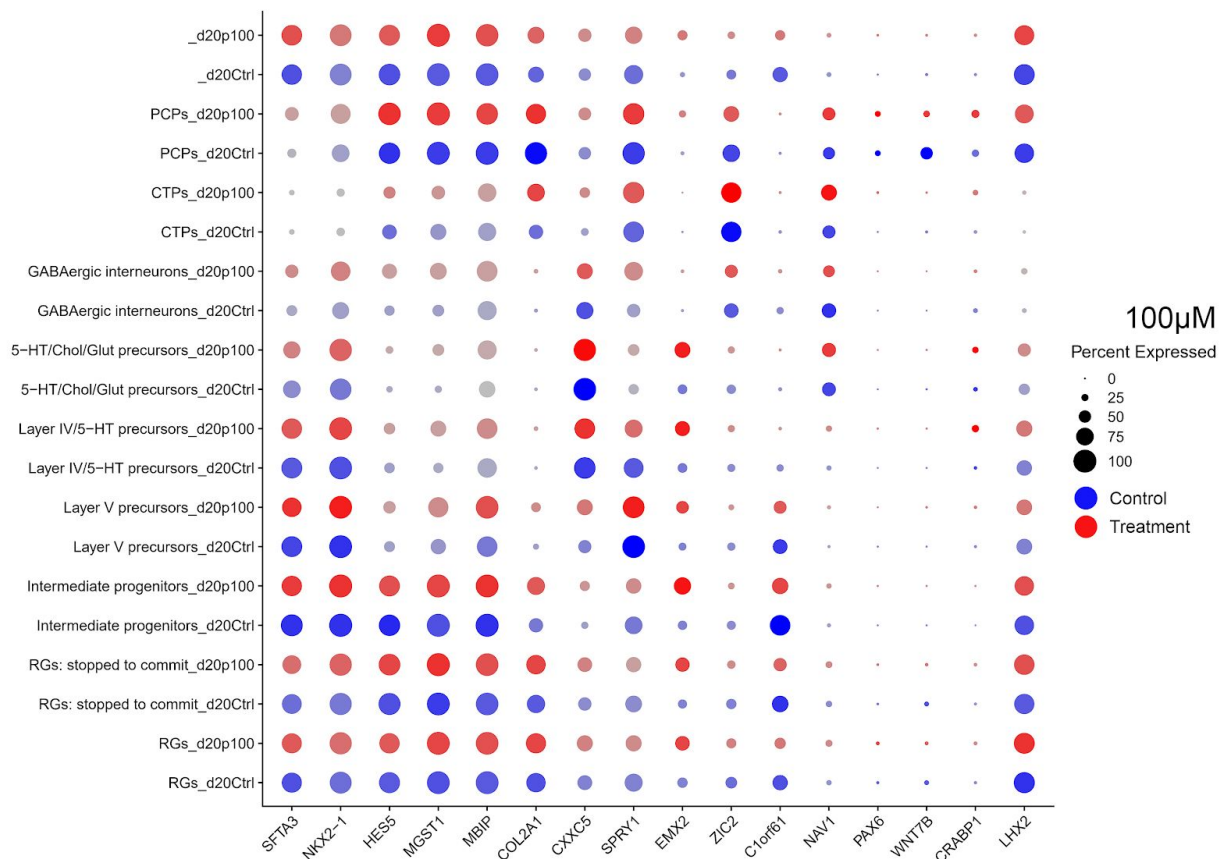
Figure 65 *Dot-plot on changed genes after P200 treatment, P100 vs. CTR.* The genes analyzed by dot-plot were commonly appearing changed in P200 cells and was also compared to P100 cells for comparison. At day 20 in P100, genes appear not to be changed in any large way. *EMX2* and *C1orf61* appear to be the two genes that visibly changed in P100.

After a more close inspection of P100-relevant parts of the scatterplots, only five genes appear changed in treatment, *EMX2, C1orf61, TPM2, BNIP3,* and *DUSP6* (Figure 66 *Dot-plot on the few genes that changed in expression in P100 vs. CTR*).
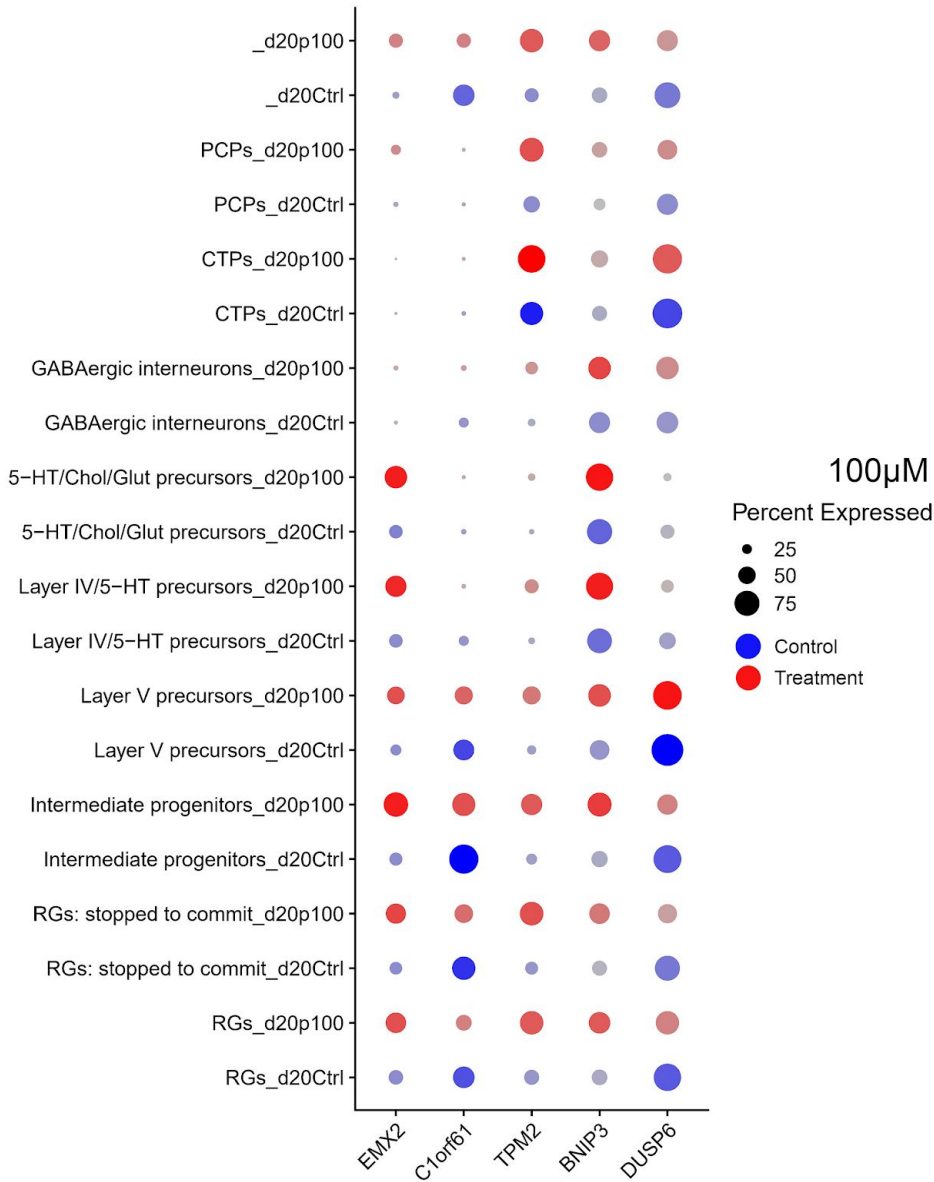
Figure 66 *Dot-plot on the few genes that changed in expression in P100 vs. CTR.* Only five genes were seen changing: *EMX2, C1orf61, TPM2, BNIP3,* and *DUSP6,* in P100 cells. They might particularly interesting if they can be related to a lower dose of paracetamol having a milder and neuroprotective effect.

P200 treatment induces changes in more genes compared to P100

In P200 cells, the larger list of genes as shown above the last figure, show bigger changes in expression, as expected for the P200 cells which the list was based on (Figure 67 *Dot-plot on changed genes in P200 vs. CTR*). *SFTA3*, *NKX2-1*, *HES5*, MGST1, *MBIP*, *COL2A1*, *CXXC5*, *SPRY1*, *EMX2, ZIC2*, *C1orf61*, *NAV1*, *PAX6*, *WNT7B*, *CRABP1*, and *LHX2* all change and have different patterns depending on the cell-type group. For the two RG-related cell-type groups *SFTA3*, *NKX2-1, HES5, MGST1*, and *C1orf61* decrease in expression, while *PAX6* appears in P200 cells, a quite striking appearance since *PAX6* was already at very low levels by day 13 in CTR, and gone by day 20, relative to P200. *COL2A1, CXXC5, ZIC2, NAV1, WNT7B*, and *SPRY1* appear to increase the most. For PCPs, *SFTA3* and *NKX2-1* appear to almost vanish although they are relatively low to start with. *HES5*, *MGST1,* and *MBIP* decrease in comparison to control, while *PAX6, NAV1,* and *WNT7B* appear to increase slightly.
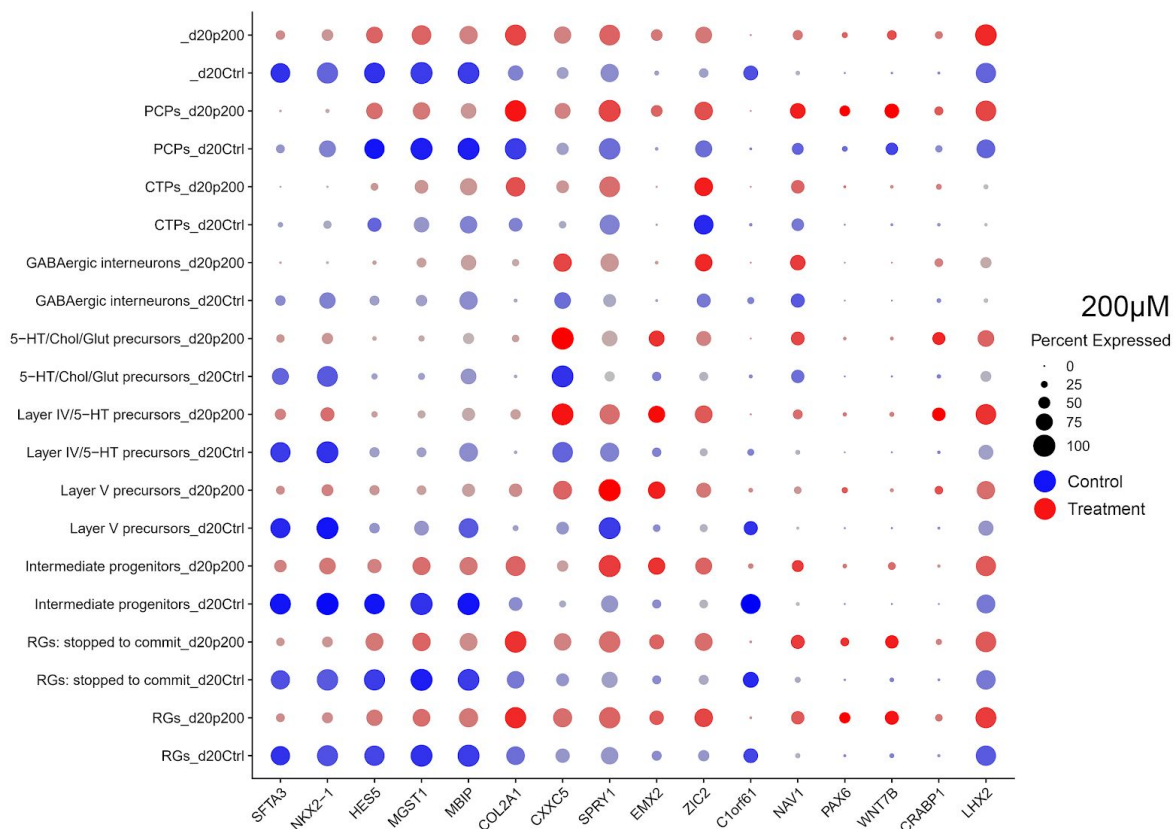
Figure 67 *Dot-plot on changed genes in P200 vs. CTR.* At day 20 in the 200μM treatment group, there were changes in expression for a collection of interesting genes: *SFTA3, NKX2-1, HES5, MGST1, MBIP, COL2A1, CXXC5, SPRY1, EMX2, ZIC2, C1orf61*, *NAV1*, *PAX6, WNT7B, CRABP1*, and *LHX2*.

134

Changes in gene expression in P100 cells are less compared to P200 cells

EMX2, TPM2, and *BNIP3* displayed increased expression levels in the P100 cells, while *C1orf61* and *DUSP6* displayed reduced expression compared to CTR. Only *BNIP3* has been discussed previously in part I of this thesis. *BNIP3* was predicted as an end-of-differentiation marker at day 0 by cytoTRACE and has been shown to have a role in both autophagy and apoptosis (a member of the apoptotic Bcl-2 family). However, *BNIP3* seemed here to be mostly implicated for autophagy in terms of general cell biology and pluripotency.

EMX2 also seemed to be expressed at higher levels in P100 and P200 cells, compared to CTR. Where the expression of *EMX2* look similarly higher in the two treatments compared to CTR, P200 show a partial region of high expression along the end of Layer V PCs, through Layer IV/5-HT PCs, and fading just before the GABAergic interneurons (Figure 67 *Dot-plot on changed genes in P200 vs. CTR,* Figure 68 *EXM2 expression in integrated P100/P200-CTR datasets, and* Figure 60 *Day 20 P200 and CTR dataset integration* for cell-types UMAP).

EMX2 is a homeodomain TF critical in CNS development. Mice lacking *EMX2* fail to develop neurons that project to the neocortex, will not develop the hippocampal dentate gyrus, and will have issues with lamination and neuronal migration (Mallamaci et al., 2000). *EMX2* has been shown to promote neurogenesis and neurite outgrowth and was suggested as exploitable for cell-based brain repair (Brancaccio et al., 2010).

.
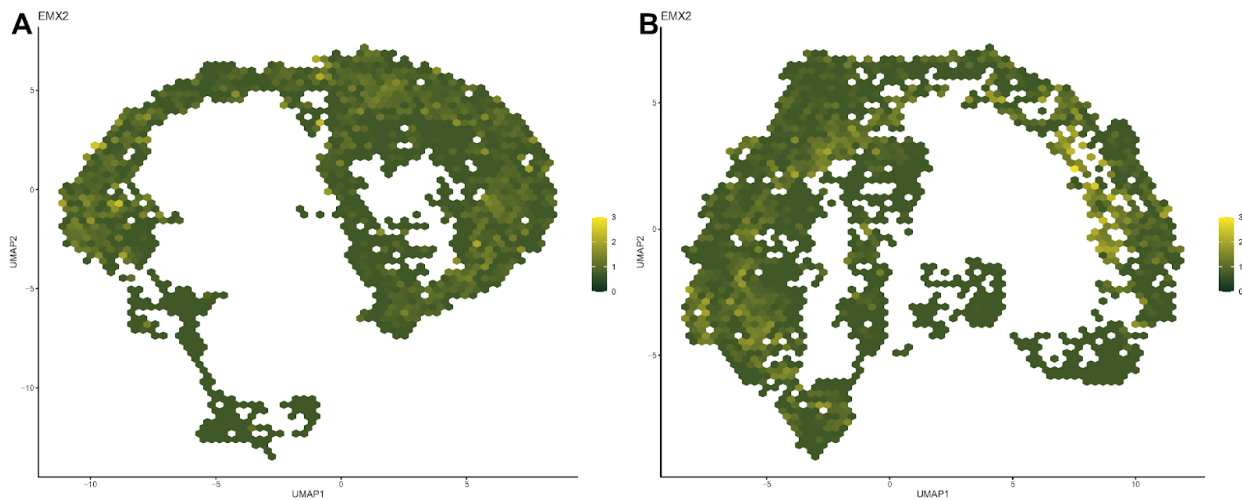
Figure 68 *EXM2 expression in integrated P100/P200-CTR datasets. EMX2* are expressed higher in both treatments compared to control, however, some areas in P200 seem to show a particularly high expression of this gene. A. P100-CTR integrated dataset shows a smooth *EMX2* expression along with the clusters except at the end towards GABAergic interneurons, where *EMX2* expression no-longer is observed. B. In P200-CTR integrated dataset, the *EMX2* expression does not seem that different to P100 at first, but along the end of Layer V PCs cluster through Layer IV/5-HT PCs, and fading just before the GABAergic interneuron cluster there is a distinct higher expression of *EMX2* compared to P100. Schex-hexagonal UMAP with gene mean expression per hexagon. Legend is a color gradient indicating a mean gene expression ranging from 0 to 3.

TMP2 was also observed to have significantly higher expression in the P100 cells, where the expression was increased for all cell-types in the dot-plot, except possibly for the 5-HT/Chol/Glut PCs, which had too low levels of *TMP2* for comparison between P100 and CTR. *TPM2* (encode protein β-Tropomyosin) is primarily known as a striated muscle protein that stabilizes actin filaments but can function in non-muscle cells as well according to GeneCards and UniProt. A link to neurodevelopment is not evident from the literature, but one study proposes that *TPM2* expression is important for cytokinesis, which seems reasonable as it affects actin filament stability (Thoms et al., 2008). Also interesting is the fact that *TPM2* is expressed in mouse stem cells, where deletion has proven difficult, as it seemed lethal (J. Hook et al., 2004; Jeff Hook et al., 2011). These types of proteins have roles in stress fiber assembly, that outside muscle cells are the main contractile bundles, with roles in morphogenesis, adhesion, and mechanosensing (Tojkander et al., 2012).

Two genes were observed to have significantly reduced expression in P100 cells, *C1orf61* and *DUSP6*. According to UniProt, *C1orf61* (Chromosome 1 Open Reading Frame 61, or Transcriptional Activator of the c-Fos Promoter), may have a function in *FOS* signaling which is important in neuronal development and remodeling. Only 7 publications are listed on GeneCards, where one presents the gene in a role as a novel activator of transcription in the brain, but under the alias *CROC-4*. *C1orf61* is described to promote transcription of the *FOS* promoter and is observed expressed in several brain regions, such as the thalamus, substantia nigra, and hippocampus (Jeffrey et al., 2000). In the same study, *C1orf61* expression is demonstrated using ICC in migrating and proliferating cells during early development in the rodent brain, which is relevant to this study and the effects of paracetamol. Furthermore, c-Fos signaling is important for learning and memory (Chung, 2015). What a variable decrease in a promoter for c-FOS could mean long-term might be interesting, since C1orf61 expression is much more downregulated P200 cells.

P200 treatment induces large changes in gene expression in comparison to P100

At day 20 in the 200μM treatment group, there were changes in expression for a larger collection of interesting genes: SFTA3, *NKX2-1*, *HES5*, MGST1, MBIP, COL2A1, CXXC5, SPRY1, EMX2, ZIC2, *C1orf61*, NAV1, *PAX6*, WNT7B, *CRABP1*, and *LHX2*.

Genes mentioned among those affected in P100 cells were *EMX2* (enhanced expression) and *C1orf61* (reduced expression). These genes were affected in P200 as well, but to a much larger extent (Figure 67 *Dot-plot on changed genes in P200 vs. CTR*). *C1orf61* expression is heavily reduced in all cell-types except a very small presence in intermediate progenitors when comparing P200 to CTR. *EMX2* was at a high level in four cell-types while having a low expression level in control cells (all cell-types except CTPs and GABAergic interneurons). *EMX2* expression seemed to increase the most for Layer V PCs, Layer IV/5-HT PCs, 5-HT/Chol/Glut PCs, and PCPs (PCPs had an almost zero expression in CTR).

NAV1, or Neuron navigator 1, is an interesting gene. It displayed increased expression in RG clusters and IPs P200 cells. The gene also had a slightly higher expression level for more mature progeny cell-types. There is limited information on NAV1 in the literature, but GeneCards report *NAV1* to have an ortholog in *C. elegans*, which was shown to be involved in axon guidance (Maes et al., 2002). GeneCards report that *NAV1* can also be alternatively spliced to several splice variants.

MGST1 is reported by GeneCards to be involved in inflammation and oxidative stress. Increasing numbers of proliferative cells and therefore enhanced cell turnover could lead to activation of genes, such as these and e.g. genes related to autophagy. Interestingly, oxidative stress was one condition reported in a previous study on paracetamol use in relation to ADHD from PharmaTox (Gervin et al., 2017). *SPRY1* is reported by UniProt to function as an *FGF* pathway antagonist. *SPRY1* has a higher expression level P200 cells and has been shown to be inversely correlated with proliferation and migration in keratinocytes (P. Wang et al., 2018) However, inactivation of *SPRY1* has also been shown to inhibit the muscle stem cell pool in aging humans, suggesting that *SPRY1* can have either effect. It does seem likely, however, that it would be pro-stem/glial in this study considering results so far, and as elaborated on in the coming discussion.

ZIC2 was observed to have an increased expression in P200 cells and has been reported to inhibit Wnt/β-Catenin signaling (Pourebrahim et al., 2011). *ZIC2* also seems to have a possible connection with c-FOS signaling, and also schizophrenia (Hatayama et al., 2011; S.-Y. Lee et al., 2004). *CRABP1* (Cellular retinoic acid-binding protein 1) has been shown upon up-regulation to reduce the differentiation potential of the human neuroblastoma cell-line SH-SY5Y, and instead increase cell proliferation. Thus, in relation to this thesis, *ZIC2* could possibly be promoting the RG pools rather than stimulating differentiation (Uhrig et al., 2008). It should be noted that this was reported in an Alzheimer's disease study where SH-SY5Y cells were overproducing amyloid precursor proteins. Another study in a chick model (avian retina) showed that retinoic acid signaling via *CRABP1* enhanced the formation of a proliferative glial cell type named as Muller glia-derived progenitor cells (Todd et al., 2018). Retinoic acid-induced proliferative

precursors with characteristics of radial glia have been shown to exist in the ganglionic eminence and subventricular zone, and enhanced *CRABP1* expression upon paracetamol exposure, via direct or other upstream effects, could enhance radial glial proliferation in this protocol (Haskell & LaMantia, 2005). Retinoic acid signaling is particularly required for olfactory development, another GWAS parameter found in association with ADHD (Gervin et al., 2017). Altered retinoic acid signaling has been shown to alter complex behavior and even potentially reduce telencephalon size and disrupt the migration of neurons (Chiang et al., 1998; Smith et al., 2001; T.-W. Wang et al., 2005).

NKX2-1 and *SFTA3* dysregulation in P200 cells after 20 days of neurodifferentiation

Expression of the forebrain marker *NKX2-1* was decreased in all cell types in a striking way for the P200 cells. *MBIP* displayed reduced expression after the same paracetamol treatment, and according to GeneCards *MBIP*-related pathways are connected to chromatin organization, which is relevant in the analysis of *NKX2-1*. In the theme of chromatin regulating genes, *CXXC5* also has increased expression in P200 cells, and according to GeneCards *CXXC5* plays a role in chromatin regulation and acetylation, while it has also been suggested as a negative-feedback regulator in Wnt-signaling (H.-Y. Kim et al., 2015). *HES5* which was used as a Notch-signaling marker in part I of this thesis had a minor decrease in expression across all cell-types. *PAX6* appeared in the radial glial and PCPs clusters when the signal had been very low from day 13 in untreated cells. *LHX2* was described as a repressor of Wnt-signaling in thesis part I, with strong expression in the RG clusters (Figure 33 *Wnt repressors LHX9 and LHX2*) in CTR. In P200 cells, *LHX2* was observed to increase, starting with Layer V PCs, then more so in Layer IV/5-HT, while a decrease was observed for 5-HT/Chol/Glut PCs. LHX2 is also expressed to a higher extent in GABAergic interneurons, but the expression levels in control cells and cells after treatment are so low that a comparison is hard to make. If more *LHX2* repressor were transcribed, then Wnt-signaling could be disturbed to a greater extent, even when it is induced by the protocol. This could explain the slight increase in *WNT7B* expression, as this could be a compensatory effect to increase the amount of Wnt ligand in the presence of more repressor. *WNT7B* is similar, but not identical, to its more well-known sibling *WNT7A*, and might thus have

other effects in this context. Furthermore, disturbed Wnt signaling has a connection with schizophrenia, bipolar disorders, and autism spectrum disorders (Mulligan & Cheyette, 2017).

The gene *SFTA3* (Surfactant Associated 3) decreased substantially in expression in all cell types. Another alias for *SFTA3* on GeneCards is *NKX2.1 Associated Non-Coding Intergenic RNA*, which is highly interesting since *NKX2-1* expression was also substantially decreased. The *SFTA3* gene is not directly called a lncRNA but instead affiliated with the lncRNA class, which is a somewhat unclear classification at first glance. Only 10 publications are listed for *SFTA3* on GeneCards, and these are related to a putative function as a surfactant. One study briefly suggests that *SFTA3* might increase in expression in an inflammatory response (Schicht et al., 2018). Another study revealed a role of *SFTA3* in interneuron specification, which is highly relevant (C. Y. Chen et al., 2018). This was a study of genes with similar expression patterns to *NKX2-1*, a forebrain marker, or more specifically for their study; medial ganglionic eminence (MGE) marker, which was further attributed as a master regulator of cortical interneuron progenitor development. *SFTA3* was found to be the strongest candidate in this connection. The study also included an analysis of hESC-derived MGE-progenitors, which is also relevant. Their *STFA3* KO hESC line demonstrated a reduced ability to differentiate down MGE-lineage but not as strong as the *NKX2-1* KO.

Part II - Discussion

It is interesting that paracetamol seems to affect cell populations and promote radial glial cell growth, as well as the formation of more PCPs instead of downstream progeny that comes after the intermediate progenitor's cluster. The path to and after intermediate progenitors seems repressed, thus driving cells via alternative routes.

Could genes *BNIP3* and *EMX2* be involved in low dose paracetamol neuroprotection?

A lower selection of genes is altered in P100 cells. *BNIP3* is interesting in being an autophagy and apoptosis-related factor. Denser cell cultures, due to the accelerated growth observed upon treatment with paracetamol could potentially be a result of a more hypoxic environment. An environment that in addition would promote autophagy over apoptosis. An increase in expression of *BNIP3* could be a sign of increased mitochondrial turnover via *BNIP3* mediated autophagy (K. E. Liu & Frazier, 2015), and interestingly, *BNIP3* expression was observed without lethal effects on mitochondria in highly respiring cells and cancer cells. This was explained by *BNIP3* phosphorylation that blocked the apoptotic effect, but not autophagy. *BNIP3* could be selectively participating in mitophagy, to degrade mitochondria that are dysfunctional, damaged, or old, which might be part of a neuroprotective paracetamol effect. In fact, it has been shown that *BNIP3* upregulation can have a pro-survival effect on cells via mitophagy (A. Singh et al., 2018).

EMX2 displayed enhanced expression in P100 and P200 cells. *EMX2* is a homeodomain TF, critical for CNS development. Mice lacking *EMX2* fails to develop neurons that project to the neocortex, will not develop hippocampal dentate gyrus, and will have defects in lamination and neuronal migration (Mallamaci et al., 2000). A slight increase in *EMX2* expression could be protective and regenerative, however, the P200 cells showed a much higher expression of *EMX2* in certain areas, and this might be more detrimental than protective.

Adverse outcomes after paracetamol use during pregnancy

c-Fos signaling is important for learning and memory (Chung, 2015) and a variable decrease for a promoter-activator for c-Fos, such as *C1orf6* - which we see at a lower expression in P200 cells -, could have long-term implications. *C1orf61* expression vanished completely in P200 cells compared to CTR, for all cell types, except for a minute expression level in the intermediate progenitors.

SFTA3, or *NKX2-1* Associated Non-Coding Intergenic RNA as it also has been called, showed substantially decreased expression in P200 cells. *SFTA3* is involved in interneuron specification and a similar observation was made in a study that showed hampered differentiation to MGE progeny by knockout of *SFTA3* and *NKX2-1* in hESC-lines. Instead, cells were following alternative differentiation paths (C. Y. Chen et al., 2018). A similar phenomenon seems to take place with paracetamol treatment of cells according to our protocol. This could explain why we observe more cells in our cultures just before part III in the protocol, for the cells that were treated with paracetamol, when part II cells are counted and made ready to be moved to part III. *NKX2-1* pathways should be active by this point in the protocol, and paracetamol is promoting the formation of more radial glial/glial/stem cells instead of progression towards MGE-like progenitor cells, by shutting down *NKX2-1* and *SFTA3* function. We can instead observe larger pools of highly proliferative cells, as a consequence of radial glial promotion and a block towards mature progeny as observed in the control study. Increased progeny formation and MGE development delay could very well explain the increased risk of relevant diagnoses at early or later stages in life, and explain why GWAS and other studies have problems identifying a more distinct profile behind diseases such as ADHD.

Loose angles

If paracetamol does not bind directly to chromatin, or to *NKX2-1*, there is still the possibility that paracetamol can activate or disturb mobile genetic elements which can be regulators of *NKX2-1*, such as LINE-1. When a drug library was designed for investigation of LINE-1 activity in regard to several mental diseases such as schizophrenia, it was also designed to be able to test paracetamol as a LINE-1 related molecule. Mobile genetic elements such as LINE-1, HERVs,

etc., have been difficult to directly investigate, due to their nature as ncRNAs. In addition, the technology has not been powerful enough, and these elements were also considered to be junk DNA for some time. More recently, however, the mobile genetic elements have been reported as possible gene regulatory phenomena have specifically taken advantage of by primate- and particularly in human evolution (Baillie et al., 2011; Faulkner & Garcia-Perez, 2017; Shpyleva et al., 2018)(Coufal et al., 2009; Cunnane & Crawford, 2014; Goodier, 2014; Grassi et al., 2020; Jönsson et al., 2019; Petri et al., 2019; Suarez et al., 2017; Terry & Devine, 2019). Of course, a direct effect of paracetamol on a master regulator such as *NKX2-1* which would have epigenetic effects further downstream could in-turn also explain activation or inactivation of mobile genetic elements in the areas affected by the changed chromatin state.

Another more loosely related view is that *SFTA3* is also a surface surfactant, primarily found in the lungs. I went and checked for symptoms in ADHD in young children and breathing problems during sleep is a prominent feature (Ren & Qiu, 2014) which might be linked to disturbed surfactant expression. Adult ADHD is also associated with asthma (Fasmer et al., 2011). Even more loosely related is that marijuana smoking seems to increase vital capacity of lungs (which does sound strange at first glance, maybe it is the inherent exercise with holding marijuana vapor in the lungs for an extended time) and people with ADHD who receives medical marijuana in regions where it is legal seems to feel that their symptoms are alleviated over time (Hollis et al., 2008; Strohbeck-Kuehner et al., 2008). Although there are many other effects with marijuana smoking that can be more likely to help with ADHD symptoms its still interesting to think a bit wider on these subjects if the hypothesis on paracetamol, *NKX2-1*, and *SFTA3* is true. There is also the matter of paracetamol brain metabolite AM404's interaction with the endocannabinoid system CB1 which might contribute to some neurodevelopmental adverse effects but it feels more rational that this would produce more specific, and easier to discover effects in e.g. GWAS studies if it was as simple as one pathway.

Part II - Conclusions

In a 2019 publication in Nature, human accelerated regions (HARs) and human gained enhancers (HGEs) were introduced as novel concepts (H. Won et al., 2019). The intention was a discussion of complex brain development systems that are beneficial to humans but also can interact with neurodevelopmental disease risk genes, trading higher cognitive function with increased risk of mental disease. This is quite an interesting concept and the publication also specifically mentions important genes, of which many are relevant to this thesis study, e.g. genes involved in patterning of dorsal-ventral telencephalon; *EMX2, PAX6, GLI3, NKX6-1,* and *NKX6-2*, as well as cortical neurogenesis involving genes; *PAX6, HES1, SOX2, GLI3*, and *TBR2*.

Paracetamol dysregulates the master regulator *NKX2-1* and the co-expressed interneuron regulator *SFTA3*, causing chromatin to be less open in areas with genes important in the development of cortical migrating neurons and progeny within the MGE of the developing forebrain. Paracetamol further decreases expression of the c-FOS signaling enhancer entity, *C1orf61*, which is a likely activator of c-FOS signaling, which in turn has important roles in neuronal migration. As previous studies have shown, regarding the *NKX2-1* function, if its ability to bind to distal regulatory elements is inhibited, this would repress alternative paths for progenitor cell differentiation within the VZ of the MGE. Both the VZ and the MGE are well in the prediction zone of cell-types that could be generated of the developed and applied protocol, as determined in part I of this thesis. The reduction of *NKX2-1* binding will result in epigenetic changes and in decreased expression of *LHX6*, which is an important co-activator of region- and lineage-specific genes in the mantle and SVZ. This would explain why *PAX6* and other important factors are suddenly seen at higher rates, the regular differentiation path in our cell culture is significantly reduced by the paracetamol treatment. Furthermore, and most importantly, this also carries implication for when paracetamol use can be a problem during pregnancy, as the formation of the MGE and the early development of the forebrain would take place during a critical time period where paracetamol use should be limited. Based on this, we could update the recommendations for paracetamol use during pregnancy until we have a clear mechanism fully worked out. If this hypothesis is correct, exactly how the paracetamol moiety

(or/and if it is metabolized by our cells - one of the metabolites) binds to *NKX2-1*, or blocks access of *NKX2-1* to binding sites on chromatin, remains to be investigated in coming experiments which would be important to be able to provide a more complete answer to questions concerning the effect of paracetamol that carries an association with adverse developmental effects, such as an increased risk of developing ADHD.

As always, risk must be balanced against benefit and it can be disadvantageous to discourage the use of paracetamol, as it is an effective analgesic that does not seem to bring the same harmful effects in pregnancy as NSAIDs do. However, the recommended use should be adjusted to reflect the increased risk of use during the *in vivo* time-frame that this study reflects.

Experimental procedures

Mycoplasma

Cell cultures were tested for mycoplasma infection.

hESC general culture, dissociation, plate coating, and freezing.

hESCs were grown in Essential 8 media (ThermoFisher, #A1517001). Signs of cell differentiation were routinely pruned away by scraping or aspiration to ensure a homogenous stem cell population at the onset of differentiation. For routine passaging and dissociation of hESCs, cells were first washed twice with 1x dPBS, then treated with 0.5 mM ethylenediaminetetraacetic acid, EDTA, (ThermoFisher, #15575020) for 4 minutes in room temperature. After aspiration of EDTA, chunks of cell colonies are de-attached by splashing media against the well floor (2-4 ml media for a 6-well plate format, using rotational movement to cover the area of the well) with a serological pipette (5ml pipette works well) using an electric pipette (any standard kind will work) on maximum fluid ejection setting. As a protocol-optimization hESCs were pre-conditioned to grow on Geltrex (ThermoFisher, #A1413302). For routine plastic coating, Geltrex was used at a 1:100 dilution in KnockOut DMEM (ThermoFisher, #10829018), where the procedure and incubation were done according to the manual from the supplier (ThermoFisher). For preserving and storing cells in liquid nitrogen, cells were frozen down in media with 10 % DMSO (Sigma-Aldrich, #D8418).

Cell counting, light microscopy, materials lists

Cells were counted using a Countess II FL Automated Cell Counter (ThermoFisher) according to supplier instructions. Routine light microscopy was done with an EVOS FL Cell Imaging System (ThermoFisher) and images were captured with a 1.3 Sony ICX445 monochrome CCD camera (part of the EVOS system). Included below is a list of common chemicals and other materials used in this protocol (Table 9 *Protocol materials list, alphabetical*) and a table describing media composition for part I, II, and III (Table 10 *Medium compositions for part I, II, and III*).

Table 9 *Protocol materials list, alphabetical*

Item	Supplier	Cat nr
Advanced DMEM/F12	ThermoFisher	12634028
ACCUTASE	STEMCELL Technologies	07920
B27 Supplement	ThermoFisher	17504044
Human bFGF2	Peprotech	100-18B
Human EGF	Peprotech	AF-100-15
Essential-8 hESC media	Life Technologies	A1517001
hESC-Qualified Geltrex	ThermoFisher	A1413302
GlutaMAX Supplement	ThermoFisher	35050061
KnockOut DMEM	ThermoFisher	10829018
LDN-193189	STEMCELL Technologies	72148
N2 Supplement	ThermoFisher	17502048
RHO/ROCK Pathway Inhibitor Y27632	STEMCELL Technologies	72304
SB431542	Sigma-Aldrich	S4317
XAV939	STEMCELL Technologies	72674
Fibronectin	ThermoFisher	33010018
Polyornithine	Sigma-Aldrich	P3655
dPBS, without Ca2+ and Mg2+	GIBCO	14190
UltraPure 0.5 M EDTA, pH 8.0	ThermoFisher	15575020
RNeasy Mini Kit	Qiagen	74106
Affinity Script 2x qPCR cDNA synthesis kit	Agilent	600559
RNase-Free DNase set (50) for use with RNasy Columns	Qiagen	79254
Penicillin Streptomycin (abbrv. P/S or Pen Strep)	ThermoFisher	15140122

Table 10 *Medium compositions for part I, II, and III*

Medium	Component	Amount (500 ml)
Part I: Induction	Advanced DMEM/F12	485 ml
	GlutaMAX Supplement	5 ml (1%)
	Pen Strep	5 ml (1%)
	N2 Supplement	5 ml (1%)
Added fresh	SB431542	10 μM final concentration
	LDN-193189	100 nM final concentration
	XAV939	2 μM final concentration
Part II: Maturation	Advanced DMEM/F12	480 ml
	GlutaMAX Supplement	5 ml (1%)
	Pen Strep	5 ml (1%)
	N2 Supplement	5 ml (1%)
	B27 Supplement	5 ml (1%)
Part III: Expansion	Advanced DMEM/F12	487.5 ml
	GlutaMAX Supplement	5 ml (1%)
	Pen Strep	5 ml (1%)
	N2 Supplement	5 ml (1%)
	B27 Supplement	2.5 ml (0.5%)
Added fresh	Human bFGF2	10 ng/ml final concentration
	Human EGF	10 ng/ml final concentration

Table 11 *Primary and secondary antibodies*

Antibody	Dilution	Manufacturer
OCT4	1:100	Santa Cruz, sc-5279
SOX2	1:200	Abcam, ab79351
PAX6	1:350	Abcam, ab195045
OTX2	1:40	R&D, AF1979
Nestin	1:200	Abcam, ab22035
β3-tubulin	1:200	Santa Cruz, sc-80005
Donkey anti rabbit, Alexa 488	1:250	Jackson Immuno, 711-545-152
Donkey anti mouse, Alexa 555	1:500	Abcam, ab150110
Donkey anti goat, Alexa 555	1:500	Abcam, ab150130

Differentiation of hESCs to neural progenitor cells for use in toxicology studies

Preparation and cell seed numbers for 12- and 24-well format

hESCs were maintained for at least two passages and allowed to reach 80-90 % confluency. A full 6-well should contain plenty of hESCs for most experiments and also enough for re-seeding hESCs for maintenance culture (if a restart of the protocol is needed). With 80-90 % confluent homogenous hESCs at protocol day 0, hESCs were washed twice with 1x dPBS then single-cell suspended using Accutase (STEMCELL Technologies, #07920), following de-attachment procedures as described by the manufacturer with following changes: hESCs are more sensitive to increased mechanical force so a longer incubation time in Accutase is preferable to more pipetting, so to attain a single-cell suspension with high viability (above 90 %), an approximate incubation of 7 minutes in Accutase was found optimal for hESCs lines HS-360 and H9, combined with approximately 10-12 resuspensions using a 1000 μl mechanical pipette. Single-cells were then transferred to Essential 8 media containing 10μM RHO/ROCK Pathway Inhibitor Y27632 (STEMCELL Technologies, #72304), centrifuged for 4 minutes at 200 x G, resuspended again in Essential 8/Rock inhibitor, counted and assessed for viability, then plated at 60,000 cells per well using a 12-well format (30,000 for a 24-well format). To ensure an even cell spread, plates were moved side to side 5 times, then up and down (horizontally) 5 times, this procedure was repeated one more time then plates were carefully moved to the cell incubator (37 °C, 5 % CO_2), while avoiding any circular motions that could cause cells to move back into the center of wells. A correct cell count is essential since the number of viable cells seeded is crucial for protocol success and reproducibility. A dedicated incubator is strongly recommended in cell labs with heavy use to avoid fluctuations in temperature and CO_2 when incubators are used by many people. Even if incubator temperature drops 1 degree for a pro-longed time (or many repeated times), hESCs can de-attach from the substrate.

Day 1 to day 7. Part I: Induction

At day 1 Essential 8/Rock inhibitor is removed and Part I: Induction was started by aspirating old media and replacing it with Part I media (Table 10 *Medium compositions for part I, II, and III*), During toxicology studies, the media would contain the compound of interest that is to be tested, pre-added to avoid local effects as cells can be exposed to a very high concentration at the initial area of addition. At day 1, cells should have a nice evenly spread grid-like layout. Part I media, with inhibitors added fresh (on the same day from frozen working-stock - avoid re-thawing!), was replaced daily between day 1 to day 6. Starting from around day 4 to 5, radial patterning became visible in the cell monolayer. By day 6 neural rosettes was clearly visible and the cell confluency high. Distinctive radial patterning and large patches with multiple rosettes covering your dish/well is an important end-point to reach before progressing. The most common issue in Part I is to underestimate or overestimate the number of cells plated at the start of your counting procedure. Seeding 10 000 cells less than recommended can have a negative impact on expected progression.

Day 7 to Day 13. Part II: Maturation

At day 7 cells are seeded in high density and use of LSX inhibitors is stopped while B27 supplement is added (ThermoFisher, #17504044) (Table 10 *Medium compositions for part I, II, and III*). In part II, Geltrex, polyornithine (Sigma-Aldrich, #P3655), and fibronectin (ThermoFisher, #33010018) are used to coat plates for cells in part II and III. First, the plates are coated with a solution containing a mix of polyornithine (20 µg/ml) and fibronectin (1 µg/ml) in ddH$_2$O which is allowed to incubate at 37 °C for 2 hours, or overnight. The polyornithine-fibronectin solution is then removed and the plates are then coated with Geltrex, incubated for 1 hour at 37 °C and were then ready to be used. They can be stored in a sealed container at 4 °C for at least a week.

On day 7 cells are split as single-cells using the same procedure as described for the hESCs at day 0 but in higher density. Recommended seed density is 450 000 cells per well for a 12-well format and 225 000 cells for a 24-well format. Cell viability is usually lower at this point,

compared to hESCs at day 0, it is not unusual to see more debris present, however as the count is made based on live cells this should not be an issue but worth mentioning as it might otherwise cause alarm that cells are unwell.

Day 13 to Day 20. Part III: Expansion

The B27 supplement is now at half concentration, compared to part II media (Table 10 *Medium compositions for part I, II, and III*), and added freshly to the media are bFGF2 (Peprotech, #100-18B) and EGF (Peprotech, #AF-100-15) where the final concentration of both should be 10 ng/ml, e.g. a 10 μg/ml stock would be used in a 1:1000 dilution. Part II cells are single-cell suspended using the same procedures as described for part II and the same seed numbers. In our experience, if cells are kept for additional passages after day 20, this seeding amount should be maintained for 4-5 passages and then the seeding amount can be re-evaluated and possibly lowered as the cells seem more stable and at this point, you might want to go by split ratios instead (e.g. 1:4) to have control over the number of days between splits.

Immunocytochemistry

Cells were washed three times with 1x dPBS, fixed with 4 % paraformaldehyde (Sigma-Aldrich, #158127) in 1x dPBS for 12 minutes, and then washed three times again. Cells were permeabilized and non-specific binding sites were blocked with 0.2% Triton X-100 (Sigma-Aldrich, #T8787) in blocking buffer containing 2 % BSA (Sigma-Aldrich, #A6003) and 0.01 % TWEEN 20 (Sigma-Aldrich, #P1379) in 1x dPBS, for 30 minutes in room temperature. Cells were washed three more times after blocking and permeabilization then incubated with specific primary antibodies (Table 11 *Primary and secondary antibodies*) in blocking buffer overnight at 4 °C. After another three washes, cells were incubated with either of the following Alexa-fluorophore conjugated secondary antibodies: Alexa Fluor 555 donkey anti-mouse (Table 10 *Primary and secondary antibodies*) or Alexa Fluor 488 anti-rabbit (Table 10 *Primary and secondary antibodies*) in blocking buffer for 60 minutes at room temperature. After a final three washes, cells were DAPI-treated and mounted with Vectashield Hardset Antifade Mounting

Medium (VECTOR Laboratories, #H-1500). Fluorescent images were captured on with an Applied Precision DeltavisionCORE system and viewed using a 100× oil objective (NA 1.4). Image deconvolution was done on the integrated software (softWoRx v04.1.2).

RNA isolation

RNA was isolated and purified with RNeasy Mini Kit (Qiagen, #74104) using an in-column DNA contamination removal step by using RNase-Free DNase (Qiagen, #79254), the procedures for isolation, purification, and DNase treatment were done according to the kit manuals. RNA was quantified using either a NanoDrop 2000 (ThermoFisher) or a Qubit 3.0 Fluorometer (ThermoFisher).

ddPCR

Samples were delivered to our Ullevål collaborators and who isolated nucleic acid and ran the ddPCR. Digital PCR obtain absolute measures of nucleic acid target sequences in a sample. Endogenous control genes were RPL30 and RAF1 (used in seperate wells) and they were found using *reffinder* (https://www.heartcure.com.au/reffinder/). cDNA was made with Qiagen Quantitect (Qiagen, #205311) according to their protocol. Input was 2 ng of material and ddPCR mastermix (BioRad, #1863024) was used according to manufacturers protocol. Taqman probes and reagents were used according per manufacturers instructions (ThermoFisher) and details on the list of probes can be arranged with our collaborators. ddPCR normalization was performed according to (Coulter, 2018).

10X Genomics single-cell sequencing

The kits acquired from 10X Genomics kits were Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3, 16 rxns (#PN-1000075), Chromium Single Cell B Chip Kit, 48 rxns (#PN-10000073), and Chromium i7 Multiplex Kit, 96 rxns (#PN-120262). Samples were delivered fresh in a heated polystyrene transport box to the lab of Robert Lyle at Oslo University

Hospital the day before isolation so that they would be ready the day after as the protocol requires a full day of work. Isolation proceeded according to 10X Genomics protocols for sample and library preparation (*Cell Preparation Guide*, n.d.) & (*Chromium Single Cell 3' Reagent Kits User Guide (v3 Chemistry)*, n.d.). Cells were prepared and resuspended by the same accutase procedure used in the differentiation protocol previously described. Target cell recovery was 1000 cells per sample. However, based on previous experience, a value of 1400 cells per sample was used for cell input calculations to correct for losses in the GEM generation. The 10X protocol use PBS as wash solution but as the hESCs and subsequent differentiated neural cells are sensitive a change to their protocol was made. To mimimize cell loss and cell aggregation a solution of 0.2 % BSA in PBS was prepared then added to the appropriate growth medium (e.g. E8 cell media) for a final working concentration of 0.04 %. The wash-resuspension media was then preheated to 37 °C and used as wash or resuspension media throughout the protocol. Samples were pooled and sequenced on two runs on a NextSeq 500 system (Illumina #SY-415-1001) using a 150 bp PE High Output kit (Illumina #FC-404-2002) for a theoretical output of approximately 50 000 reads per cell.

A note on 10X Genomics protocol versions

Protocol version 3.0 revision A was used for library preparation. 10X Genomics keeps their protocols updated and the current revision of their library preparation user guide is revision C. However, 10X Genomics does not host old revisions on its website. 10X Genomics recommends using their latest revisions but a digital copy of rev A can be downloaded via this link if one is interested in making comparisons:

https://drive.google.com/file/d/1tkDo8ujhX8kicG90GbNhZDRQ9SCYZ4CB/view?usp=sharing, (file stored on the author's Google Drive).

CellRanger

Output from the 10X Genomics platform was run through CellRanger (3.1.0) on the SAGA supercomputer clusters for computing read counts from neurodifferentiation samples (two experimental repeats, timepoints; day 0, day 7, day 13, and day 20), and the paracetamol

153

toxicology study (two experimental repeats, timepoints day 7 (not included in this thesis), and day 20, and two treatments; 100μM- and 200μM paracetamol. See table below for clarification. Reference genoms GRCh38-3.0.0 was used for read alignment. We have tested aggregating the repeated experiments both in CellRanger with the *aggr* command, and downstream integration after quality control in Rstudio.

Table listing samples run in CellRanger

Timepoint (two repeated experiments)	Protocol/control	100μM paracetamol	200μM paracetamol
Day 0	✓	✓	✓
Day 7	✓	✓	✓
Day 13	✓	✓	✓
Day 20	✓	✓	✓

scRNA-seq

scRNA analysis was done mainly in R (ver 4.0) using R-studio (ver 1.3). Packages used in this pipeline was AnnotationDbi (1.50.0), colorspace (1.4-1), dplyr (1.0.0), DropletUtils (1.8.0), EnsDb.Hsapiens.v86 (2.99.0), Matrix (1.2-18), org.Hs.eg.db (3.11.4), scater (1.16.1), scClustViz (1.3.8), scran (1.16.0), shiny (1.5.0), and Seurat (3.1.5.90). The work flow in the pipeline was as follows; 1. Loading raw data from CellRanger, from experimental repeat A and B.

2. Quality control, total reads, gene library, etc.

3. Inspection of mitochondrial content and cell-cycle prediction for each cell. Mitochondrial reads were allowed much more lenient that other studies since mitochondrial reads seems to be a poor indicated of quality sometimes(O'Flanagan et al., 2019) and much less is known about mitochondrial content in human neurodevelopment and embryonic stem cells.

4. Multi-dimensional filtering and manual filtering of cells, with an emphasis on lesser is better (Luecken & Theis, 2019; D. J. McCarthy et al., 2017).

5. Integration of dataset A and B with SCTransform and Seurat (Butler et al., 2018; Hafemeister & Satija, 2019).

6. Clustering and analysis in scClustViz (Innes & Bader, 2018).

7. Re-iterating and challenging analysis methods, as is common (Luecken & Theis, 2019).

Any cells filtered out was written as tabular output in txt format to keep records. Script was written in Markdown format (Bauer, 2020), thus an html file, experimental record of sorts, was produced for each timepoint with used code and settings, as well as plots from QC and filtering. For more information if you need to set up your own analysis check out (Amezquita et al., 2020) and (Luecken & Theis, 2019), for an excellent start that will save you alot of hard googling and trial-and-error.

For this thesis, I preferred to inspect each experimental dataset individually first, perform QC and filtering, then integrate them using Seurat via SCTransform, which was developed by Satijalab (Hafemeister & Satija, 2019) as a better alternative with heterogeneous datasets, and the SCTransform algorithm is also utilized in the code and Seurat modules that integrate the repeat runs as it improves comparison and common identity between the two datasets, then normalizes and scales the data. The established integrated dataset can then smoothly by added upon with similar methods, e.g. another treatment study at the same timepoint, or for the "all timepoints" datasets used in this thesis.

CytoTRACE analysis was done in R, initially using ver 0.2.1, but due to some bugs and after interaction with the author from Stanford (Gulati et al., 2020), who is really nice and answers e-mails very fast, a new version was released a couple of weeks later, and analysis was then updated to use 0.3.2. For testing this tool, there is an online version of this tool at (https://cytotrace.stanford.edu/) with demo datasets and possibility to upload your own.

PAGA analysis and pseudotime analysis (Wolf et al., 2019) was done with scanpy (1.5.1) (Wolf et al., 2018) in miniconda3 (4.8.3) using python (2.8.1).

Figures and other images

Figures were either created directly as output from R-studio, or they made in Adobe Illustrator (2020), and FIJI/ImageJ (1.8.0_112) - with a special shoutout to a very handy FIJI plugin called

EZFig by Benoit Aigouy (Aigouy & Mirouse, 2013).

scRNA-seq, cells filtered by isOutlier and manual cutoffs

isOutlier is set to a MADs threshold of 6 for mitochondrial content outliers. By default, MAD is set to 3, but threshold but can be increased to reduce the number of cells filtered. The default value for MAD, 3, were used for outlier detection in library size (total counts per cell) and features (genes per cell). Filtering settings for isOutlier, the MAD threshold, are set the same for both run A and run B, i.e. 6 for mitochondrial outlier detection, and 3 for library and feature outlier detection.

MADs thresholding is part of the isOutlier() function from scran to be able to detect outliers from by weighing several parameters at the same time. Manual cutoffs for min and max on library size and features were also set. These cutoffs were determined by inspecting histograms on library size and features, e.g. histograms shown in SI Day 0 Figure II. The aim is to inspect the distribution and then cutoff low amounts of cells that spread in either direction. Cells removed manually were usually very few. The idea was first that it would improve scaling of the datasets, to manually prune the histogram, but since so few cells were removed it was mostly a good opportunity to take time to inspect the data. Both MAD- and manual filtering will be re-visited at a later date. How to filter, and on what metrics, seems to vary highly on the study/paper, and there is no consensus yet or any golden standard software that do it (Luecken & Theis, 2019).

Day 0 cell filtering (SI Day 0 Figures I and II)

Filtering result for run B showed more cells kept compared to run A, which could indicate that run A had a lower quality. This may additionally explain why the runs A and B did not integrate that well on day 0, as seen in the UMAP for day 0 (SI Day 0 Figure I, C.). The difference between day 0 integration, compared to other time points, was in the main text explained as being caused by minor differences becoming relatively large when integrating two very homogenous datasets. The aforementioned low quality of run A was discovered the day before

thesis submission but the matter will be looked at and reported on later.

Run A

647 single-cells were kept for downstream analysis while 486 cells were removed of which 465 was removed by isOutlier. The relative high amount of cells removed might be indicative that the day 0 cells were more damaged. Which also explain higher % mitochondria. IsOutlier removed; mitochondrial outliers, > 25.87 %, $= 216$ cells, library outliers $= 178$ cells, and features outliers $= 71$ cells. Manual cutoffs for library (counts/cell) set to keep cells between 7 000 to 50 000. Below the lower cutoff, 0 cells were removed. Above the higher cutoff, 18 cells were removed. Manual cutoffs for features (genes/cell) set to keep cells between 2 000 and 7 500. Below the lower cutoff, 0 cells were removed. Above the higher cutoff, 0 cells were removed.

Run B

1029 single-cells were kept for downstream analysis, 329 cells were removed of which 279 was removed by *isOutlier*. Retained cells in run B are higherin number compared to run A. IsOutlier removed; mitochondrial outliers, > 21.53 %, $= 130$ cells, library outliers $= 123$ cells, and features outliers $= 26$ cells. Manual cutoffs for library (counts/cell) set to keep cells between 7 000 to 50 000. Below the lower cutoff, 0 cells were removed. Above the higher cutoff, 50 cells were removed. Manual cutoffs for features (genes/cell) set to keep cells between 2 000 and 7 500. Below the lower cutoff, 0 cells were removed. Above the higher cutoff, 0 cells were removed. Library max cutoff for run B should have been set relatively higher than run A, similar to max feature cutoff, which is set higher for run B (8000 compared to 7500 for set A).

Day 7 cell filtering (SI Day 7 Figures I and II)

Run B histograms were shifted towards the right on the x-asis, compared to run A. This is another reason why manual inspection of the individual experiments can give an idea of how the data looks like. Because of the shift to the right, run B manual cutoffs were slightly higher than for set A, but they are still set in relation to where the distribution curve starts, and ends.

Run A

1204 single-cells were kept for downstream analysis, 222 cells were removed of which 163 was removed by *isOutlier*. IsOutlier removed; mitochondrial outliers, > 16.78 %, = 146 cells, library outliers = 16 cells, and features outliers = 1 cell. Manual cutoffs for library (counts/cell) set to keep cells between 4 000 to 30 000. Below the lower cutoff, 27 cells were removed. Above the higher cutoff, 15 cells were removed. Manual cutoffs for features (genes/cell) set to keep cells between 1 500 and 7 000. Below the lower cutoff, 0 cells were removed. Above the higher cutoff, 0 cells were removed.

Run B

1060 single-cells were kept for downstream analysis, 315 cells were removed of which 257 was removed by *isOutlier*. IsOutlier removed; mitochondrial outliers, > 18.02 %, = 187 cells, library outliers = 69 cells, and features outliers = 1 cell. Manual cutoffs for library (counts/cell) set to keep cells between 6 000 to 45 000. Below the lower cutoff, 24 cells were removed. Above the higher cutoff, 34 cells were removed. Manual cutoffs for features (genes/cell) set to keep cells between 2 500 and 8 000. Below the lower cutoff, 0 cells were removed. Above the higher cutoff, 0 cells were removed.

Day 13 cell filtering (SI Day 13 Figures I and II)

Run A and run B on day 13 are very similar in terms of quality and levels in library size and features. The biggest difference are that less cells were run through the 10x machine for run A compared to run B, 976 cells compared to 1 407. For all samples, our collaborators were aiming to get atleast 1 000 cells sequenced.

Run A

801 single-cells were kept for downstream analysis, 175 cells were removed of which 130 was removed by *isOutlier*. IsOutlier removed; mitochondrial outliers, > 13.5 %, $= 91$ cells, library outliers $= 38$ cells, and features outliers $= 1$ cell. Manual cutoffs for library (counts/cell) set to keep cells between 6 000 to 40 000. Below the lower cutoff, 6 cells were removed. Above the higher cutoff, 39 cells were removed. Manual cutoffs for features (genes/cell) set to keep cells between 2 000 and 8 000. Below the lower cutoff, 0 cells were removed. Above the higher cutoff, 0 cells were removed.

Run B

1157 single-cells were kept for downstream analysis, 250 cells were removed of which 210 was removed by *isOutlier*. IsOutlier removed; mitochondrial outliers, > 15.29 %, $= 157$ cells, library outliers $= 50$ cells, and features outliers $= 3$ cell. Manual cutoffs for library (counts/cell) set to keep cells between 6 000 to 40 000. Below the lower cutoff, 10 cells were removed. Above the higher cutoff, 30 cells were removed. Manual cutoffs for features (genes/cell) set to keep cells between 2 000 and 8 000. Below the lower cutoff, 0 cells were removed. Above the higher cutoff, 0 cells were removed.

Day 20 cell filtering (SI Day 20 Figures I and II)

Run A and B were similar at day 20.

Run A

1510 single-cells were kept for downstream analysis, 292 cells were removed of which 245 was removed by *isOutlier*. IsOutlier removed; mitochondrial outliers, > 17.53 %, = 94 cells, library outliers = 98 cells, and features outliers = 53 cell. Manual cutoffs for library (counts/cell) set to keep cells between 4 000 to 35 000. Below the lower cutoff, 16 cells were removed. Above the higher cutoff, 15 cells were removed. Manual cutoffs for features (genes/cell) set to keep cells between 1 850 and 7 500. Below the lower cutoff, 16 cells were removed. Above the higher cutoff, 0 cells were removed.

Run B

1328 single-cells were kept for downstream analysis, 282 cells were removed of which 253 was removed by *isOutlier*. IsOutlier removed; mitochondrial outliers, > 19.24 %, = 115 cells, library outliers = 111 cells, and features outliers = 27 cell. Manual cutoffs for library (counts/cell) set to keep cells between 4 000 to 35 000. Below the lower cutoff, 9 cells were removed. Above the higher cutoff, 20 cells were removed. Manual cutoffs for features (genes/cell) set to keep cells between 1 850 and 7 500. Below the lower cutoff, 0 cells were removed. Above the higher cutoff, 0 cells were removed.

Supplemental material

Supplemental material I: Neurodifferentiation
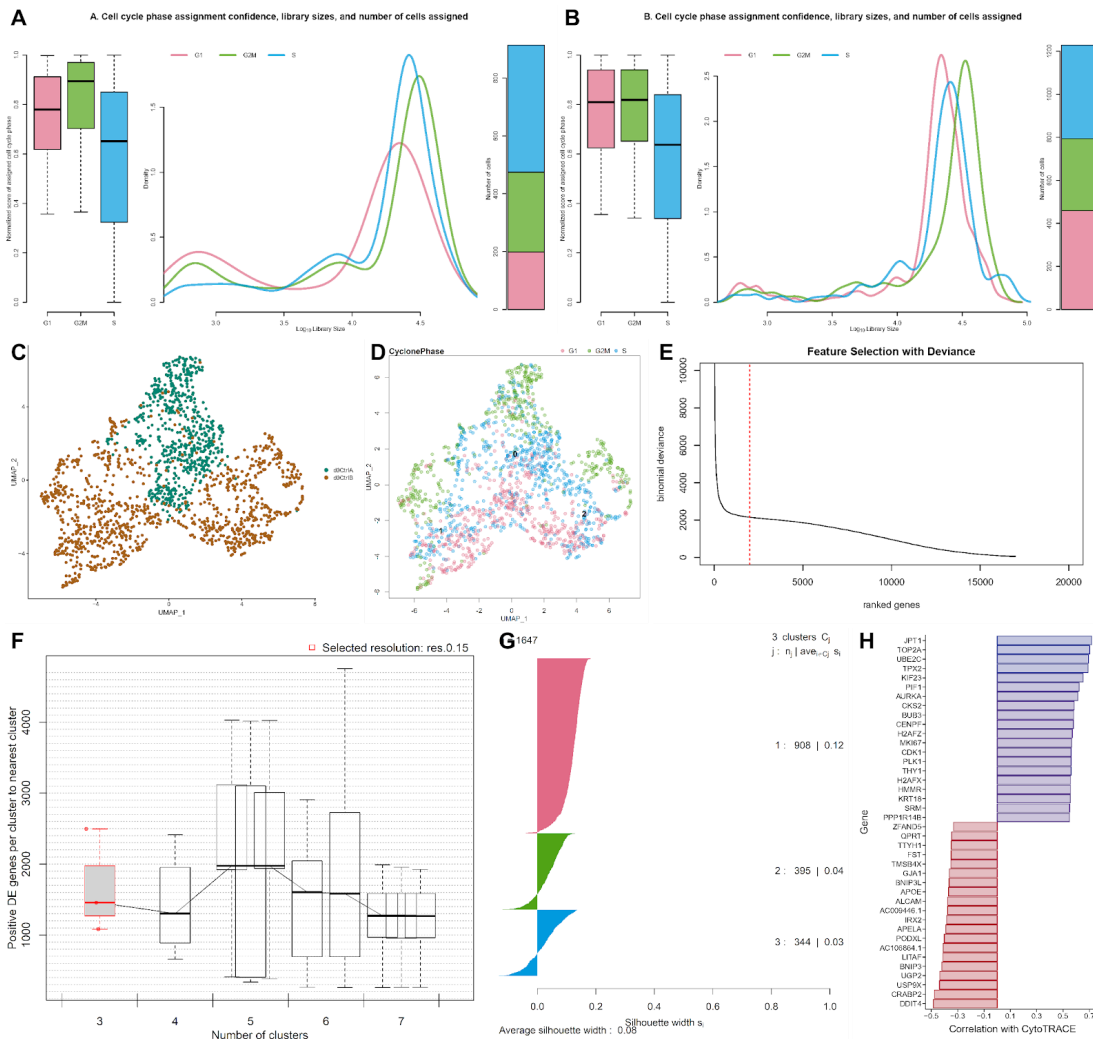
SI Day 0 Figure I *Cell-cycle, integration, scClustViz, and cytoTRACE* **A & B.** Cell-cycle assignments using Cyclone() package from SCRAN on sequencing run A (A.) and run B (B.). The plot shows the cell-cycle assignment scores and what phase cells are assigned as where red is G1-, green is G2M-, and blue is S-phase. **C.** Integration of experimental runs with Seurat depicted as a UMAP where clusters separate per run, likely due to tiny differences as this is a very homogeneous dataset. Green is sequencing run A and brown is sequencing run B. **D.** Cell-cycle assignments, UMAP. **E.** Deviance feature selection, showing 1500 genes used for all time-points. The deviance gene statistic is based on a multinomial null model that assumes each gene has a constant rate. Genes with a large metric are likely to be more useful in analysis. Low deviance features are instead discarded to speed up downstream analyses. A different approach is to use Seurat's 2000 variable genes (both were tested). **F.** Differentially expressed genes (DE genes) per cluster, compared to neighboring clusters, in scClustViz. A "good" clustering solution is based on having a minimum number of differential genes expressed to its neighbors, while all cluster silhouettes have as good a positive number as possible, seen in **G.** H. Extended list for genes used by CytoTRACE differentiation prediction.

163

SI Day 0 Figure II *QC, mitochondrial content, and filtering* Sequencing run A are represented in panels A to F, while sequencing run B is represented in panels G to L. **A.** Mitochondria % outliers removal by R-package scran with a subsequent histogram **B.** showing the cut-off point. In current practice, cells with extremely high mitochondria content are removed as they are considered bad, in our case this is done conservatively. **C & D**. Histograms on total read counts per cell before and after manual fine-tuning for sequencing run A. **E & F**. Histogram on features (genes) detected per cell before and after manual fine-tuning for sequencing run B. **G to L**. Same context as A to F but representing sequencing run B.

164

SI Day 7 Figure I *Cell-cycle, integration, scClustViz, and cytoTRACE* **A & B.** Cell-cycle assignments using Cyclone() package from SCRAN on sequencing run A (A.) and run B (B.). The plot shows the cell-cycle assignment scores and what phase cells are assigned as where red is G1-, green is G2M-, and blue is S-phase. **C.** Integration of experimental runs with Seurat depicted as a UMAP. Green is sequencing run A and brown is sequencing run B. **D.** Cell-cycle assignments, UMAP. **E.** Deviance feature selection, showing 1500 genes used for all time-points. The deviance gene statistic is based on a multinomial null model that assumes each gene has a constant rate. Genes with a large metric are likely to be more useful in analysis. Low deviance features are instead discarded to speed up downstream analyses. A different approach is to use Seurat's 2000 variable genes (both were tested). **F.** Differentially expressed genes (DE genes) per cluster, compared to neighboring clusters, in scClustViz. A "good" clustering solution is based on having a minimum number of differential genes expressed to its neighbors, while all cluster silhouettes have as good a positive number as possible, seen in **G.** H. Extended list for genes used by CytoTRACE differentiation prediction.

SI Day 7 Figure II *QC, mitochondrial content, and filtering* Sequencing run A are represented in panels A to F, while sequencing run B is represented in panels G to L. **A.** Mitochondria % outliers removal by R-package scran with a subsequent histogram **B.** showing the cut-off point. In current practice, cells with extremely high mitochondria content are removed as they are considered bad, in our case this is done conservatively. **C & D**. Histograms on total read counts per cell before and after manual fine-tuning for sequencing run A. **E & F**. Histogram on features (genes) detected per cell before and after manual fine-tuning for sequencing run B. **G to L**. Same context as A to F but representing sequencing run B.

166

SI Day 13 Figure I *Cell-cycle, integration, scClustViz, and cytoTRACE* **A & B.** Cell-cycle assignments using Cyclone() package from SCRAN on sequencing run A (A.) and run B (B.). The plot shows the cell-cycle assignment scores and what phase cells are assigned as where red is G1-, green is G2M-, and blue is S-phase. **C.** Integration of experimental runs with Seurat depicted as a UMAP. Green is sequencing run A and brown is sequencing run B. **D.** Cell-cycle assignments, UMAP. **E.** Deviance feature selection, showing 1500 genes used for all time-points. The deviance gene statistic is based on a multinomial null model that assumes each gen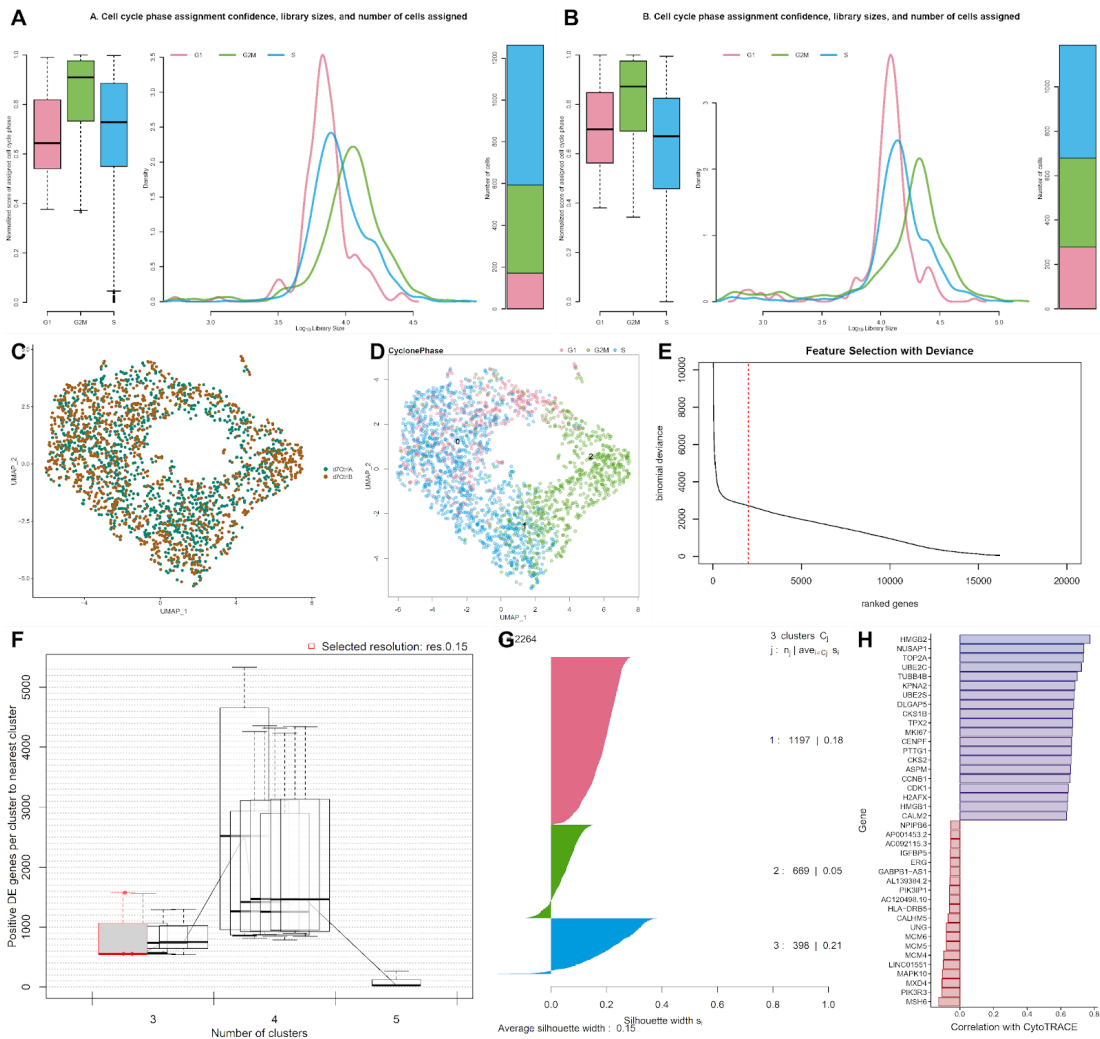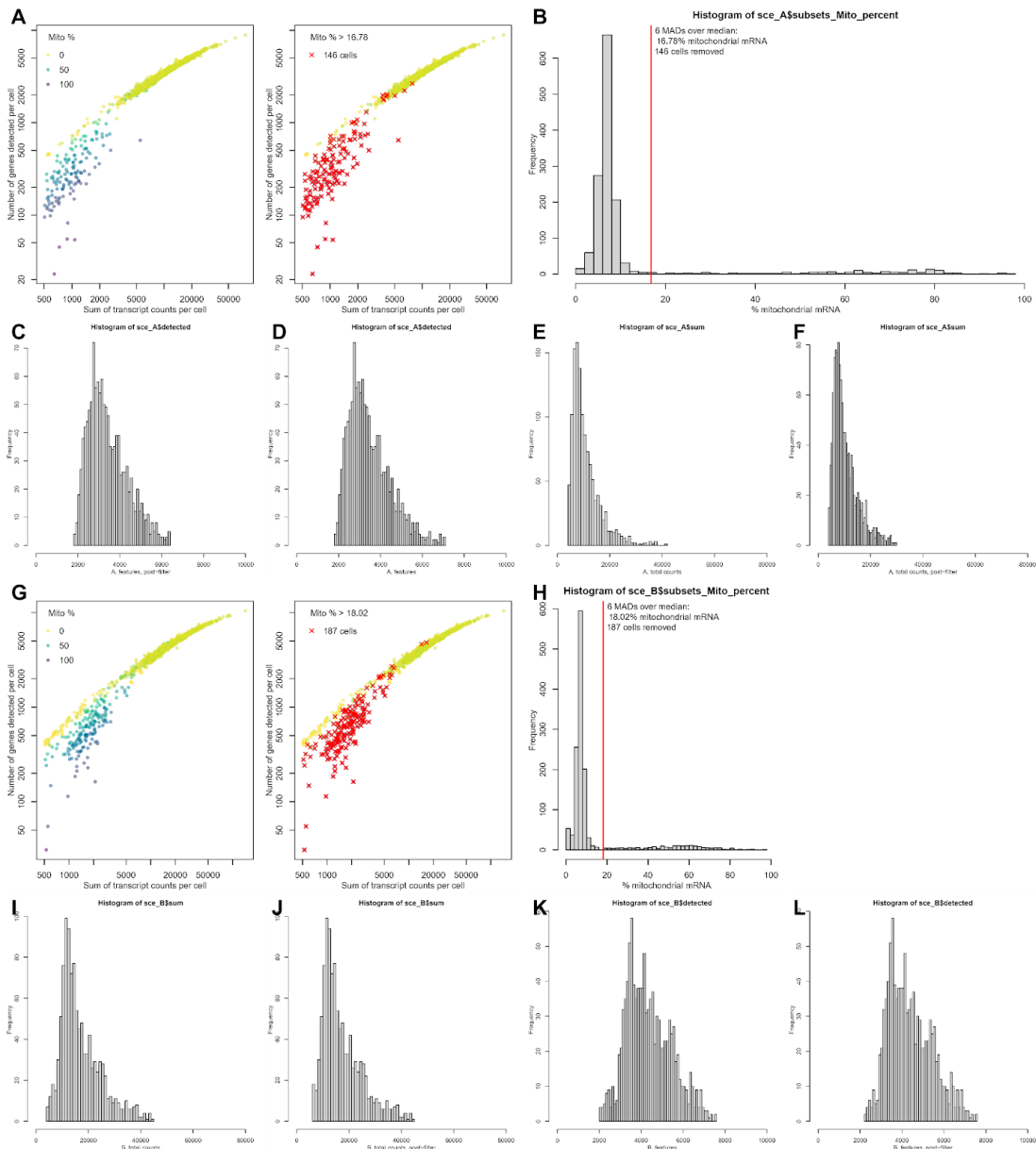e has a constant rate. Genes with a large metric are likely to be more useful in analysis. Low deviance features are instead discarded to speed up downstream analyses. A different approach is to use Seurat's 2000 variable genes (both were tested). **F.** Differentially expressed genes (DE genes) per cluster, compared to neighboring clusters, in scClustViz. A "good" clustering solution is based on having a minimum number of differential genes expressed to its neighbors, while all cluster silhouettes have as good a positive number as possible, seen in **G.** H. Extended list for genes used by CytoTRACE differentiation prediction.

SI Day 13 Figure II *QC, mitochondrial content, and filtering* Sequencing run A are represented in panels A to F, while sequencing run B is represented in panels G to L. **A.** Mitochondria % outliers removal by R-package scran with a subsequent histogram **B.** showing the cut-off point. In current practice, cells with extremely high mitochondria content are removed as they are considered bad, in our case this is done conservatively. **C & D**. Histograms on total read counts per cell before and after manual fine-tuning for sequencing run A. **E & F**. Histogram on features (genes) detected per cell before and after manual fine-tuning for sequencing run B. **G to L**. Same context as A to F but representing sequencing run B.

SI Day 20 Figure I *Cell-cycle, integration, scClustViz, and cytoTRACE* **A & B.** Cell-cycle assignments using Cyclone() package from SCRAN on sequencing run A (A.) and run B (B.). The plot shows the cell-cycle assignment scores and what phase cells are assigned as where red is G1-, green is G2M-, and blue is S-phase. **C.** Integration of experimental runs with Seurat depicted as a UMAP. Green is sequencing run A and brown is sequencing run B. **D.** Cell-cycle assignments, UMAP. **E.** Deviance feature selection, showing 1500 genes used for all time-points. The deviance gene statistic is based on a multinomial null model that assumes each gene has a constant rate. Genes with a large metric are likely to be more useful in analysis. Low deviance features are instead discarded to speed up downstream analyses. A different approach is to use Seurat's 2000 variable genes (both were tested). **F.** Differentially expressed genes (DE genes) per cluster, compared to neighboring clusters, in scClustViz. A "good" clustering solution is based on having a minimum number of differential genes expressed to its neighbors, while all cluster silhouettes have as good a positive number as possible, seen in **G.** H. Extended list for genes used by CytoTRACE differentiation prediction.

169

SI Day 20 Figure II *QC, mitochondrial content, and filtering* Sequencing run A are represented in panels A to F, while sequencing run B is represented in panels G to L. **A.** Mitochondria % outliers removal by R-package scran with a subsequent histogram **B.** showing the cut-off point. In current practice, cells with extremely high mitochondria content are removed as they are considered bad, in our case this is done conservatively. **C & D**. Histograms on total read counts per cell before and after manual fine-tuning for sequencing run A. **E & F**. Histogram on features (genes) detected per cell before and after manual fine-tuning for sequencing run B. **G to L**. Same context as A to F but representing sequencing run B.
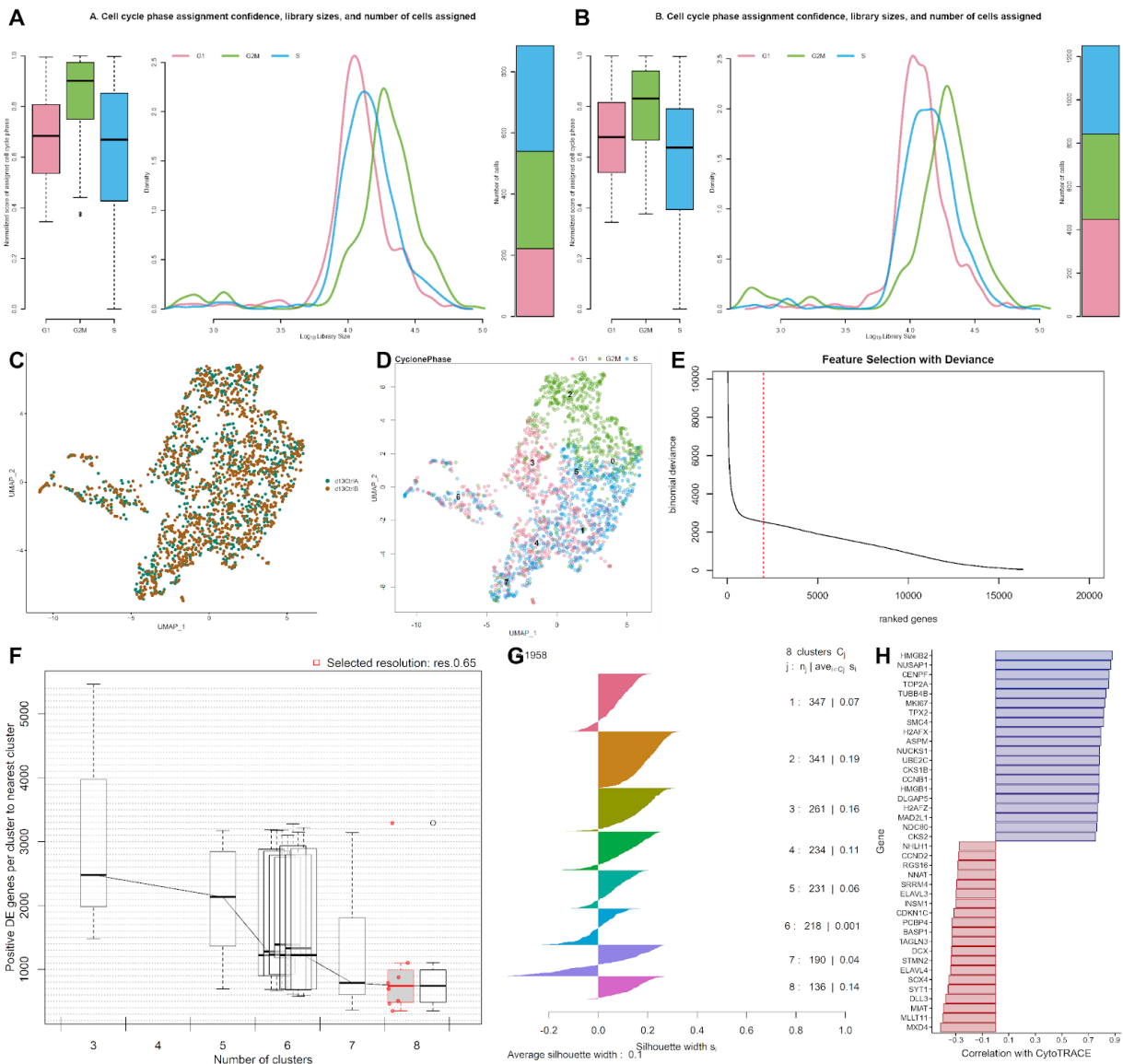
Supplemental material II: Paracetamol study

Supplemental II: scatterplots with treatment-induced changes in cell-type groups between treatment and control.
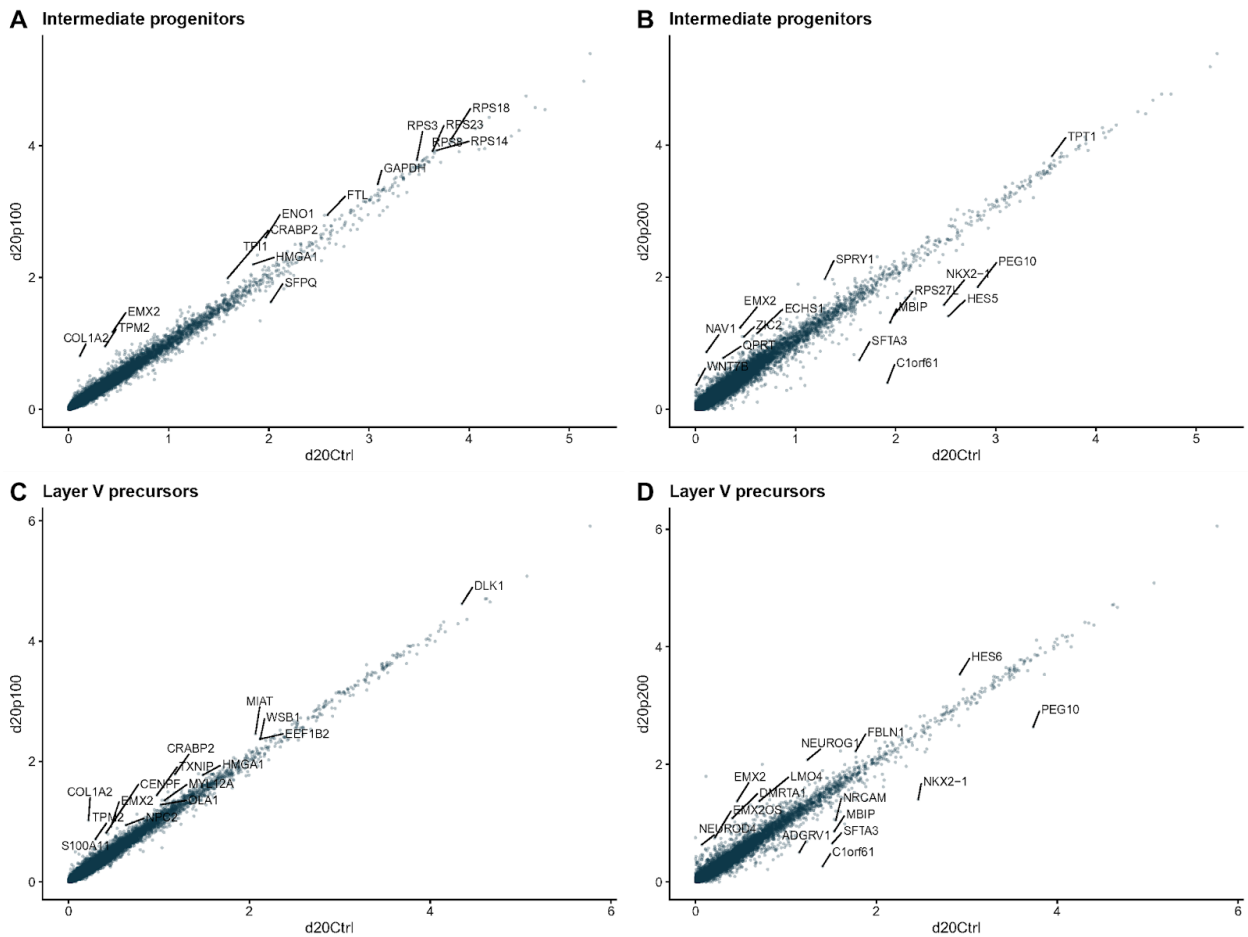
SII Figure I *Genes changed for IPs and Layer V precursors after paracetamol treatment.* **A.** Intermediate progenitors (IPs), P100. **B.** IPs, P200. **C.** Layer V precursors, P100. **D.** Layer V precursors, P200. All plots are treatment (y-axis) vs control (x-axis). Axes are relative gene expression and genes outside the diagonal are different (lower or higher in their expression) for either treatment or control.
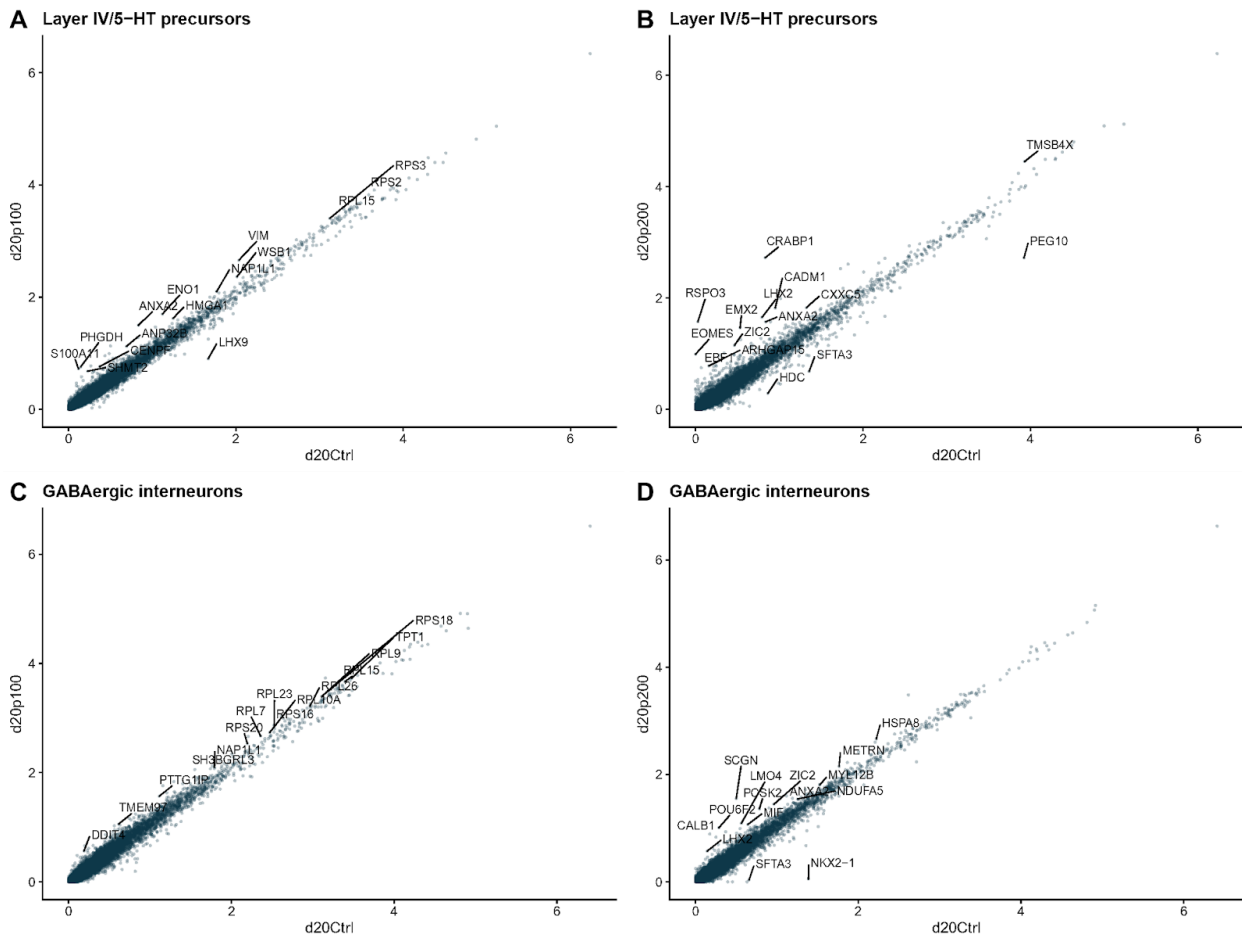
SII Figure II *Genes changed for Layer IV/5-HT and GABAergic interneurons after paracetamol treatment.* **A.** *Layer IV/5-HT*, P100. **B.** *Layer IV/5-HT*, P200. **C.** *GABAergic interneurons*, P100. **D.** *GABAergic interneurons*, P200. All plots are treatment (y-axis) vs control (x-axis). Axes are relative gene expression and genes outside the diagonal are different (lower or higher in their expression) for either treatment or control.
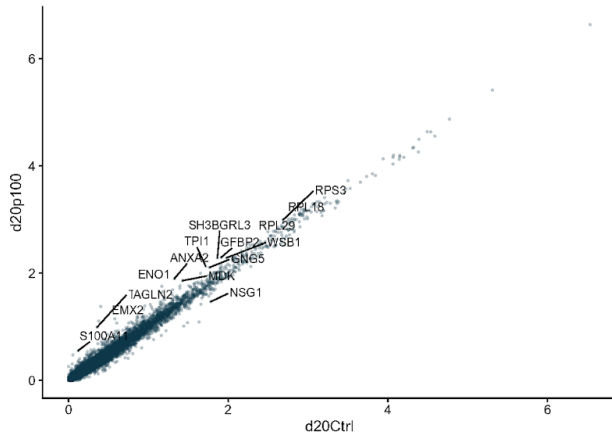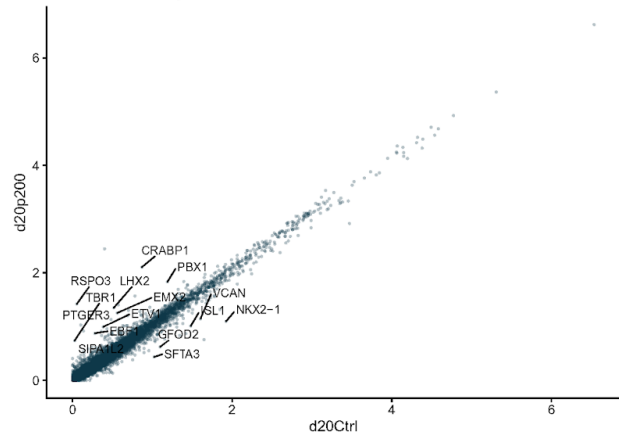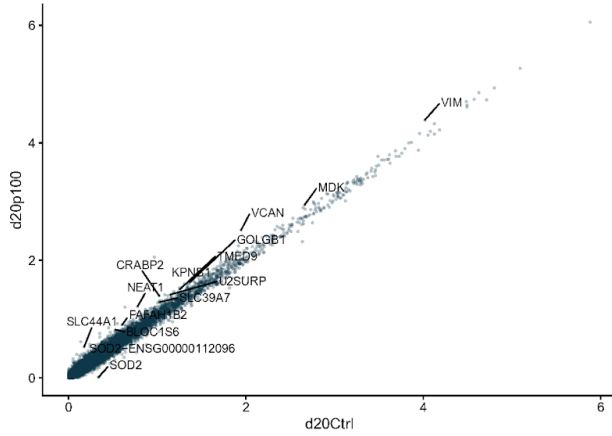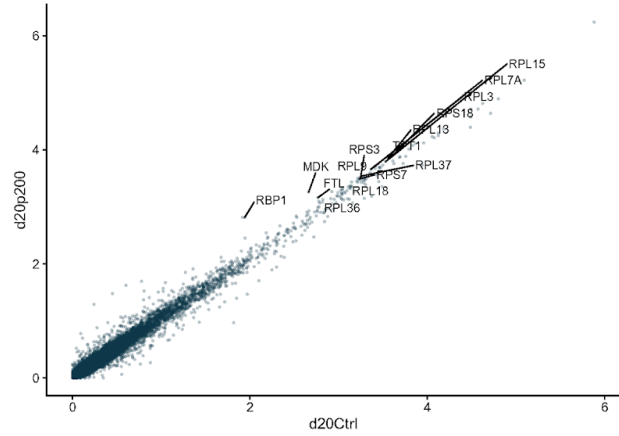
SII Figure III *Genes changed for 5-HT/Chol/Glut precursors and corticothalamic precursors (CTPs) after paracetamol treatment.* **A.** *5-HT/Chol/Glut*, P100. B. *5-HT/Chol/Glut precursors*, P200. **C.** Corticothalamic precursors (*CTPs*), P100. **D.** Corticothalamic precursors (*CTPs*), P200. All plots are vs CTR (x-axis). All plots are treatment (y-axis) vs control (x-axis). Axes are relative gene expression and genes outside the diagonal are different (lower or higher in their expression) for either treatment or control.

SIII Code for processing scRNA-seq data

Scripts and code took a long time to develop and are not officially published as of now. However if the committee wishes so, I will happily arrange so that my scripts can be demonstrated. That includes any other data as well.

References

Abaci, H. E., Truitt, R., Luong, E., Drazer, G., & Gerecht, S. (2010). Adaptation to oxygen deprivation in cultures of human pluripotent stem cells, endothelial progenitor cells, and umbilical vein endothelial cells. *American Journal of Physiology. Cell Physiology*, *298*(6), C1527–C1537.

Abranches, E., Silva, M., Pradier, L., Schulz, H., Hummel, O., Henrique, D., & Bekman, E. (2009). Neural Differentiation of Embryonic Stem Cells In Vitro: A Road Map to Neurogenesis in the Embryo. *PLoS ONE*, *4*(7), e6286.

Acampora, D., Di Giovannantonio, L. G., & Simeone, A. (2013). Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development* , *140*(1), 43–55.

Adjei, A. A., Gaedigk, A., Simon, S. D., Weinshilboum, R. M., & Leeder, J. S. (2008). Interindividual variability in acetaminophen sulfation by human fetal liver: Implications for pharmacogenetic investigations of drug-induced birth defects. *Birth Defects Research. Part A, Clinical and Molecular Teratology*, *82*(3), 155–165.

Aigouy, B., & Mirouse, V. (2013). ScientiFig: a tool to build publication-ready scientific figures. *Nature Methods*, *10*(11), 1048.

Ajram, L. A., Horder, J., Mendez, M. A., Galanopoulos, A., Brennan, L. P., Wichers, R. H., Robertson, D. M., Murphy, C. M., Zinkstok, J., Ivin, G., Heasman, M., Meek, D., Tricklebank, M. D., Barker, G. J., Lythgoe, D. J., Edden, R. A. E., Williams, S. C., Murphy, D. G. M., & McAlonan, G. M. (2017). Shifting brain inhibitory balance and connectivity of the prefrontal cortex of adults with autism spectrum disorder. *Translational Psychiatry*, *7*(5), e1137.

Amezquita, R. A., Lun, A. T. L., Becht, E., Carey, V. J., Carpp, L. N., Geistlinger, L., Marini, F., Rue-Albrecht, K., Risso, D., Soneson, C., Waldron, L., Pagès, H., Smith, M. L., Huber, W., Morgan, M., Gottardo, R., & Hicks, S. C. (2020). Orchestrating single-cell analysis with Bioconductor. *Nature Methods*, *17*(2), 137–145.

Amundsen, S., Nordeng, H., Nezvalová-Henriksen, K., Stovner, L. J., & Spigset, O. (2015). Pharmacological treatment of migraine during pregnancy and breastfeeding. *Nature Reviews. Neurology*, *11*(4), 209–219.

Andrade, S. E., Gurwitz, J. H., Davis, R. L., Chan, K. A., Finkelstein, J. A., Fortman, K., McPhillips, H., Raebel, M. A., Roblin, D., Smith, D. H., Yood, M. U., Morse, A. N., & Platt, R. (2004). Prescription drug use in pregnancy. *American Journal of Obstetrics and Gynecology*, *191*(2), 398–407.

Anwar, M., Tambalo, M., Ranganathan, R., Grocott, T., & Streit, A. (2017). A gene network regulated by FGF signalling during ear development. *Scientific Reports*, *7*(1), 6162.

Austdal, L. P. E., Bjørnstad, S., Mathisen, G. H., Aden, P. K., Mikkola, I., Paulsen, R. E., & Rakkestad, K. E. (2016). Glucocorticoid Effects on Cerebellar Development in a Chicken Embryo Model: Exploring Changes in PAX6 and Metalloproteinase-9 After Exposure to Dexamethasone. *Journal of Neuroendocrinology*, *28*(12). https://doi.org/10.1111/jne.12438

Avella-Garcia, C. B., Julvez, J., Fortuny, J., Rebordosa, C., García-Esteban, R., Galán, I. R., Tardón, A., Rodríguez-Bernal, C. L., Iñiguez, C., Andiarena, A., Santa-Marina, L., & Sunyer, J. (2016). Acetaminophen use in pregnancy and neurodevelopment: attention function and autism spectrum

symptoms. *International Journal of Epidemiology*, *45*(6), 1987–1996.

Ayanlaja, A. A., Xiong, Y., Gao, Y., Ji, G., Tang, C., Abdikani Abdullah, Z., & Gao, D. (2017). Distinct Features of Doublecortin as a Marker of Neuronal Migration and Its Implications in Cancer Cell Mobility. *Frontiers in Molecular Neuroscience*, *10*, 199.

Ayoub, N., Jeyasekharan, A. D., Bernal, J. A., & Venkitaraman, A. R. (2008). HP1-beta mobilization promotes chromatin changes that initiate the DNA damage response. *Nature*, *453*(7195), 682–686.

Ayoub, S. S., & Flower, R. J. (2019). Loss of hypothermic and anti-pyretic action of paracetamol in cyclooxygenase-1 knockout mice is indicative of inhibition of cyclooxygenase-1 variant enzymes. *European Journal of Pharmacology*, *861*, 172609.

Bagherpoor, A. J., Dolezalova, D., Barta, T., Kučírek, M., Sani, S. A., Ešner, M., Kunova Bosakova, M., Vinarský, V., Peskova, L., Hampl, A., & Štros, M. (2017). Properties of Human Embryonic Stem Cells and Their Differentiated Derivatives Depend on Nonhistone DNA-Binding HMGB1 and HMGB2 Proteins. *Stem Cells and Development*, *26*(5), 328–340.

Baillie, J. K., Barnett, M. W., Upton, K. R., Gerhardt, D. J., Richmond, T. A., De Sapio, F., Brennan, P. M., Rizzu, P., Smith, S., Fell, M., Talbot, R. T., Gustincich, S., Freeman, T. C., Mattick, J. S., Hume, D. A., Heutink, P., Carninci, P., Jeddeloh, J. A., & Faulkner, G. J. (2011). Somatic retrotransposition alters the genetic landscape of the human brain. *Nature*, *479*(7374), 534–537.

Bal-Price, A., Pistollato, F., Sachana, M., Bopp, S. K., Munn, S., & Worth, A. (2018). Strategies to improve the regulatory assessment of developmental neurotoxicity (DNT) using in vitro methods. *Toxicology and Applied Pharmacology*. https://doi.org/10.1016/j.taap.2018.02.008

Baranzini, S. E., Srinivasan, R., Khankhanian, P., Okuda, D. T., Nelson, S. J., Matthews, P. M., Hauser, S. L., Oksenberg, J. R., & Pelletier, D. (2010). Genetic variation influences glutamate concentrations in brains of patients with multiple sclerosis. *Brain: A Journal of Neurology*, *133*(9), 2603–2611.

Barral, S., Beltramo, R., Salio, C., Aimar, P., Lossi, L., & Merighi, A. (2014). Phosphorylation of histone H2AX in the mouse brain from development to senescence. *International Journal of Molecular Sciences*, *15*(1), 1554–1573.

Basak, O., & Taylor, V. (2007). Identification of self-replicating multipotent progenitors in the embryonic nervous system by high Notch activity and Hes5 expression. *The European Journal of Neuroscience*, *25*(4), 1006–1022.

Basavarajappa, B. S., Nixon, R. A., & Arancio, O. (2009). Endocannabinoid system: emerging role from neurodevelopment to neurodegeneration. *Mini Reviews in Medicinal Chemistry*, *9*(4), 448–462.

Baser, A., Skabkin, M., Kleber, S., Dang, Y., Gülcüler Balta, G. S., Kalamakis, G., Göpferich, M., Ibañez, D. C., Schefzik, R., Lopez, A. S., Bobadilla, E. L., Schultz, C., Fischer, B., & Martin-Villalba, A. (2019). Onset of differentiation is post-transcriptionally controlled in adult neural stem cells. *Nature*. https://doi.org/10.1038/s41586-019-0888-x

Bauer, P. C. (2020). *Writing a reproducible paper in R Markdown*. https://files.osf.io/v1/resources/q395s/providers/osfstorage/5e89c5e6d697350058bdec0c?format=pdf&action=download&direct&version=1

Bayard, F., Nymberg Thunell, C., Abé, C., Almeida, R., Banaschewski, T., Barker, G., Bokde, A. L. W., Bromberg, U., Büchel, C., Quinlan, E. B., Desrivières, S., Flor, H., Frouin, V., Garavan, H., Gowland, P., Heinz, A., Ittermann, B., Martinot, J.-L., Martinot, M.-L. P., … IMAGEN Consortium. (2018). Distinct brain structure and behavior related to ADHD and conduct disorder traits. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-018-0202-6

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I. W. H., Ng, L. G., Ginhoux, F., & Newell, E. W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology*. https://doi.org/10.1038/nbt.4314

Behan, Á. T., Byrne, C., Dunn, M. J., Cagney, G., & Cotter, D. R. (2009). Proteomic analysis of membrane microdomain-associated proteins in the dorsolateral prefrontal cortex in schizophrenia and bipolar disorder reveals alterations in LAMP, STXBP1 and BASP1 protein expression. *Molecular Psychiatry*, *14*(6), 601–613.

Behl, M., Hsieh, J.-H., Shafer, T. J., Mundy, W. R., Rice, J. R., Boyd, W. A., Freedman, J. H., Hunter, E. S., 3rd, Jarema, K. A., Padilla, S., & Tice, R. R. (2015). Use of alternative assays to identify and prioritize organophosphorus flame retardants for potential developmental and neurotoxicity. *Neurotoxicology and Teratology*, *52*(Pt B), 181–193.

Belloni, M., Uberti, D., Rizzini, C., Jiricny, J., & Memo, M. (1999). Induction of two DNA mismatch repair proteins, MSH2 and MSH6, in differentiated human neuroblastoma SH-SY5Y cells exposed to doxorubicin. *Journal of Neurochemistry*, *72*(3), 974–979.

Ben-David, U., Cowell, I. G., Austin, C. A., & Benvenisty, N. (2015). Controlling the Survival of Human Pluripotent Stem Cells by Small Molecule-Based Targeting of T opoisomerase II A lpha. *Stem Cells*, *33*(3), 1013–1019.

Bertolini, A., Ferrari, A., Ottani, A., Guerzoni, S., Tacchi, R., & Leone, S. (2006). Paracetamol: new vistas of an old drug. *CNS Drug Reviews*, *12*(3-4), 250–275.

Bhattacharyya, A., McMillan, E., Chen, S. I., Wallace, K., & Svendsen, C. N. (2009). A critical period in cortical interneuron neurogenesis in down syndrome revealed by human neural progenitor cells. *Developmental Neuroscience*, *31*(6), 497–510.

Bjørnstad, S., Austdal, L., Roald, B., Glover, J., & Paulsen, R. (2015). Cracking the Egg: Potential of the Developing Chicken as a Model System for Nonclinical Safety Studies of Pharmaceuticals. *The Journal of Pharmacology and Experimental Therapeutics*, *355*(3), 386–396.

Blough, E. R., & Wu, M. (2011). Acetaminophen: beyond pain and Fever-relieving. *Frontiers in Pharmacology*, *2*, 72.

Bollati, V., & Baccarelli, A. (2010). Environmental epigenetics. *Heredity*, *105*(1), 105–112.

Bond, J., Roberts, E., Mochida, G. H., Hampshire, D. J., Scott, S., Askham, J. M., Springell, K., Mahadevan, M., Crow, Y. J., Markham, A. F., Walsh, C. A., & Woods, C. G. (2002). ASPM is a major determinant of cerebral cortical size. *Nature Genetics*, *32*(2), 316–320.

Bornehag, C.-G., Reichenberg, A., Hallerback, M. U., Wikstrom, S., Koch, H. M., Jonsson, B. A., & Swan, S. H. (2018). Prenatal exposure to acetaminophen and children's language development at 30 months. *European Psychiatry: The Journal of the Association of European Psychiatrists*. https://doi.org/10.1016/j.eurpsy.2017.10.007

Bradley, C. A. (2020). Single-cell genomics illuminates human forebrain development [Review of *Single-cell genomics illuminates human forebrain development*]. *Nature Reviews. Genetics*, *21*(1), 2–3.

Brancaccio, M., Pivetta, C., Granzotto, M., Filippis, C., & Mallamaci, A. (2010). Emx2 and Foxg1 inhibit gliogenesis and promote neuronogenesis. *Stem Cells*, *28*(7), 1206–1218.

Brandlistuen, R., Ystrom, E., Nulman, I., Koren, G., & Nordeng, H. (2013). Prenatal paracetamol exposure and child neurodevelopment: a sibling-controlled cohort study. *International Journal of Epidemiology*, *42*(6), 1702–1713.

Bremer, L., Goletzke, J., Wiessner, C., Pagenkemper, M., Gehbauer, C., Becher, H., Tolosa, E., Hecher, K., Arck, P. C., Diemert, A., & Tiegs, G. (2017). Paracetamol Medication During Pregnancy: Insights on Intake Frequencies, Dosages and Effects on Hematopoietic Stem Cell Populations in Cord Blood From a Longitudinal Prospective Pregnancy Cohort. *Ebiomedicine*. https://doi.org/10.1016/j.ebiom.2017.10.023

Breuss, M. W., Leca, I., Gstrein, T., Hansen, A. H., & Keays, D. A. (2017). Tubulins and brain development - The origins of functional specification. *Molecular and Cellular Neurosciences*, *84*, 58–67.

Brodie, B. B., & Axelrod, J. (1948). The estimation of acetanilide and its metabolic products, aniline, N-acetyl p-aminophenol and p-aminophenol (free and total conjugated) in biological fluids and tissues. *The Journal of Pharmacology and Experimental Therapeutics*, *94*(1), 22–28.

Bronicki, L. M., & Jasmin, B. J. (2013). Emerging complexity of the HuD/ELAVl4 gene; implications for neuronal development, function, and dysfunction. *RNA* , *19*(8), 1019–1037.

Budday, S., Steinmann, P., & Kuhl, E. (2015). Physical biology of human brain development. *Frontiers in Cellular Neuroscience*, *9*, 257.

Bujalska, M. (2004). Effect of nitric oxide synthase inhibition on antinociceptive action of different doses of acetaminophen. *Polish Journal of Pharmacology*, *56*(5), 605–610.

Burke, E. E., Chenoweth, J. G., Shin, J. H., Collado-Torres, L., Kim, S.-K., Micali, N., Wang, Y., Colantuoni, C., Straub, R. E., Hoeppner, D. J., Chen, H.-Y., Sellers, A., Shibbani, K., Hamersky, G. R., Diaz Bustamante, M., Phan, B. N., Ulrich, W. S., Valencia, C., Jaishankar, A., … Jaffe, A. E. (2020). Dissecting transcriptomic signatures of neuronal differentiation and maturation using iPSCs. *Nature Communications*, *11*(1), 462.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, *36*(5), 411–420.

Cahn, A., & Hepp, P. (1886). Das antifebrin, ein neues fiebermittel. *Centralblatt Für Klinische Medizin*, *7*, 561–564.

Caiazzo, M., Giannelli, S., Valente, P., Lignani, G., Carissimo, A., Sessa, A., Colasante, G., Bartolomeo, R., Massimino, L., Ferroni, S., Settembre, C., Benfenati, F., & Broccoli, V. (2015). Direct conversion of fibroblasts into functional astrocytes by defined transcription factors. *Stem Cell Reports*, *4*(1), 25–36.

Castelo-Branco, G., Wagner, J., Rodriguez, F. J., Kele, J., Sousa, K., Rawal, N., Pasolli, H. A., Fuchs, E., Kitajewski, J., & Arenas, E. (2003). Differential regulation of midbrain dopaminergic neuron development by Wnt-1, Wnt-3a, and Wnt-5a. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(22), 12747–12752.

Castelvecchi, D. (2018). How human embryonic stem cells sparked a revolution (vol 555, pg 428, 2018). *Nature*, *557*(7703), 18–18.

Cell Preparation Guide. (n.d.). 10X Genomics. https://support.10xgenomics.com/single-cell-gene-expression/sample-prep/doc/demonstrated-protocol-single-cell-protocols-cell-preparation-guide

Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M., & Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature Biotechnology*, *27*(3), 275–280.

Chambers, S. M., Mica, Y., Lee, G., Studer, L., & Tomishima, M. J. (2013). Dual-SMAD inhibition/WNT activation-based methods to induce neural crest and derivatives from human pluripotent stem cells. In *Human Embryonic Stem Cell Protocols* (pp. 329–343). Springer.

Chandrasekharan, N. V., Dai, H., Roos, K. L. T., Evanson, N. K., Tomsik, J., Elton, T. S., & Simmons, D. L. (2002). COX-3, a cyclooxygenase-1 variant inhibited by acetaminophen and other analgesic/antipyretic drugs: cloning, structure, and expression. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(21), 13926–13931.

Chan, W.-L., Yuo, C.-Y., Yang, W.-K., Hung, S.-Y., Chang, Y.-S., Chiu, C.-C., Yeh, K.-T., Huang, H.-D., & Chang, J.-G. (2013). Transcribed pseudogene ψPPM1K generates endogenous siRNA to suppress oncogenic cell growth in hepatocellular carcinoma. *Nucleic Acids Research*, *41*(6), 3734–3747.

Chen, C., Lee, G. A., Pourmorady, A., Sock, E., & Donoghue, M. J. (2015). Orchestration of Neuronal Differentiation and Progenitor Pool Expansion in the Developing Cortex by SoxC Genes. *Journal of Neuroscience*, *35*(29), 10629–10642.

Chen, C. Y., Anderson, N. C., Becker, S., Schicht, M., Stoddard, C., Bräuer, L., Paulsen, F., & Grabel, L. (2018). Examining the role of the surfactant family member SFTA3 in interneuron specification. *PloS One*, *13*(11), e0198703.

Chen, G. G., Fiori, L. M., Moquin, L., Gratton, A., Mamer, O., Mechawar, N., & Turecki, G. (2010). Evidence of altered polyamine concentrations in cerebral cortex of suicide completers. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *35*(7), 1477–1484.

Chiang, M.-Y., Misner, D., Kempermann, G., Schikorski, T., Giguère, V., Sucov, H. M., Gage, F. H., Stevens, C. F., & Evans, R. M. (1998). An Essential Role for Retinoid Receptors RARβ and RXRγ In Long-Term Potentiation and Depression. *Neuron*, *21*(6), 1353–1361.

Cho, Y.-H., Han, K.-M., Kim, D., Lee, J., Lee, S.-H., Choi, K.-W., Kim, J., & Han, Y.-M. (2014). Autophagy Regulates Homeostasis of Pluripotency-Associated Proteins in hESCs. *Stem Cells*, *32*(2), 424–435.

Chromium Single Cell 3' Reagent Kits User Guide (v3 Chemistry). (n.d.). [FILE LINK: https://assets.ctfassets.net/an68im79xiti/5VgAXpSNhFJBN6XYGIQsY1/008faef1ea0df36bc650505b06900d53/CG000201_TechNote_Chromium_Single_Cell_3___v3_Reagent__Workflow___Software_Updates_RevA.pdf]. 10X Genomics. URL: https://support.10xgenomics.com/single-cell-gene-expression/library-prep/doc/technical-note-chromium-single-cell-3-v3-reagent-workflow-and-software-updates

Chung, L. (2015). A Brief Introduction to the Transduction of Neural Activity into Fos Signal. *Development & Reproduction*, *19*(2), 61–67.

Cicchillitti, L., Penci, R., Di Michele, M., Filippetti, F., Rotilio, D., Donati, M. B., Scambia, G., & Ferlini, C. (2008). Proteomic characterization of cytoskeletal and mitochondrial class III beta-tubulin. *Molecular Cancer Therapeutics*, *7*(7), 2070–2079.

Clarke, M. F., & Fuller, M. (2006). Stem cells and cancer: two faces of eve. *Cell*, *124*(6), 1111–1115.

Cohen, D. R., Cheng, C. W., Cheng, S. H., & Hui, C.-C. (2000). Expression of two novel mouse Iroquois homeobox genes during neurogenesis. *Mechanisms of Development*, *91*(1-2), 317–321.

Colleoni, S., Galli, C., Gaspar, J. A., Meganathan, K., Jagtap, S., Hescheler, J., Sachinidis, A., & Lazzari, G. (2012). Characterisation of a Neural Teratogenicity Assay Based on Human ESCs Differentiation

Following Exposure to Valproic Acid. *Current Medicinal Chemistry*, *19*(35), 6065–6071.

Collin, J., Queen, R., Zerti, D., Dorgau, B., Hussain, R., Coxhead, J., Cockell, S., & Lako, M. (2019). Deconstructing retinal organoids: single cell RNA-seq reveals the cellular components of human pluripotent stem cell-derived retina. *Stem Cells*, *37*(5), 593–598.

Combes, R., Barratt, M., & Balls, M. (2003). An overall strategy for the testing of chemicals for human hazard and risk assessment under the EU REACH system. *Alternatives to Laboratory Animals: ATLA*, *31*(1), 7–19.

Conti, L., & Cattaneo, E. (2010). Neural stem cell systems: physiological players or in vitro entities? *Nature Reviews. Neuroscience*, *11*(3), 176–187.

Coufal, N. G., L, G.-P. J. L., Peng, G. E., Yeo, G. W., Mu, Y., Lovci, M. T., Morell, M., Ks, O. 'shea, Moran, J. V., & Gage, F. H. (2009). L1 retrotransposition in human neural progenitor cells. *Nature*, *460*(7259), 1127–1131.

Couillard-Despres, S., Winner, B., Schaubeck, S., Aigner, R., Vroemen, M., Weidner, N., Bogdahn, U., Winkler, J., Kuhn, H.-G., & Aigner, L. (2005). Doublecortin expression levels in adult brain reflect neurogenesis. *European Journal of Neuroscience*, *21*(1), 1–14.

Coulter, S. J. (2018). Mitigation of the effect of variability in digital PCR assays through use of duplexed reference assays for normalization. *BioTechniques*, *65*(2), 86–91.

Courtney, N. A., Bao, H., Briguglio, J. S., & Chapman, E. R. (2019). Synaptotagmin 1 clamps synaptic vesicle fusion in mammalian neurons independent of complexin. *Nature Communications*, *10*(1), 4076.

Cui, C.-P., Zhang, Y., Wang, C., Yuan, F., Li, H., Yao, Y., Chen, Y., Li, C., Wei, W., Liu, C. H., He, F., Liu, Y., & Zhang, L. (2018). Dynamic ubiquitylation of Sox2 regulates proteostasis and governs neural progenitor cell differentiation. *Nature Communications*, *9*(1), 4648.

Cunnane, S. C., & Crawford, M. A. (2014). Energetic and nutritional constraints on infant brain development: Implications for brain expansion during human evolution. *Journal of Human Evolution*, *77*, 88–98.

D'Aiuto, L., Di Maio, R., Heath, B., Raimondi, G., Milosevic, J., Watson, A. M., Bamne, M., Parks, W. T., Yang, L., Lin, B., Miki, T., Mich-Basso, J. D., Arav-Boger, R., Sibille, E., Sabunciyan, S., Yolken, R., & Nimgaonkar, V. (2012). Human induced pluripotent stem cell-derived models to investigate human cytomegalovirus infection in neural cells. *PloS One*, *7*(11), e49700.

Danielsson, F., Peterson, M. K., Caldeira Araújo, H., Lautenschläger, F., & Gad, A. K. B. (2018). Vimentin Diversity in Health and Disease. *Cells* , *7*(10). https://doi.org/10.3390/cells7100147

Darr, A. J. (2009). *FGF2 Maintains the Proliferation of Neural Progenitors by Actively Suppressing the CKI p27Kip1 through Regulation of Cks1b Transcription.* http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.821.7765&rep=rep1&type=pdf

Dehay, C., Kennedy, H., & Kosik, K. S. (2015). The outer subventricular zone and primate-specific cortical complexification. *Neuron*, *85*(4), 683–694.

de Lombares, C., Heude, E., Alfama, G., Fontaine, A., Hassouna, R., Vernochet, C., de Chaumont, F., Olivo-Marin, C., Ey, E., Parnaudeau, S., Tronche, F., Bourgeron, T., Luquet, S., Levi, G., & Narboux-Nême, N. (2019). Dlx5 and Dlx6 expression in GABAergic neurons controls behavior, metabolism, healthy aging and lifespan. *Aging*, *11*(17), 6638–6656.

Devhare, P., Meyer, K., Steele, R., Ray, R. B., & Ray, R. (2017). Zika virus infection dysregulates human neural stem cell growth and inhibits differentiation into neuroprogenitor cells. *Cell Death & Disease*,

8(10), e3106.

Domínguez-Frutos, E., Vendrell, V., Alvarez, Y., Zelarayan, L. C., López-Hernández, I., Ros, M., & Schimmang, T. (2009). Tissue-specific requirements for FGF8 during early inner ear development. *Mechanisms of Development*, *126*(10), 873–881.

Donovan, S. L., Mamounas, L. A., Andrews, A. M., Blue, M. E., & McCasland, J. S. (2002). GAP-43 is critical for normal development of the serotonergic innervation in forebrain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *22*(9), 3543–3552.

Duggal, G., Warrier, S., Ghimire, S., Broekaert, D., Van der Jeught, M., Lierman, S., Deroo, T., Peelman, L., Van Soom, A., Cornelissen, R., Menten, B., Mestdagh, P., Vandesompele, J., Roost, M., Slieker, R. C., Heijmans, B. T., Deforce, D., De Sutter, P., De Sousa Lopes, S. C., & Heindryckx, B. (2015). Alternative Routes to Induce Naïve Pluripotency in Human Embryonic Stem Cells. *Stem Cells*, *33*(9), 2686–2698.

Dupont, S., Mamidi, A., Cordenonsi, M., Montagner, M., Zacchigna, L., Adorno, M., Martello, G., Stinchfield, M. J., Soligo, S., Morsut, L., Inui, M., Moro, S., Modena, N., Argenton, F., Newfeld, S. J., & Piccolo, S. (2009). FAM/USP9x, a Deubiquitinating Enzyme Essential for TGFβ Signaling, Controls Smad4 Monoubiquitination. *Cell*, *136*(1), 123–135.

Edri, R., Yaffe, Y., Ziller, M. J., Mutukula, N., Volkman, R., David, E., Jacob-Hirsch, J., Malcov, H., Levy, C., Rechavi, G., Gat-Viks, I., Meissner, A., & Elkabetz, Y. (2015). Analysing human neural stem cell ontogeny by consecutive isolation of Notch active neural progenitors. *Nature Communications*, *6*, 6500.

Ehashi, T., Suzuki, N., Ando, S., Sumida, K., & Saito, K. (2014). Effects of valproic acid on gene expression during human embryonic stem cell differentiation into neurons. *The Journal of Toxicological Sciences*, *39*(3), 383–390.

Eriksson, M., Taskinen, M., & Leppä, S. (2007). Mitogen activated protein kinase-dependent activation of c-Jun and c-Fos is required for neuronal differentiation but not for growth and stress response in PC12 cells. *Journal of Cellular Physiology*, *210*(2), 538–548.

Falck, B., & Hillarp, N.-Å. (1959). ON THE CELLULAR LOCALIZATION OF CATECHOL AMINES IN THE BRAIN. *Cells Tissues Organs*, *38*(3), 277–279.

Farlow, J. L., Lin, H., Sauerbeck, L., Lai, D., Koller, D. L., Pugh, E., Hetrick, K., Ling, H., Kleinloog, R., van der Vlies, P., Deelen, P., Swertz, M. A., Verweij, B. H., Regli, L., Rinkel, G. J. E., Ruigrok, Y. M., Doheny, K., Liu, Y., Broderick, J., … FIA Study Investigators. (2015). Lessons learned from whole exome sequencing in multiplex families affected by a complex genetic disorder, intracranial aneurysm. *PloS One*, *10*(3), e0121104.

Fasmer, O. B., Halmøy, A., Eagan, T. M., Oedegaard, K. J., & Haavik, J. (2011). Adult attention deficit hyperactivity disorder is associated with asthma. *BMC Psychiatry*, *11*, 128.

Fathi, A., Hatami, M., Hajihosseini, V., Fattahi, F., Kiani, S., Baharvand, H., & Salekdeh, G. H. (2011). Comprehensive gene expression analysis of human embryonic stem cells during differentiation into neural cells. *PloS One*, *6*(7), e22856.

Fatima, A., Irmak, D., Noormohammadi, A., Rinschen, M. M., Das, A., Leidecker, O., Schindler, C., Sánchez-Gaya, V., Wagle, P., Pokrzywa, W., Hoppe, T., Rada-Iglesias, A., & Vilchez, D. (2020). The ubiquitin-conjugating enzyme UBE2K determines neurogenic potential through histone H3 in human embryonic stem cells. *Communications Biology*, *3*(1), 262.

Faulkner, G. J., & Garcia-Perez, J. L. (2017). L1 Mosaicism in Mammals: Extent, Effects, and Evolution.

Trends in Genetics: TIG, *33*(11), 802–816.

Fernando, R. N., Eleuteri, B., Abdelhady, S., Nussenzweig, A., Andäng, M., & Ernfors, P. (2011). Cell cycle restriction by histone H2AX limits proliferation of adult neural stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(14), 5837–5842.

Ferrari, C., Lega, C., Vernice, M., Tamietto, M., Mende-Siedlecki, P., Vecchi, T., Todorov, A., & Cattaneo, Z. (2016). The Dorsomedial Prefrontal Cortex Plays a Causal Role in Integrating Social Impressions from Faces and Verbal Descriptions. *Cerebral Cortex*, *26*(1), 156–165.

Flower, R. J., & Vane, J. R. (1972). Inhibition of prostaglandin synthetase in brain explains the anti-pyretic activity of paracetamol (4-acetamidophenol). *Nature*, *240*(5381), 410–411.

Foudah, D., Monfrini, M., Donzelli, E., Niada, S., Brini, A. T., Orciani, M., Tredici, G., & Miloso, M. (2014). Expression of neural markers by undifferentiated mesenchymal-like stem cells from different sources. *Journal of Immunology Research*, *2014*, 987678.

Freytag, S., & Lister, R. (2020). schex avoids overplotting for large single-cell RNA-sequencing datasets. *Bioinformatics* , *36*(7), 2291–2292.

Fritsche, E., Barenys, M., Klose, J., Masjosthusmann, S., Nimtz, L., Schmuck, M., Wuttke, S., & Tigges, J. (2018). Current Availability of Stem Cell-Based In Vitro Methods for Developmental Neurotoxicity (DNT) Testing. *Toxicological Sciences: An Official Journal of the Society of Toxicology*, *165*(1), 21–30.

Galán, A., Montaner, D., Eugenia Póo, M., Valbuena, D., Ruiz, V., Aguilar, C., Dopazo, J., & Simón, C. (2010). Functional Genomics of 5- to 8-Cell Stage Human Embryos by Blastomere Single-Cell cDNA Analysis. *PLoS ONE*, *5*(10), e13615.

Gao, J., Yin, X., Yu, X., Dai, C., & Zhou, F. (2019). Long noncoding LINC01551 promotes hepatocellular carcinoma cell proliferation, migration, and invasion by acting as a competing endogenous RNA of microRNA‑122‑5p to regulate ADAM10 expression. *Journal of Cellular Biochemistry*, *120*(10), 16393–16407.

Garza-Manero, S., Sindi, A. A. A., Mohan, G., Rehbini, O., Jeantet, V. H. M., Bailo, M., Latif, F. A., West, M. P., Gurden, R., Finlayson, L., Svambaryte, S., West, A. G., & West, K. L. (2019). Maintenance of active chromatin states by HMGN2 is required for stem cell identity in a pluripotent stem cell model. *Epigenetics & Chromatin*, *12*(1), 73.

Genschow, E., Spielmann, H., Scholz, G., Seiler, A., Brown, N., Piersma, A., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S., & Becker, K. (2002). The ECVAM International Validation Study on In Vitro Embryotoxicity Tests: Results of the Definitive Phase and Evaluation of Prediction Models. *Alternatives to Laboratory Animals: ATLA*, *30*(2), 151–176.

Gervin, K., Nordeng, H., Ystrom, E., Ted, R.-K., & Lyle, R. (2017). Long-term prenatal exposure to paracetamol is associated with DNA methylation differences in children diagnosed with ADHD. *Clinical Epigenetics*, *9*(1), 77.

Ghanem, C. I., Pérez, M. J., Manautou, J. E., & Mottino, A. D. (2016). Acetaminophen from liver to brain: New insights into drug pharmacological action and toxicity. *Pharmacological Research: The Official Journal of the Italian Pharmacological Society*, *109*, 119–131.

Ghazizadeh, Z., Fattahi, F., Mirzaei, M., Bayersaikhan, D., Lee, J., Chae, S., Hwang, D., Byun, K., Tabar, M. S., Taleahmad, S., Mirshahvaladi, S., Shabani, P., Fonoudi, H., Haynes, P. A., Baharvand, H., Aghdami, N., Evans, T., Lee, B., & Salekdeh, G. H. (2018). Prospective Isolation of ISL1+ Cardiac Progenitors from Human ESCs for Myocardial Infarction Therapy. *Stem Cell Reports*, *10*(3),

848–859.

Glinka, A., Wu, W., Delius, H., Monaghan, A. P., Blumenstock, C., & Niehrs, C. (1998). Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature*, *391*(6665), 357–362.

Goodier, J. L. (2014). Retrotransposition in tumors and brains. *Mobile DNA*, *5*, 11.

Götz, M., & Huttner, W. B. (2005). The cell biology of neurogenesis. *Nature Reviews. Molecular Cell Biology*, *6*(10), 777–788.

Graham, G. G., Davies, M. J., Day, R. O., Mohamudally, A., & Scott, K. F. (2013). The modern pharmacology of paracetamol: therapeutic actions, mechanism of action, metabolism, toxicity and recent pharmacological findings. *Inflammopharmacology*, *21*(3), 201–232.

Grassi, D. A., Brattås, P. L., Jönsson, M. E., Atacho, D., Karlsson, O., Nolbrant, S., Parmar, M., & Jakobsson, J. (2020). Profiling of lincRNAs in human pluripotent stem cell derived forebrain neural progenitor cells. *Heliyon*, *6*(1), e03067.

Guhad, F. (2005). Introduction to the 3Rs (refinement, reduction and replacement). *Contemporary Topics in Laboratory Animal Science / American Association for Laboratory Animal Science*, *44*(2), 58–59.

Gulati, G. S., Sikandar, S. S., Wesche, D. J., Manjunath, A., Bharadwaj, A., Berger, M. J., Ilagan, F., Kuo, A. H., Hsieh, R. W., Cai, S., Zabala, M., Scheeren, F. A., Lobo, N. A., Qian, D., Yu, F. B., Dirbas, F. M., Clarke, M. F., & Newman, A. M. (2020). Single-cell transcriptional diversity is a hallmark of developmental potential. *Science*, *367*(6476), 405–411.

Guo, G. L., Moffit, J. S., Nicol, C. J., Ward, J. M., Aleksunes, L. A., Slitt, A. L., Kliewer, S. A., Manautou, J. E., & Gonzalez, F. J. (2004). Enhanced acetaminophen toxicity by activation of the pregnane X receptor. *Toxicological Sciences: An Official Journal of the Society of Toxicology*, *82*(2), 374–380.

Hafemeister, C., & Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. In *bioRxiv* (p. 576827). https://doi.org/10.1101/576827

Handbook, E. (2011). Fundamental Techniques in Cell Culture Laboratory Handbook. *The European Collection of Cell Cultures. [ONLINE] Available at: Http://www. Sigmaaldrich. Com/life-Science/cellculture/learning-Center/ecacc-Handbook. Html. [Accessed 09 July 2014]*.

Hansen, D. V., Lui, J. H., Parker, P. R. L., & Kriegstein, A. R. (2010). Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature*, *464*(7288), 554–561.

Hart, R. P. (2019). Strategies for Integrating Single-Cell RNA Sequencing Results With Multiple Species. In *bioRxiv* (p. 671115). https://doi.org/10.1101/671115

Haskell, G. T., & LaMantia, A.-S. (2005). Retinoic acid signaling identifies a distinct precursor population in the developing and adult forebrain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(33), 7636–7647.

Hatayama, M., Ishiguro, A., Iwayama, Y., Takashima, N., Sakoori, K., Toyota, T., Nozaki, Y., Odaka, Y. S., Yamada, K., Yoshikawa, T., & Aruga, J. (2011). Zic2 hypomorphic mutant mice as a schizophrenia model and ZIC2 mutations identified in schizophrenia patients. *Scientific Reports*, *1*(1). https://doi.org/10.1038/srep00016

Hay, D. C., Zhao, D., Fletcher, J., Hewitt, Z. A., McLean, D., Urruticoechea-Uriguen, A., Black, J. R., Elcombe, C., Ross, J. A., Wolf, R., & Cui, W. (2008). Efficient differentiation of hepatocytes from human embryonic stem cells exhibiting markers recapitulating liver development in vivo. *Stem Cells*,

26(4), 894–902.

Heinricher, M. M. (2005). Nociceptin/orphanin FQ: pain, stress and neural circuits. *Life Sciences*, *77*(25), 3127–3132.

Henke, R. M., Meredith, D. M., Borromeo, M. D., Savage, T. K., & Johnson, J. E. (2009). Ascl1 and Neurog2 form novel complexes and regulate Delta-like3 (Dll3) expression in the neural tube. *Developmental Biology*, *328*(2), 529–540.

Hettige, N. C., & Ernst, C. (2019). FOXG1 Dose in Brain Development. *Frontiers in Pediatrics*, *7*, 482.

Heuer, H., Ehrchen, J., Bauer, K., & Schäfer, M. K. H. (1998). Region-specific expression of thyrotrophin-releasing hormone-degrading ectoenzyme in the rat central nervous system and pituitary gland. *The European Journal of Neuroscience*, *10*(4), 1465–1478.

He, Z., Han, D., Efimova, O., Guijarro, P., Yu, Q., Oleksiak, A., Jiang, S., Anokhin, K., Velichkovsky, B., Grünewald, S., & Khaitovich, P. (2017). Comprehensive transcriptome analysis of neocortical layers in humans, chimpanzees and macaques. *Nature Neuroscience*, *20*(6), 886–895.

Hinz, B., Cheremina, O., & Brune, K. (2008). Acetaminophen (paracetamol) is a selective cyclooxygenase-2 inhibitor in man. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, *22*(2), 383–390.

Hirabayashi, Y., Itoh, Y., Tabata, H., Nakajima, K., Akiyama, T., Masuyama, N., & Gotoh, Y. (2004). The Wnt/beta-catenin pathway directs neuronal differentiation of cortical neural precursor cells. *Development* , *131*(12), 2791–2801.

Hoch, R. V., Clarke, J. A., & Rubenstein, J. L. R. (2015). Fgf signaling controls the telencephalic distribution of Fgf-expressing progenitors generated in the rostral patterning center. *Neural Development*, *10*(1). https://doi.org/10.1186/s13064-015-0037-7

Hoch, R. V., Lindtner, S., Price, J. D., & Rubenstein, J. L. R. (2015). OTX2 Transcription Factor Controls Regional Patterning within the Medial Ganglionic Eminence and Regional Identity of the Septum. *Cell Reports*, *12*(3), 482–494.

Hollis, C., Groom, M. J., Das, D., Calton, T., Bates, A. T., Andrews, H. K., Jackson, G. M., & Liddle, P. F. (2008). Different psychological effects of cannabis use in adolescents at genetic high risk for schizophrenia and with attention deficit/hyperactivity disorder (ADHD). *Schizophrenia Research*, *105*(1-3), 216–223.

Holm, P. C., Mader, M. T., Haubst, N., Wizenmann, A., Sigvardsson, M., & Götz, M. (2007). Loss- and gain-of-function analyses reveal targets of Pax6 in the developing mouse telencephalon. *Molecular and Cellular Neurosciences*, *34*(1), 99–119.

Ho, L., Tan, S. Y. X., Wee, S., Wu, Y., Tan, S. J. C., Ramakrishna, N. B., Chng, S. C., Nama, S., Szczerbinska, I., Chan, Y.-S., Avery, S., Tsuneyoshi, N., Ng, H. H., Gunaratne, J., Ray Dunn, N., & Reversade, B. (2015). ELABELA Is an Endogenous Growth Factor that Sustains hESC Self-Renewal via the PI3K/AKT Pathway. *Cell Stem Cell*, *17*(4), 435–447.

Hooker, C. W., & Hurlin, P. J. (2006). Of Myc and Mnt. *Journal of Cell Science*, *119*(Pt 2), 208–216.

Hook, J., Lemckert, F., Qin, H., Schevzov, G., & Gunning, P. (2004). Gamma tropomyosin gene products are required for embryonic development. *Molecular and Cellular Biology*, *24*(6), 2318–2323.

Hook, J., Lemckert, F., Schevzov, G., Fath, T., & Gunning, P. (2011). Functional identity of the gamma tropomyosin gene: Implications for embryonic development, reproduction and cell viability. *Bioarchitecture*, *1*(1), 49–59.

Huang, B., Ning, S., Zhang, Q., Chen, A., Jiang, C., Cui, Y., Hu, J., Li, H., Fan, G., Qin, L., & Liu, J.

(2017). Bisphenol A Represses Dopaminergic Neuron Differentiation from Human Embryonic Stem Cells through Downregulating the Expression of Insulin-like Growth Factor 1. *Molecular Neurobiology*, *54*(5), 3798–3812.

Huang, N., Wu, Z., Hong, H., Wang, X., Yang, F., & Li, H. (2019). Overexpression of CKS2 is associated with a poor prognosis and promotes cell proliferation and invasion in breast cancer. *Molecular Medicine Reports*, *19*(6), 4761–4769.

Huang, S.-M. A., Mishina, Y. M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G. A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., Hild, M., Shi, X., Wilson, C. J., Mickanin, C., Myer, V., Fazal, A., Tomlinson, R., Serluca, F., Shao, W., … Cong, F. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature*, *461*(7264), 614–620.

Hu, J. S., Vogt, D., Sandberg, M., & Rubenstein, J. L. (2017). Cortical interneuron development: a tale of time and space. *Development* , *144*(21), 3867–3878.

Hurd, T. W., Culbert, A. A., Webster, K. J., & Tavaré, J. M. (2002). Dual role for mitogen-activated protein kinase (Erk) in insulin-dependent regulation of Fra-1 (fos-related antigen-1) transcription and phosphorylation. *Biochemical Journal*, *368*(Pt 2), 573–580.

Hur, E. E., Edwards, R. H., Rommer, E., & Zaborszky, L. (2009). Vesicular glutamate transporter 1 and vesicular glutamate transporter 2 synapses on cholinergic neurons in the sublenticular gray of the rat basal forebrain: a double-label electron microscopic study. *Neuroscience*, *164*(4), 1721–1731.

Innes, B. T., & Bader, G. D. (2018). scClustViz - Single-cell RNAseq cluster assessment and visualization. *F1000Research*, *7*. https://doi.org/10.12688/f1000research.16198.2

Isoda, M., & Noritake, A. (2013). What makes the dorsomedial frontal cortex active during reading the mental states of others? *Frontiers in Neuroscience*, *7*, 232.

Jeffrey, P. L., Capes-Davis, A., Dunn, J. M., Tolhurst, O., Seeto, G., Hannan, A. J., & Lin, S. L. (2000). CROC-4: a novel brain specific transcriptional activator of c-fos expressed from proliferation through to maturation of multiple neuronal cell types. *Molecular and Cellular Neurosciences*, *16*(3), 185–196.

Jones, A. R., Overly, C. C., & Sunkin, S. M. (2009). The Allen Brain Atlas: 5 years and beyond. *Nature Reviews. Neuroscience*, *10*(11), 821–828.

Jones, L., López-Bendito, G., Gruss, P., Stoykova, A., & Molnár, Z. (2002). Pax6 is required for the normal development of the forebrain axonal connections. *Development* , *129*(21), 5041–5052.

Jönsson, M. E., Ludvik Brattås, P., Gustafsson, C., Petri, R., Yudovich, D., Pircs, K., Verschuere, S., Madsen, S., Hansson, J., Larsson, J., Månsson, R., Meissner, A., & Jakobsson, J. (2019). Activation of neuronal genes via LINE-1 elements upon global DNA demethylation in human neural progenitors. *Nature Communications*, *10*(1), 3182.

Josh Huang, Z., & Paul, A. (2018). Diversity of GABAergic interneurons and diversification of communication modules in cortical networks. In *bioRxiv* (p. 490797). https://doi.org/10.1101/490797

Jossin, Y., & Cooper, J. A. (2011). Reelin, Rap1 and N-cadherin orient the migration of multipolar neurons in the developing neocortex. *Nature Neuroscience*, *14*(6), 697.

Jóźwiak-Bebenista, M., & Nowak, J. Z. (2014). Paracetamol: mechanism of action, applications and safety concern. *Acta Poloniae Pharmaceutica*, *71*(1), 11–23.

Kadereit, S., Zimmer, B., van Thriel, C., Hengstler, J. G., & Leist, M. (2012). Compound selection for in vitro modeling of developmental neurotoxicity. *Frontiers in Bioscience* , *17*, 2442–2460.

Kaminsky, Z., & Payne, J. (2014). Seeing the future: epigenetic biomarkers of postpartum depression.

Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology, *39*(1), 233–234.

Kang, L., Yao, C., Khodadadi-Jamayran, A., Xu, W., Zhang, R., Banerjee, N. S., Chang, C.-W., Chow, L. T., Townes, T., & Hu, K. (2016). The Universal 3D3 Antibody of Human PODXL Is Pluripotent Cytotoxic, and Identifies a Residual Population After Extended Differentiation of Pluripotent Stem Cells. *Stem Cells and Development*, *25*(7), 556–568.

Kanton, S., Boyle, M. J., He, Z., Santel, M., Weigert, A., Sanchís-Calleja, F., Guijarro, P., Sidow, L., Fleck, J. S., Han, D., Qian, Z., Heide, M., Huttner, W. B., Khaitovich, P., Pääbo, S., Treutlein, B., & Camp, J. G. (2019). Organoid single-cell genomic atlas uncovers human-specific features of brain development. *Nature*, *574*(7778), 418–422.

Kaplan, M. P., Chin, S. S., Fliegner, K. H., & Liem, R. K. (1990). Alpha-internexin, a novel neuronal intermediate filament protein, precedes the low molecular weight neurofilament protein (NF-L) in the developing rat brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *10*(8), 2735–2748.

Kearsey, S. E., Maiorano, D., Holmes, E. C., & Todorov, I. T. (1996). The role of MCM proteins in the cell cycle control of genome duplication. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *18*(3), 183–190.

Kee, N., Volakakis, N., Kirkeby, A., Dahl, L., Storvall, H., Nolbrant, S., Lahti, L., Björklund, Å. K., Gillberg, L., Joodmardi, E., Sandberg, R., Parmar, M., & Perlmann, T. (2017). Single-Cell Analysis Reveals a Close Relationship between Differentiating Dopamine and Subthalamic Nucleus Neuronal Lineages. *Cell Stem Cell*, *20*(1), 29–40.

Kelley, M., Wu, D., & Fay, R. R. (2006). *Development of the Inner Ear*. Springer Science & Business Media.

Kesler, S. R., Schwartz, C., Stevenson, R. E., & Reiss, A. L. (2009). The impact of spermine synthase (SMS) mutations on brain morphology. *Neurogenetics*, *10*(4), 299–305.

Khalil, A., Kamar, A., & Nemer, G. (2020). Thalidomide-Revisited: Are COVID-19 Patients Going to Be the Latest Victims of Yet Another Theoretical Drug-Repurposing? *Frontiers in Immunology*, *11*, 1248.

Kim, E. J., Battiste, J., Nakagawa, Y., & Johnson, J. E. (2008). Ascl1 (Mash1) lineage cells contribute to discrete cell populations in CNS architecture. *Molecular and Cellular Neurosciences*, *38*(4), 595–606.

Kim, E. J., Leung, C. T., Reed, R. R., & Johnson, J. E. (2007). In vivo analysis of Ascl1 defined progenitors reveals distinct developmental dynamics during adult neurogenesis and gliogenesis. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(47), 12764–12774.

Kim, H.-Y., Yoon, J.-Y., Yun, J.-H., Cho, K.-W., Lee, S.-H., Rhee, Y.-M., Jung, H.-S., Lim, H. J., Lee, H., Choi, J., Heo, J.-N., Lee, W., No, K. T., Min, D., & Choi, K.-Y. (2015). CXXC5 is a negative-feedback regulator of the Wnt/β-catenin pathway involved in osteoblast differentiation. *Cell Death and Differentiation*, *22*(6), 912–920.

Kis, B., Snipes, J. A., & Busija, D. W. (2005). Acetaminophen and the Cyclooxygenase-3 Puzzle: Sorting out Facts, Fictions, and Uncertainties. *Journal of Pharmacology and Experimental Therapeutics*, *315*(1), 1–7.

Klim, J. R., Williams, L. A., Limone, F., Guerra San Juan, I., Davis-Dusenbery, B. N., Mordes, D. A.,

Burberry, A., Steinbaugh, M. J., Gamage, K. K., Kirchner, R., Moccia, R., Cassel, S. H., Chen, K., Wainger, B. J., Woolf, C. J., & Eggan, K. (2019). ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nature Neuroscience*, *22*(2), 167–179.

Klimova, L., Antosova, B., Kuzelova, A., Strnad, H., & Kozmik, Z. (2015). Onecut1 and Onecut2 transcription factors operate downstream of Pax6 to regulate horizontal cell development. *Developmental Biology*, *402*(1), 48–60.

Kobak, D., & Berens, P. (2019). The art of using t-SNE for single-cell transcriptomics. *Nature Communications*, *10*(1), 5416.

Koehn, L., Habgood, M., Huang, Y., Dziegielewska, K., & Saunders, N. (2019). Determinants of drug entry into the developing brain. *F1000Research*, *8*, 1372.

Kohtz, J. D., & Fishell, G. (2004). Developmental regulation of EVF-1, a novel non-coding RNA transcribed upstream of the mouse Dlx6 gene. *Gene Expression Patterns: GEP*, *4*(4), 407–412.

Kreitzer, F. R., Salomonis, N., Sheehan, A., Huang, M., Park, J. S., Spindler, M. J., Lizarraga, P., Weiss, W. A., So, P.-L., & Conklin, B. R. (2013). A robust method to derive functional neural crest cells from human pluripotent stem cells. *American Journal of Stem Cells*, *2*(2), 119–131.

Krug, A. K., Kolde, R., Gaspar, J. A., Rempel, E., Balmer, N. V., Meganathan, K., Vojnits, K., Baquié, M., Waldmann, T., Ensenat-Waser, R., Jagtap, S., Evans, R. M., Julien, S., Peterson, H., Zagoura, D., Kadereit, S., Gerhard, D., Sotiriadou, I., Heke, M., … Sachinidis, A. (2013). Human embryonic stem cell-derived test systems for developmental neurotoxicity: a transcriptomics approach. *Archives of Toxicology*, *87*(1), 123–143.

Kučírek, M., Bagherpoor, A. J., Jaroš, J., Hampl, A., & Štros, M. (2019). HMGB2 is a negative regulator of telomerase activity in human embryonic stem and progenitor cells. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, *33*(12), 14307–14324.

Kudo, L. C., Karsten, S. L., Chen, J., Levitt, P., & Geschwind, D. H. (2007). Genetic Analysis of Anterior–Posterior Expression Gradients in the Developing Mammalian Forebrain. *Cerebral Cortex*, *17*(9), 2108–2122.

Kufer, T. A., Silljé, H. H. W., Körner, R., Gruss, O. J., Meraldi, P., & Nigg, E. A. (2002). Human TPX2 is required for targeting Aurora-A kinase to the spindle. *The Journal of Cell Biology*, *158*(4), 617–623.

Kundakovic, M., & Jaric, I. (2017). The Epigenetic Link between Prenatal Adverse Environments and Neurodevelopmental Disorders. *Genes*, *8*(3). https://doi.org/10.3390/genes8030104

Kurtz, A., Seltmann, S., Müller, R., Dewender, J., Mah, N., & Bultjer, N. (n.d.). *KIe009-A · Cell Line · hPSCreg*. Retrieved June 20, 2020, from http://hpscreg.eu/

Ladher, R. K., Wright, T. J., Moon, A. M., Mansour, S. L., & Schoenwolf, G. C. (2005). FGF8 initiates inner ear induction in chick and mouse. *Genes & Development*, *19*(5), 603–613.

Lancaster, M. A., Renner, M., Martin, C. A., & Nature, W. D. (2013). Cerebral organoids model human brain development and microcephaly. *Nature*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3817409

Larcher, L., Norris, J. W., Lejeune, E., Buratti, J., Mignot, C., Garel, C., Keren, B., Schwartz, C. E., & Whalen, S. (2020). The complete loss of function of the SMS gene results in a severe form of Snyder-Robinson syndrome. *European Journal of Medical Genetics*, *63*(4), 103777.

Lee, D.-F., Su, J., Ang, Y.-S., Carvajal-Vergara, X., Mulero-Navarro, S., Pereira, C. F., Gingold, J.,

Wang, H.-L., Zhao, R., Sevilla, A., Darr, H., Williamson, A. J. K., Chang, B., Niu, X., Aguilo, F., Flores, E. R., Sher, Y.-P., Hung, M.-C., Whetton, A. D., … Lemischka, I. R. (2012). Regulation of embryonic and induced pluripotency by aurora kinase-p53 signaling. *Cell Stem Cell*, *11*(2), 179–194.

Lee, H.-K., Yang, Y., Su, Z., Hyeon, C., Lee, T.-S., Lee, H.-W., Kweon, D.-H., Shin, Y.-K., & Yoon, T.-Y. (2010). Dynamic Ca2+-dependent stimulation of vesicle fusion by membrane-anchored synaptotagmin 1. *Science*, *328*(5979), 760–763.

Lee, S. M., Chin, L.-S., & Li, L. (2012). Charcot-Marie-Tooth disease-linked protein SIMPLE functions with the ESCRT machinery in endosomal trafficking. *The Journal of Cell Biology*, *199*(5), 799–816.

Lee, S.-Y., Lee, H.-S., Moon, J.-S., Kim, J.-I., Park, J.-B., Lee, J.-Y., Park, M. J., & Kim, J. (2004). Transcriptional regulation of Zic3 by heterodimeric AP-1(c-Jun/c-Fos) during Xenopus development. *Experimental & Molecular Medicine*, *36*(5), 468–475.

Lei, C.-Y., Wang, W., Zhu, Y.-T., Fang, W.-Y., & Tan, W.-L. (2016). The decrease of cyclin B2 expression inhibits invasion and metastasis of bladder cancer. *Urologic Oncology*, *34*(5), 237.e1–e10.

Liang, W., Miao, S., Zhang, B., He, S., Shou, C., Manivel, P., Krishna, R., Chen, Y., & Shi, Y. E. (2015). Synuclein γ protects Akt and mTOR and renders tumor resistance to Hsp90 disruption. *Oncogene*, *34*(18), 2398–2405.

Liew, Z., Ritz, B., Rebordosa, C., Lee, P.-C., & Olsen, J. (2014). Acetaminophen Use During Pregnancy, Behavioral Problems, and Hyperkinetic Disorders. *JAMA Pediatrics*, *168*(4), 313–320.

Li, F., Su, M., Zhao, H., Xie, W., Cao, S., Xu, Y., Chen, W., Wang, L., Hou, L., & Tan, W. (2019). HnRNP-F promotes cell proliferation by regulating TPX2 in bladder cancer. *American Journal of Translational Research*, *11*(11), 7035–7048.

Li, J., Settivari, R., LeBaron, M. J., & Marty, M. S. (2019). An industry perspective: A streamlined screening strategy using alternative models for chemical assessment of developmental neurotoxicity. *Neurotoxicology*, *73*, 17–30.

Li, J.-Y., Patterson, M., Mikkola, H. K. A., Lowry, W. E., & Kurdistani, S. K. (2012). Dynamic distribution of linker histone H1.5 in cellular differentiation. *PLoS Genetics*, *8*(8), e1002879.

Li, P., Wu, M., Lin, Q., Wang, S., Chen, T., & Jiang, H. (2018). Key genes and integrated modules in hematopoietic differentiation of human embryonic stem cells: a comprehensive bioinformatic analysis. *Stem Cell Research & Therapy*, *9*(1), 301.

Liu, K. E., & Frazier, W. A. (2015). Phosphorylation of the BNIP3 C-Terminus Inhibits Mitochondrial Damage and Cell Death without Blocking Autophagy. *PloS One*, *10*(6), e0129667.

Liu, S. J., Nowakowski, T. J., Pollen, A. A., Lui, J. H., Horlbeck, M. A., Attenello, F. J., He, D., Weissman, J. S., Kriegstein, A. R., Diaz, A. A., & Lim, D. A. (2016). Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biology*, *17*, 67.

Liu, Y., Shin, S., Zeng, X., Zhan, M., Gonzalez, R., Mueller, F.-J., Schwartz, C. M., Xue, H., Li, H., Baker, S. C., Chudin, E., Barker, D. L., McDaniel, T. K., Oeser, S., Loring, J. F., Mattson, M. P., & Rao, M. S. (2006). Genome wide profiling of human embryonic stem cells (hESCs), their derivatives and embryonal carcinoma cells to develop base profiles of U.S. Federal government approved hESC lines. *BMC Developmental Biology*, *6*, 20.

Li, X.-J., Zhang, X., Johnson, M. A., Wang, Z.-B., Lavaute, T., & Zhang, S.-C. (2009). Coordination of sonic hedgehog and Wnt signaling determines ventral and dorsal telencephalic neuron types from human embryonic stem cells. *Development*, *136*(23), 4055–4063.

Luecken, M. D., & Theis, F. J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. *Molecular Systems Biology*, *15*(6), e8746.

Luijsterburg, M. S., Acs, K., Ackermann, L., Wiegant, W. W., Bekker-Jensen, S., Larsen, D. H., Khanna, K. K., Van Attikum, H., Mailand, N., & Dantuma, N. P. (2012). A new non-catalytic role for ubiquitin ligase RNF8 in unfolding higher-order chromatin structure. *The EMBO Journal*, *31*(11), 2511–2527.

Luz, A. L., & Tokar, E. J. (2018). Pluripotent Stem Cells in Developmental Toxicity Testing: A Review of Methodological Advances. *Toxicological Sciences: An Official Journal of the Society of Toxicology*, *165*(1), 31–39.

Maaten, L. van der, & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research: JMLR*, *9*(Nov), 2579–2605.

Maes, T., Barceló, A., & Buesa, C. (2002). Neuron navigator: a human gene family with homology to unc-53, a cell guidance gene from Caenorhabditis elegans. *Genomics*, *80*(1), 21–30.

Magnus, P., Irgens, L. M., Haug, K., Nystad, W., Skjaerven, R., Stoltenberg, C., & MoBa Study Group. (2006). Cohort profile: the Norwegian Mother and Child Cohort Study (MoBa). *International Journal of Epidemiology*, *35*(5), 1146–1150.

Maharaj, D. S., Saravanan, K. S., Maharaj, H., Mohanakumar, K. P., & Daya, S. (2004). Acetaminophen and aspirin inhibit superoxide anion generation and lipid peroxidation, and protect against 1-methyl-4-phenyl pyridinium-induced dopaminergic neurotoxicity in rats. *Neurochemistry International*, *44*(5), 355–360.

Maharaj, H., Maharaj, D. S., & Daya, S. (2006). Acetylsalicylic acid and acetaminophen protect against oxidative neurotoxicity. *Metabolic Brain Disease*, *21*(2-3), 180–190.

Male GPs are less likely to assess cardiovascular risk in female patients. (2016). *European Heart Journal*, *37*(45), 3370–3371.

Mallamaci, A., Mercurio, S., Muzio, L., Cecchi, C., Pardini, C. L., Gruss, P., & Boncinelli, E. (2000). The lack of Emx2 causes impairment of Reelin signaling and defects of neuronal migration in the developing cerebral cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *20*(3), 1109–1118.

Manning, C. S., Biga, V., Boyd, J., Kursawe, J., Ymisson, B., Spiller, D. G., Sanderson, C. M., Galla, T., Rattray, M., & Papalopulu, N. (2019). Quantitative single-cell live imaging links HES5 dynamics with cell-state and fate in murine neurogenesis. *Nature Communications*, *10*(1), 2835.

Ma, N.-X., Yin, J.-C., & Chen, G. (2019). Transcriptome Analysis of Small Molecule–Mediated Astrocyte-to-Neuron Reprogramming. *Frontiers in Cell and Developmental Biology*, *7*, 82.

Maroof, A. M., Keros, S., Tyson, J. A., Ying, S.-W., Ganat, Y. M., Merkle, F. T., Liu, B., Goulburn, A., Stanley, E. G., Elefanty, A. G., Widmer, H. R., Eggan, K., Goldstein, P. A., Anderson, S. A., & Studer, L. (2013). Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. *Cell Stem Cell*, *12*(5), 559–572.

Marshall, J. J., & Mason, J. O. (2019). Mouse vs man: Organoid models of brain development & disease. *Brain Research*, *1724*, 146427.

Martin, J., Taylor, M. J., Rydell, M., Riglin, L., Eyre, O., Lu, Y., Lundström, S., Larsson, H., Thapar, A., & Lichtenstein, P. (2018). Sex-specific manifestation of genetic risk for attention deficit hyperactivity disorder in the general population. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. https://doi.org/10.1111/jcpp.12874

Martinsson-Ahlzén, H.-S., Liberal, V., Grünenfelder, B., Chaves, S. R., Spruck, C. H., & Reed, S. I. (2008). Cyclin-dependent kinase-associated proteins Cks1 and Cks2 are essential during early embryogenesis and for cell cycle progression in somatic cells. *Molecular and Cellular Biology*, *28*(18), 5698–5709.

Masjosthusmann, S., Becker, D., Petzuch, B., Klose, J., Siebert, C., Deenen, R., Barenys, M., Baumann, J., Dach, K., Tigges, J., Hübenthal, U., Köhrer, K., & Fritsche, E. (2018). A transcriptome comparison of time-matched developing human, mouse and rat neural progenitor cells reveals human uniqueness. *Toxicology and Applied Pharmacology*, *354*, 40–55.

Maslov, A. Y., Barone, T. A., Plunkett, R. J., & Pruitt, S. C. (2004). Neural stem cell detection, characterization, and age-related changes in the subventricular zone of mice. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(7), 1726–1733.

Mayer, C., Hafemeister, C., Bandler, R. C., Machold, R., Batista Brito, R., Jaglin, X., Allaway, K., Butler, A., Fishell, G., & Satija, R. (2018). Developmental diversification of cortical inhibitory interneurons. *Nature*, *555*(7697), 457–462.

McCarthy, D. J., Campbell, K. R., Lun, A. T. L., & Wills, Q. F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* . https://academic.oup.com/bioinformatics/article-abstract/33/8/1179/2907823

McCarthy, S. (2016). Gender inequality: Unconscious and systematic bias remains a problem in emergency medicine. *Emergency Medicine Australasia*, *28*(3), 344–346.

McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, *3*(29), 861.

Medina, L. (2009). Evolution and Embryological Development of Forebrain. In *Encyclopedia of Neuroscience* (pp. 1172–1192). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-29678-2_3112

Meganathan, K., Jagtap, S., Wagh, V., Winkler, J., Gaspar, J., Hildebrand, D., Trusch, M., Lehmann, K., Hescheler, J., Schlüter, H., & Sachinidis, A. (2012). Identification of Thalidomide-Specific Transcriptomics and Proteomics Signatures during Differentiation of Human Embryonic Stem Cells. *PloS One*, *7*(8), e44228.

Micali, N., Kim, S.-K., Diaz-Bustamante, M., Stein-O'Brien, G., Seo, S., Shin, J.-H., Rash, B. G., Ma, S., Wang, Y., Olivares, N. A., Arellano, J. I., Maynard, K. R., Fertig, E. J., Cross, A. J., Bürli, R. W., Brandon, N. J., Weinberger, D. R., Chenoweth, J. G., Hoeppner, D. J., … McKay, R. D. (2020). Variation of Human Neural Stem Cells Generating Organizer States In Vitro before Committing to Cortical Excitatory or Inhibitory Neuronal Fates. *Cell Reports*, *31*(5), 107599.

Miller, J. A., Ding, S.-L., Sunkin, S. M., Smith, K. A., Ng, L., Szafer, A., Ebbert, A., Riley, Z. L., Royall, J. J., Aiona, K., Arnold, J. M., Bennet, C., Bertagnolli, D., Brouner, K., Butler, S., Caldejon, S., Carey, A., Cuhaciyan, C., Dalley, R. A., … Lein, E. S. (2014). Transcriptional landscape of the prenatal human brain. *Nature*, *508*(7495), 199–206.

Mizutani, K.-I., Yoon, K., Dang, L., Tokunaga, A., & Gaiano, N. (2007). Differential Notch signalling distinguishes neural stem cells from intermediate progenitors. *Nature*, *449*(7160), 351–355.

Mohamed, J. S., Sheik Mohamed, J., Gaughwin, P. M., 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*(2), 324–337.

Molyneaux, B. J., Goff, L. A., Brettler, A. C., Chen, H.-H., Hrvatin, S., Rinn, J. L., & Arlotta, P. (2015).

DeCoN: genome-wide analysis of in vivo transcriptional dynamics during pyramidal neuron fate selection in neocortex. *Neuron*, *85*(2), 275–288.

Mulligan, K. A., & Cheyette, B. N. R. (2017). Neurodevelopmental Perspectives on Wnt Signaling in Psychiatry. *Molecular Neuropsychiatry*, *2*(4), 219–246.

Nazıroğlu, M., Cihangir Uğuz, A., Koçak, A., & Bal, R. (2009). Acetaminophen at Different Doses Protects Brain Microsomal Ca2+-ATPase and the Antioxidant Redox System in Rats. *The Journal of Membrane Biology*, *231*(2), 57–64.

Nelson, B. R., Hodge, R. D., Bedogni, F., & Hevner, R. F. (2013). Dynamic interactions between intermediate neurogenic progenitors and radial glia in embryonic mouse neocortex: potential role in Dll1-Notch signaling. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(21), 9122–9139.

O'Flanagan, C. H., Campbell, K. R., Zhang, A. W., Kabeer, F., Lim, J. L. P., Biele, J., Eirew, P., Lai, D., McPherson, A., Kong, E., Bates, C., Borkowski, K., Wiens, M., Hewitson, B., Hopkins, J., Pham, J., Ceglia, N., Moore, R., Mungall, A. J., … The CRUK IMAXT Grand Challenge Team. (2019). Dissociation of solid tumor tissues with cold active protease for single-cell RNA-seq minimizes conserved collagenase-associated stress responses. *Genome Biology*, *20*(1), 210.

O'Hearn, E. E., Hwang, H. S., Holmes, S. E., Rudnicki, D. D., Chung, D. W., Seixas, A. I., Cohen, R. L., Ross, C. A., Trojanowski, J. Q., Pletnikova, O., Troncoso, J. C., & Margolis, R. L. (2015). Neuropathology and Cellular Pathogenesis of Spinocerebellar Ataxia Type 12. *Movement Disorders: Official Journal of the Movement Disorder Society*, *30*(13), 1813–1824.

O'Hearn, E., Holmes, S. E., & Margolis, R. L. (2012). Spinocerebellar ataxia type 12. *Handbook of Clinical Neurology*, *103*, 535–547.

Ohtsuka, T., Ishibashi, M., Gradwohl, G., Nakanishi, S., Guillemot, F., & Kageyama, R. (1999). Hes1 and Hes5 as notch effectors in mammalian neuronal differentiation. *The EMBO Journal*, *18*(8), 2196–2207.

O'Leary, D. D., & Sahara, S. (2008). Genetic regulation of arealization of the neocortex. *Current Opinion in Neurobiology*, *18*(1), 90–100.

Osborne, L., Clive, M., Kimmel, M., Gispen, F., Guintivano, J., Brown, T., Cox, O., Judy, J., Meilman, S., Braier, A., Beckmann, M. W., Kornhuber, J., Fasching, P. A., Goes, F., Payne, J. L., Binder, E. B., & Kaminsky, Z. (2016). Replication of Epigenetic Postpartum Depression Biomarkers and Variation with Hormone Levels. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *41*(6), 1648–1658.

Ostvold, A. C., Holtlund, J., & Laland, S. G. (1985). A novel, highly phosphorylated protein, of the high-mobility group type, present in a variety of proliferating and non-proliferating mammalian cells. *European Journal of Biochemistry*, *153*(3), 469–475.

Ouwenga, R., Lake, A. M., Aryal, S., Lagunas, T., Jr, & Dougherty, J. D. (2018). The Differences in Local Translatome across Distinct Neuron Types Is Mediated by Both Baseline Cellular Differences and Post-transcriptional Mechanisms. *eNeuro*, *5*(6). https://doi.org/10.1523/ENEURO.0320-18.2018

Ozair, M. Z., Kintner, C., & Brivanlou, A. H. (2013). Neural induction and early patterning in vertebrates. *Wiley Interdisciplinary Reviews. Developmental Biology*, *2*(4), 479–498.

Paeschke, K., Bochman, M. L., Garcia, P. D., Cejka, P., Friedman, K. L., Kowalczykowski, S. C., & Zakian, V. A. (2013). Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature*, *497*(7450), 458–462.

Paina, S., Garzotto, D., DeMarchis, S., Marino, M., Moiana, A., Conti, L., Cattaneo, E., Perera, M., Corte, G., Calautti, E., & Merlo, G. R. (2011). Wnt5a is a transcriptional target of Dlx homeogenes and promotes differentiation of interneuron progenitors in vitro and in vivo. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(7), 2675–2687.

Pal, R., Mamidi, M. K., Das, A. K., & Bhonde, R. (2012). Diverse effects of dimethyl sulfoxide (DMSO) on the differentiation potential of human embryonic stem cells. *Archives of Toxicology*, *86*(4), 651–661.

Pal, R., Totey, S., Mamidi, M. K., Bhat, V. S., & Totey, S. (2009). Propensity of human embryonic stem cell lines during early stage of lineage specification controls their terminal differentiation into mature cell types. *Experimental Biology and Medicine* , *234*(10), 1230–1243.

Pan, C., & Fan, Y. (2016). Role of H1 linker histones in mammalian development and stem cell differentiation. *Biochimica et Biophysica Acta*, *1859*(3), 496–509.

Parish, E. V., Mason, J. O., & Price, D. J. (2016). Expression of Barhl2 and its relationship with Pax6 expression in the forebrain of the mouse embryo. *BMC Neuroscience*, *17*(1), 76.

Park, S. W., Nhieu, J., Persaud, S. D., Miller, M. C., Xia, Y., Lin, Y.-W., Lin, Y.-L., Kagechika, H., Mayo, K. H., & Wei, L.-N. (2019). A new regulatory mechanism for Raf kinase activation, retinoic acid-bound Crabp1. *Scientific Reports*, *9*(1), 10929.

Patani, R., Compston, A., Puddifoot, C. A., Wyllie, D. J. A., Hardingham, G. E., Allen, N. D., & Chandran, S. (2009). Activin/Nodal inhibition alone accelerates highly efficient neural conversion from human embryonic stem cells and imposes a caudal positional identity. *PLoS One*, *4*(10), e7327.

Pederick, D. T., Richards, K. L., Piltz, S. G., Kumar, R., Mincheva-Tasheva, S., Mandelstam, S. A., Dale, R. C., Scheffer, I. E., Gecz, J., Petrou, S., Hughes, J. N., & Thomas, P. Q. (2018). Abnormal Cell Sorting Underlies the Unique X-Linked Inheritance of PCDH19 Epilepsy. *Neuron*, *97*(1), 59–66.e5.

Peng, Y., Xu, R.-H., Mei, J.-M., Li, X.-P., Yan, D., Kung, H.-F., & Phang, J. M. (2002). Neural inhibition by c-Jun as a synergizing factor in bone morphogenetic protein 4 signaling. *Neuroscience*, *109*(4), 657–664.

Person, F., Wilczak, W., Hube-Magg, C., Burdelski, C., Möller-Koop, C., Simon, R., Noriega, M., Sauter, G., Steurer, S., Burdak-Rothkamm, S., & Jacobsen, F. (2017). Prevalence of βIII-tubulin (TUBB3) expression in human normal tissues and cancers. *Tumour Biology: The Journal of the International Society for Oncodevelopmental Biology and Medicine*, *39*(10), 1010428317712166.

Petri, R., Brattås, P. L., Sharma, Y., Jönsson, M. E., Pircs, K., Bengzon, J., & Jakobsson, J. (2019). LINE-2 transposable elements are a source of functional human microRNAs and target sites. *PLoS Genetics*, *15*(3), e1008036.

Peyre, E., Silva, C. G., & Nguyen, L. (2015). Crosstalk between intracellular and extracellular signals regulating interneuron production, migration and integration into the cortex. *Frontiers in Cellular Neuroscience*, *9*, 129.

Pickering, G., Estève, V., Loriot, M.-A., Eschalier, A., & Dubray, C. (2008). Acetaminophen reinforces descending inhibitory pain pathways. *Clinical Pharmacology and Therapeutics*, *84*(1), 47–51.

Popovitchenko, T., Park, Y., Page, N. F., Luo, X., Krsnik, Z., Liu, Y., Salamon, I., Stephenson, J. D., Kraushar, M. L., Volk, N. L., Patel, S. M., Wijeratne, H. R. S., Li, D., Suthar, K. S., Wach, A., Sun, M., Arnold, S. J., Akamatsu, W., Okano, H., … Rasin, M.-R. (2020). Translational derepression of Elavl4 isoforms at their alternative 5′ UTRs determines neuronal development. *Nature Communications*, *11*(1), 1674.

Pourebrahim, R., Houtmeyers, R., Ghogomu, S., Janssens, S., Thelie, A., Tran, H. T., Langenberg, T., Vleminckx, K., Bellefroid, E., Cassiman, J.-J., & Tejpar, S. (2011). Transcription factor Zic2 inhibits Wnt/β-catenin protein signaling. *The Journal of Biological Chemistry*, *286*(43), 37732–37740.

Pressler, R., & Auvin, S. (2013). Comparison of Brain Maturation among Species: An Example in Translational Research Suggesting the Possible Use of Bumetanide in Newborn. *Frontiers in Neurology*, *4*, 36.

Quintana-Urzainqui, I., Kozić, Z., Mitra, S., Tian, T., Manuel, M., Mason, J. O., & Price, D. J. (2018). Tissue-Specific Actions of Pax6 on Proliferation and Differentiation Balance in Developing Forebrain Are Foxg1 Dependent. *iScience*, *10*, 171–191.

Raciti, M., & Ceccatelli, S. (2017). Epigenetic mechanisms in developmental neurotoxicity. *Neurotoxicology and Teratology*. https://doi.org/10.1016/j.ntt.2017.12.002

Radio, N. M., Breier, J. M., Reif, D. M., Judson, R. S., Martin, M., Houck, K. A., Mundy, W. R., & Shafer, T. J. (2015). Use of Neural Models of Proliferation and Neurite Outgrowth to Screen Environmental Chemicals in the ToxCast Phase I Library. *Applied In Vitro Toxicology*, *1*(2), 131–139.

Raj, B., & Blencowe, B. J. (2015). Alternative Splicing in the Mammalian Nervous System: Recent Insights into Mechanisms and Functional Roles. *Neuron*, *87*(1), 14–27.

Ratié, L., Ware, M., Jagline, H., David, V., & Dupé, V. (2014). Dynamic expression of Notch-dependent neurogenic markers in the chick embryonic nervous system. *Frontiers in Neuroanatomy*, *8*, 158.

Rattner, J. B., Rao, A., Fritzler, M. J., Valencia, D. W., & Yen, T. J. (1993). CENP-F is a ca 400 kDa kinetochore protein that exhibits a cell-cycle dependent localization. *Cell Motility and the Cytoskeleton*, *26*(3), 214–226.

Re, A., Workman, C. T., Waldron, L., Quattrone, A., & Brunak, S. (2014). Lineage-specific interface proteins match up the cell cycle and differentiation in embryo stem cells. *Stem Cell Research*, *13*(2), 316–328.

Ren, Z., & Qiu, A. (2014). Sleep-related breathing disorder is associated with hyperactivity in preschoolers. *Singapore Medical Journal*, *55*(5), 257–260.

Richmond, L. M. (2019). Gender Bias, Discrimination Common in Academic Psychiatry. *Psychiatric News*, *54*(13). https://doi.org/10.1176/appi.pn.2019.6b23

Rifes, P., Isaksson, M., Rathore, G. S., Aldrin-Kirk, P., Møller, O. K., Barzaghi, G., Lee, J., Egerod, K. L., Rausch, D. M., Parmar, M., Pers, T. H., Laurell, T., & Kirkeby, A. (2020). Modeling neural tube development by differentiation of human embryonic stem cells in a microfluidic WNT gradient. *Nature Biotechnology*. https://doi.org/10.1038/s41587-020-0525-0

Roberts, T. C., & Morris, K. V. (2013). Not so pseudo anymore: pseudogenes as therapeutic targets. *Pharmacogenomics*, *14*(16), 2023–2034.

Roberts, T. C., Morris, K. V., & Wood, M. J. A. (2014). The role of long non-coding RNAs in neurodevelopment, brain function and neurological disease. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1652). https://doi.org/10.1098/rstb.2013.0507

Rolls, A., Shechter, R., London, A., Ziv, Y., Ronen, A., Levy, R., & Schwartz, M. (2007). Toll-like receptors modulate adult hippocampal neurogenesis. *Nature Cell Biology*, *9*(9), 1081–1088.

Rose, J. E., Behm, F. M., Drgon, T., Johnson, C., & Uhl, G. R. (2010). Personalized smoking cessation: interactions between nicotine dose, dependence and quit-success genotype score. *Molecular Medicine* , *16*(7-8), 247–253.

Rosenbaum, J. N., Duggan, A., & García-Añoveros, J. (2011). Insm1promotes the transition of olfactory progenitors from apical and proliferative to basal, terminally dividing and neuronogenic. *Neural Development*, *6*(1), 6.

Roth, M., Bonev, B., Lindsay, J., Lea, R., Panagiotaki, N., Houart, C., & Papalopulu, N. (2010). FoxG1 and TLE2 act cooperatively to regulate ventral telencephalon formation. *Development*, *137*(9), 1553–1562.

Rust, W., Balakrishnan, T., & Zweigerdt, R. (2009). Cardiomyocyte enrichment from human embryonic stem cell cultures by selection of ALCAM surface expression. *Regenerative Medicine*, *4*(2), 225–237.

Saliba, S. W., Bonifacino, T., Serchov, T., Bonanno, G., de Oliveira, A. C. P., & Fiebich, B. L. (2019). Neuroprotective Effect of AM404 Against NMDA-Induced Hippocampal Excitotoxicity. *Frontiers in Cellular Neuroscience*, *13*, 566.

Sarkar, P., & M., B. (2011). Role of Signaling Pathways and Epigenetic Factors in Lineage Determination During Human Embryonic Stem Cell Differentiation. *Embryonic Stem Cells - Differentiation and Pluripotent Alternatives*. https://doi.org/10.5772/22787

Satija Lab. (n.d.). Retrieved July 3, 2020, from https://satijalab.org/seurat/v3.0/immune_alignment.html

Sato, K., Yabe, I., Fukuda, Y., Soma, H., Nakahara, Y., Tsuji, S., & Sasaki, H. (2010). Mapping of autosomal dominant cerebellar ataxia without the pathogenic PPP2R2B mutation to the locus for spinocerebellar ataxia 12. *Archives of Neurology*, *67*(10), 1257–1262.

Schicht, M., Garreis, F., Hartjen, N., Beileke, S., Jacobi, C., Sahin, A., Holland, D., Schröder, H., Hammer, C. M., Paulsen, F., & Bräuer, L. (2018). SFTA3 – a novel surfactant protein of the ocular surface and its role in corneal wound healing and tear film surface tension. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-28005-9

Schulpen, S., de Jong, E., de la Fonteyne, L., de Klerk, A., & Piersma, A. H. (2015). Distinct gene expression responses of two anticonvulsant drugs in a novel human embryonic stem cell based neural differentiation assay protocol. *Toxicology in Vitro: An International Journal Published in Association with BIBRA*, *29*(3), 449–457.

Schult, D., Hölsken, A., Buchfelder, M., Schlaffer, S.-M., Siegel, S., Kreitschmann-Andermahr, I., Fahlbusch, R., & Buslei, R. (2015). Expression pattern of neuronal intermediate filament α-internexin in anterior pituitary gland and related tumors. *Pituitary*, *18*(4), 465–473.

Seki, T., Hori, T., Miyata, H., Maehara, M., & Namba, T. (2019). Analysis of proliferating neuronal progenitors and immature neurons in the human hippocampus surgically removed from control and epileptic patients. *Scientific Reports*, *9*(1), 18194.

Shang, Z., Chen, D., Wang, Q., Wang, S., Deng, Q., Wu, L., Liu, C., Ding, X., Wang, S., Zhong, J., Zhang, D., Cai, X., Zhu, S., Yang, H., Liu, L., Fink, J. L., Chen, F., Liu, X., Gao, Z., & Xu, X. (2018). Single-cell RNA-seq reveals dynamic transcriptome profiling in human early neural differentiation. *GigaScience*, *7*(11). https://doi.org/10.1093/gigascience/giy117

Shao, X., Liao, J., Lu, X., Xue, R., Ai, N., & Fan, X. (2020). scCATCH: Automatic Annotation on Cell Types of Clusters from Single-Cell RNA Sequencing Data. *iScience*, *23*(3), 100882.

Shimizu, T., Nakazawa, M., Kani, S., Bae, Y.-K., Shimizu, T., Kageyama, R., & Hibi, M. (2010). Zinc finger genes Fezf1 and Fezf2 control neuronal differentiation by repressing Hes5 expression in the forebrain. *Development*, *137*(11), 1875–1885.

Shinde, V., Klima, S., Sureshkumar, P. S., Meganathan, K., Jagtap, S., Rempel, E., Rahnenführer, J.,

Hengstler, J. G., Waldmann, T., Hescheler, J., Leist, M., & Sachinidis, A. (2015). Human Pluripotent Stem Cell Based Developmental Toxicity Assays for Chemical Safety Screening and Systems Biology Data Generation. *Journal of Visualized Experiments: JoVE*, *100*, e52333.

Shin, S. S., Bales, J. W., Yan, H. Q., Kline, A. E., Wagner, A. K., Lyons-Weiler, J., & Dixon, C. E. (2013). The effect of environmental enrichment on substantia nigra gene expression after traumatic brain injury in rats. *Journal of Neurotrauma*, *30*(4), 259–270.

Shpyleva, S., Melnyk, S., Pavliv, O., Pogribny, I., & Jill James, S. (2018). Overexpression of LINE-1 Retrotransposons in Autism Brain. *Molecular Neurobiology*, *55*(2), 1740–1749.

Singh, A., Azad, M., Shymko, M. D., Henson, E. S., Katyal, S., Eisenstat, D. D., & Gibson, S. B. (2018). The BH3 only Bcl-2 family member BNIP3 regulates cellular proliferation. *PloS One*, *13*(10), e0204792.

Singh, S., Park, B., Jankipram, S.-G., Road, K., Lucknow, & India. (2019). Functional association between NUCKS1 gene and Parkinson disease: A potential susceptibility biomarker. *Bioinformation*, *15*(8), 548–556.

Smith, D., Wagner, E., Koul, O., McCaffery, P., & Dräger, U. C. (2001). Retinoic acid synthesis for the developing telencephalon. *Cerebral Cortex* , *11*(10), 894–905.

Sofer, A., Lei, K., Johannessen, C. M., & Ellisen, L. W. (2005). Regulation of mTOR and cell growth in response to energy stress by REDD1. *Molecular and Cellular Biology*, *25*(14), 5834–5845.

Sohn, H.-M., Kim, H. Y., Park, S., Han, S.-H., & Kim, J.-H. (2017). Isoflurane decreases proliferation and differentiation, but none of the effects persist in human embryonic stem cell-derived neural progenitor cells. *Journal of Anesthesia*, *31*(1), 36–43.

Soltani, M. H., Pichardo, R., Song, Z., Sangha, N., Camacho, F., Satyamoorthy, K., Sangueza, O. P., & Setaluri, V. (2005). Microtubule-Associated Protein 2, a Marker of Neuronal Differentiation, Induces Mitotic Defects, Inhibits Growth of Melanoma Cells, and Predicts Metastatic Potential of Cutaneous Melanoma. *The American Journal of Pathology*, *166*(6), 1841–1850.

Somel, M., Franz, H., Yan, Z., Lorenc, A., Guo, S., Giger, T., Kelso, J., Nickel, B., Dannemann, M., Bahn, S., Webster, M. J., Weickert, C. S., Lachmann, M., Pääbo, S., & Khaitovich, P. (2009). Transcriptional neoteny in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(14), 5743–5748.

Sotthibundhu, A., McDonagh, K., von Kriegsheim, A., Garcia-Munoz, A., Klawiter, A., Thompson, K., Chauhan, K. D., Krawczyk, J., McInerney, V., Dockery, P., Devine, M. J., Kunath, T., Barry, F., O'Brien, T., & Shen, S. (2016). Rapamycin regulates autophagy and cell adhesion in induced pluripotent stem cells. *Stem Cell Research & Therapy*, *7*(1), 166.

Sousa, A. M. M., Zhu, Y., Raghanti, M. A., Kitchen, R. R., Onorati, M., Tebbenkamp, A. T. N., Stutz, B., Meyer, K. A., Li, M., Kawasawa, Y. I., Liu, F., Perez, R. G., Mele, M., Carvalho, T., Skarica, M., Gulden, F. O., Pletikos, M., Shibata, A., Stephenson, A. R., … Sestan, N. (2017). Molecular and cellular reorganization of neural circuits in the human lineage. *Science*, *358*(6366), 1027–1032.

Stafford, N. (2009). Female doctors are more likely to follow heart failure guidelines than male counterparts. *BMJ*, *338*(feb02 2), b424–b424.

Stahlberg, A., Bengtsson, M., Hemberg, M., & Semb, H. (2009). Quantitative transcription factor analysis of undifferentiated single human embryonic stem cells. *Clinical Chemistry*, *55*(12), 2162–2170.

Stark, A. M., Doukas, A., Hugo, H.-H., Hedderich, J., Hattermann, K., Maximilian Mehdorn, H., & Held-Feindt, J. (2015). Expression of DNA mismatch repair proteins MLH1, MSH2, and MSH6 in

recurrent glioblastoma. *Neurological Research*, *37*(2), 95–105.

Stergiakouli, E., Thapar, A., & Smith, G. (2016). Association of Acetaminophen Use During Pregnancy With Behavioral Problems in Childhood: Evidence Against Confounding. *AMA Pediatr.* https://doi.org/10.1001/jamapediatrics.2016.1775

Storm, E. E., Garel, S., Borello, U., Hebert, J. M., Martinez, S., McConnell, S. K., Martin, G. R., & Rubenstein, J. L. R. (2006). Dose-dependent functions of Fgf8 in regulating telencephalic patterning centers. *Development* , *133*(9), 1831–1844.

Strohbeck-Kuehner, P., Skopp, G., & Mattern, R. (2008). Cannabis improves symptoms of ADHD. *Cannabinoids*, *3*(1), 1–3.

Ström, S., Holm, F., Bergström, R., Strömberg, A.-M., & Hovatta, O. (2010). Derivation of 30 human embryonic stem cell lines—improving the quality. *In Vitro Cellular & Developmental Biology - Animal*, *46*(3-4), 337–344.

Suarez, N. A., Macia, A., & Muotri, A. R. (2017). LINE-1 retrotransposons in healthy and diseased human brain. *Developmental Neurobiology*. https://doi.org/10.1002/dneu.22567

Sugino, K., Clark, E., Schulmann, A., Shima, Y., Wang, L., Hunt, D. L., Hooks, B. M., Tränkner, D., Chandrashekar, J., Picard, S., Lemire, A. L., Spruston, N., Hantman, A. W., & Nelson, S. B. (2019). Mapping the transcriptional diversity of genetically and anatomically defined cell populations in the mouse brain. *eLife*, *8*. https://doi.org/10.7554/eLife.38619

Sunkin, S. M., Ng, L., Lau, C., Dolbeare, T., Gilbert, T. L., Thompson, C. L., Hawrylycz, M., & Dang, C. (2013). Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Research*, *41*(Database issue), D996–D1008.

Sünwoldt, J., Bosche, B., Meisel, A., & Mergenthaler, P. (2017). Neuronal Culture Microenvironments Determine Preferences in Bioenergetic Pathway Use. *Frontiers in Molecular Neuroscience*, *10*, 305.

Suzuki, I. K., & Vanderhaeghen, P. (2015). Is this a brain which I see before me? Modeling human neural development with pluripotent stem cells. *Development* , *142*(18), 3138–3150.

Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, *126*(4), 663–676.

Tantirigama, M. L. S., Oswald, M. J., Duynstee, C., Hughes, S. M., & Empson, R. M. (2014). Expression of the Developmental Transcription Factor Fezf2 Identifies a Distinct Subpopulation of Layer 5 Intratelencephalic-Projection Neurons in Mature Mouse Motor Cortex. *Journal of Neuroscience*, *34*(12), 4303–4308.

Teotia, P., Niu, M., & Ahmad, I. (2020). Mapping developmental trajectories and subtype diversity of normal and glaucomatous human retinal ganglion cells by single-cell transcriptome analysis: Single-cell transcriptome analysis of hRGCs. *Stem Cells*, *37*, 31.

Terry, D. M., & Devine, S. E. (2019). Aberrantly High Levels of Somatic LINE-1 Expression and Retrotransposition in Human Neurological Disorders. *Frontiers in Genetics*, *10*, 1244.

Theil, T., Dominguez-Frutos, E., & Schimmang, T. (2008). Differential requirements for Fgf3 and Fgf8 during mouse forebrain development. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *237*(11), 3417–3423.

Thompson, J., Waldie, K. E., Wall, C. R., Murphy, R., Mitchell, E. A., & Study group, A. (2014). Associations between Acetaminophen Use during Pregnancy and ADHD Symptoms Measured at Ages 7 and 11 Years. *PLoS ONE*, *9*(9), e108210.

Thoms, J. A. I., Loch, H. M., Bamburg, J. R., Gunning, P. W., & Weinberger, R. P. (2008). A

tropomyosin 1 induced defect in cytokinesis can be rescued by elevated expression of cofilin. *Cell Motility and the Cytoskeleton*, *65*(12), 979–990.

Todd, L., Suarez, L., Quinn, C., & Fischer, A. J. (2018). Retinoic Acid-Signaling Regulates the Proliferative and Neurogenic Capacity of Müller Glia-Derived Progenitor Cells in the Avian Retina. *Stem Cells*, *36*(3), 392–405.

Tojkander, S., Gateva, G., & Lappalainen, P. (2012). Actin stress fibers--assembly, dynamics and biological roles. *Journal of Cell Science*, *125*(8), 1855–1864.

Toledo, C. M., Herman, J. A., Olsen, J. B., Ding, Y., Corrin, P., Girard, E. J., Olson, J. M., Emili, A., DeLuca, J. G., & Paddison, P. J. (2014). BuGZ is required for Bub3 stability, Bub1 kinetochore function, and chromosome alignment. *Developmental Cell*, *28*(3), 282–294.

Trevino, A. E., Sinnott-Armstrong, N., Andersen, J., Yoon, S.-J., Huber, N., Pritchard, J. K., Chang, H. Y., Greenleaf, W. J., & Paşca, S. P. (2020). Chromatin accessibility dynamics in a model of human forebrain development. *Science*, *367*(6476). https://doi.org/10.1126/science.aay1645

Uhrig, M., Brechlin, P., Jahn, O., Knyazev, Y., Weninger, A., Busia, L., Honarnejad, K., Otto, M., & Hartmann, T. (2008). Upregulation of CRABP1 in human neuroblastoma cells overproducing the Alzheimer-typical Aβ 42 reduces their differentiation potential. *BMC Medicine*, *6*(1), 1–12.

Uzquiano, A., Gladwyn-Ng, I., Nguyen, L., Reiner, O., Götz, M., Matsuzaki, F., & Francis, F. (2018). Cortical progenitor biology: key features mediating proliferation versus differentiation. *Journal of Neurochemistry*, *146*(5), 500–525.

Vadivelu, N., Kai, A. M., Kodumudi, V., Sramcik, J., & Kaye, A. D. (2018). The Opioid Crisis: a Comprehensive Overview. *Current Pain and Headache Reports*, *22*(3), 16.

Valenzuela, D. M., Economides, A. N., Rojas, E., Lamb, T. M., Nuñez, L., Jones, P., Lp, N. Y., Espinosa, R., 3rd, Brannan, C. I., & Gilbert, D. J. (1995). Identification of mammalian noggin and its expression in the adult nervous system. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *15*(9), 6077–6084.

van Heusden, F., Macey-Dare, A., Krajeski, R. N., Sharott, A., & Ellender, T. J. (2019). Distinct neural progenitor pools in the ventral telencephalon generate diversity in striatal spiny projection neurons. In *bioRxiv* (p. 770057). https://doi.org/10.1101/770057

van Thriel, C., Westerink, R. H. S., Beste, C., Bale, A. S., Lein, P. J., & Leist, M. (2012). Translating neurobehavioural endpoints of developmental neurotoxicity tests into in vitro assays and readouts. *Neurotoxicology*, *33*(4), 911–924.

Villanea, F. A., Perry, G. H., Gutiérrez-Espeleta, G. A., & Dominy, N. J. (2012). ASPM and the evolution of cerebral cortical size in a community of New World monkeys. *PloS One*, *7*(9), e44928.

Vong, K. I., Leung, C. K. Y., Behringer, R. R., & Kwan, K. M. (2015). Sox9 is critical for suppression of neurogenesis but not initiation of gliogenesis in the cerebellum. *Molecular Brain*, *8*(1), 25.

Waldmann, T., Rempel, E., Balmer, N. V., König, A., Kolde, R., Gaspar, J. A., Henry, M., Hescheler, J., Sachinidis, A., Rahnenführer, J., Hengstler, J. G., & Leist, M. (2014). Design principles of concentration-dependent transcriptome deviations in drug-exposed differentiating stem cells. *Chemical Research in Toxicology*, *27*(3), 408–420.

Walker, T. L., Yasuda, T., Adams, D. J., & Bartlett, P. F. (2007). The doublecortin-expressing population in the developing and adult brain contains multipotential precursors in addition to neuronal-lineage cells. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(14), 3734–3742.

Wang, J., Jenjaroenpun, P., Bhinge, A., Angarica, V. E., Del Sol, A., Nookaew, I., Kuznetsov, V. A., & Stanton, L. W. (2017). Single-cell gene expression analysis reveals regulators of distinct cell subpopulations among developing human neurons. *Genome Research*, *27*(11), 1783–1794.

Wang, J., Zhang, Y., Hou, J., Qian, X., Zhang, H., Zhang, Z., Li, M., Wang, R., Liao, K., Wang, Y., Li, Z., Zhong, D., Wan, P., Dong, L., Liu, F., Wang, X., Wan, Y., Xiao, W., & Zhang, W. W. (2016). Ube2s regulates Sox2 stability and mouse ES cell maintenance. *Cell Death and Differentiation*, *23*(3), 393–404.

Wang, L., Yu, L., Zhang, T., Wang, L., Leng, Z., Guan, Y., & Wang, X. (2014). HMGB1 enhances embryonic neural stem cell proliferation by activating the MAPK signaling pathway. *Biotechnology Letters*, *36*(8), 1631–1639.

Wang, P., Zhou, Y., Yang, J.-Q., Landeck, L., Min, M., Chen, X.-B., Chen, J.-Q., Li, W., Cai, S.-Q., Zheng, M., & Man, X.-Y. (2018). The role of Sprouty1 in the proliferation, differentiation and apoptosis of epidermal keratinocytes. *Cell Proliferation*, *51*(5), e12477.

Wang, T.-W., Zhang, H., & Parent, J. M. (2005). Retinoic acid regulates postnatal neurogenesis in the murine subventricular zone-olfactory bulb pathway. *Development* , *132*(12), 2721–2732.

Wardle, M., Majounie, E., Muzaimi, M. B., Williams, N. M., Morris, H. R., & Robertson, N. P. (2009). The genetic aetiology of late-onset chronic progressive cerebellar ataxia. A population-based study. *Journal of Neurology*, *256*(3), 343–348.

Ward, R. M. (2001). Difficulties in the study of adverse fetal and neonatal effects of drug therapy during pregnancy. *Seminars in Perinatology*, *25*(3), 191–195.

Ware, M., Hamdi-Rozé, H., Le Friec, J., David, V., & Dupé, V. (2016). Regulation of downstream neuronal genes by proneural transcription factors during initial neurogenesis in the vertebrate brain. *Neural Development*, *11*(1), 22.

Werler, M. M., Mitchell, A. A., Hernandez-Diaz, S., & Honein, M. A. (2005). Use of over-the-counter medications during pregnancy. *American Journal of Obstetrics and Gynecology*, *193*(3 Pt 1), 771–777.

Wersinger, C., & Sidhu, A. (2009). Partial regulation of serotonin transporter function by gamma-synuclein. *Neuroscience Letters*, *453*(3), 157–161.

Wilson, S. W., & Houart, C. (2004). Early steps in the development of the forebrain. *Developmental Cell*, *6*(2), 167–181.

Winham, S. J., Cuellar-Barboza, A. B., Oliveros, A., McElroy, S. L., Crow, S., Colby, C., Choi, D.-S., Chauhan, M., Frye, M., & Biernacka, J. M. (2014). Genome-wide association study of bipolar disorder accounting for effect of body mass index identifies a new risk allele in TCF7L2. *Molecular Psychiatry*, *19*(9), 1010–1016.

Wolf, F. A., Angerer, P., & Theis, F. J. (2018). SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology*, *19*(1), 15.

Wolf, F. A., Hamey, F. K., Plass, M., Solana, J., Dahlin, J. S., Göttgens, B., Rajewsky, N., Simon, L., & Theis, F. J. (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biology*, *20*(1), 59.

Won, H., Huang, J., Opland, C. K., Hartl, C. L., & Geschwind, D. H. (2019). Human evolved regulatory elements modulate genes involved in cortical expansion and neurodevelopmental disease susceptibility. *Nature Communications*, *10*(1), 2396.

Won, J. Y., Nam, E.-C., Yoo, S. J., Kwon, H. J., Um, S. J., Han, H. S., Kim, S. H., Byun, Y., & Kim, S.

Y. (2004). The effect of cellular retinoic acid binding protein-I expression on the CYP26-mediated catabolism of all-trans retinoic acid and cell proliferation in head and neck squamous cell carcinoma. *Metabolism*, *53*(8), 1007–1012.

Xue, K., Ng, J.-H., & Ng, H.-H. (2011). Mapping the networks for pluripotency. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1575), 2238–2246.

Xu, Z., Robitaille, A. M., Berndt, J. D., Davidson, K. C., Fischer, K. A., Mathieu, J., Potter, J. C., Ruohola-Baker, H., & Moon, R. T. (2016). Wnt/β-catenin signaling promotes self-renewal and inhibits the primed state transition in naïve human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(42), E6382–E6390.

Yamada, M., Clark, J., & Iulianella, A. (2014). MLLT11/AF1q is differentially expressed in maturing neurons during development. *Gene Expression Patterns: GEP*, *15*(2), 80–87.

Yan, L., Li, S., Xu, C., Zhao, X., Hao, B., Li, H., & Qiao, B. (2013). Target protein for Xklp2 (TPX2), a microtubule-related protein, contributes to malignant phenotype in bladder carcinoma. *Tumour Biology: The Journal of the International Society for Oncodevelopmental Biology and Medicine*, *34*(6), 4089–4100.

Yoon, K.-J., Koo, B.-K., Im, S.-K., Jeong, H.-W., Ghim, J., Kwon, M.-C., Moon, J.-S., Miyata, T., & Kong, Y.-Y. (2008). Mind bomb 1-expressing intermediate progenitors generate notch signaling to maintain radial glial cells. *Neuron*, *58*(4), 519–531.

Ystrom, E., Gustavson, K., Brandlistuen, R., Knudsen, G., Magnus, P., Susser, E., Smith, G., Stoltenberg, C., Surén, P. al, H\r aberg, S. E., Hornig, M., Lipkin, I. W., Nordeng, H., & Ted, R.-K. (2017). Prenatal Exposure to Acetaminophen and Risk of ADHD. *Pediatrics*, *140*(5), e20163840.

Yuan, F., Fang, K.-H., Cao, S.-Y., Qu, Z.-Y., Li, Q., Krencik, R., Xu, M., Bhattacharyya, A., Su, Y.-W., Zhu, D.-Y., & Liu, Y. (2015). Efficient generation of region-specific forebrain neurons from human pluripotent stem cells under highly defined condition. *Scientific Reports*, *5*, 18550.

Yum, S. W., Zhang, J., Mo, K., Li, J., & Scherer, S. S. (2009). A novel recessive Nefl mutation causes a severe, early-onset axonal neuropathy. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *66*(6), 759–770.

Yu, P. B., Deng, D. Y., Lai, C. S., Hong, C. C., Cuny, G. D., Bouxsein, M. L., Hong, D. W., McManus, P. M., Katagiri, T., Sachidanandan, C., Kamiya, N., Fukuda, T., Mishina, Y., Peterson, R. T., & Bloch, K. D. (2008). BMP type I receptor inhibition reduces heterotopic [corrected] ossification. *Nature Medicine*, *14*(12), 1363–1369.

Zablotsky, B., Black, L. I., Maenner, M. J., Schieve, L. A., Danielson, M. L., Bitsko, R. H., Blumberg, S. J., Kogan, M. D., & Boyle, C. A. (2019). Prevalence and Trends of Developmental Disabilities among Children in the United States: 2009-2017. *Pediatrics*, *144*(4). https://doi.org/10.1542/peds.2019-0811

Zeng, J., & Yu, H. (2018). A Scalable Distributed Louvain Algorithm for Large-Scale Graph Community Detection. *2018 IEEE International Conference on Cluster Computing (CLUSTER)*. https://doi.org/10.1109/cluster.2018.00044

Zeng, Y., Kurokawa, Y., Win-Shwe, T.-T., Zeng, Q., Hirano, S., Zhang, Z., & Sone, H. (2016). Effects of PAMAM dendrimers with various surface functional groups and multiple generations on cytotoxicity and neuronal differentiation using human neural progenitor cells. *The Journal of Toxicological Sciences*, *41*(3), 351–370.

Zhang, S., Bell, E., Zhi, H., Brown, S., Imran, S. A. M., Azuara, V., & Cui, W. (2019). OCT4 and PAX6

determine the dual function of SOX2 in human ESCs as a key pluripotent or neural factor. *Stem Cell Research & Therapy*, *10*(1), 122.

Zhang, S., Li, J., Lea, R., Vleminckx, K., & Amaya, E. (2014). Fezf2 promotes neuronal differentiation through localised activation of Wnt/β-catenin signalling during forebrain development. *Development*, *141*(24), 4794–4805.

Zhao, X., Kuja-Panula, J., Rouhiainen, A., Chen, Y.-C., Panula, P., & Rauvala, H. (2011). High mobility group box-1 (HMGB1; amphoterin) is required for zebrafish brain development. *The Journal of Biological Chemistry*, *286*(26), 23200–23213.

Zhou, C., Yang, X., Sun, Y., Yu, H., Zhang, Y., & Jin, Y. (2016). Comprehensive profiling reveals mechanisms of SOX2-mediated cell fate specification in human ESCs and NPCs. *Cell Research*, *26*(2), 171–189.

Ziller, M. J., Edri, R., Yaffe, Y., Donaghey, J., Pop, R., Mallard, W., Issner, R., Gifford, C. A., Goren, A., Xing, J., Gu, H., Cacchiarelli, D., Tsankov, A. M., Epstein, C., Rinn, J. L., Mikkelsen, T. S., Kohlbacher, O., Gnirke, A., Bernstein, B. E., … Meissner, A. (2014). Dissecting neural differentiation regulatory networks through epigenetic footprinting. *Nature*, *518*(7539), 355–359.