Differentiation potential, lineage commitment and gene expression profile of human cortical neural progenitor cells derived from pluripotent stem cells

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

Von

Zahra Ehsaei

aus, Iran

Basel, 2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen

Fakultät auf Antrag von

Prof. Verdon Taylor

Prof. Raphael Guzman

Basel, 18.09.2018

Prof. Dr. Martin Spiess, Dekan

Contents

Summary

The human cerebral cortex is composed of the variety of neurons and glial cells that are organized into six different layers. During development, this complex structure originates from a simple neuroepithelium. As neurogenesis continues the neural stem and progenitor cells residing in the ventricular zone (VZ) and subventricular zone (SVZ), generate cortical projection neurons in an inside-out order. The neurons within each layer have particular functions, gene expression patterns and morphologies. These neurons are born at defined stages during development. The developmental mechanisms that regulate neural progenitor fate specification during cerebral cortex development remained elusive. Stem cell based systems allow at least partial recapitulation of the important aspects of human cortical neurogenesis in a simple and accessible cell culture manner. These systems have been successfully used to understand specific mechanisms associated with human cortical development and disorders.

Here, I studied the fate potential of human cortical neural progenitors derived from 2-dimensional (2D) *in vitro* corticogenesis. I also investigated how neural stem cells/progenitors generate the great diversity of neurons during *in vitro* cortical differentiation. I propose that fate potential of the neural progenitor (NP) pool changes during human cortical development. I employed retrovirus birth dating to investigate timing of neurogenesis. I found that not only is the early progenitor pool multipotent and generates both deep and upper layer neurons but also the late progenitors are capable to give rise to deep and upper layer neurons. It has been suggested that both intrinsic and extrinsic factors mediate fate specification of neural progenitors during development. I first tried to find the transcriptional program, which regulates the competence of NPCs during differentiation. I isolated the NPs based on $CD184^{+}$, $CD24^{+}$, $CD271^{-}$, $CD44^{-}$ expression at sequential stages during *in vitro* corticogenesis and studied the transcriptional profile of these cells. The analysis confirmed the dynamic transcriptional program of NPs over the course of differentiation. Late sorted NPs give rise to more upper layer neurons

and surprisingly to a high proportion of deep layer neurons in comparison to early progenitors. Moreover, NPs (isolated at early and late stages of differentiation) when co-cultured with the cells from different differentiation stage alter the proportions and types of generated neurons in response to environmental signals.

This thesis is organized in five chapters. The first chapter provides an introduction to early human cerebral cortex development as well as a summary of common strategies for *in vitro* cortical differentiation derived from human pluripotent stem cells (hPSCs). The chapter ends with the aims of the project. The second chapter contains the main results that are presented as a research manuscript. **"Human pluripotent stem cell derived neural progenitors display two modes of neural fate determination"**. Chapter two also includes my collaboration in a research project in our lab entitled **"Neurodegeneration associated TDP-43 induces p53-mediated cell death of stem cells and neurons"**

In chapter two I also present my result during collaboration in the project "Multigene delivery in primary and stem cells". The two papers entitled **"Highly efficient baculovirus-mediated multigene"** and "**Baculovirus-based genome editing in primary cells"** are attached as an appendix. In Chapter three and four, I summarize and discussed our results and refer to the limitations of our system. Chapter five contains the detailed methods and protocols that have been used in the thesis.

CHAPTER 1
Introduction

This chapter aims to provide background information to the reader about how the early brain is formed and, in more detail how the cerebral cortex develops during embryonic and early fetal periods. Furthermore, the approaches to recapitulate key aspects of neurogenesis *in vitro* by the use of pluripotent stem cells is introduced in the form of review paper published in the Neuropsychiatry. I wrote the text and prepared the figures.

"What is perhaps the most intriguing question of all is whether the brain is powerful enough to solve the problem of its own creation."

Gregor Eichele (1992)

1-1 Introduction

The adult human brain is one of the most fascinating structures and the most complex organ in the human body. It is made up of approximately 86 billion of information processing cells, called neurons and a similar number of signal regulating and supporting cells, called glia cells[1,2]. Neurons of the brain coordinate body sensation, movement and function, and are the executive unit and basis of cognition. In vertebrates the brain consists of six major structures: the medulla oblongata, pons, cerebellum, midbrain, diencephalon and cerebrum[3].

The development of the human brain is a prolong process that starts from the third week of gestation and continues throughout life[4,5]. Prenatal development is divided into two periods: the embryonic and the fetal period. In these periods the gross anatomy of the CNS develops and the majority of neurons in the cerebral cortex is generated[4,5]. However, the generation of some neurons continues after birth in certain brain regions such as in the striatum and the dentate gyrus.

The cerebral cortex is the thin highly neuron rich layer of the brain that covers the outer portion of the cerebrum. The cerebral cortex is a highly organized structure playing a key role in memory, cognition, perception, language and other fundamental processes. It is composed of different cell types including neurons and glia cells. Cortical neurons are classified into two sub-types: glutamatergic projection neurons (80%) [2,6] and are GABAergic interneurons (20%)[7,8].

In this chapter I outlined the major process of cerebral cortex development that occur during embryonic and early fetal periods.

1-2 Brain development in the embryonic period

The primary structures of the brain and the central nervous system are established by the end of embryonic period. The embryonic period in human, from conception till GW8, is an important period in neocortex development. The fundamental processes that occur in the

embryonic period include neural tube formation, formation of primary vesicles and secondary vesicles[4].

1-2-1 Neural tube formation

During early embryogenesis the blastocyst is composed of a trophoblast and the ICM (inner cell mass). Oct4, Sox2 and Nanog expression is required to maintain pluripotency of the ICM[9]. During development a paracrine milieu exposed to the cells will change the ratio of expression of these genes and ICM cells start to differentiate into cells of the different germ layers[10].

By the end of the third week after gestation the embryo transforms from a two-layered structure into a three-layered during the gastrulation process. Gastrulation begins with the appearance of a slit-like opening, the primitive streak, in the upper layer of the embryo at embryonic day 13[4,10]. The upper cell layer the so-called epiblast differentiates into three germ layers (mesoderm, ectoderm, endoderm) that will eventually generate all cells of the embryo. OCT4 expression alone, Without SOX2, induces the differentiation of cells towards mesoderm and endoderm lineages. The endoderm the most inner layer gives rise to the gut, lungs, and liver. The mesoderm, the middle layer, gives rise to connective tissue, muscles, bones and the vascular system. However, cells that express SOX2 only, give rise to the ectoderm by activating genes responsible for the synthesis retinoic acid and FGFs. The ectoderm, the outer layer, gives rise to the central and peripheral nervous system[10].

During and after gastrulation, the neuroectoderm is formed from the dorsal-most part of the ectoderm during the process of neurulation. Neural induction is initiated by inhibition of suppressive signals such as BMP inhibitors (Noggin, Chordin)[11-13]. Several studies showed that inhibition of Activin and Nodal pathways also play an important role in neural induction[14]. Additionally, inhibition of the Wnt pathway has also been shown to be crucial for neural induction[14] (Figure 1A). It has been demonstrated that the vertebrate ectoderm gives rise to three major domains based on the level of BMP exposure. Cells that are exposed to high levels of BMP become the surface ectoderm (primarily epidermis). Cells, which are exposed to mediate levels of BMP, differentiate into neural crest cell lineages, which give rise to the peripheral nervous system, bones and muscles of the face, pigment cells and chromaffine cells. Finally, cells which are exposed to very low levels of BMP-signals convert to the cells of the neural plate, which further develops into the central nervous system[10]. Neural stem cells (progenitors) within the neural plate express the sox family transcription factors Sox1, Sox2 and

Sox3, which have been shown to activate genes for neural plate formation. Additionally they inhibit the formation of epidermis and neural crest by inhibiting the transcription and signaling of BMPs[15].

During the third week of human embryonic development the neural plate folds into a tubular structure and forms neural tube along the dorsal side of the embryo. The neural tube eventually develops into brain and spinal cord.

1-2-2 Patterning of neural tube

Soon after the neural tube starts closing it shows regionalization along anterior-posterior (A-P) and dorsal- ventral (D-V) axes. The patterning of the neural tube is orchestrated by morphogen gradients. FGFs, Wnts and RA specify the A-P patterning. During the forth week of embryonic development the anterior region of neural tube expands and differentiates into three primary brain vesicles. The posterior region of neural tube gives rise to the spinal cord[3,16]. The primary vesicles are the anterior lying prosencephalon (future forebrain), the mesencephalon (future midbrain) and the rhombencephalon (future hindbrain)[3,16]. In the absence of morphogenes or in the presence of FGF and Wnts inhibitors respectively, in the most anterior part of neural tube, this anterior domain becomes specified into forebrain tissue. Wnt1 and Fgf8 are produced from the cells in the midbrain-hindbrain boundary and specify the midbrain and hindbrain identity.

The D-V axis of the neural tube is determined by Wnt, BMP and SHH (sonic hedgehog) signaling pathways[16-18]. D-V patterning is majorly established by SHH as being responsible for ventral induction and antagonizing Wnt signalling and BMP signaling as dorsalization factor. SHH is secreted from the notochord and floor plate and plays a prominent role in ventral patterning of the telencephalon, whereas BMPs and Wnts derived from roof plate induce dorsal fates.

Later in development the primary vesicles divide further to specify different parts of the future brain. The prosencephalon gives rise to the telencephalon and diencephalon, and the hindbrain vesicle subdivides into the metencephalon and myelencephalon[4,16]. The mesencephalon, which gives rise to the future midbrain does not divide further. The mammalian telencephalon has two main subdivision: the dorsal (pallium) and the ventral (subpalium) telencephalon. The cerebral cortex originates from the dorsal part the pallium, whereas the subpalium gives rise to the basal ganglia[16]. The neocortex is the largest part of cerebral cortex

and is considered as being the prominent information-processing network of the brain[19] (Figure 1).

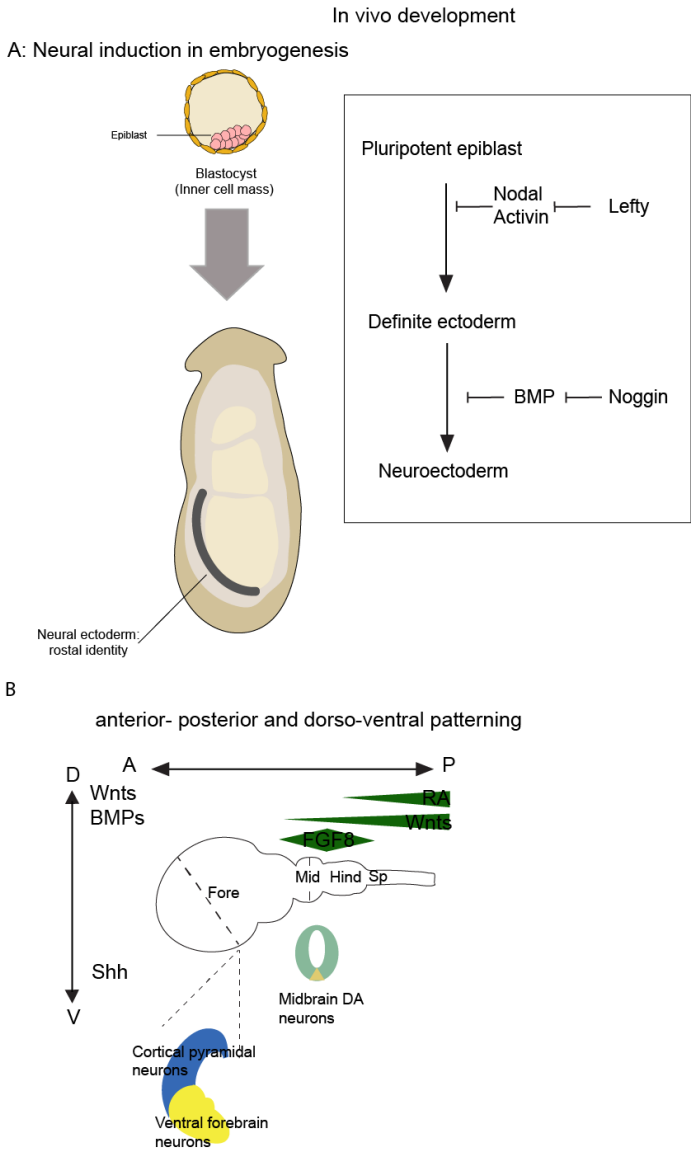


Figure 1:

A: Neural induction occurs by inhibition of BMP/Nodal signaling in vivo, lead to generation of neural ectoderm with anterior identity.
B: The patterning within central nervous system. The neural tube is subsequently patterned by extrinsic morphogens along the anterior- posterior and dorso-ventral axes in to different domain (modified from Ikuo K. Suzuki, 2015 and Nicolas Gaspard, 2010)

1-3 Brain development in the fetal stage

The fetal stage starts from the ninth gestational week until the end of gestation. The major events in the fetal period are the generation of neurons, their migration and differentiation. These events over the course of fetal periods result in 40 fold increase in the weight of brain[20].

1-3-1 Neurons production

The vast majority of the human brain neurons are generated by mid-gestation[21,22]. The layer of neuroepithelial cells that lies in the wall of the neural tube form the ventricular zone (VZ). These neuroepithelial cells are the neural stem/progenitor cells that generate all types of neurons and macroglia (astrocytes and oligodedrocytes) of the future CNS[23]. The pool of neural progenitor cells of the VZ at the end of gastrulation is not big enough to ensure for the production of billions neurons in human brain[4]. Therefore, the neural progenitor pool is being expanded by symmetric cell division starting from the end of gastrulation until E42 in human. This expansion of the neuronal stem cell pool leads to the production of two identical neural progenitor cells[4]. Thereafter, another proliferative zone called subventricular zone (SVZ) appears above the VZ. The SVZ continues expanding during early and mid fetal stages of development[19]. During early development the NE cells in ventricular zone divide, expand and give rise to radial glia (RG) cells another type of neural progenitor cells. RG cells, which reside in the VZ are called apical or inner RG cells (iRG), whereas the basal or outer RG cells (ORG) populate the SVZ[7,19]. From E51 onwards the RG and ORG division mode switches from symmetric to asymmetric cell division leading to the production of neurons and intermediate progenitor cells (IPCs). In contrast to the intermediate progenitors, which remain in the proliferative zone and continue to divide newly generated post mitotic neurons migrate out of the proliferate zone and reside in layers above the SVZ. The massive amplification of ORG cells in SVZ leads to a subdivision of the SVZ to iSVZ and oSVZ in mammalian such as human and macaque. In contrast in mice and other rodents the ORG cells are an infrequent population within germinal zone[19,24].

In humans cortical neurogenesis is finished approximately by E191[25]. Glial cell differentiation, that follows neurogenesis is peaking around birth and continues also after birth[23].

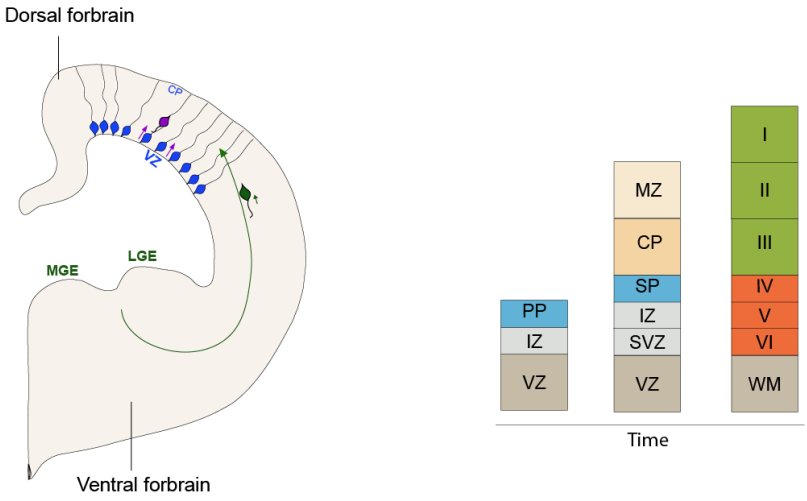
1-3-2 Neuron migration and differentiation

Cell migration is an important step in tissue formation during development. The excitatory projection neurons are generated from progenitors in the dorsal telencephalon and migrate radially into the cortex. These neurons use the radial glial fiber to migrate out of the VZ and SVZ towards the developing cortical plate. For this type of migration, the radial glial fibers provide a scaffold for the radially migrating neurons[24,26,27]. In contrast, GABAergic interneurons are produced in domains of the ventral telencephalon (ganglionic eminence) and migrate towards cortex by tangential migration[28,29] (figure 2A).

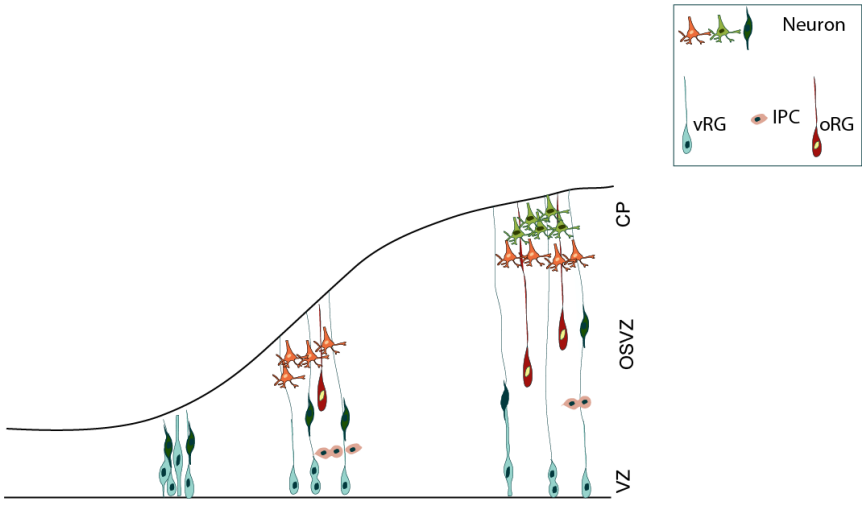
1-3-3 Layering of neocortex

Before onset of neurogenesis, a population of early born pioneer neurons migrate from the outside of the neocortocal primordium and form the early marginal zone also called preplate (PP) right above the VZ. Then, newly born projection neurons migrate radially from the VZ to form the cortical plate (CP) between the MZ and SP. MZ and SP have important role in fetal brain development. However, major parts of these regions disappear by end of the fetal period. The Reelin producing Cajal-Retzius cells (CR), which represent the main cell population within the MZ play an important role in the positioning of newborn radially migrating neurons from the VZ/SVZ into the different cortical layers above. Reelin produced by CR cells is thought to instruct neurons to migrate into the layers above the VZ/SVZ and to find their correct position. Thereby, forming the different layers of the cerebral cortex[30,31]. The first neurons, which populate the CP are the cortical neurons of layer VI, subsequently later born upper layer neurons migrate towards the CP[32]. This migration of neurons leads to the formation of a 6-layered developing neocortex.

Next, glutamatergic projection neurons extend long distance axons to other intracortical, subcortical or subcerebral structures. These neurons are categorized based on their morphology, their position in a specific layer of cortex, their transcriptional pattern and their function. Layer VI is composed of corticothalamic neurons, which project to the thalamus. subcerebral projection neurons reside in layer V and project to spinal cord, spinal cord, pons and superior colliculus. The upper layers neurons extend to intracortical projections, as the layer II and III neurons connect two cerebral hemispheres through projection across the corpus callosum[32,33] (figure 2B,C).



A: Different mode of neural migration ro the neocortex B: layering of neocortex thruoghout time.

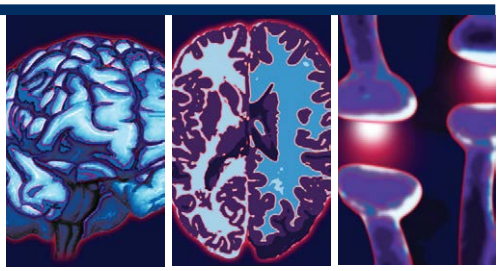


C: cerebral cortex development

Figure 2:

A: Different way of neural migration to the neocortex: Glutamatergic projection neurons (purple neurons) migrate from VZ to cortical plate through the radial glial fiber. Inhibitory neurons (green neurons) migrate from ganglionic eminences by tangential migration.
B: layering of neocortex over time. Early born pioneer neurons migrate from VZ and form preplate (PP) above the VZ. Then, newly born projection neurons populate the cortical plate (CP) between the MZ and SP. The early-generated neurons populate the deeper layers of neocortex while the later born neurons position within more superficial layers.
C: Schematic illustration of the human neocortical development. Cerebral cortex develops from neural stem cells in the VZ and SVZ. The neural stem cells give rise to different subtypes of neurons in inside-out order. (Modified from Stiles J, 2010)

Review



Pluripotent Stem Cell Based Cultures to Study Key Aspects of Human Cerebral Cortex Development

Zahra Ehsaei[1], Ginetta Collo[1,2], Verdon Taylor[†]

ABSTRACT

The human brain is a highly organized structure and the cerebral cortex in particular has expanded massively in size during evolution. The cerebral cortex is arranged into layers of specialized neuron subtypes formed during development by orchestrated stem cell maintenance, expansion, fate commitment and differentiation. The cortical neural stem cells generate billions of neurons in a systematic fashion. The mechanisms and their interplay that control most aspects of human brain development are unclear. This is partially due to the ethical and practical challenges associated with analyzing fetal human development. Recent progress into understanding the formation of the human brain has taken advantage of *in vitro* modeling of corticogenesis using pluripotent cells. Human pluripotent stem cells and procedures developed for their differentiation provided previously unavailable opportunities to study the mechanisms involved in development of the cerebral cortex. These human cell culture models can be applied to address specific biological questions and have been successfully utilized to investigate mechanisms associated, not only with normal brain development, but also neuropsychiatric disorders. Here, we review the recent literature that uses these cell culture models to study human corticogenesis. Then, we discuss the challenges and limitations of the current models.

Keywords

Neurogenesis, Neurons, Embryonic stems cells, Induced pluripotent stem cells, iPSC, Human cortex, Differentiation

Introduction

The cerebral cortex of mammals is a unique and complex structure composed of diverse neural cell types organized into precise networks [1]. Despite its complexity, the cerebral cortex originates from a simple anlage consisting of a polarized sheet of pseudostratified epithelial cells, the neural epithelium [2]. The mechanisms controlling cerebral cortex development have been of major interests for neurobiologists, and have been intensively studied over the past decades. Decoding the key aspects of human cerebral cortex development could enable new therapeutic approaches to be developed for the prevention and treatment of devastating brain disorders.

Animal models have provided significant insights into the development, structure and function of the brain and about the mechanisms leading to neurological disorders [3]. Despite the basic similarities between cerebral cortex development in different species, the human neocortex has some key and unique cellular and molecular characteristics that cannot be captured by classic animal models including an expanded outer sub ventricular zone (OSVZ), specific progenitors and an increased number of neurons [4,5].

[1]Department of Biomedicine, University of Basel, Mattenstrasse 28, CH-4058, Basel, Switzerland

[2]Department of Molecular and Translational Medicine, University of Brescia, Viale Europa 11, 25123, Brescia, Italy

[†]Author for correspondence: Verdon Taylor, Department of Biomedicine, University of BaseL, Mattenstrasse 28, CH-4058, Basel, Switzerland, email: verdon.taylor@unibas.ch verdon.taylor@unibas.ch

Over the last years, there has been a huge effort to find ways to study human brain development. Significant progress has been made towards decoding the mechanisms controlling the generation of the human cerebral cortex by using embryonic tissues from aborted fetuses [4]. However, when using fetal-derived tissue one is faced with many challenges including difficulties in accessibility and ethical issues. Hence, the wide application of human fetal tissues for research is rather restricted. The advent of human pluripotent stem cells, embryonic stem cells (ESCs) and induced pluripotent cells (iPSCs), opened new opportunities to study organogenesis. Human pluripotent cells are useful for studying developmental processes and also disease mechanism and for drug screening to find new treatment strategies [6].

In this review, we outline the basic process of cerebral cortex development in humans. Then, we will review the different strategies currently available for studying the mechanism of human cortical differentiation from pluripotent stem cells with a particular focus on the formation of the six-layered isocortex with its different subtypes of excitatory neurons. Finally, we discuss the advantages and limitations of the current protocols.

Mechanisms of Cerebral Cortex Development

The cerebral cortex is organized into layers forming an isocortex. The laminar cytoarchitecture of the cerebral cortex is preserved among mammals. The six different layers (layer I-VI) contain different subtypes of neurons with specific functions, connections and gene expression **(Figure 1)**. The neuron types can be simply divided into glutamatergic excitatory projection neurons that transmit information over longer distances and GABAergic inhibitory interneurons that

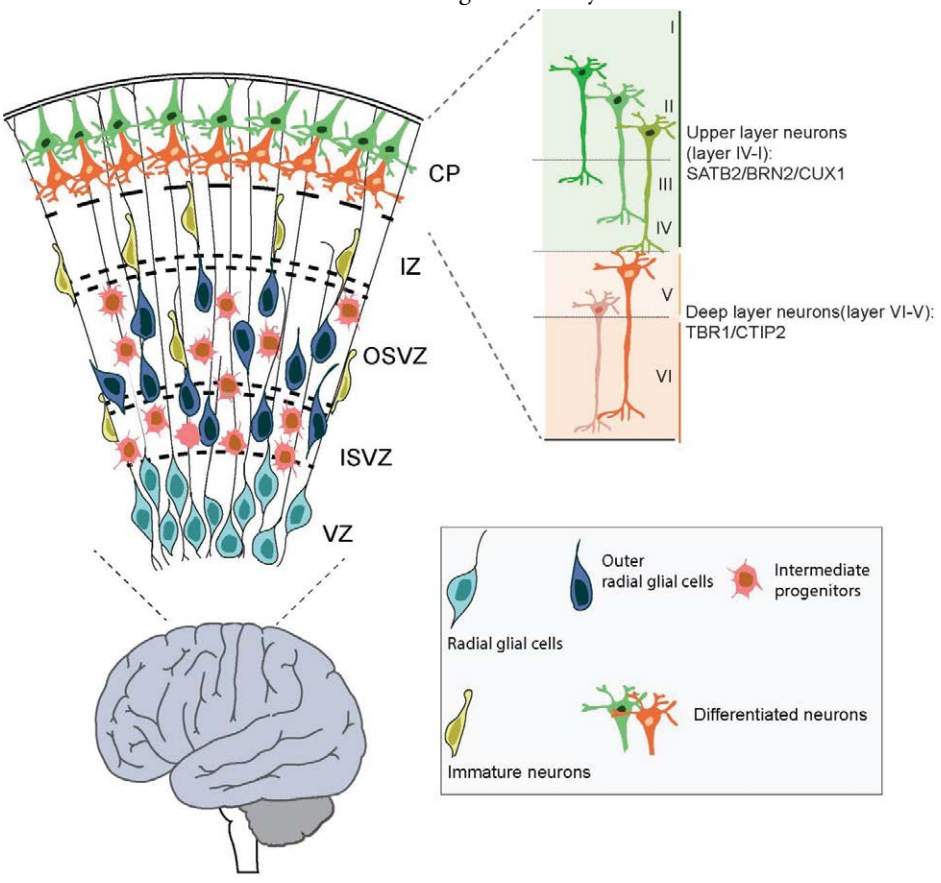


Figure 1: Human cerebral cortex development.
During human cortical development, neurons in the cerebral cortex are generated sequentially in an inside-out order from stem cells and progenitors residing in the ventricular zone (VZ) and subventricular zone (SVZ) of the neural tube [3]. In humans and other primates, the subventricular zone is expanded and divided into the inner subventricular zone (ISVZ) and outer subventricular zone (OSVZ) [5]. The cerebral cortex is organized into six layers (layer I-VI). Projection neurons in each layer express distinct markers and have different morphologies and connection patterns. Deep layers (VI and V) are composed of TBR1+ and CTIP2+ corticothalamic projection and subcortical projection neurons, respectively. Superficial layers contain SATB2+, CUX1+, BRN2+ callosal projection neurons [1]. Intermediate zone (IZ), cortical plate (CP).

modulate local information flux [3,7]. However, within this basic segregation, neurons are divided into subtypes. For example, deep cortical layers (VI and V) contain mainly TBR1$^+$ and CTIP2$^+$ corticothalamic projection neurons and subcortical projection neurons, which project and carry information from the cerebral cortex to subcortical structures including the thalamus. The more superficial layers (IV-II) contain SATB2$^+$, CUX1$^+$, BRN2$^+$ callosal projection neurons that project to contralateral brain regions thereby transmitting information from one brain hemisphere to the other [1].

During development, the excitatory neurons are produced in an inside-out laminar order from neural stem cells and progenitors that reside in proliferative zones, the ventricular zone (VZ) and subventricular zone (SVZ) of dorsal telencephalic region of the neural tube [8,9]. The deep layer neurons are generated at early stages of neurogenesis while the late-born neurons populate the upper layers. Conversely, the interneurons are generated from progenitors in the ventral telencephalon and migrate into the forming cerebral cortex [10]. Astrocytes and oligodendrocytes are produced at the end of embryogenesis and after birth from both the dorsal and ventral telencephalic regions [11]. In this review our main focus will be on the generation of excitatory cortical neurons.

Early during cortical development, neuroepithelial (NE) cells form the walls of the neural tube. As neurogenesis commences, VZ NEs transform into radial glial cells, which become the major cortical progenitors in the VZ. Through symmetric and asymmetric cell divisions, radial glia expands the progenitor pool and generate neurons. During the course of cortical neurogenesis, radial glia generates intermediate progenitor populations that expand the precursor pool [3]. The length of cortical neurogenesis differs between species. In humans, cortical neurogenesis is protracted, beginning at gestation week (GW) 5 and ending around gestation week (GW) 30 [12]. In mouse, cortical neurogenesis starts at embryonic day 11 and ends around embryonic day 18, shortly before birth [12].

The basal progenitors are the major type of progenitors in SVZ of rodents. If they divide, they mostly divide symmetrically and give rise to two neurons [3,5]. In humans and other primates, the SVZ is massively expanded compared to in rodents and becomes the major source of cortical neuron progenitors. The human SVZ is so large that it is divided in two regions, the inner subventricular zone (ISVZ) and outer subventricular zone (OSVZ) and contains additional neural progenitor cell types, the outer radial glial cells (ORG) [5]. The expanded progenitor pools and their extended proliferation phases contribute to the increased number of neurons and subsequent folding and gyrification of the primate cerebral cortex [4,13,14].

Pluripotent Stem Cells: Application and Promise

For decades, pluripotent stem cells have been available for mice and have been a major driving force to study gene function by knockout [15,16]. Pluripotent cells can generate all cell-types of the embryo. However, in other species, the isolation of pluripotent cells has been a challenge. In 1998, the first ESCs were generated from the inner cell mass of an early human embryo paving the way for the generation of specialized human cells and tissues in the laboratory and analysis of the mechanisms controlling early development in human [17]. Due to their high telomerase activity, ESCs are able to divide unlimited times and therefore present an unlimited source of human cells [17].

However, when using human ESCs, one faces major challenges, not least ethical problems as the cells are isolated form human embryos. In 2007, technology for the generation of so-called iPSCs (induced pluripotent stem cells) from human skin fibroblasts was developed [18]. Forced expression of a combination of four transcription factors (Oct3/4, Sox2, Klf4, c-Myc) drives somatic human cells back in their development to an ESC-like state. These four transcription factors play fundamental roles in maintaining the characteristics of pluripotent stem cells by suppressing the genes required for differentiation, and activating genes that promote the pluripotent state [19]. iPSCs are also capable of renewing and presumably differentiating to all cell-types of the body [18,20]. iPSCs share similar gene expression, telomerase activity and epigenetic status of pluripotency genes with ESCs [18,20]. The somatic origin of iPSCs circumvents many of the ethical issues limiting ESC and fetal tissue-based research. Pluripotent cell technology has massively increased the potential for basic research into human biology, regenerative medicine, disease modeling, drug

discovery, and stem cell-based therapy [17]. Moreover, iPSC technology enables disease-relevant cell types to be generated from patients and these can be applied to investigate disease etiology and pathological mechanism. Advances in gene editing techniques, including TALEN [21,22], Zinc finger [23] and CRISPR [24-27] provide new tools to manipulate pluripotent cells and address mechanisms involved in brain development and disease. The gene editing tools commonly applied to generate mutated iPSCs and ESCs can also be used to repair the mutations in these cells and their progeny [6].

Different Strategies for *In vitro* Differentiation of Cortical Neurons

Since the isolation of the first human pluripotent cells, protocols have been developed and continually improved to generate robust culture system that, at least in part, recapitulate *in vivo* corticogenesis [28]. Cortical neurogenesis is controlled by a combination of intrinsic and extrinsic signals [29]. Correspondingly, *in vitro* studies have aimed to mimic these extrinsic signals by the addition of cocktails of growth factors and small molecules to the culture medium, promoting cell type specification and differentiation in a temporal fashion [30]. Here, we will summarize recent progress towards the generation of cortical excitatory neurons from pluripotent stem cells **(Figure 2)**. Basically, the differentiation methods can be grouped into two categories: Adherent culture systems for differentiation **(Table 1)** and differentiation from 3D complex culture structures **(Table 2)**.

Adherent Pluripotent Stem Cell Cortical Differentiation Protocols

Pluripotent cells resemble the most primitive cells of the blastocyst. The first step in the adherent cortical neuronal differentiation protocols is to push the pluripotent stem cells towards the ectodermal/neuroectodermal lineages. This step mimics the lineage commitment steps of gastrulation during early development but restricts the fate of the pluripotent cells away from mesoderm and endoderm and favors ectodermal differentiation. Subsequently, ectodermal cells are promoted to adopt neuroectodermal stem cell fates similar to the process of neurulation. Pluripotent cell-derived neuroectodermal cells form NEs that self-organize into rosette-like structures. Rosettes resemble the cellular organization of the early neural tube. Following expansion of NEs, switching the culture conditions and cocktail of factors promotes differentiation to definitive neural cell-types and specific neuron populations. Early adherent cortical differentiation protocols relied on autocrine fibroblast growth factor (FGF) signaling to promote the differentiation of neuroectodermal progenitors that later produced neurons and glial cells [31].

More recently, the differentiation of pluripotent stem cells to dorsal cortical neural fates was enhanced by culturing the cells at reduced cell-density in the presence of cyclopamine, an antagonist of the sonic hedgehog pathway [32]. Sonic hedgehog promotes ventral cell fates in the developing [19] progenitor differentiation. One significant advance due to this protocol was that long-term culture of the dorsal forebrain progenitors led to the generation of both deep and upper layer neurons in a sequential fashion, similar to cortical differentiation *in vivo* [3,33]. However, the method favored early neuron subtypes and upper layer neuron differentiation was limited [34].

Subsequently, the initial dorsal neural fate induction step was improved for human pluripotent stem cells by using a combination of BMP and TGF-β receptor inhibitors, Noggin (or the Alk2 and Alk3 BMP receptor inhibitor LDN193189) and SB431542 (an Alk5 TGF-β/Activin receptor inhibitor), during early stage of differentiation [35]. This also resulted in a more homogeneous and rapid neural conversion. As both TGF-β and BMP signaling pathways act through downstream SMAD transcriptional effectors [36], the protocol is often referred to as the dual-SMAD inhibitor procedure. Inhibition of TGF-β and BMP signaling increases the efficiency of neuroectodermal fate differentiation by preventing endodermal and mesodermal differentiation [35]. Accordingly, many adherent culture protocols now use dual TGF-β and BMP signaling inhibitors in order to generate neuronal subtypes without the need to pass through embryoid body (EB) formation or culturing the pluripotent cells in the presence of feeder or stromal cells [11,37].

Retinoic acid, a metabolite of vitamin A, has long been known to promote neural differentiation [38]. Hence, logical progression of the dorsal cortical neuron differentiation procedure from pluripotent cells saw the combination of dual TGF-β and BMP signaling inhibition and

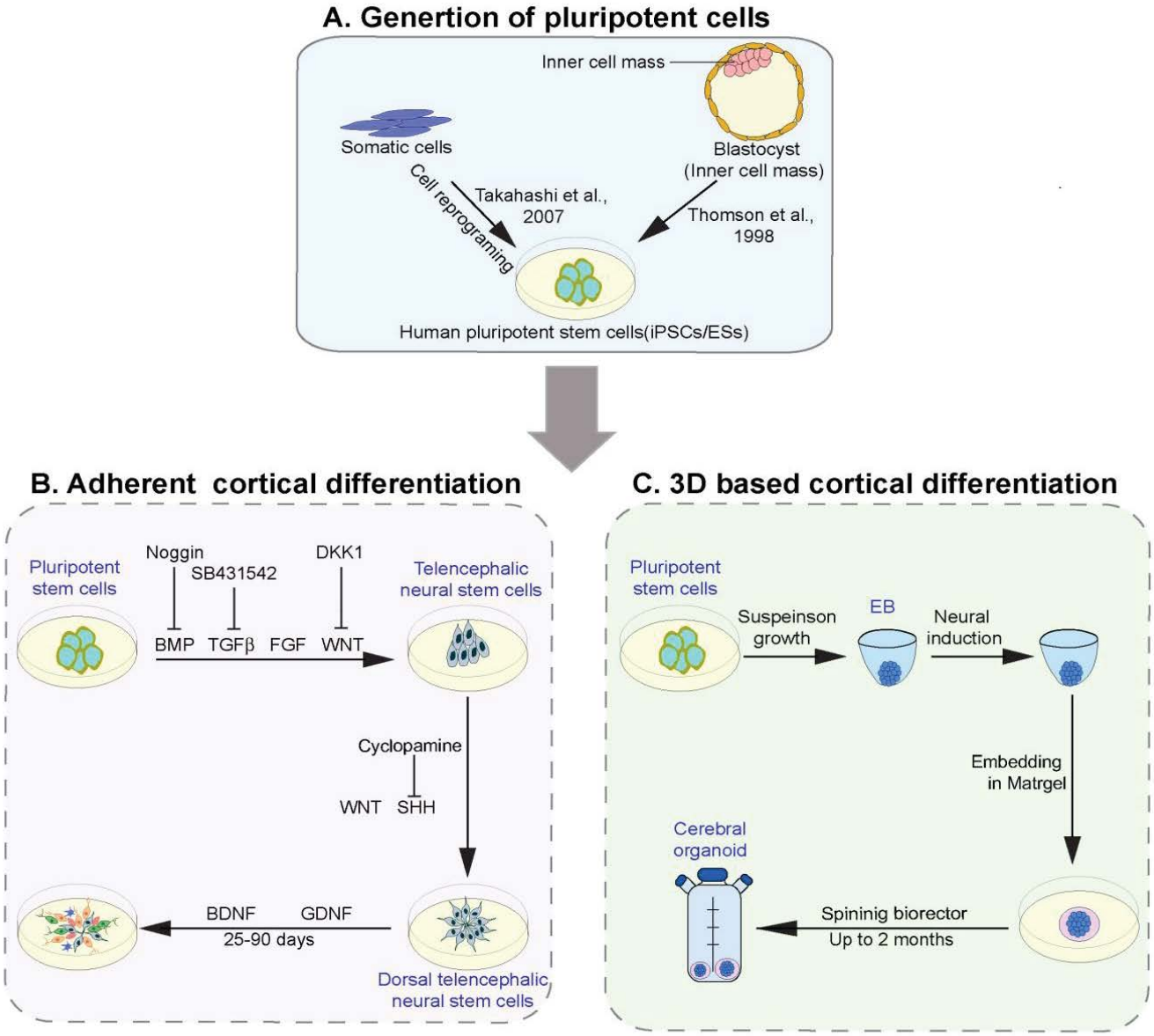


Figure 2: Strategies of *in vitro* cortical neurons differentiation derived from pluripotent stem cells (PSCs).

A) Two major sources of pluripotent stem cells: ESCs are derived from blastocyst, a week old embryos, and iPSCs derived from cell reprograming of somatic cells. B) Adherent differentiation of pluripotent stem cells to cortical neurons: Blockade of TGFβ and BMP signaling using chemical inhibitors (SB, LDN) trigger the differentiation of pluripotent stem cells to anterior neural stem cells [35]. Inhibition of WNT pathway also increases the efficiency of telencephalic neural progenitors in expense of neural crest cells [28]. The endogenous FGF signaling promotes neural progenitors proliferation in this process [39]. The dorsal telencephalic fate is induced by Wnt stimulation, blockade of Shh or through intrinsic default mechanisms. It has been reported the Shh inhibition by adding cyclopamine increase the efficiency of differentiation to dorsal forebrain progenitors [34]. RA also has been considered important in differentiation of radial glial cells to neurons [38]. Growing cells in the differentiation media containing neurotrophic factors such as GDNF and BDNF enhanced maturation and survival of derived cells [28]. C) The 3D cortical neuron differentiation start with EB formation. The EBs transferred to neural induction medium [48]. This medium can be supplemented with external factors such as dual SMAD inhibitors. Then the deriving cell aggregates embedded in to Matrigel and transfer to spinning bioreactor and culture in differentiation media [46]. ESCs (embryonic stem cells), induced pluripotent stem cells (iPSCs), retinoic acid (RA), embryoid body (EB).

treatment with retinoic acid [11]. Growing cells at high density under these conditions increased the efficiency of differentiation of mouse and human pluripotent stem cells into cortical neural progenitors [11]. When these dorsal neural progenitors' cells are differentiated in the presence of brain-derived neurotrophic factor and glial-derived neurotrophic factor over 90 days, deep and upper layer neurons are generated in sequential order. Moreover, the neurons that are generated by this procedure acquire electrophysiological activity during the

Table 1: Comparison of the adherent protocols to generate cortical neurons.

Cell source	Culture condition	Outcome	Pros and Cons	Author
Mouse ESCs	-Serum free medium without inductive signals and growth factors -Autocrine FGF	Neuroectodermal progenitors → neurons and glial cells	+First differentiation strategy without EB formation or feeder cells -Unclear which cortical neuron subtypes generated -The progenitors were able to gave rise to TH^+ and GABAergic neurons after cultured in defined condition	Ying et al. [31]
Mouse ESCs	-Cells cultured at low density -Cyclopamine added as SHH inhibitor -Differentiation media supplemented with N2/B27	Dorsal forebrain progenitors → different cortical neurons subtypes	+Sequential generation of deep and upper layer neurons -Low number of upper layer neurons	Gaspard et al. [34]
Human ESCs/iPSCs	-Cells cultured at early stages of differentiation in presence of Noggin and SB431542	Neural rosette	+Rapid and efficient neural induction lead to neural rosette formation - Used to generate midbrain dopamine and spinal motoneurons	Chambers et al. [35]
Human ESCs/iPSCs	-Cells cultured at high density in presence of Noggin and SB431542 at early stage -Growing media contain retinoic acid -Differentiation media contain BDNF and GDNF	Neural rosette → different cortical neurons subtypes and astrocytes	+Sequential generation of deep and upper layer neurons +-The number of upper layer neurons more than previous methods -Do not fully mimic cytoarchitecture of *in vivo*	Pauklin et al. [37]
Human ESCs/iPSCs	-Cells cultured at low density in presence of Noggin at early stages -Differentiation media without any growth factors	Cortical neural progenitors → different cortical neurons subtypes and astrocytes	+Sequential generation of deep and upper layer neurons +The derived neurons integrated to existing circuits after transplantation -Do not fully mimic cytoarchitecture of *in vivo*	Espuny-Camacho et al. [37]

Table 2: Comparison of 3D based differentiation protocols to generate cortical neurons.

Cell source	Culture condition	Outcome	Pros and Cons	Author
Mouse ESCs	- EB cultured in presence of retinoic acid	EB → neurons	+First study demonstrated retinoic acid enhanced neural fate -Unclear which cortical neuron subtypes generated	Bain et al. [41]
Mouse ECCs	-EB cultured in serum free media containing FGF2	EB → neurons	+Self organized neuronal rosette +Neurons, astrocytes and oligodendrocytes -Unclear which cortical neuron subtypes generated	Zhang et al. [43]
Mouse/human ESCs	-EB cultured in serum free media (SFEB/SFEBq) -Addition of Noggin and Wnt during early stages	EB → large rosette →different cortical neurons	+Rosette containing forebrain neural progenitors +Generation of $Reelin^+$, deep and upper neurons -No upper layer neurons production from human ESCs - Did not give rise to inside out order structure	Watanabe et al. [44] Eiraku et al. [45]
Human ESCs/iPSCs	-EB cultured in serum free media -EB Embedded in matrigel -Growing in spinning bioreactor	EB → cerebral organoid	+$Reelin^+$ neurons, deep and upper layer neurons -Do not form six layer structure -Heterogeneous structures with batch to batch variability -Viability of cells after long term culture	Lancaster et al. [46]
Human iPSCs	-Growing EB in presence of dual SMAD inhibitors at early stage	EB → cortical spheroids (without embedding in matrigel)	+Generation of deep and upper layer neurons +Generation of astrocytes -No laminar structure -Cell survival after protracted culture	Pasca et al. [51]
Human iPSCs	- EB cultured in presence of inductive factors at early stage - Growing in miniaturized spinning bioreactor	EB → region specific organoids	+Region specific organoids (forebrain, midbrain or hypothalamic) +Cost effective and more reproducible -No lamination of organoids	Qian et al. [53]

differentiation [11]. Importantly, the number of upper layer neurons generated in the presence of dual-SMAD inhibitors, retinoic acid, brain-derived neurotrophic factor and glial-derived neurotrophic factor is significantly higher than with previous methods.

Interestingly, and in contrast to mouse cells, corticoneurogenesis from human pluripotent

cells does not require blockade of ventral differentiation using sonic hedgehog antagonist [8,34,37]. Deep and upper layer neurons are generated in culture in a timeframe similar to corticogenesis *in vivo* and the neurons acquire some degree of maturity over time. Interestingly, the *in vitro* generated human neurons survived xenografting and integrated into the existing circuitry of host mouse brain [37]. This suggests that both paracrine but also intrinsic programs of differentiation and maturation control cortical neuron development even *in vitro*.

3D Based Differentiation Protocols

Many standard methods of *in vitro* differentiation of pluripotent cells, including human ESCs and iPSCs, require the formation of EBs. EBs resembles the gastrulation stage of the embryo and is formed by growing pluripotent cells in aggregates in the presence of serum. EBs give-rise to all three germinal layers of the embryo [39,40]. EBs are then cultured in serum free medium containing supplements and growth factors to select for ectodermal cells and subsequently neuroectodermal cells which go on to generate different types of neurons.

As with the adherent culture systems, the presence of retinoic acid during EB differentiation enhances neural fate determination and differentiation. Neurons generated from retinoic acid treated EBs display more mature electrophysiological properties [41]. The formation of NE cells is further enhanced by treatment of EBs with FGF-2 in serum-free medium [42]. These EB-derived NEs generate rosettes and give-rise to neurons, astrocytes and oligodendrocytes after removal of FGF2 [42]. Importantly, *in vitro* generated human neural precursors integrate into the mouse brain when xenographted and give-rise to neurons and astrocytes [43].

One of the challenges of the EB differentiation procedure is the variability and relatively uncontrolled differentiation process. Subsequently, the Sasai group refined the 3D culture system of differentiation to neural tissue with two, serum free mouse EB differentiation systems, SFEB [44] and SFEB-q [45]. These culture paradigms enhance selection for increased telencephalic precursor differentiation in polarized neuroepithelium like structures. Both SFEB and SFEB-q use additions of Wnt and Nodal inhibitors at early stages of differentiation [44]. In the SFEB cultures, Cajal Retzius Reelin$^+$ neurons are generated first, as *in vivo*, and these are followed by TBR1$^+$ and CTIP2$^+$ deep layer neurons and, subsequently, SATB2$^+$, BRN2$^+$ and CUX1$^+$ upper layer neurons [45]. However, as with most culture systems, the neurons do not form a laminated structure. The SFEB protocols for mouse ESCs were applied to differentiate human ESCs to cortical neurons [45]. Although human pluripotent stem cells differentiated into dorsal telencephalic progenitors, late born upper layer neurons were not generated, even after longer differentiation periods [45]. This suggested that, in addition to the time component, mouse and human cortical differentiation might not be entirely the same and have species-specific mechanism of commitment and differentiation.

In an attempt to develop a culture system that more closely recapitulates human brain development, the EB procedure was developed further [46]. EBs were induced to a neuroectodermal fate and then embedded in the extracellular matrix Matrigel [47]. The resulting 3D aggregates formed self-organizing 3D structures, which showed apical basal polarity of the dorsal telencephalon. These structures are referred to as cerebral organoids. When transferred to and cultured further in spinning bioreactors, cerebral organoids continued to grow and differentiate, and ultimately reached 4 mm in diameter [48].

Cerebral organoids contain fluid-filled cavities that resemble ventricles of the postnatal brain and develop VZ and OSVZ like structures [14,48]. Cerebral organoids also contain tissues of other brain regions and further refinement has resulted in cultures that contain a continuous neuroectoderm consisting of forebrain, retinal, midbrain, and hindbrain tissues [46,49]. Upon differentiation and with age, mature organoids contain Reelin$^+$ neurons, early born cortical neurons (TBR1$^+$, CTIP2$^+$) and some late born neurons (SATB2$^+$, BRN2$^+$), although upper layer neurons are rather rare and their lamination is not complete [46,50]. However, neurons in these cerebral oganoids are electrophysiologically active [46].

3D organoids hold much promise for the analysis of brain development and disease and new protocols can generate specific subregions of the brain [51-53]. Adaptation of the method and combination of small molecule inhibitors including those blocking BMP and TGFβ signaling increase the formation of functional deep and upper layer cortical neurons and glial cells [51,53]. The initial cerebral organoid

system was expensive and not very reproducible. However, technical advances have now managed to miniaturize the cultures reducing medium costs and increasing reproducibility [53].

To What Degree Do *In vitro* Differentiation Models Recapitulate *In vivo* Differentiation (Progress and Limitations of Current Techniques)

Modeling the differentiation process of cerebral cortex development *in vitro* from beginning to end is not straightforward due to inherent complexity of the cerebral cortex and lack of a complete understanding of the cellular and molecular mechanisms involved. However, recent pluripotent stem cell culture technologies recapitulate the key molecular and cellular phenotypes of *in vivo* differentiation. Taken together, the current *in vitro* corticogenesis models recapitulate three important aspects of *in vivo* cortical development; 1) Conversion of pluripotent stem cells to neural progenitors under the control of intrinsic mechanisms and inductive factors. 2) Induction of diverse neuronal subtypes and glial cells. 3) Maturation of pluripotent stem cell-derived neurons.

Each culture method has certain advantages and limitations. The adherent models of *in vitro* corticogenesis are robust, cost effective and reproducible (in comparison to 3D based cultures) [52]. Moreover, analyzing and controlling the differentiation of pluripotent cells to specific neuron subtypes is relatively easy. Adding morphogens and growth factors to the culture medium or induction of specific genes can trigger fate specification during differentiation and this is helped considerably by the accessibility of the adherent system. In addition, the adherent systems can be used to guide differentiation to specific neural cell types that can be isolated as relatively homogenous populations for clinical studies, drug discovery and eventually transplantation. However, adherent differentiation culture systems also have some limitations. 1) They do not fully simulate the complex cytoarchitecture and diversity of neuronal cell types *in vivo*. 2) Functional, complex neuronal circuits tend not to be fully formed as the cultures lack, for example, ventral-derived interneurons. However, these culture systems have been used successfully to study diseases including, Alzheimer [54,55], schizophrenia [56,57], autism spectrum disorder [58] and bipolar disorder [59,60].

Conversely, while 3D cultures, and particularly cerebral organoid, recapitulated some key features of cortex development, these models also have limitations; 1) They also fail to fully mimic cytoarchitecture of cerebral cortex, and the formation of upper layer neurons, and the six-layered isocortex. 2) Growing EBs to initiate the cultures without external control of fate leads to heterogeneous structures with significant batch-to-batch variability [53]. 3) Although different regions of the cerebral organoids resemble different brain regions, they are randomly organized within the tissue [49]. As with the adherent culture systems, ventral forebrain structures tend to be scarce as a result of intrinsic patterning. However, when generated, inhibitory interneurons are able to migrate through the organoids [49]. To address this migration of ventral inhibitory neurons, Bagley et al. generated ventral organoids using a ventralizing medium containing sonic hedgehog activators and WNT inhibitors. They were able to reconstruct the dorsoventral axis of the telencephalon by co-culturing ventral and dorsal organoids and found that interneurons migrated from the ventral to the dorsal parts of chimeric organoids [61]. Potentially, the major limitation for cerebral organoids at present is their rapid decrease in survival over time. Cell death, especially in the center of the organoids, increases dramatically potentially due to the lack of vascularization and penetration of nutrients and gases. Co-culture with other cell types such as endothelial cells or engineer a circulatory system might circumvent these issues in the future [48]. Brain organoids hold promise for modeling the neurodegenerative and psychiatric diseases in a 3D environment and cerebral organoids have already provided important insights into human disease mechanisms including microcephaly [46], autism [62], Zika virus infection [63,64], and Miller Dieker syndrome (MDS) [65,66], autism [62] and schizophrenia [67].

Outlook and Future Technology

A major step in the generation of cerebral cortex in a dish has now been taken with the development of the 3D organoids. One question arose whether the *in vitro* differentiation models can reflect the substantial differences between human and other species. Importantly, pluripotent based models reflect the differences in timing of neurogenesis *in vivo*.

The development of cortical culture systems from human pluripotent cells uncovered that

differentiation of human cortical neurons is protracted compared to mouse. Although this could have been expected, as during murine embryogenesis cortical neurons are generated over a 6-day period, in humans the process takes more than 100 days (GW5-30) [12,30]. The fact that similar species-specific time scales are maintained *in vitro* suggests some form of inherent or self-regulatory mechanism exists that "times" neurogenesis. Comparison of adherent and organoid cultures from mouse, human and other primates revealed that the proliferative capacity of progenitors is regulated cell-autonomously and differs across species [68]. This is not due to aberrant progenitor cell specification *in vitro* as outer radial glial (ORG) cells derived from human pluripotent stem cells have similar molecular characteristics to ORG cells isolated from human fetal brains [69]. In addition, the ORG cells in the OSV-like region of organoids expressed the prominent markers of ORG including HOPX, FAM107A, and PTPRZ1 [53,65]. Moreover, none of the current cerebral organoid models recapitulated folding and gyrification of human brain. Recently, deletion of PTEN, an anti-proliferative protein, induced neural progenitor proliferation, which led to larger and partially folded human cerebral organoids. However, in contrast, deletion of PTEN from mouse cerebral organoids did not induce folding demonstrating species-specific regulatory mechanism that is conserved and active *in vitro* [70].

The fact that human cortical neurogenesis *in vitro* is very protracted presents challenges for analysis of gene function and drug screening. Hence, accelerated differentiation protocols are being developed by combining different combinations and temporal regimes of small molecules, to produced cortical neurons from pluripotent stem cells in a precise and prompt way [71]. The simultaneous inhibition of progenitor maintenance and targeted differentiation with small molecules, reduces the differentiation time to generate, for example, functional deep layer (layer VI) neurons [71].

Another approach that is being explored to circumvent the time problem is directed differentiation using transcription factor-mediated differentiation [72,73]. Direct programing of pluripotent or even somatic cells to specific neuronal subtypes is an attractive prospect to generate homogeneous neuron populations. For example, conversion of somatic cells (fibroblast) to neurons by expression the transcription factors Ascl1, Brn2 (Pou3f2) and Myt1l [74]. In this case, the procedure bypasses expansion of the pluripotent state and directly induces neurons.

Interestingly, somatic cell aging hallmarks are maintained in neurons derived by direct programing, while many of the age associated epigenetic marks are reset in neurons differentiated from iPSCs and, therefore, they do not retain age-associated characteristics [75,76]. Importantly, the current cultures give-rise to relatively immature neurons that are at an early to mid-stage of cortical development [53,61,77]. Comparison of gene expression profiles of human fetal brain confirmed that organoid-derived cells resemble the early stages of brain development [51]. In the future, it will be important to model neurons of late stages of development or even the adult brain. Towards this goal, induced aging of neurons derived from iPSCs has been attempted by expressing Progerin [78] or by inducing the cells with cellular stressors [6,79]. Modeling of aging in culture still remains to be a major challenge and it will be interesting to study the disease associated mechanisms in an aged brain model in the future [6].

Author Contributions

Conceptualization Z.E., G.C. and V.T.; Writing – Original Draft Z.E., G.C. and V.T.; Writing – Review and Editing Z.E., G.C. and V.T.; Funding Acquisition V.T.; Project Administration G.C. and V.T.; Resources V.T.

Acknowledgments

We thank the members of the Taylor lab for critical reading of the manuscript. This work was supported by the Swiss National Science Foundation, the SystemsX.ch project NeuroStemX, and the University of Basel.

References

1. Molyneaux BJ, Arlotta P, Menezes JR, *et al.* Neuronal subtype specification in the cerebral cortex. *Nat. Rev. Neurosci* 8(6), 427-437 (2007).
2. Price DJ. Building brains : an introduction to neural development (2011).
3. Franco SJ, Muller U. Shaping our minds: stem and progenitor cell diversity in the mammalian neocortex. *Neuron* 77(1), 19-34 (2013).
4. Hansen DV, Lui JH, Parker PR, *et al.* Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* 464(7288), 554-561 (2010)
5. Lui JH, Hansen DV, Kriegstein AR. Development and evolution of the human neocortex. *Cell* 146(1), 18-36 (2011).
6. Shi Y, Inoue H, Wu JC, *et al.* Induced pluripotent stem cell technology: a decade of progress. *Nat. Rev. Drug. Discov* 16(2), 115-130 (2017)
7. Shen Q, Wang Y, Dimos JT, *et al.* The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nat. Neurosci* 9(6), 743-751 (2006).
8. Gaspard N, Vanderhaeghen P. Laminar fate specification in the cerebral cortex. *F1000. Biol. Rep* 36(1), (2011).
9. Kriegstein A, Alvarez-Buylla A. The glial nature of embryonic and adult neural stem cells. *Annu. Rev. Neurosci* 32(1), 149-184 (2009)
10. Wonders CP, Anderson SA. The origin and specification of cortical interneurons. *Nat. Rev. Neurosci* 7(9), 687-696 (2006).
11. Shi Y, Kirwan P, Livesey FJ. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nat. Protoc* 7(10), 1836-1846 (2012).
12. Lodato S, Shetty AS, Arlotta P. Cerebral cortex assembly: generating and reprogramming projection neuron diversity. *Trends. Neurosci* 38(2), 117-125 (2015).
13. Fietz SA, Huttner WB. Cortical progenitor expansion, self-renewal and neurogenesis-a polarized perspective. *Curr. Opin. Neurobiol* 21(1), 23-35 (2011).
14. Lewitus E, Kelava I, Huttner WB. Conical expansion of the outer subventricular zone and the role of neocortical folding in evolution and development. *Front. Hum. Neurosci* 7424 (2013).
15. Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292(5819), 154-156 (1981).
16. Bradley A, Evans M, Kaufman MH, *et al.* Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* 309(5965), 255-256 (1984).
17. Thomson JA, Itskovitz-Eldor J, Shapiro SS, *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* 282(5391), 1145-1147 (1998).
18. Takahashi K, Tanabe K, Ohnuki M, *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131(5), 861-872 (2007).
19. Stadtfeld M, Hochedlinger K. Induced pluripotency: history, mechanisms, and applications. *Genes. Dev* 24(20), 2239-2263 (2010).
20. Takahashi K, Okita K, Nakagawa M, *et al.* Induction of pluripotent stem cells from fibroblast cultures. *Nat. Protoc* 2(12), 3081-3089 (2007).
21. Maeder ML, Gersbach CA. Genome-editing Technologies for Gene and Cell Therapy. *Mol. Ther* 24(3), 430-446 (2016).
22. Hockemeyer D, Wang H, Kiani S, *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol* 29(8), 731-734 (2011).
23. Hockemeyer D, Soldner F, Beard C, *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol* 27(9), 851-857 (2009).
24. Xue H, Wu J, Li S, *et al.* Genetic Modification in Human Pluripotent Stem Cells by Homologous Recombination and CRISPR/Cas9 System, Humana Press, New York (2014).
25. Cong L, Ran FA, Cox D, *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121), 819-823 (2013).
26. Powell SK, Gregory J, Akbarian S, *et al.* Application of CRISPR/Cas9 to the study of brain development and neuropsychiatric disease. *Mol. Cell. Neurosci* 82(1), 157-166 (2017).
27. Malankhanova TB, Malakhova AA, Medvedev SP, *et al.* Modern Genome Editing Technologies in Huntington's Disease Research. *J. Huntingtons. Dis* 6(1), 19-31 (2017).
28. Hansen DV, Rubenstein JL, Kriegstein AR. Deriving excitatory neurons of the neocortex from pluripotent stem cells. *Neuron* 70(4), 645-660 (2011).
29. Tiberi L, Vanderhaeghen P, van den Ameele J. Cortical neurogenesis and morphogens: diversity of cues, sources and functions. *Curr. Opin. Cell. Biol* 24(2), 269-276 (2012).
30. van den Ameele J, Tiberi L, Vanderhaeghen P, *et al.* Thinking out of the dish: what to learn about cortical development using pluripotent stem cells. *Trends. Neurosci* 37(6), 334-342 (2014).
31. Ying QL, Stavridis M, Griffiths D, *et al.* Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat. Biotechnol* 21(2), 183-186 (2003).
32. Chen JK, Taipale J, Cooper MK, *et al.* Inhibition of Hedgehog signaling by direct binding of cyclopamine to Smoothened. *Genes. Dev* 16(21), 2743-2748 (2002).
33. Qian X, Shen Q, Goderie SK, *et al.* Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. *Neuron* 28(1), 69-80 (2000).
34. Gaspard N, Bouschet T, Herpoel A, *et al.* Generation of cortical neurons from mouse embryonic stem cells. *Nat. Protoc* 4(10), 1454-1463 (2009).
35. Chambers SM, Fasano CA, Papapetrou EP, *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol* 27(3), 275-280 (2009).
36. Pauklin S, Vallier L. Activin/Nodal signalling in stem cells. *Development* 142(4), 607-619 (2015).
37. Espuny-Camacho I, Michelsen KA, Gall D, *et al.* Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo. *Neuron* 77(3), 440-456 (2013).
38. Siegenthaler JA, Ashique AM, Zarbalis K, *et al.* Retinoic acid from the meninges regulates cortical neuron generation. *Cell* 139(3), 597-609 (2009).
39. Smukler SR, Runciman SB, Xu S, *et al.* Embryonic stem cells assume a primitive neural stem cell fate in the absence of extrinsic influences. *J. Cell. Biol* 172(1), 79-90 (2006).
40. Desbaillets I, Ziegler U, Groscurth P, *et al.* Embryoid Bodies: An In Vitro Model of Mouse Embryogenesis. *Exp. Physiology* 85(6), 645-651 (2000).
41. Bain G, Kitchens D, Yao M, *et al.* Embryonic stem cells express neuronal properties in vitro. *Dev. Biol* 168(2), 342-357 (1995).
42. Kelly CM, Zietlow R, Dunnett SB, *et al.* The effects of various concentrations of FGF-2 on the proliferation and neuronal yield of murine embryonic neural precursor cells in vitro. *Cell. Transplant* 12(3), 215-223 (2003).
43. Zhang SC, Wernig M, Duncan ID, *et al.* In vitro differentiation of transplantable neural precursors from human embryonic stem cells. *Nat. Biotechnol* 19(12), 1129-1133 (2001).
44. Watanabe K, Kamiya D, Nishiyama A, *et al.* Directed differentiation of telencephalic precursors from embryonic stem cells. *Nat. Neurosci* 8(3), 288-296 (2005).
45. Eiraku M, Watanabe K, Matsuo-Takasaki M, *et al.* Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell. Stem. Cell* 3(5), 519-532 (2008).

46. Lancaster MA, Renner M, Martin CA, *et al.* Cerebral organoids model human brain development and microcephaly. *Nature* 501(7467), 373-379 (2013).

47. Sato T, Vries RG, Snippert HJ, *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* 459(7244), 262-265 (2009).

48. Kelava I, Lancaster MA. Stem Cell Models of Human Brain Development. *Cell. Stem. Cell* 18(6), 736-748 (2016).

49. Renner M, Lancaster MA, Bian S, *et al.* Self-organized developmental patterning and differentiation in cerebral organoids. *EMBO. J* 36(10), 1316-1329 (2017).

50. Lancaster MA, Knoblich JA. Organogenesis in a dish: modeling development and disease using organoid technologies. *Science* 345(6194), 1247125 (2014).

51. Pasca AM, Sloan SA, Clarke LE, *et al.* Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. *Nat. Methods* 12(7), 671-678 (2015).

52. Qian X, Nguyen HN, Jacob F, *et al.* Using brain organoids to understand Zika virus-induced microcephaly. *Development* 144(6), 952-957 (2017).

53. Qian X, Nguyen HN, Song MM, *et al.* Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. *Cell* 165(5), 1238-1254 (2016).

54. Israel MA, Yuan SH, Bardy C, *et al.* Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature* 482(7384), 216-220 (2012).

55. Yang J, Li S, He XB, *et al.* Induced pluripotent stem cells in Alzheimer's disease: applications for disease modeling and cell-replacement therapy. *Mol. Neurodegener* 11(1), 39 (2016).

56. Hoffman GE, Hartley BJ, Flaherty E, *et al.* Transcriptional signatures of schizophrenia in hiPSC-derived NPCs and neurons are concordant with post-mortem adult brains. *Nat. Commun* 8(1), 2225 (2017).

57. Brennand KJ, Simone A, Jou J, *et al.* Modelling schizophrenia using human induced pluripotent stem cells. *Nature* 473(7346), 221-225 (2011).

58. Young-Pearse TL, Morrow EM. Modeling developmental neuropsychiatric disorders with iPSC technology: challenges and opportunities. *Curr. Opin. Neurobiol* 3666-3673 (2016).

59. Madison JM, Zhou F, Nigam A, *et al.* Characterization of bipolar disorder patient-specific induced pluripotent stem cells from a family reveals neurodevelopmental and mRNA expression abnormalities. *Mol. Psychiatry* 20(6), 703-717 (2015).

60. Watmuff B, Berkovitch SS, Huang JH, *et al.* Disease signatures for schizophrenia and bipolar disorder using patient-derived induced pluripotent stem cells. *Mol. Cell. Neurosci* 7396-103 (2016).

61. Bagley JA, Reumann D, Bian S, *et al.* Fused cerebral organoids model interactions between brain regions. *Nat. Methods* 14(7), 743-751 (2017).

62. Mariani J, Coppola G, Zhang P, *et al.* FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. *Cell* 162(2), 375-390 (2015).

63. Cugola FR, Fernandes IR, Russo FB, *et al.* The Brazilian Zika virus strain causes birth defects in experimental models. *Nature* 534(7606), 267-271 (2016).

64. Dang J, Tiwari SK, Lichinchi G, *et al.* Zika Virus Depletes Neural Progenitors in Human Cerebral Organoids through Activation of the Innate Immune Receptor TLR3. *Cell. Stem. Cell* 19(2), 258-265 (2016).

65. Bershteyn M, Nowakowski TJ, Pollen AA, *et al.* Human iPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia. *Cell. Stem. Cell* 20(4), 435-449 (2017).

66. Iefremova V, Manikakis G, Krefft O, *et al.* An Organoid-Based Model of Cortical Development Identifies Non-Cell-Autonomous Defects in Wnt Signaling Contributing to Miller-Dieker Syndrome. *Cell. Rep* 19(1), 50-59 (2017).

67. Stachowiak EK. Cerebral organoids reveal early cortical mal development in schizophrenia—computational anatomy and genomics, role of FGFR1. *Trans. Psychiatry* (2017).

68. Otani T, Marchetto MC, Gage FH, *et al.* 2D and 3D Stem Cell Models of Primate Cortical Development Identify Species-Specific Differences in Progenitor Behavior Contributing to Brain Size. *Cell. Stem. Cell* 18(4), 467-480 (2016).

69. Pollen AA, Nowakowski TJ, Chen J, *et al.* Molecular identity of human outer radial glia during cortical development. *Cell* 163(1), 55-67 (2015).

70. Li Y, Muffat J, Omer A, *et al.* Induction of Expansion and Folding in Human Cerebral Organoids. *Cell. Stem. Cell* 20(3), 385-396 (2017).

71. Qi Y, Zhang XJ, Renier N, *et al.* Combined small-molecule inhibition accelerates the derivation of functional cortical neurons from human pluripotent stem cells. *Nat. Biotechnol* 35(2), 154-163 (2017).

72. Pang ZP, Yang N, Vierbuchen T, *et al.* Induction of human neuronal cells by defined transcription factors. *Nature* 476(7359), 220-223 (2011).

73. Vierbuchen T, Wernig M. Direct lineage conversions: unnatural but useful? *Nat Biotechnol* 29(10), 892-907 (2011).

74. Vierbuchen T, Ostermeier A, Pang ZP, *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463(7284), 1035-1041 (2010).

75. Mertens J, Paquola ACM, Ku M, *et al.* Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects. *Cell. Stem. Cell* 17(6), 705-718 (2015).

76. Tang Y, Liu ML, Zang T, *et al.* Direct Reprogramming Rather than iPSC-Based Reprogramming Maintains Aging Hallmarks in Human Motor Neurons. *Front. Mol. Neurosci* 10359 (2017).

77. Di Lullo E, Kriegstein AR. The use of brain organoids to investigate neural development and disease. *Nat. Rev. Neurosci* 18(10), 573-584 (2017).

78. Miller JD, Ganat YM, Kishinevsky S, *et al.* Human iPSC-based modeling of late-onset disease via progerin-induced aging. *Cell. Stem. Cell* 13(6), 691-705 (2013).

79. Liu GH, Qu J, Suzuki K, *et al.* Progressive degeneration of human neural stem cells caused by pathogenic LRRK2. *Nature* 491(7425), 603-607 (2012).

Aims of the study

In this study I used hiPSCs to study human cortical development. *In vitro* corticogenesis recapitulates *in vivo* cortical development with the appropriate sequential generation of deep and upper layer neurons, their numbers, and the maturation of cortical projection neurons. The present study aims to explore the mechanisms regulating *in vitro* cortical differentiation of neural progenitor cells (NPs) derived from pluripotent stem cells.

The aims of the study are:

- To examine the differentiation potential of NPs during *in vitro* differentiation.
- To identify the differences in transcriptional profiles between NPs isolated at different stages of *in vitro* cortical differentiation.
- To assess the role of extrinsic cues in regulating NPs potency.

To achieve these aims I performed five sets of experiments:

1. Establish the model of human in vitro cortical differentiation. 2D differentiation of hiPSCs to NPs and then cortical neurons.
2. Retrovirus-mediated labeling of NPs at different time points in culture to follow NPs progenies and analyze their cell fate.
3. FACS-isolation of NPs based on neural stem-related cell surface markers to evaluate NP fate potential at specific stages.
4. Gene expression profiling of FACS-isolated NPs to identify key fate determinants during cortical differentiation.
5. Heterochronic culture experiments using early and late NPs to unravel environmental influence on NP differentiation.

CHAPTER 2
Results

Human pluripotent stem cell derived neural progenitors display two modes of neuronal fate determination

Zahra Ehsaei[1], Chiara Rolando[1], Zahra Karimaddini[2], Maysam Mansouri[3], Dagmar Iber[2], Philipp Berger[3], Ginetta Collo[1,4*] and Verdon Taylor[1*]

[1] *Department of Biomedicine, University of Basel, Mattenstrasse 28, CH-4058, Basel, Switzerland*

[2] *Department of Biosystems Science and Engineering (D-BSSE), ETH Zurich, Mattenstrasse 26, 4058 Basel, Switzerland*

[3] *Paul Scherrer Institute, Biomolecular Research, Molecular Cell Biology, CH-5232 Villigen, Switzerland*

[4] *Department of Molecular and Translational Medicine, University of Brescia, Viale Europa 11, 25123, Brescia, Italy*

* Correspondence should be addressed to Verdon.taylor@unibas.ch, Ginetta.collo@unibas.ch

Abstract

How the pools of neural stem and progenitor cells in the developing human brain generate neuronal diversity in the cerebral cortex is currently poorly understood. Neurogenesis in the developing cerebral cortex is a precisely controlled process. Stem cells pass through sequential phases of expansion, neurogenesis, and gliogenesis. Two basic models of stem cell-mediated neurogenesis have been proposed, the common progenitor model and multi-progenitor model. Here we studied human cortical neurogenesis *in vitro* from induced pluripotent stem cells (hiPSCs). In contrast to the cell fate restriction of cortical neural stem cells (NSCs) of mice, the human cortical NSC pool retain multipotency throughout cortical neurogenesis *in vitro*. At the onset of neurogenesis from hiPSCs, deep-layer neuron production is favored and upper-layer neuronal fate is instructively acquired over time. Subsequently, deep-layer neuronal fate determination is repressed but not lost. Later stage NSCs reacquire deep-layer fate and are reprogrammed upon grafting into a promiscuous environment *in vitro*, indicating that extrinsic signals contributed to neural progenitor fate acquisition in humans. However, neural rosette forming potential, an early ventricular radial glial (vRG) character, is lost and NSCs gain outer radial glial (oRG) identity during late phase of *in vitro* differentiation. Conversely, early stage NSCs acquire upper-layer fate potential over time and the transition coincides with acquisition of the oRG phenotype. Thus, extrinsic signals contributed to human NSC fate repression and acquisition and the environment of the developing brain programs fate. Our findings shed light on the mechanisms on neural specification and revealed a population of neural stem cells *in vitro* with a previously unknown plasticity and identity.

Introduction

Development of the cerebral cortex requires a fine coordination of neural stem and progenitor expansion, fate specification and cellular differentiation[34]. During cortical neurogenesis, neural progenitor cells (NPs) in ventricular zone (VZ) and subventricular zone (SVZ) generate different cortical projection neurons in an inside-out order. Early born neurons populate the deep cortical layers (VI-V) and are followed by neurons that settle in the superficial layers (IV-II)[23,35,36]. The mechanisms controlling NP fate determination and neurogenesis during mammalian cortical development remain to be elucidated[34,35]. Pioneer transplantation studies in ferret suggest that mammalian NP multipotency changes over time during development, and their neurogenic fate potential becomes progressively restricted over the course of corticogenesis[37]. Heterochronic grafting experiments indicate that early NPs are able to comprehend and respond to signals of the cortex at later stages and adopt the appropriate fate[37]. Conversely, later-stage NPs are no longer able to generate early neuron subtypes even when grafted into the early primordium of the cerebral cortex[38]. Hence, it is generally accepted that mammalian NPs are multipotent, able to generate all cell fates derived from the NP pool in the dorsal cortex, and their fate becomes successively more restricted with time and developmental progression[39]. However, although this data has been supported by genetic labeling and fate tracing in mice, it has also been proposed that the progenitor pool of the early cerebral cortex may contain fate restricted NPs that have a restricted differentiation potential[40].

During human cortical development, the fate potential of the NPs is less clear, and how the potential fates of human cortical NPs changes over time is not known. The human cerebral primordium contains vRG constituting the majority of the ventricular zone[19] . These vRG are comparable to the RGs in mice. However, the human cortex contains a second more expansive progenitor zone called the outer SVZ, which is populated by oRG, the major expansive and neurogenic population[7,19,41]. Several RNA-Seq analyses have been performed on human fetal brain cells to decode potential transcriptional regulatory mechanisms involved in human cerebral cortex development[42-46]. NPs derived from hiPSCs can generate different subtypes of cortical neurons in a sequential order thereby recapitulating mouse and presumably human corticogenesis[47-49]. Transcriptional analysis of human embryonic stem cell derived cortical cultures identified heterogeneity in the cultured neural progenitors[50-52]. However, the mechanisms that control the ability of human vRGs and oRGs to produce neuronal subtypes at different stages of corticogenesis, especially the late-

born cell types, remain unclear and a detailed understanding of each stage of human cortical differentiation is missing.

Analysis of human cortical development is a major challenge. Unlike in mouse where development can be analyzed in detail due to accurate timing of development, inbred genetic backgrounds and freely available tissue, human fetal cortical tissue for research purposes is rare and a major challenge to analyze both due to logistical and ethical barriers. Hence, *in vitro* differentiation protocols have boosted analysis of human cortical development. The commonly used protocols can be divided into those that use fetal-derived neural progenitors as a starting material or human pluripotent cell-derived NPs. The human pluripotent cell-derived culture systems include adherent cultures and organoid, three-dimensional self-organizing cultures.

Here, we performed an in-depth analysis of neural progenitor fate acquisition during human cortical neurogenesis *in vitro*. We performed cortical differentiation from hiPSCs for up to 110 days. We show sequential generation of cortical neuron subtypes and performed progenitor lineage-tracing to elucidate NP potential. By retroviral marking of NP, we show the periods of neuronal subtype progenitor activity and cell cycle exit and analyzed cell fate restriction over time. We sequentially isolated NPs during the course of cortical neurogenesis *in vitro* and analyzed their differentiation potential and changes in transcriptional profiles over developmental time. Finally, we performed heterochronic grafting of hiPSC-derived NP *in vitro* to address fate restriction and acquisition. Our results suggest key differences in the mechanism of cortical neurogenesis between human and mouse development.

Result

Differentiation of hiPSCs to cortical neurons over 110 days shows similar dynamic to *in vivo* cortical development

To address the differentiation potential of hiPSC-derived NPs over time, we utilized directed differentiation of hiPSCs to cortical neurons subtypes in an adherent culture system (Figure 1A)[53,54]. Human hiPSCs were cultured under defined conditions in the presence of dual TGF-□ signaling inhibitors and induced to undergo neural induction into dorsal cortical progenitors (Figure 1A) [53,55]. After 6 days of expansion, at day 18 of differentiation, NPs generated rosette-like structures and 95.3 +/- 0.2% of the cells expressed the neural stem/progenitor cell marker PAX6 (Figure S1A, S1B). NPs were induced to differentiate over a period of 110 days and monitored for the progenitor and neuronal marker expression (Figure 1A and S1).

At day 25 of differentiation, most NPs expressed PAX6, SOX2 and NESTIN (Figure S1B,C). However, the proportion of cells that expressed PAX6 rapidly decreased over time. By day 40, $PAX6^+$ progenitors accounted for less than 14.6 +/- 1.9% of the cells (Figure S1B). Although the proportion of $PAX6^+$ cells continued to decline with age, $PAX6^+$ neural stem/progenitors remained in the cultures even after 110 days of differentiation (Figure S1B).

In mice, NSCs generate neurons either directly of via an intermediate progenitor cell type called basal progenitors. By day 25, $TBR2^+$ basal progenitor-like cells had appeared around the rosettes indicating a progression in the cultures towards neuronal commitment and differentiation (Figure S1D). We analyzed the number of different neuron subtypes generated in the cultures at different stage of culture by immunostaining for neuronal-layer specific markers. TBR1 and CTIP2 are markers of deep layer (layer V and VI) projection neurons while BRN2 and SATB2 are expressed by upper layer (layer II-IV) projection neurons (Figure 1A). At day 31, $TBR1^+$ and $CTIP2^+$ deep layer neurons were already present in the cultures. After peaking at day 31, the proportion of $TBR1^+$ neurons ($MAP2^+$) remained constant (~13.2% of total cells) during the rest of the differentiation to day 110 (Figure 1C,S1E). Conversely, $CTIP2^+$ neurons ($MAP2^+$) increased over time reaching a plateau of ~22.2 +/-2.8% of total cells between day 60 and day 90 of differentiation (Figure 1E,F, 1SF).

Upper layer neurons express the transcription factor BRN2 or SATB2. Interestingly, $BRN2^+$ neurons ($MAP2^+$) were present from early stages of differentiation as early as day 31 (Figure 1E,F, and S1G). Whereas the other neuronal markers were almost exclusively expressed by $MAP2^+$ neurons, a significant number of $BRN2^+$ cells were

MAP2$^-$. The proportion of BRN2$^+$ and BRN2$^+$MAP2$^+$ cells remained constant throughout the differentiation period and, as the total number of the cells increased over time. This indicated that BRN2$^+$ cells were continually formed during the culture period. In contrast to BRN2 expression, SATB2$^+$ neurons (MAP2$^+$) were absent before day 40 of differentiation and only appeared in significant numbers from day 60 (Figure 1E,G and S1H). The proportion of SATB2$^+$MAP2$^+$ neurons continually increased over time (Figure 1G).

After the neurogenic phase of differentiation, NSCs enter the gliogenic period[23,56]. Therefore, we also address the differentiation of S100□$^+$ astrocytes in the cultures. Astrocytes were absent in the cultures until day 60 and were only detected in significant numbers after day 90 (Figure S1I and data not shown). We did not detect expression of the oligodendrocyte markers Sox10, Olig2 or NG2 up to 100 days of differentiation (data not shown).

In summary, hiPSCs generated different neuronal populations and astrocytes over a prolonged period and the generation of deep and upper layer neurons and astrocytes occurred in a sequential fashion *in vitro*. Therefore, the cultures mimicked the putative transition of neural progenitors *in vivo* from expansion through neurogenesis to gliogenesis. In addition, the sequential generation of deep and upper layer neurons from human NPs showed striking similarities to mouse cortical development.

The NPs population generate different types of cortical neurons during *in vitro* corticogenesis

Having established a human *in vitro* corticogenesis model that sequentially generates different neuronal subtypes of the deep and superficial layers, we addressed when the progenitors of the specific neuronal subtypes exit cell cycle. First, we performed genetic lineage tracing of mitotic progenitors by infection with a retrovirus (RV) expressing EGFP. We infected cells at day 28 (day of infection 28 (Di)) and analyzed their neuronal progeny at different stages (day 30, 40, 60, 80, 90 and 110; Figure 2A). 2 days post-infection (analysis at day 30 (Da30)), 91.5 +/- 2% of the GFP$^+$ cells expressed PAX6, confirming that the majority of labeled cells were NPs (Figure S2A, S2B or Figure 2B). As expected after labeling mitotic progenitors, the number of GFP expressing progeny increased continually over the culture period from Da40 to Da110 (Figure S2C or Figure 2B,C). Therefore, we analyzed the appearance of deep and upper layer neurons during the course of the culture from the infected progenitors. The percentage of GFP-labeled TBR1$^+$MAP2$^+$ deep-layer neurons remained constant over time from Da40-110 (Figure 2B,C). Taking the >2-

fold increase in GFP^+ progeny over the course of the culture, $TBR1^+$ neurons were not only generated early but continually over the culture period. This also suggests that on average, 16.9% of the differentiated progeny produced at any time in the culture differentiate into $TBR1^+$ neurons. Similarly, $GFP^+CTIP2^+MAP2^+$ and $SATB2^+MAP2^+$ neurons increased over the entire culture period (Figure 2D,E). Therefore, the $PAX6^+$ progenitor population at day 28 is multipotent and generates deep and upper layer neurons. However, unexpectedly, the RV labeled progenitors for deep layer neuronal types do not become restricted over time.

The results suggested that the progenitors of all neuronal subtypes are mitotically active at day 28 in the cultures as they can be infected and labeled with RVs. This could be explained by all neurons having a common progenitor at day 28. However, an alternative hypothesis could be that the neuronal subtypes have different and potentially fate restricted progenitors that are all in an expansion state at day 28 but that could exit cell cycle at different times to initiate terminal differentiation to mature $MAP2^+$ neurons.

In order to address when the progenitors of the different neuronal subtypes exit cell cycle, we performed RV-genetic labeling at different stages of the cultures and analyzed the fates of their progeny (Figure 2F). Infection of progenitors at Di60 and Di80 resulted in an increase in progeny over time consistent with a continued expansion of a progenitor pool at these later stages (Figure S2E).

Infection at Di60 onwards resulted in a reduction in the production of $TBR1^+MAP2^+$ neurons with time suggesting cell cycle exit of most of their progenitors between day 28 and day 60 (Figure 2G,H). This reduction in $TBR1^+$ neuron production was particularly evident at Di90. Similarly, the production of $CTIP2^+$ neurons decreased from progenitors labeled at Di60 and Di80, however, Di90 labeled progenitors produced proportionally more $CTIP2^+$ neurons than progenitors at Di60 or Di80 (Figure 2G and 2I). This reduction was not due to reduced GFP^+ progeny in the later cultures as the density of RV infected cells at Da110 was similar for Di60-Di90 (Figure S2E). In contrast to the early neuron subtypes $Tbr1^+$ and $CTIP2^+$, SATB2 neuron progenitors were mitotically active over the entire culture period and contributed a constant proportion of offspring irrespective of the date of RV-labeling (Figure 2G and 2J). However, $SATB2^+$ neurons only appeared late in the cultures suggesting that the terminal commitment and differentiation was delayed.

Hence, progenitors labeled at day 60 gave rise to deep and upper layer neurons and late progenitors (at day 80 and 90) switched their fate towards the generation of

upper layer SATB2$^+$ neurons but surprisingly did not lose Tbr1 and CTIP2 neuron potential completely. This is in contrast to late progenitors from the mouse that become fate restricted over time[57]. Although, progenitors in the culture at day 90 generated less TBR1$^+$ neurons in comparison to day 28, the potential for deep neuron formation was not completely lost (Figure 2H). This suggests that although neural progenitor fate for the production of TBR1$^+$ deep layer reduces over time; early fate may not be lost but rather repressed. In addition, these data suggest that neurons progenitors progressively acquire SATB2 fate over time (Figure 2J).

FACS purified NPs retain radial glial identity during *in vitro* differentiation

To address developmental changes in NP gene expression, we sorted progenitors from the cultures based on CD184, CD24, CD44 and CD271 expression (Figure 3A,B)[58]. Fluorescent-assisted cell sorted (FACS) CD184$^+$/CD24$^+$/CD44$^-$/CD271$^-$ NPs isolated at different stages of differentiation (day 25, 40, 60, 70, 80, 90) expressed PAX6 and SOX2 (Figure 3C,D, S3A,B). RNA-Seq analysis of the sorted NPs isolated from independent cultures confirmed expression of canonical radial glial markers including Vimentin, HES1, PAX6 and SOX2 (Figure 3E). As a control, we checked the expression of positive selection markers (CD184 and CD24) and negative selection marker (CD44 and CD271) (Figure 3E). Therefore, the sorting procedure isolated cortical NPs from the cultures between day 25 and 90.

In order to address the fate potential, we replated the sorted cells on to poly-Lysine/Laminin coated plates and continued differentiation for 20 days. Sorted NPs from all ages of culture generated TBR1, CTIP2, BRN2 and SATB2$^+$ neurons, although SATB2$^+$ upper layer neurons were rarely generated by the day 25 NPs (Figure 3F and data not shown). These data indicate that the CD184$^+$/CD24$^+$/CD44$^-$/CD271$^-$ sorting strategy allowed isolation of a population of dorsal cortical progenitors from hiPCS-derived cortical differentiation cultures that showed multi-neuronal fate potential.

To address the transcriptional profile of the NPs and compare those of NPs of different stages, we performed whole genome analysis of variable gene expression. t-distributed stochastic neighbor embedding (tSNE) analysis clustered the cells of the same differentiation time together (Figure 3F). In this way, we identified 4 different clusters of early (day 25), intermediate (day 40), late-intermediate (day 60) and late (day 70, 80, 90).

As expected, cells of the same differentiation time clustered together on a principle component analysis (PCA) of variant gene expression (Figure 3G-I). However, the

PCA analysis also revealed distinct differences in gene expression between the NPs of different ages and a time-dependent shift in gene expression (Figure 3G-I). The principle components clearly separated the NPs based on age and stage of differentiation (Figure 3G).

PC1 genes separated the samples based on the differences between early (sorted at day 25 and 40) and late NPs (sorted at day 60, 70, 80, 90) (Figure 3G). We analyzed the RNA-Seq data for genes that were highly regulated between NPs of different stages by pairwise comparisons. Among the genes differentially regulated between early and late progenitors were the outer radial glial markers, including HOPX, MOXD1, FAM107A and TNC (Figure 3H)[45]. MOXD1, FAM107A and TNC were up regulated by the late NPs compared to the early NPs (Figure 3H). Indeed, the number of differentially regulated genes of every two consecutive time points revealed that the maximum change is between day 40 and 60 (Figure 3SI) and some ORG correlated genes are differentially expressed between day 40 and 60 (Figure 3SJ). Hence, the NP population in the cultures showed a transition from ventricular NP to outer radial glial-like gene expression by day 60. Interestingly, this coincides with the transition in neurogenesis to SATB2 neuron production (Figure 4E).

Late NPs generate upper layer neurons in comparison to early NPs

To examine the characteristics and the intrinsic fate potential of NPs at different stages of *in vitro* cortical differentiation (day 25, 40, 60, 80 and 90) we replated $CD184^+/CD24^+/CD44^-/CD271^-$ NPs on poly-Lysine/Laminin coated plates (Figure 4A). After 20 days of differentiation we analyzed production of neuronal subtypes. Early NPs (day 25 and 40) generated predominantly $TBR1^+$ and $CTIP2^+$ neurons ($MAP2^+$) (Figure 4B-Dlox). Less than 3.5 +/- 1.6% of the neurons were $SATB2^+$ (Figure 4B,E). Analysis of neurogenic potential of the later stage NPs confirmed the analysis of neurogenesis in the cultures *in situ* as day 40-90 NPs continued to generate $TBR1^+$ and $CTIP2^+$ neurons ($MAP2^+$) (Figure 4C-D). In addition, NP potential to generate SATB2+ neurons increased dramatically with age (Figure 4E). These results indicated that NPs at early stages generated deep layer neurons at the expense of upper layer neurons. However, with time, both the total neurogenic potential and the ability to produce $SATB2^+$ upper layer neurons increased (Figure 4C-E). Hence, in contrast to the fate restriction reported for mouse cortical progenitors, hiPSC-derived NPs show continued multipotency and a gradual acquisition of upper neuronal fate potential. In addition, these results correlate with previous analyses indicating that human outer radial glial cells but not ventricular NPs generate the majority of upper layer neurons[59,60].

The local environment affects behavior and neuronal potency of NPs *in vitro*

The human cortical differentiation cultures became increasingly dense and complex with time as different neuronal subpopulations were generated. Grafting experiments in ferrets indicated that early NPs are multipotent and able to respond to local fate determining cues of a host cortex. We asked whether neural progenitor fate acquisition followed an intrinsic developmental program or whether extrinsic signals such as environmental cues can affect the fate of hiPSC-derived NPs. We addressed whether this increasing complexity contributed to changes in NP fate.

Grafting experiments *in vivo* are not possible for human NPs and xenografting of human NPs into mouse hosts may result in bias due to differences in species-specific signaling and developmental timing. Therefore, to address environmental regulation of NP fate and differentiation of the hiPSC-derived NPs, we performed *in vitro* grafting experiments. We performed isochronic and heterochronic coculture experiments by plating FACSed donor NPs from one culture onto host cultures. In order to discriminate the donor NPs from the host cells, we labeled the host hiPSCs by insertion of an EGFP coding region into the HMGA1 locus using CRISPR/CAS9 as described previously[61] (Figure S5A-D). We sorted $GFP^+/CD184^+/CD24^+/CD44^-/CD271^-$ NPs from donor cultures at day 25 and day 90 of differentiation and plated these onto host feeder cultures (Figure 5A).

Initially, we controlled the spontaneous differentiation potential of these NPs of each age and plated the cells on poly-L-lysine/Laminin as a substrate (Figure S4H). The aim was to address to what extent is the differentiation and potential of the NPs controlled by their local environment and potential paracrine signals, and what aspects are do to cell autonomous changes in fate that are not modulated by neighboring cells.

We performed in vitro grafting experiments by plating day 25 $EGFP^+/CD184^+/CD24^+/CD44^-/CD271^-$. Day 25 $EGFP^+/CD184^+/CD24^+/CD44^-/CD271^-$ NPs generated rosette-like structures within 20 days when plated on day 25 cells as a feeder layer and about 44.4 +/- 5.1 % of donor $EGFP^+$ cells expressed $PAX6^+$ (Figure 5B). This was similar to when plated directly on poly-l-lysine/Laminin as a substrate (Figure 5D). Approximately 29.4 +/- 1.9% and 7.4 +/- 1.5% of the donor $EGFP^+/CD184^+/CD24^+/CD44^-/CD271^-$ NPs-derived cells (day 25) expressed the neuronal markers TBR1 and CTIP2 respectively but only 0.1 % SATB2 after 20 days of differentiation (Figure 5D,E).

Conversely, when day 25 EGFP$^+$/CD184$^+$/CD24$^+$/CD44$^-$/CD271$^-$ NPs were plated on day 90 feeder cells, they failed to form neural rosettes and only 10.6 +/- 0.2% retained expression of PAX6 after 20 days (Figure 5C). The reduction in PAX6$^+$ progenitors was associated with an increase in the proportion of donor cells that expressed the deep layer neuron markers TBR1$^+$ (37.9+/- 3%) and CTIP2$^+$ (19 +/- 3.6%) (Figure 5D, E). Although differentiation of day 25 donor cells to CTIP2$^+$ neurons was similar on day 90 feeder cells to plating on poly-L-Lysine/Laminin substrate alone, differentiation towards CTIP2 fate was repressed by day 25 feeder cells (compare Figure 4D and 5E). Conversely, differentiation towards TBR1$^+$ neurons was enhanced by the day 90 feeder cells compared to day 25 feeders and to differentiation on poly-L-Lysine/Laminin (compare Figure 4E and 5E).

The proportion of day 25 donor-derived cells (GFP$^+$) that expressed the upper layer marker SATB2 (2.2 +/- 0.2%) remained relatively low in the grafting experiments but was significantly increased when day 90 cells were used as feeders compared to day 25 cells (Figure 5D,E). Differentiation of day 25 donor cells to SATB2$^+$ neurons was not higher on day 90 feeder cells than after plating on poly-L-Lysine/Laminin substrate and differentiating for 20 days (compare Figure 4E and 5E).

We performed the reciprocal experiment by plating day 90 donor EGFP$^+$/CD184$^+$/CD24$^+$/CD44$^-$/CD271$^-$ NPs onto day 25 or day 90 host cells (Figure 5A). Day 90 NPs did not form rosette-like structures even on day 25 feeders and only rarely seemed to integrate into the rosettes formed by the feeders themselves (Figure 5F). When plated onto day 90 feeders only around 12.2 +/- 0.9% of the donor day 90 NPs retained expression of PAX6 (Figure 5F,G). However, the differentiation of the day 90 NPs to TBR1$^+$ neurons was repressed on day 25 feeders (13.2 +/- 0.9%) compared to plating on to day 90 feeders (38.3 +/- 1.6%) and to plating onto poly-L-Lysine/Laminin (compare Figure S4G and 5H, I). The proportion of CTIP2$^+$ neurons generated by day 90 donor cells on day 25 and day 90 feeders was similar (16.7 +/- 4.7% versus 23.7 +/- 2.9%) suggestion that neither environment significantly affected differentiation fate (Figure 5I). However, the differentiation of day 90 progenitors to SATB2$^+$ neurons was significantly reduced on day 25 feeder cells compared to day 90 feeders and to plating on poly-L-Lysine/Laminin indicating that the environment provided by the day 25 cultures repressed upper neuronal differentiation (compare Figure S4E and 5H,I).

Discussion

Here we found that the timing and cell fate determination of human cortical NSCs is a combination of induced cell fate acquisition and fate repression. We transcriptionally profiled the NPs throughout cortical differentiation and identified genes that distinguish early NPs from late NPs. Comparison of the transcriptional profile of NPs at sequential time points of differentiation showed that oRG markers start to be prominent from day 60 of differentiation and late sorted NPs acquire the ORG characteristic[45,54]. Human iPSC-derived NPs undergo what appears to be a normal developmental transition from vRG to oRG NPs over time in culture. This transition is associated with expression of oRG markers including FAM107A, TNC, HOPX, LIFR and PTPRZ1, all of which are associated with late human NSC maintenance and expansion. In addition, the vRG to oRG transition is associated with the sequential production of TBR1, CTIP2, BRN2 and SATB2 deep and upper layer neurons, respectively, in the cultures. However, it is currently unclear how RG fate is controlled during human cortical development[45] (Pollen et al. REFS). Our isolation of NPs from different stages of corticogenesis *in vitro* revealed that the population retains the potential to generate TBR1 positive neurons and acquire SATB2 differentiation potential from day 60 onwards. However, genetic labeling *in situ* revealed, as expected, that TBR1 and CTIP2 positive neuron production diminishes in the cultures over time whereas $SATB2^+$ neuron production increases. This is consistent with birth dating studies using H^3-thymidine incorporation in primate and mouse showing sequential deep and upper layer cortical differentiation *in vitro*[62]. However, in contrast to the fate restriction of mouse NPs during cortical neurogenesis, human NPs do not lose deep neuronal fate but progressively acquire SATB2 differentiation potential. This is also supported by the finding that isolated early NPs generate mainly TBR1 and $CTIP2^+$ neurons at the expense of $SATB2^+$ upper layer neurons in the absence of feeder cell-derived signals. These findings are consistent with xenografting experiments where transplanted human NPs integrated into host mouse brains and generated mostly $SATB2^+$ upper layer neurons[47].

However, it remained unclear how upper neuronal fates potential is controlled. Here we show that SATB2 fate is acquired by NPs over time in the cultures. Early NPs are unable to generate $SATB2^+$ neurons within a 20 day period, even when exposed to the environment of the late neurogenic period at day 90 of differentiation *in vitro*. Furthermore, grafting of early NP onto late stage feeder cells did not result in an immediate switch in fate to SATB2. Rather, early NPs continued to generate few $SATB2^+$ neurons. This suggests that SATB2 neuron differentiation is not immediately

directed by signals produced by the later cultures and implies that progenitors must undergo an intrinsic developmental maturation. This maturation is likely due to exposure to accumulating signals and cue in the developing cortex. The co culture experiments suggest that the progeny of the NPs are likely the source of these instructive cues but intrinsic changes within the NPs themselves and how they respond to certain signals may also play a role. Addition analyses will be required to identify and study these signals.

Conversely, we show that early neuronal fate potential of human NPs is not lost with developmental age. When day 90 NPs are grafted onto a day 25 feeder layer that represents the expansion and early neurogenesis phases of cortical development, their neurogenic differentiation is repressed. This suggests that regulation of cortical development involves signals that not only promote upper layer fate but also repress differentiation. However, and unlike day 25 NPs, day 90 NPs do not form neural rosettes or even integrate into the polarized rosettes formed by the feeder cells. The day 90 NP population have an oRG-like transcriptional profile and, therefore, it seems that even when placed into and environment that resemble the VZ of the developing cortex, they do not transform back to a vRG-like phenotypes, at least in terms of the rosette-forming behavior.

The repressive effect on neurogenesis of early cortical cultures was most evident when day 90 NPs, which have already acquired an intrinsic ability to generate upper layer neurons, were plated onto day 25 feeder cells. SATB2^{+} neuron production was completely suppressed. When placed in a neutral environment on poly-L-lysine/Laminin, and isolated from signals propagated in the cultures, day 90 NPs predominantly generated deep layer neurons (TBR1 and CTIP2^{+}). This suggests that TBR1 neuronal fate is intrinsic to NPs and the population retains the ability to generate TBR1 positive neurons irrespective of developmental stage.

Moreover, although day 90 NPs did not revert to the same rosette forming potential as day 25 NPs on early feeders, their neuronal differentiation was reprogrammed to that of early NPs on early feeders. Therefore, late human NPs are able respond to the local signals and adopt early potential *in vitro*. This is somewhat in contrast to the mouse and ferret NPs which loose potential with age[37,57]. Here we show that human NPs display a great cellular heterogeneity and complexity during cortical neurogenesis[53].

Our findings imply that human cortical NPs use different strategies to assure the sequential and ordered formation of cortical neurons. Firstly, early NPs have an

intrinsic potential to generate deep layer neurons but not upper layer neurons. Upper layer neuronal fate must be acquired over time and is not the direct consequence of cues that shift differentiation from early to late fate. In addition, our grafting experiments indicate that the environment of the early cerebral cortex repress neuronal differentiation, presumably enabling the NP population to expand and generate sufficient precursors to ensure enough neurons can be generated. We also show that once upper neuronal fate is acquired by the NP pool, they do not lose the ability to generate deep layer neurons. Rather, the environment of the later neurogenic cortex is promiscuous for SATB2 neuron production. This bimodal regulation of cortical neurogenesis during human cortical development is assumably necessary to control the massive expansion and prolonged neuron production in humans compared to rodents and none primate mammals. The challenge is now to identify the extrinsic and intrinsic signals that control the various transitions and competence states of human cortical NPs.

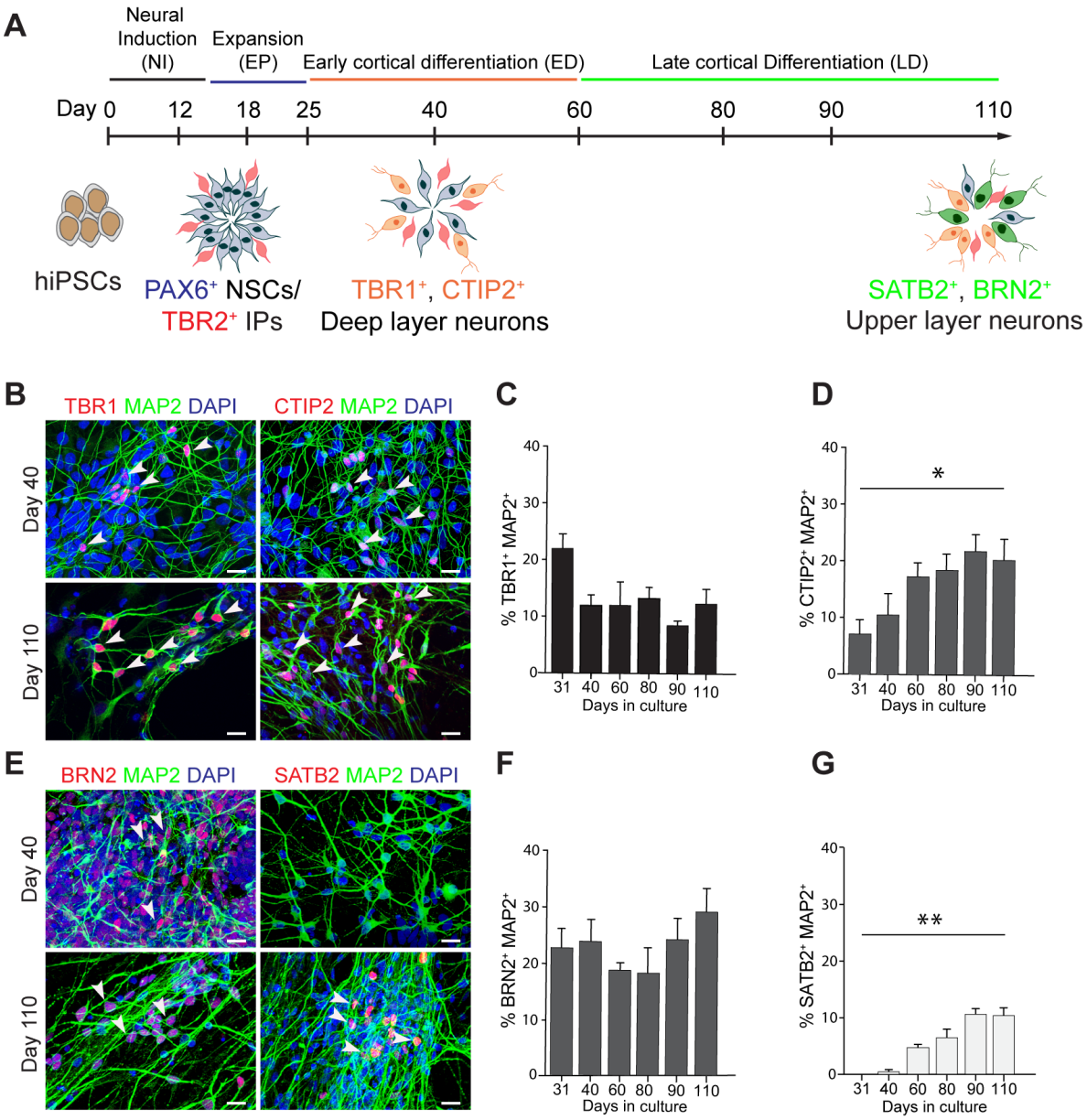
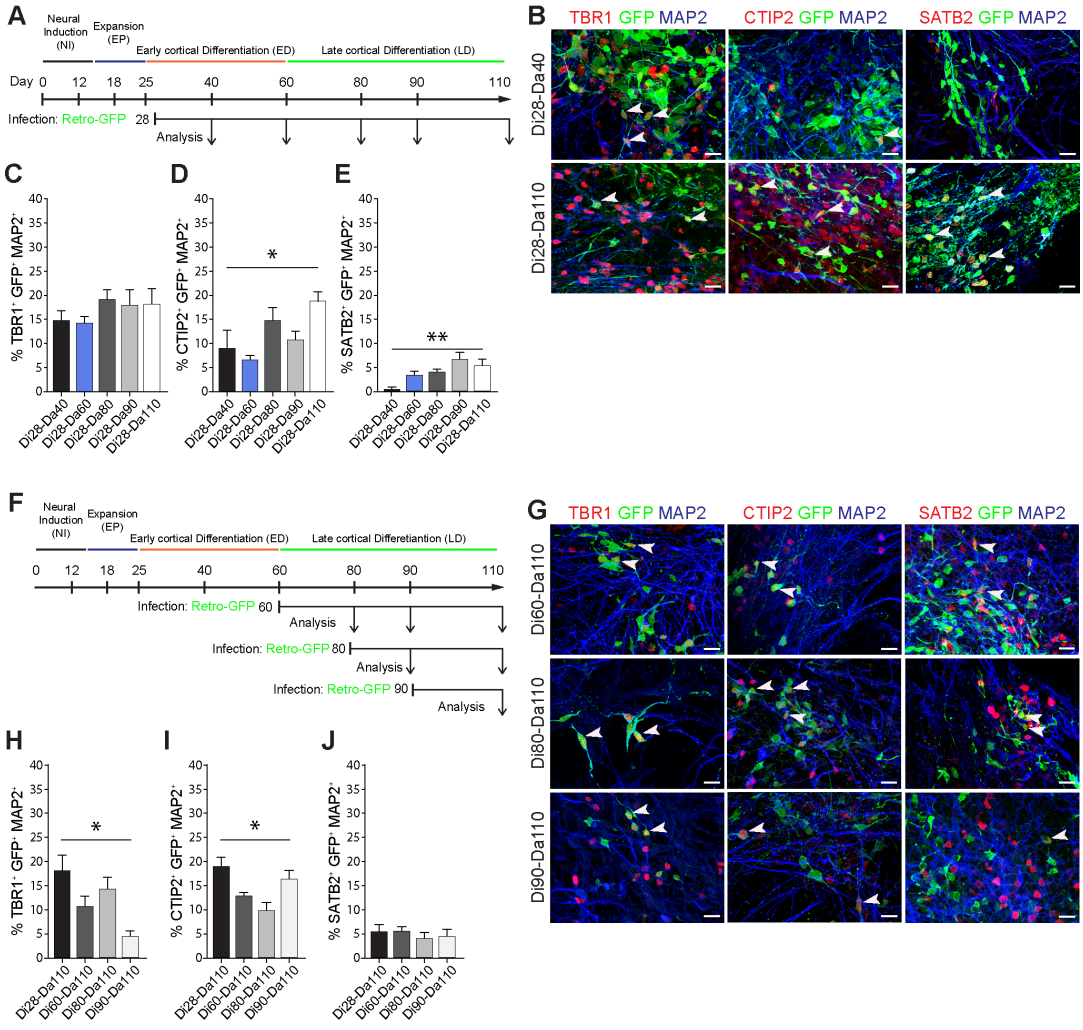


Figure 1: ***In vitro*** **Differentiation of hiPSC to different subtypes of cortical neurons**

A. Diagram of the differentiation experimental paradigms of hiPSCs to cortical neurons.
B. Representative pictures of layer V and VI neurons expressing TBR1 and CTIP2 at day 40 and 110 of differentiation (Scale bar: 20 µm).
C-D. Quantification of proportion of TBR1 and CTIP2 expressing neurons over time (n=3 experiment for each marker). Neurons of layer V and VI were generated at early stage. Kruskal-Wallis with Dunn post hoc test: *p < 0.05, **p < 0.01
E. Representative pictures of BRN2 and SATB2-expressing neurons of layer II and III at day 40, 110 of differentiation. (Scale bars: 20 µm).
F-G. Quantification of proportion BRN2 and SATB2 upper layer neurons over time (n=3 experiment for each marker). Neurons of layer II and III (SATB2-positive neurons) are produced after deep layer neurons. Kruskal-Wallis test:*p < 0.05.

Figure 2: Birth dating of neuronal subtypes using retrovirus labeling of neural progenitors

A. Diagram of birth dating experiments using a GFP-expressing retrovirus to lineage-trace day 28 NPs *in vitro*.
B. Representative images of infected cells at day 28 and analysis at day 40 and 110 of differentiation. (Scale bars: 20 µm).
C-E. Quantification of TBR1, CTIP2, SATB2 expressing neurons generated from labeled progenitors at different stages of differentiation (day 28, 60, 80, 90) at day 110.
F. Diagram of birth dating experiments using a GFP-expressing retrovirus to lineage-trace day 6 , 80, 90 NPs *in vitro*.
G. Representative images of infection at day 60, 80, 90 and analysis at day 110.
H-J. Quantification of TBR1, CTIP2, and SATB2 expressing neurons generated from labeled progenitors at different stages of differentiation (day 28, 60, 80, 90) at day 110.



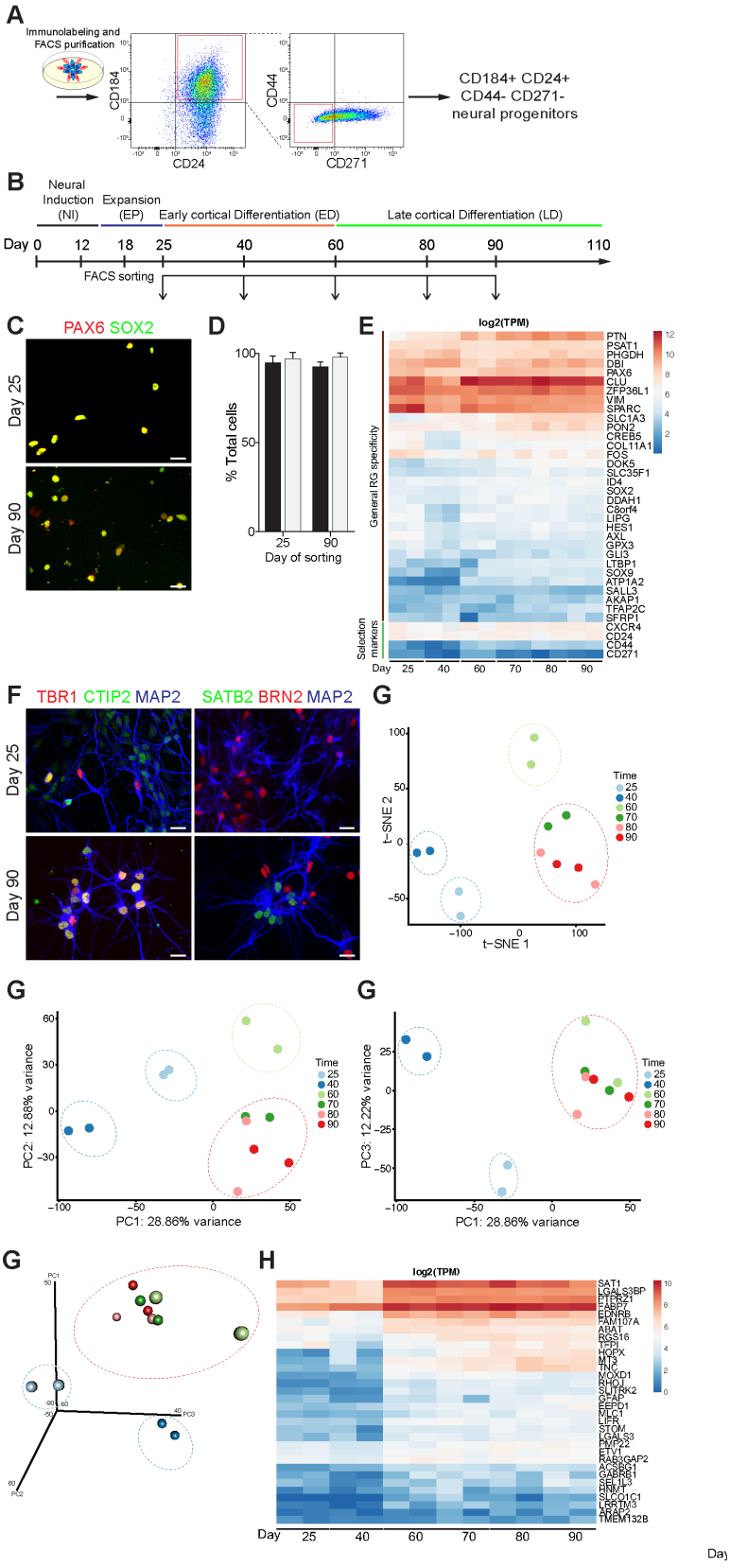


Figure 3: Characterization of FACS isolated neural progenitors based on CD184^{+}, CD24^{+}, CD44^{-}, CD271^{-}

A. Strategy of FACS sorting of neural progenitors at day 25 of differentiation
B. Sequential isolation of neural stem/progenitors based on CD184, CD24, CD44, CD271.
C. Representative images of sorted cells at day 25 and 90 and stained after few hours for PAX6, SOX2 (Scale bars: 20 µm).
D. Over 90 percent of cells are expressing PAX6 and SOX2 as prominent markers of cortical neural stem cells
E. Gene expression profile of neural stem cell at day 25 of neural differentiation using RNA-seq.
F. Representative images of differentiation of sorted neural progenitors after 20 days of differentiation (Scale bars: 20 µm).
G. Principal component analysis revealed that neural progenitors are changing from day 25 to day 90 of differentiation
H. Enrichment of ORG markers throughout the differentiation.

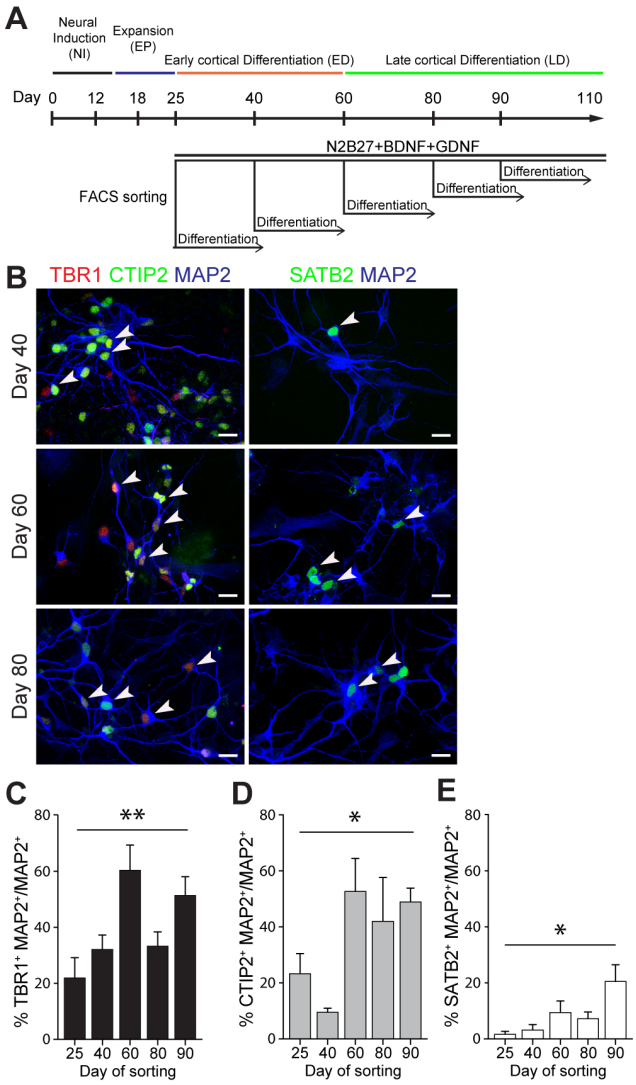


Figure 4: late progenitors generated more SATB2 neurons compared with early progenitors

(A). Diagram of sequential isolation of neural stem/progenitors based on CD184, CD24, CD44, CD271.

(B). Immunostaining for deep and upper layer markers in combination with MAP2 neuronal marker after 20 days differentiation

(C-E). Bar charts represent proportion of deep and upper layer neurons derived from early and late progenitors after 20 days of differentiation. Kruskal-Wallis with

Dunn post hoc test: *p < 0.05, **p < 0.01

Data are mean ± SEM. The scale bars represent 20 μm in (B).

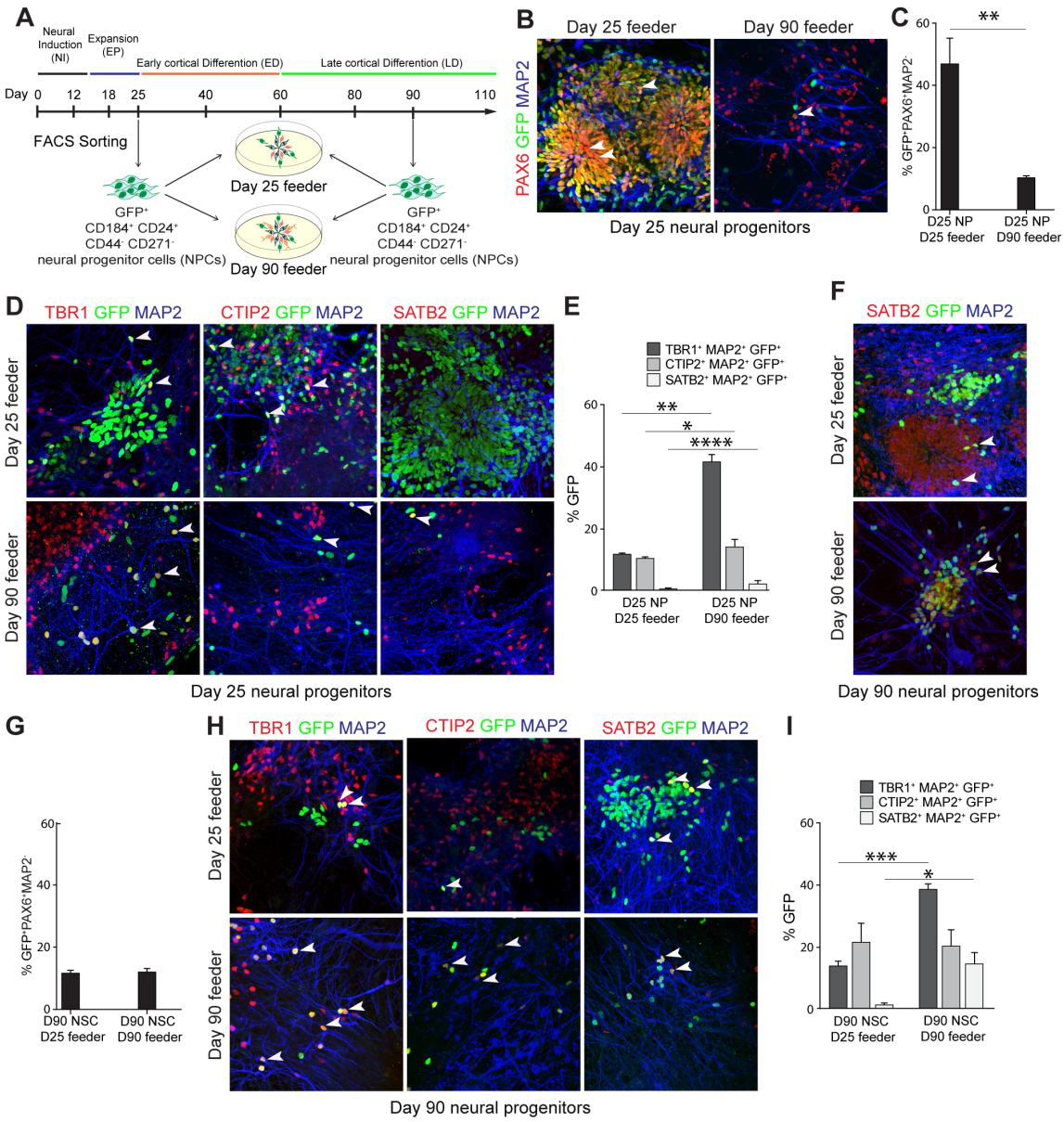
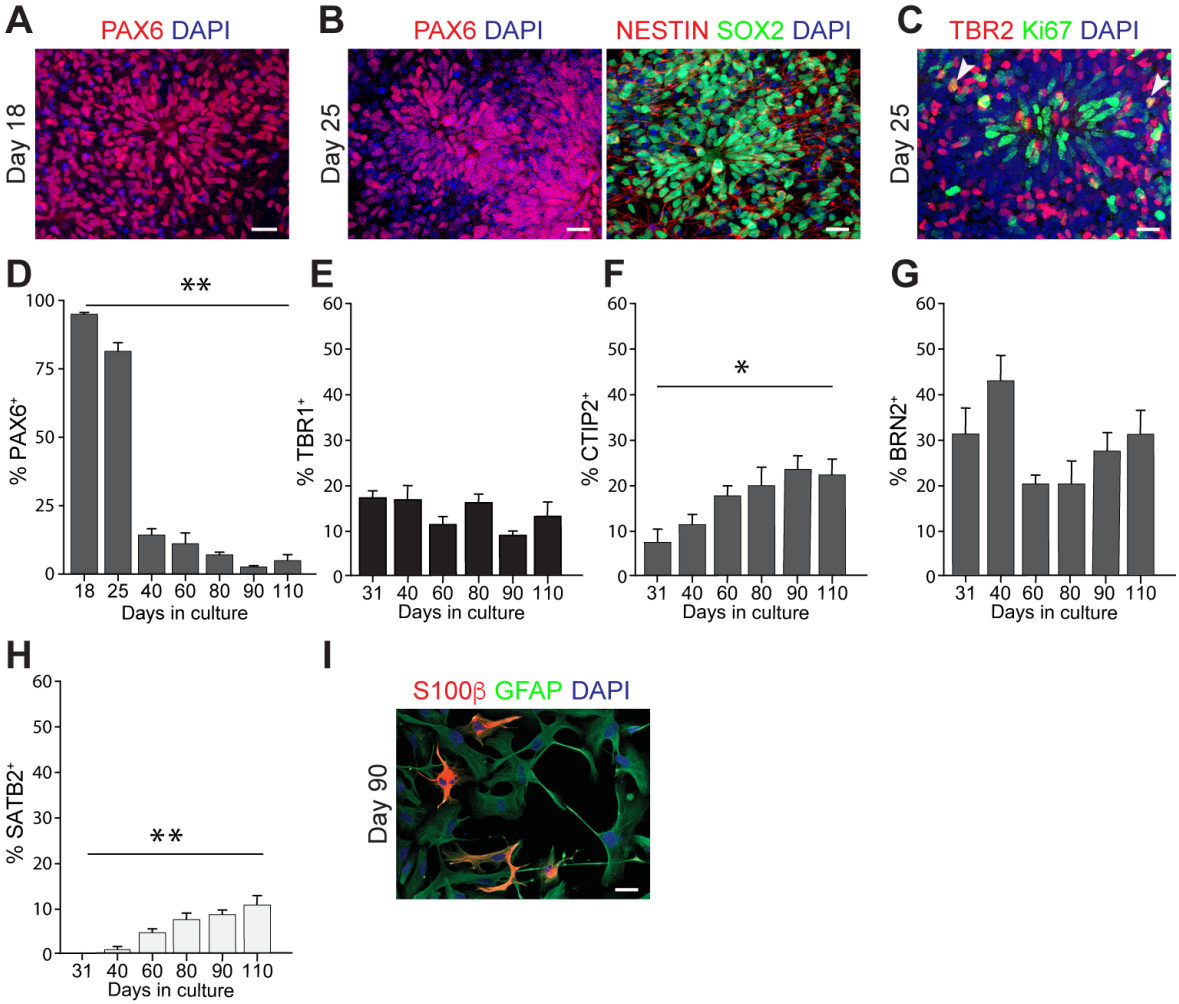


Figure 5: The early stage environment reverts the differentiation potential of late progenitors

(A) Experimental design for isolation of neural progenitors at day 25 and 90 and plate them on cortical cell layers (day 25 and 90cortical differentiation culture derived from hiPSCs).
(B) Representative image of plated progenitors and stained for PAX6 on the feeder layer of day 25 and 90.
(C) Quantification of proportion of PAX6$^+$ cells derived from labeled progenitors. Two-sided Student's t test: *p < 0.05, **p < 0.01.
(D) Representative images of sorted progenitors and stained for TBR1, CTIP2 and SATB2 on the feeder layer of day 25 and 90.
(E) Quantification of proportion of TBR1, CTIP2 and SATB2 neurons derived from labeled progenitors. Two-sided Student's t test: *p < 0.05, **p < 0.01, ***p < 0.001.
(F) Representative images of plated progenitors and stained for PAX6 on the feeder layer of day 25 and 90.
(G) Quantification of proportion of PAX6 neurons derived from labeled progenitors.
(H). Representative images of sorted progenitors and stained for TBR1, CTIP2 and SATB2 on the feeder layer of day 25 and 90.
(I). Quantification of proportion of TBR1, CTIP2 and SATB2 neurons derived from labeled progenitors. Two-sided Student's t test: *p < 0.05, **p < 0.01, ***p < 0.001.
Data are mean ± SEM. The scale bars represent 20 µm in (B,D,F,H).

Supplementary Information: Supplementary Figures and Legends

Figure S1: Differentiation of hiPSC to different subtypes of cortical neurons



A) Representative images of neural stem cells derived from directed differentiation of pluripotent stem cell to cortical neurons at day 18 of differentiation.
B) Immunostaining for PAX6, SOX2 and NESTIN as primary cortical neural stem cell markers.
C) Immunostaining for TBR2 as basal progenitor markers at day 25 of differentiation.
D) The proportion of PAX6$^+$ cells over the time of differentiation. Bar chart represent over 95% of cells are PAX6$^+$ neural stem cells at day 18.
E-H) Proportion of cells expressing TBR1, CTIP2, BRN2, SATB2 at different stages of differentiation (day 40, 80, 90,110). Quantification from 3 independent experiments.
I) Immunofluorescence image of Astrocytes appeared at late stage of differentiation.
Mean +/- SEM (N=3). (Scale bars: 20

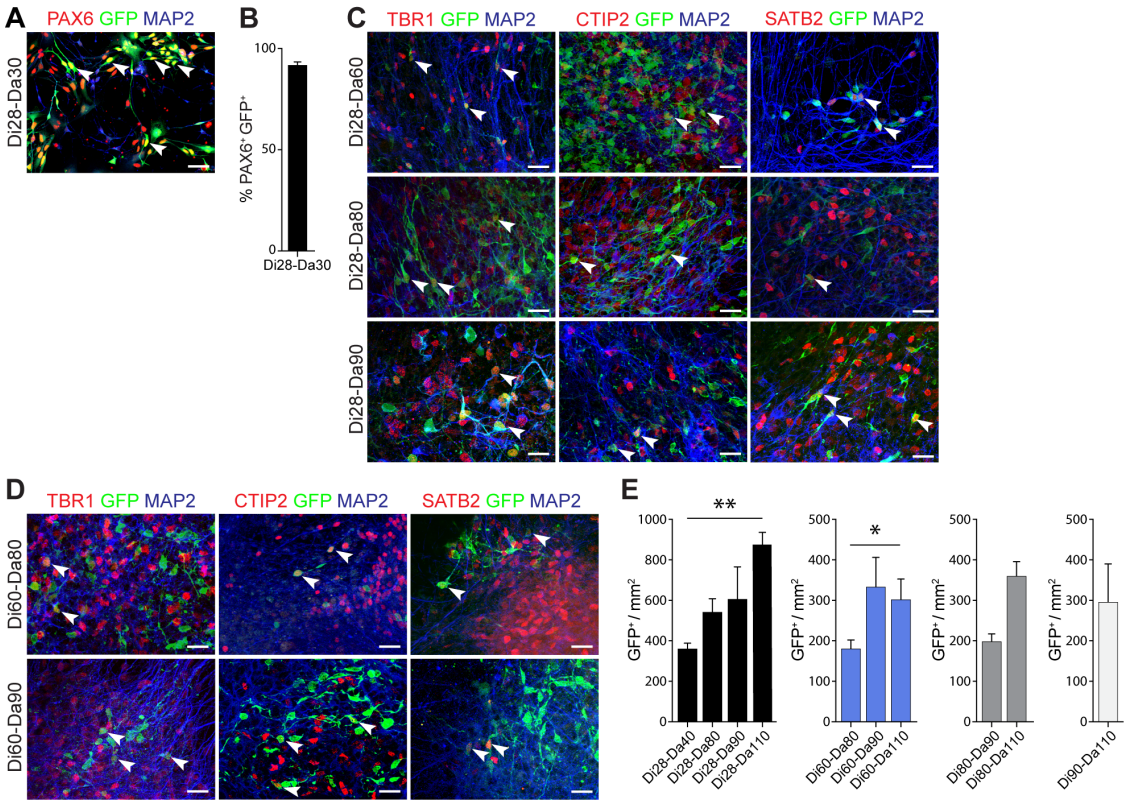


Figure S2: **labeling neural progenitors using retrovirus**

A-B) Representative image of retrovirus infected cells at day 28 and analyzed at day 30. The quantification of GFP labeled cells, which are PAX6^{+} neural stem cells.
C) Immunofluorescence images of infection at day 28 and fixed at day 60, 80 and 90.
D) Representative images of infection at day 60 and fixed at day 80 and 90 and 110.
E) Quantification of the number of GFP^{+} cells/mm^{2} after infection at different time points and differentiation.

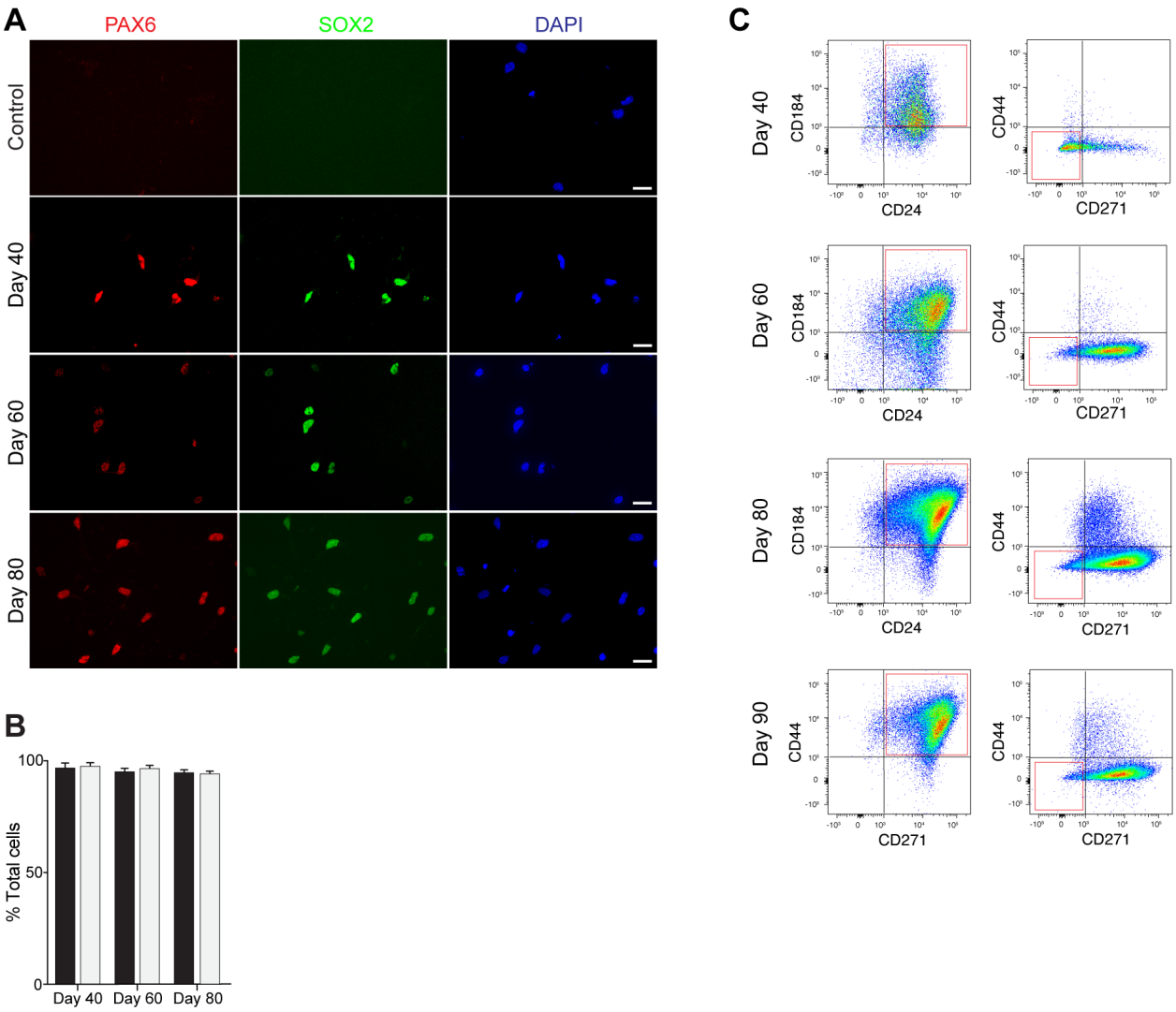


Figure S3: **Isolation the neural progenitors at different stages of differentiation by FCAS**

A) Immunostaining of sorted neural progenitors for PAX6 and SOX2.
B) Quantification of sorted cells (day 40, 60 and 80) which express PAX6 and SOX2.
C) The FACS plot of sorting the neural progenitors at different time points: day 40, 60, 80 and 90.

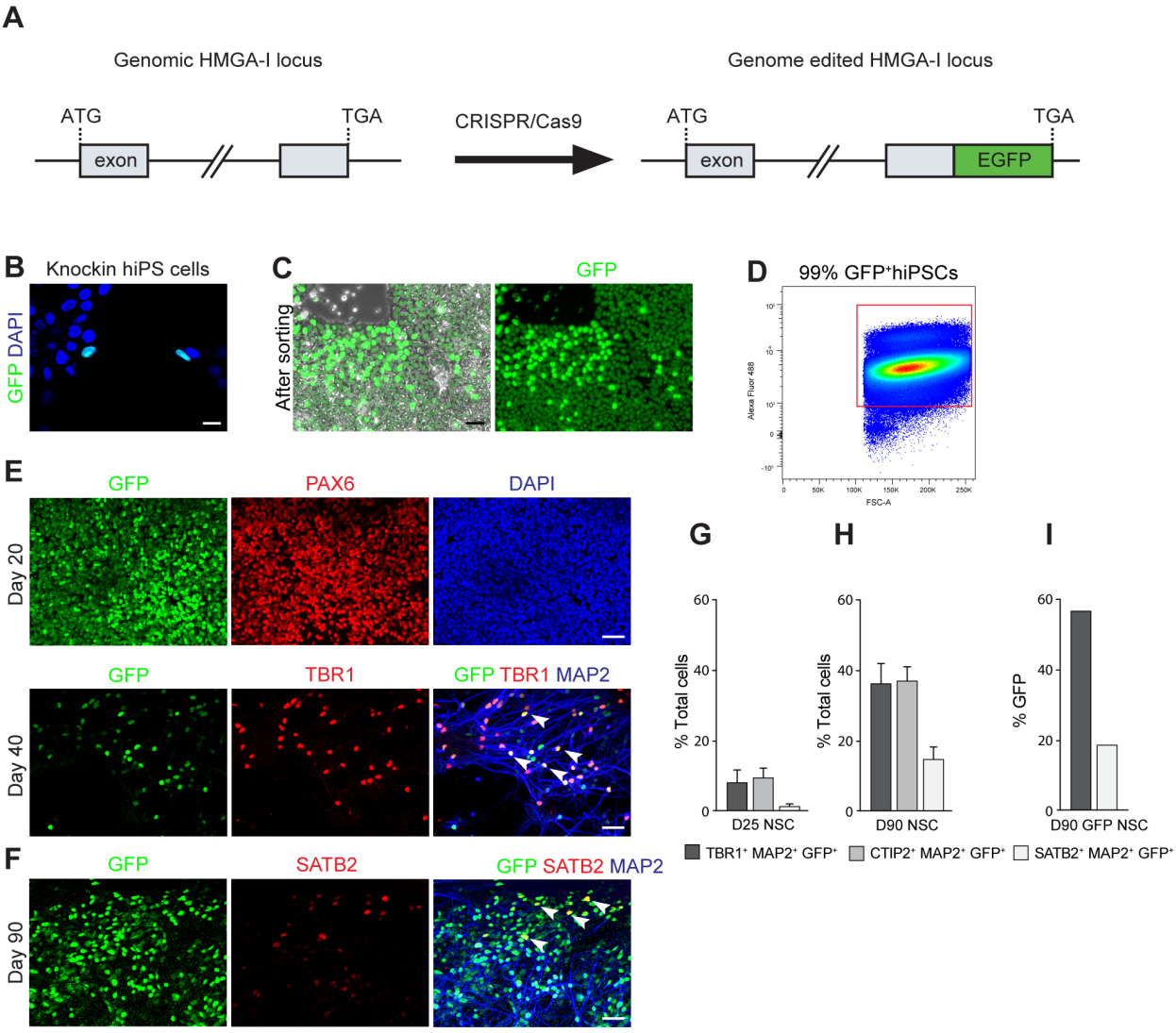


Figure S4:

(A) Plasmid and targeting strategy used in this study in order to label hips cells in HMGA1 locus.
(B) Edited cells after transduction.
(C-D) After FACS sorting and expansion of labeled hiPSCs most of cells were GFP^{+}.
(E) Representative images of differentiation of GFP labeled hiPSCs to neural stem cells expressing PAX6 at day 20 and deep layer neurons at day 40 and upper layer neurons at day 90.
(G-I) Quantification of differentiated neurons from sorted progenitors at day 25 and 90.

References

1 Tomassy, G. S., Lodato, S., Trayes-Gibson, Z. & Arlotta, P. Development and regeneration of projection neuron subtypes of the cerebral cortex. *Sci Prog* **93**, 151-169 (2010).

2 Franco, S. J. & Muller, U. Shaping our minds: stem and progenitor cell diversity in the mammalian neocortex. *Neuron* **77**, 19-34, doi:10.1016/j.neuron.2012.12.022 (2013).

3 Pinto, L. & Gotz, M. Radial glial cell heterogeneity--the source of diverse progeny in the CNS. *Prog Neurobiol* **83**, 2-23, doi:10.1016/j.pneurobio.2007.02.010 (2007).

4 Qian, X., Goderie, S. K., Shen, Q., Stern, J. H. & Temple, S. Intrinsic programs of patterned cell lineages in isolated vertebrate CNS ventricular zone cells. *Development* **125**, 3143-3152 (1998).

5 McConnell, S. K. & Kaznowski, C. E. Cell cycle dependence of laminar determination in developing neocortex. *Science* **254**, 282-285 (1991).

6 Frantz, G. D. & McConnell, S. K. Restriction of late cerebral cortical progenitors to an upper-layer fate. *Neuron* **17**, 55-61 (1996).

7 Desai, A. R. & McConnell, S. K. Progressive restriction in fate potential by neural progenitors during cerebral cortical development. *Development* **127**, 2863-2872 (2000).

8 Franco, S. J. *et al.* Fate-restricted neural progenitors in the mammalian cerebral cortex. *Science (New York, N.Y.)* **337**, 746-749, doi:10.1126/science.1223616 (2012).

9 Lui, J. H., Hansen, D. V. & Kriegstein, A. R. Development and evolution of the human neocortex. *Cell* **146**, 18-36, doi:10.1016/j.cell.2011.06.030 (2011).

10 Zecevic, N., Chen, Y. & Filipovic, R. Contributions of cortical subventricular zone to the development of the human cerebral cortex. *The Journal of comparative neurology* **491**, 109-122, doi:10.1002/cne.20714 (2005).

11 Hansen, D. V., Lui, J. H., Parker, P. R. & Kriegstein, A. R. Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* **464**, 554-561, doi:10.1038/nature08845 (2010).

12 Johnson, M. B. *et al.* Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat Neurosci* **18**, 637-646, doi:10.1038/nn.3980 (2015).

13 Lui, J. H. *et al.* Radial glia require PDGFD-PDGFRbeta signalling in human but not mouse neocortex. *Nature* **515**, 264-268, doi:10.1038/nature13973 (2014).

14 Florio, M. *et al.* Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science* **347**, 1465-1470, doi:10.1126/science.aaa1975 (2015).

15 Pollen, A. A. *et al.* Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55-67, doi:10.1016/j.cell.2015.09.004 (2015).

16 Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199-206, doi:10.1038/nature13185 (2014).

17 Espuny-Camacho, I. *et al.* Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo. *Neuron* **77**, 440-456, doi:10.1016/j.neuron.2012.12.011 (2013).

18 Gaspard, N. *et al.* Generation of cortical neurons from mouse embryonic stem cells. *Nat Protoc* **4**, 1454-1463, doi:10.1038/nprot.2009.157 (2009).

19 Gaspard, N. *et al.* An intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature* **455**, 351-357, doi:10.1038/nature07287 (2008).

20 van de Leemput, J. *et al.* CORTECON: a temporal transcriptome analysis of in vitro human cerebral cortex development from human embryonic stem cells. *Neuron* **83**, 51-68, doi:10.1016/j.neuron.2014.05.013 (2014).

21 Edri, R. *et al.* Analysing human neural stem cell ontogeny by consecutive isolation of Notch active neural progenitors. *Nat Commun* **6**, 6500, doi:10.1038/ncomms7500 (2015).
22 Yao, Z. *et al.* A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development. *Cell Stem Cell* **20**, 120-134, doi:10.1016/j.stem.2016.09.011 (2017).
23 Shi, Y., Kirwan, P., Smith, J., Robinson, H. P. & Livesey, F. J. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nat Neurosci* **15**, 477-486, S471, doi:10.1038/nn.3041 (2012).
24 Bershteyn, M. *et al.* Human iPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia. *Cell Stem Cell* **20**, 435-449 e434, doi:10.1016/j.stem.2016.12.007 (2017).
25 Chambers, S. M. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* **27**, 275-280, doi:10.1038/nbt.1529 (2009).
26 Qian, X. *et al.* Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. *Neuron* **28**, 69-80, doi:10.1016/S0896-6273(00)00086-6 (2000).
27 Shen, Q. *et al.* The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nat Neurosci* **9**, 743-751, doi:10.1038/nn1694 (2006).
28 Yuan, S. H. *et al.* Cell-surface marker signatures for the isolation of neural stem cells, glia and neurons derived from human pluripotent stem cells. *PLoS One* **6**, e17540, doi:10.1371/journal.pone.0017540 (2011).
29 Lukaszewicz, A. *et al.* G1 phase regulation, area-specific cell cycle control, and cytoarchitectonics in the primate cortex. *Neuron* **47**, 353-364, doi:10.1016/j.neuron.2005.06.032 (2005).
30 Smart, I. H., Dehay, C., Giroud, P., Berland, M. & Kennedy, H. Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. *Cerebral cortex (New York, N.Y. : 1991)* **12**, 37-53 (2002).
31 Mansouri, M., Ehsaei, Z., Taylor, V. & Berger, P. Baculovirus-based genome editing in primary cells. *Plasmid* **90**, 5-9, doi:10.1016/j.plasmid.2017.01.003 (2017).
32 Rakic, P. Neurons in rhesus monkey visual cortex: systematic relation between time of origin and eventual disposition. *Science* **183**, 425-427 (1974).

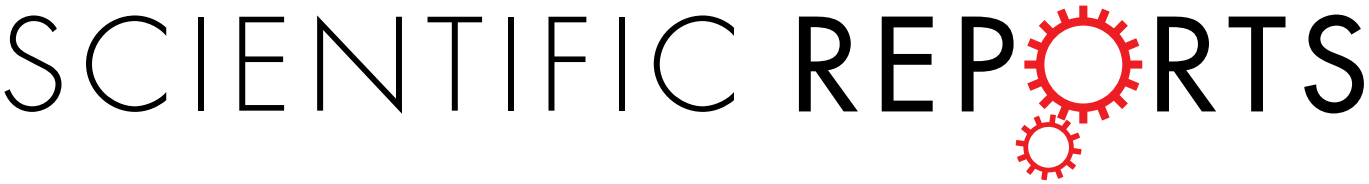
2-2 "TDP-43 induces p53-mediated cell death of cortical stem cells and neurons"

TDP43 is nucleotide-binding protein regulating cellular process including gene transcription, RNA splicing, mRNA stability and microRNA biogenesis. Previous studies showed that point mutations and accumulation of TDP43 protein aggregates is associated with neuro-degenerative diseases such as Amyotrophic lateral sclerosis (ALS), Frontotemporal lobar degeneration (FTD) and Alzheimer's disease. Although TDP 43 is considered the trigger of neurodegeneration in these diseases, molecular mechanisms underlying TDP 43 dysfunction remain to be elusive.

In present study Vogt. et al investigated the role of TDP-43 during telencephalic development. Vogt. et al reported that the overexpression of wild type and mutant TDP43, and the TDP-43A315T ALS- associated mutation, induces p53 dependent cell death of neural stem cells and cortical neurons. Moreover, the early lethality of this mutation was rescued by pharmacological inhibition of p53.

Using human iPSC derived cortical differentiation cultures we confirmed the toxicity of TDP-43 overexpression. TDP43 wt and mutant overexpression induced cell death of human iPSCs derived cortical neurons. Furthermore, the over expression of TDP43 led to up-regulation of proapoptotic genes including TRP53, CDKN1A, BBC3 and BAX in the human cells. Importantly, this effect can be rescued by the p53 inhibitor (PFT-alpha). In accordance, we observed p53-dependent apoptosis in cortical differentiated culture derive from the patient hiPSCs (TDP-43^{G298S} mutant). Indeed we showed the over expression of proapoptotic genes including Trp53, Cdkn1a, Bbbc3 and Bax in the human cells derived from pluripotent cells.

For this publication I performed the experiments summarized in figures 6 and 7 and I created the corresponding figures. Furthermore, I generated the human iPSCs derived neural differentiation culture and was involved in editing the manuscript.



OPEN

TDP-43 induces p53-mediated cell death of cortical progenitors and immature neurons

Received: 27 November 2017
Accepted: 9 May 2018
Published online: 25 May 2018

Miriam A. Vogt[1,7], Zahra Ehsaei[1], Philip Knuckles[2], Adrian Higginbottom[3], Michaela S. Helmbrecht[4], Tilo Kunath[5], Kevin Eggan[6], Luis A. Williams[6], Pamela J. Shaw[3], Wolfgang Wurst[4], Thomas Floss[4], Andrea B. Huber[4,8] & Verdon Taylor[1]

TAR DNA-binding protein 43 (TDP-43) is a key player in neurodegenerative diseases including frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS). Accumulation of TDP-43 is associated with neuronal death in the brain. How increased and disease-causing mutant forms of TDP-43 induce cell death remains unclear. Here we addressed the role of TDP-43 during neural development and show that reduced TDP-43 causes defects in neural stem/progenitor cell proliferation but not cell death. However, overexpression of wild type and TDP-43^{A315T} proteins induce p53-dependent apoptosis of neural stem/progenitors and human induced pluripotent cell (iPS)-derived immature cortical neurons. We show that TDP-43 induces expression of the proapoptotic BH3-only genes *Bbc3* and *Bax*, and that p53 inhibition rescues TDP-43 induced cell death of embryonic mouse, and human cortical neurons, including those derived from TDP-43^{G298S} ALS patient iPS cells. Hence, an increase in wild type and mutant TDP-43 induces p53-dependent cell death in neural progenitors developing neurons and this can be rescued. These findings may have important implications for accumulated or mutant TDP-43 induced neurodegenerative diseases.

The nucleotide binding protein TDP-43 regulates a multitude of cellular processes including gene transcription, RNA splicing, mRNA stability, localization and translation, and microRNA biogenesis[1]. TDP-43 is associated with most cases of FTLD and ALS. Point mutations and elevated TDP-43 levels cause FTLD-TDP coupled with neuronal death. Affected neurons have nuclear depletion and cytoplasmic accumulation of TDP-43. Interestingly, accumulation of TDP-43 in neurons is characteristic of many neurodegenerative diseases in addition to FTLD-TDP and ALS, including Alzheimer's disease. TDP-43 seems to play roles in the onset and progression of neurodegeneration but mechanistically how it triggers and contributes to disease and cell death is not known.

TDP-43 was originally identified as a factor binding to the TAR DNA of human immunodeficiency virus where it is implicated in transcriptional regulation[2]. It belongs to the family of heterogeneous nuclear ribonucleoproteins (hnRNP) and is ubiquitously expressed. Deletion of the *Tardbp* (the gene encoding TDP-43) in mice leads to early embryonic lethality between E3.5 and E6.5[3,4] indicating an important role during early development. TDP-43 contains two RNA-recognition motifs (RRMs), RRM-1 and RRM-2, and a glycine-rich region at its C-terminus[5]. The RRM-1 is necessary and sufficient for nucleic acid binding to single stranded RNA at GU-rich sequences[6]. The C-terminus of TDP-43 is necessary for the formation of hnRNP-rich complexes[7] and contains most of the TDP-43 point mutations identified in FTLD and ALS patients. TDP-43 is localized predominantly nuclear in cells and has both a nuclear localization sequence (NLS) and a predicted nuclear export sequence (NES) and seems to be continuously shuttled between the two cellular compartments[8].

[1]Department of Biomedicine, University of Basel, Mattenstrasse 28, 4058, Basel, Switzerland. [2]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058, Basel, Switzerland. [3]Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield, S10 2HQ, UK. [4]Helmholtz Zentrum München, Ingolstädter Landstrasse 1, 85764, Neuherberg, Germany. [5]MRC Centre for Regenerative Medicine, The University of Edinburgh, 5 Little France Drive, Edinburgh, EH16 4UU, UK. [6]Harvard Stem Cell Institute, Harvard University, Howard Hughes Medical Institute, 7 Divinity Avenue, Cambridge, MA, 02138, USA. [7]Present address: Ludwig-Maximilians University Munich, Feodor-Lynen-Strasse 17, 81377, München, Germany. [8]Present address: ETH Zurich, Department of Biosystems Science and Engineering, Mattenstrasse 26, 4058, Basel, Switzerland. Correspondence and requests for materials should be addressed to V.T. (email: verdon.taylor@unibas.ch)

TDP-43 is one of the main components of cytoplasmic inclusions, which are a characteristic feature of many neurodegenerative disorders. Apoptotic neurons that display cytoplasmic inclusions show a partial loss of TDP-43 in the nucleus[9], which was suggested to drive, at least in part, disease pathogenesis. However, the cause and function of TDP-43 aggregates remains to be shown unequivocally. In mice, robust cytoplasmic TDP-43 aggregation is associated with dramatic neuron loss and features of human pathology, which can be reversed by increased clearance of TDP-43[10]. Interestingly, mislocalization of TDP-43 to the cytoplasm of mouse neurons is sufficient to induce apoptosis even in the absence of aggregation, suggesting that cytoplasmic TDP-43 aggregates may not be necessary to induce cell death and early mortality in mice[9,11–13].

Aberrant TDP-43 causes pleiotropic effects in cells and results in extensive changes in splicing and RNA metabolism[14]. Cross-linked immunoprecipitation and RNA sequencing (CLIP-Seq) revealed that TDP-43 can bind thousands of RNAs via a UG-rich consensus sequence in the 3′ untranslated regions of target RNAs[15–17]. Whereas the RNAs bound by TDP43 in the mouse brain are relatively consistent in the different analyses, TDP-43 targets vary considerably between cell types[16,17].

Aggregates in diseased neurons contain hyper-phosphorylated and fragmented TDP-43 protein. Interestingly, TDP-43 can be cleaved by caspases[18], and other factors of the apoptosis pathway including Bim, Bax and Bcl may be involved in TDP-43-induced cell death[19]. Components of the proapoptotic pathway are downstream targets of p53 and elevated p53 levels have been detected in affected neurons of ALS patients[20,21]. However, the absence of p53 in a transgenic mouse model for ALS ($hSOD1^{G93A}$) did not rescue apoptosis, suggesting that cell death in these animals occurred in a p53-independent manner[22,23]. Although aberrant TDP-43 expression is associated with stress responses[24], a causal link between p53 and TDP-43 induced cell death has not been reported. TDP-43 is expressed in the developing and adult brain, therefore, we addressed the role of TDP-43 during development of the telencephalon by gain- and loss-of-function experiments. We thereby hoped to gain insights into TDP-43 functions in the formation and maintenance of the nervous system.

Here we show that expression of TDP-43 and its mutant form $TDP\text{-}43^{A315T}$ results in p53-mediated apoptosis in neural stem/progenitor cells and immature neurons of the developing mouse telencephalon. In addition, we observed cell death of cortical neurons derived from human iPS cells following TDP-43 expression and found that this neuronal death could also be rescued by p53-inhibition. Expression of the proapoptotic BH3-only genes *Bbc3* and *Bax* was increased in mice and human neural cells as a result of aberrant TDP-43 expression, supporting a role for p53 in the TDP-43 induced cell death. Furthermore, we show that TDP-43 is associated with the mRNA of *Cdkn1a* and increases Cdkn1a levels, likely explaining the altered neural stem/progenitor cell cycle regulation following TDP-43 and $TDP\text{-}43^{A315T}$ expression.

Results

TDP-43 controls cell cycle, neurogenesis and is toxic for neural progenitors. *Tardbp* is expressed by neural progenitors in the developing central nervous system (Supplementary Fig. 1a)[3]. In the developing telencephalon at e14.5, TDP-43 protein is prominent in ventricular zone stem/progenitor cells including by those in M-phase of the cell cycle at the ventricular surface where expression partially overlaps with phospho-Histone H3 (pH3) (Fig. 1a, arrowheads, Supplementary Fig. 1b). TDP-43 protein expression is also prominent in differentiating neurons in the cortical plate (Fig. 1b).

To address the functions of TDP-43 during forebrain development, we knocked-down (KD) TDP-43 *in vivo* targeting the ventricular zone stem and progenitor cells by *in utero* electroporation (IUE) with hairpin RNAs targeting its mRNA (shTDP-43). TDP-43 mRNA and protein levels were reduced in neural progenitors by up to ~70% 48 hours after transfection (Fig. 1c, Supplementary Fig. 2a–d). The percentage of transfected cells that expressed detectable levels of TDP-43 *in vivo* reduced from 93.9 +/− 4.2% in controls to 3.7 +/− 2.5% in shTDP-43 transfected embryos (Supplementary Fig. 2c,d). TDP-43 KD increased the number of progenitors in M-phase within 24 hours ($pH3^+$: Fig. 1d,e), as well as those entering S-phase (BrdU labeled) (Fig. 1f). These effects were consistent with the expression of TDP-43 by mitotic cells in the developing telencephalon. TDP-43 KD also reduced the number of $Pax6^+$ ventricular zone progenitors and increased the generation of $Tbr2^+$ basal progenitors slightly ($p = 0.07$) (Fig. 1d,g,h and Supplementary Fig. 2e–h). After 48 hours, the increase in $BrdU^+$ cells in the ventricular zone following TDP-43 KD was still evident but the formation of $Tbr2^+$ basal progenitors had returned to normal and there were no obvious effects on cell survival (Supplementary Fig. 2i,j and data not shown).

Conversely, the moderate (2.4-fold) increase in wild type TDP-43 expression induced by IUE (Fig. 2a,b) resulted in an increase in active caspase-3 expressing cells and rapid apoptotic death of most neural progenitors *in vivo* (Fig. 2c,d). TDP-43 expression reduced not only the absolute number but also the proportion of $Pax6^+$ cells remaining in the developing telencephalon (Fig. 2e). The few surviving TDP-43 overexpressing cells that were undergoing mitosis were displaced from their normal position at the ventricular lining (Fig. 2f,g). Furthermore, the differentiation of neural stem/progenitor cells into basal progenitors ($Tbr2^+$) was reduced and generation of newborn neurons ($Tbr1^+$) in the cortical plate was blocked by TDP-43 overexpression indicating cell death and impaired neuronal differentiation (Fig. 2h,i and Supplementary Fig. 3). These data suggested that overexpression of wild type TDP-43 in neural progenitors of the developing telencephalon is toxic, induces changes in cell proliferation and blocks neurogenesis.

Mutant $TDP\text{-}43^{A315T}$ induces p53-dependent cell death. Autosomal dominant mutations in human TDP-43 are sufficient to cause familial ALS and FTLD. Disease causing mutations occur predominantly in the C-terminal prion-domain of TDP-43[25]. The mechanisms of mutant TDP-43 induced toxicity remain unclear but are associated with an increase in TDP-43 protein levels[26]. We examined the effects of mutant TDP-43 expression in neural progenitors *in vivo*. Mutant human $TDP\text{-}43^{A315T}$ (Fig. 2a), like wild type TDP-43, induced cell death and caspase-3 activation when expressed by IUE in neural progenitors *in vivo* (Fig. 2c). $TDP\text{-}43^{A315T}$ expression

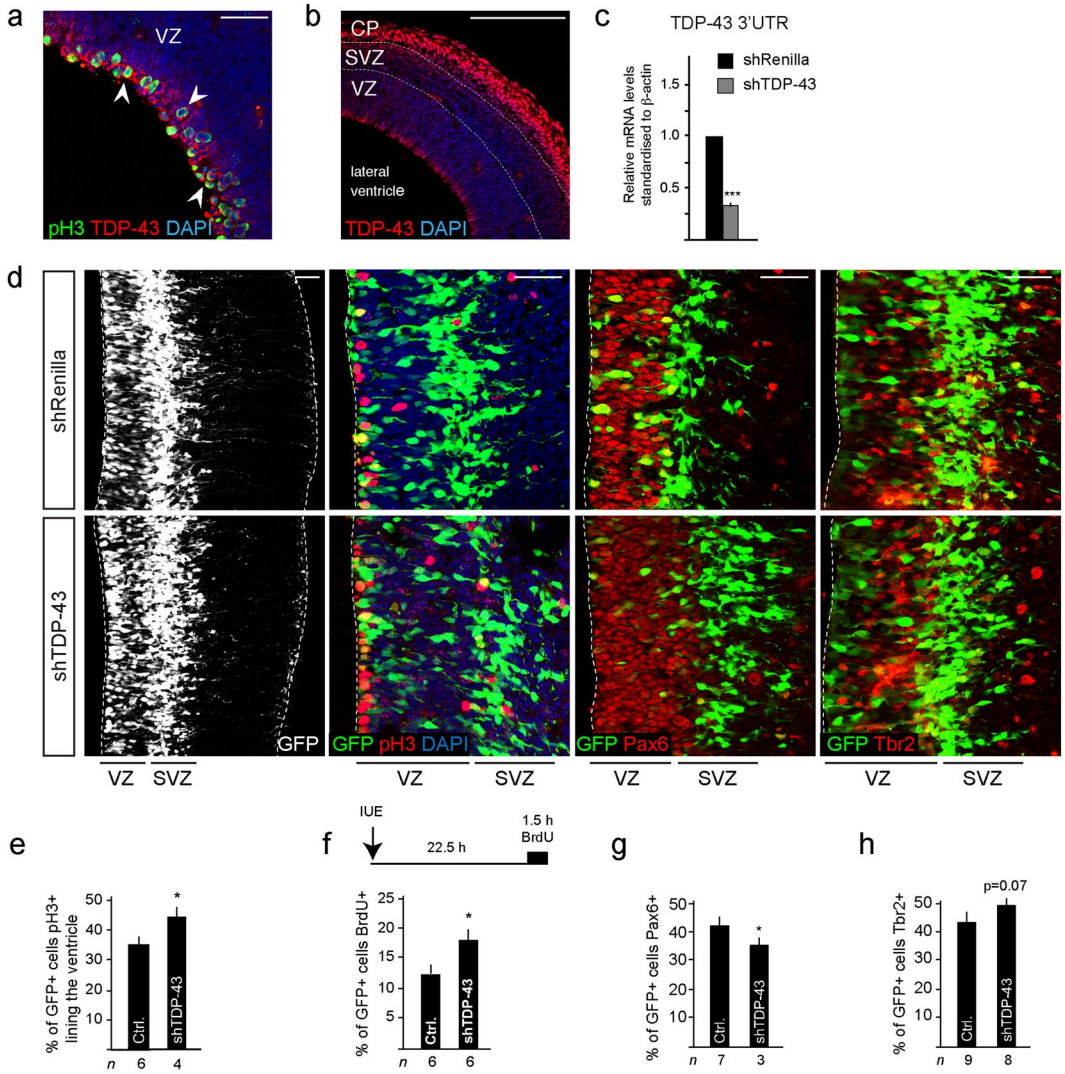


Figure 1. TDP-43 knockdown (KD) decreases Pax6^{+} neural stem/progenitors and affects cell cycle. (**a**) TDP-43 is expressed by ventricular zone (VZ) progenitors including by cells in mitosis (arrows). (**b**) TDP-43 is highly expressed by neurons in the cortical plate (CP) at e14.5. (**c**) TDP-43 KD (shTDP-43) by *in utero* electroporation (IUE) of neural stem/progenitors at e13.5 and analyzed at e14.5 reduces TDP-43 mRNA levels by 70% ($^{***}P < 0.001$) compared to control (shRenilla). (**d**) TDP-43 KD (GFP^{+}) cells survive in the VZ and differentiate to generate Tbr2^{+} basal progenitors in the subventricular zone (SVZ) similar to control transfected cells (shRenilla). (**e**) The proportion of cells lining the ventricle that are pH3^{+} increases following TDP-43 KD (shTDP-43) compared to control cells (shRenilla: Ctrl.), suggesting accumulation in M-phase of cell cycle. Values are shown as the proportion of transfected cells (GFP^{+}). (**f**) Scheme of the BrdU labeling procedure in e13.5 mice with a BrdU pulse 22.5 hours post IUE and 1.5 hours prior to killing. The proportion of GFP^{+} cells in the VZ that are BrdU labeled following TDP-43 KD (shTDP-43) increases compared to control cells (shRenilla: Ctrl.), suggesting increased S-phase entry. Values are shown as the proportion of transfected cells (GFP^{+}). (**g**) The proportion of GFP^{+} cells that are Pax6^{+} following TDP-43 KD (shTDP-43) reduces compared to control cells (shRenilla: Ctrl.). Values are shown as the proportion of transfected cells (GFP^{+}). (**h**) The proportion of GFP^{+} cells that are Tbr2^{+} following TDP-43 KD (shTDP-43) is unaffected compared to control cells (shRenilla: Ctrl.). Values are shown as the proportion of transfected cells (GFP^{+}). Scale bars in a and d = 25 µm, in b = 100 µm. Dashed line marks the telencephalic vesicle lining. tTest $^{*}P < 0.05$, $^{***} < 0.001$. ventricular zone (VZ), subventricular zone (SVZ), cortical plate (CP).

resulted in a loss of Pax6^{+} progenitors in the ventricular zone and aberrant differentiation into basal progenitors (Tbr2^{+}) (Fig. 2e,h). These data indicate that mutant TDP-43^{A315T} partially phenocopied the effects seen after overexpression of wild type TDP-43.

TDP-43 induced cell death is p53-dependent. p53-dependent neuronal cell death occurs in several neurodegenerative diseases[27]. Given that TDP-43 induces cell death, the link between TDP-43 cleavage and caspases, and demonstration of elevated p53 in neurons of ALS patients[20,21], we investigated whether TDP-43

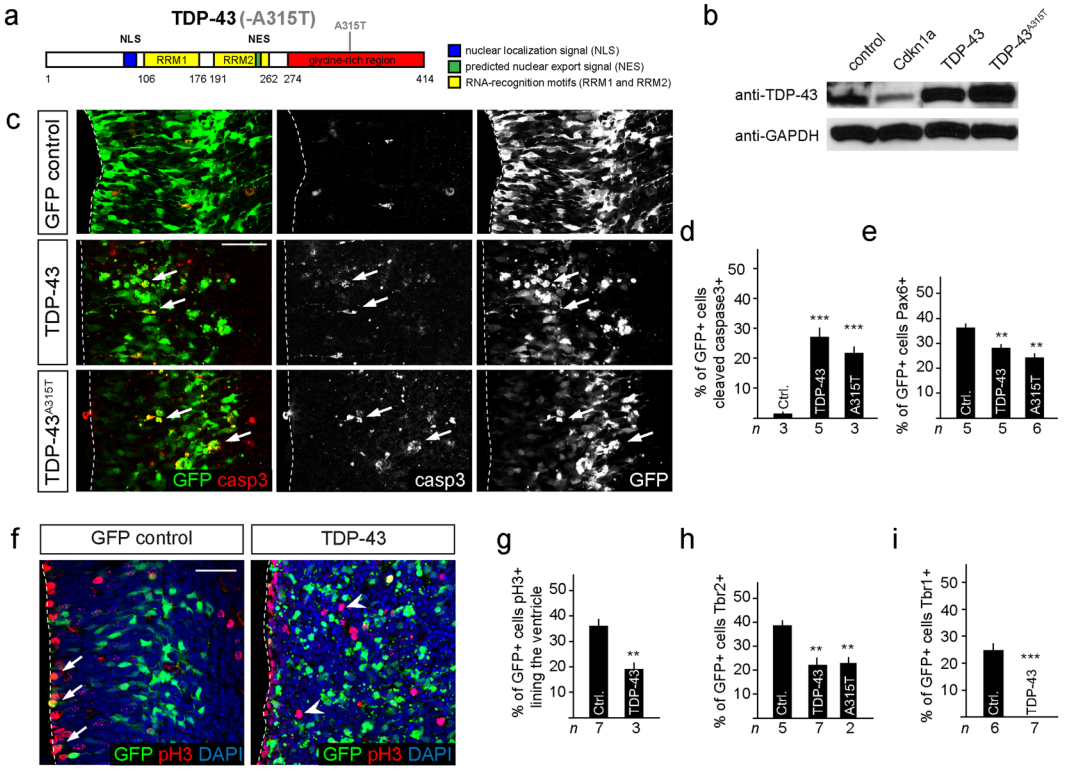


Figure 2. Expression of TDP-43 and TDP-43^{A315T} induce p53-dependent apoptosis of neural progenitors. (**a**) Scheme of human TDP-43 and TDP-43^{A315T} with a single point mutation in the glycine-rich C-terminal region, nuclear localization signal (NLS), nuclear export signal (NES), RNA recognition motif (RRM). (**b**) Transfection of neural progenitors results in a 2.4-fold expression of TDP-43 and 2.8-fold for TDP-43^{A315T} compared to endogenous levels. Expression of Cdkn1 reduces endogenous TDP-43 levels suggesting a reciprocal interaction with TDP-43. (**c**) Expression of TDP-43 and TDP-43^{A315T} activates caspase-3 (arrows) and drives cells in the ventricular zone into apoptosis within 24 hours. (**d**) TDP-43 and TDP-43^{A315T} expression induce caspase-3 activity and apoptosis in transfected cells (GFP$^+$). Values are shown as the proportion of transfected cells (GFP$^+$). (**e**) TDP-43 and TDP-43^{A315T} expression reduce the proportion of Pax6$^+$GFP$^+$ progenitors compared to control transfected cells (Ctrl.). Values are shown as the proportion of transfected cells (GFP$^+$). (**f**,**g**) TDP-43 expression causes mislocalization of GFP$^+$pH3$^+$ cells away from the ventricular lining (GFP control: arrows) to more basal locations (TDP-43: arrowheads) and reduces the proportion of ventricular progenitors in M-phase (pH3$^+$) compared to control transfected cells (Ctrl.). Values are shown as the proportion of transfected cells (GFP$^+$). (**h**) TDP-43 and TDP-43^{A315T} expression reduce differentiation to Tbr2$^+$ basal progenitors compared to control transfected cells (Ctrl.). Values are shown as the proportion of transfected cells (GFP$^+$). (**i**) TDP-43 overexpressing progenitors fail to generate Tbr1$^+$ neurons in the cortical plate. Scale bar in c and f = 25 µm. Dashed line marks the telencephalic vesicle lining. tTest $^*P < 0.05$, $^{**} < 0.01$, $^{***} < 0.001$.

induced apoptosis of neural progenitors cells depends on p53 activity. In contrast to expression in control (p53^{wt}) embryos, IUE-mediated TDP-43 overexpression in p53-deficient embryos (p53$^{-/-}$) did not disrupt neural progenitors in the ventricular zone or affect their morphology (Fig. 3a).

To separate potential cell autonomous and non-autonomous effects of TDP-43 overexpression, we performed a mosaic deletion of *Trp53* combined with IUE of TDP-43. We analyzed mice carrying *Hes5::CreER^{T2}* and *Rosa26* floxed STOP transcription GFP Cre reporter (*Rosa-CAG::GFP*) alleles, to lineage trace the Cre-expressing neural stem/progenitor cells and their progeny, which were either wild type or homozygous for floxed *Trp53* alleles. We induced deletion of *Trp53* from ventricular zone progenitors in embryos at e11 by Tamoxifen induction[28] and expressed TDP-43 by IUE in these embryos at e13.5 (Supplementary Fig. 4a,b). Under these mosaic conditions, TDP-43 overexpressing cells that were wild type for *Trp53* (GFP$^-$) underwent apoptosis (caspase-3) (Fig. 3b,c). Conversely, TDP-43 overexpressing cells where *Trp53* had been deleted (GFP$^+$) were viable and showed a normal morphology (Fig. 3b,c). By this conditional approach, loss of p53 also rescued the reduction in Pax6$^+$ progenitors and the proliferation defects (pH3$^+$ and BrdU labeling) induced by TDP-43 overexpression (Fig. 3d–f). This suggested that the cell death caused by TDP-43 is a p53-dependent cell autonomous response and not the result of disruption of the neural progenitor zones in the developing telencephalon.

Similarly, injection of the pharmacological inhibitor of p53, Pifithrin-α (PFT-α)[29,30], into p53^{wt} mothers rescued TDP-43 overexpressing cells and blocked caspase-3 activation (Fig. 4a,b). We analyzed PFT-α rescued cells in greater detail and found that the number and integrity of ventricular zone progenitors (Pax6$^+$) and basal progenitors (Tbr2$^+$) were comparable to controls (Fig. 4b,c and data not shown). Mitotically active cells (pH3$^+$)

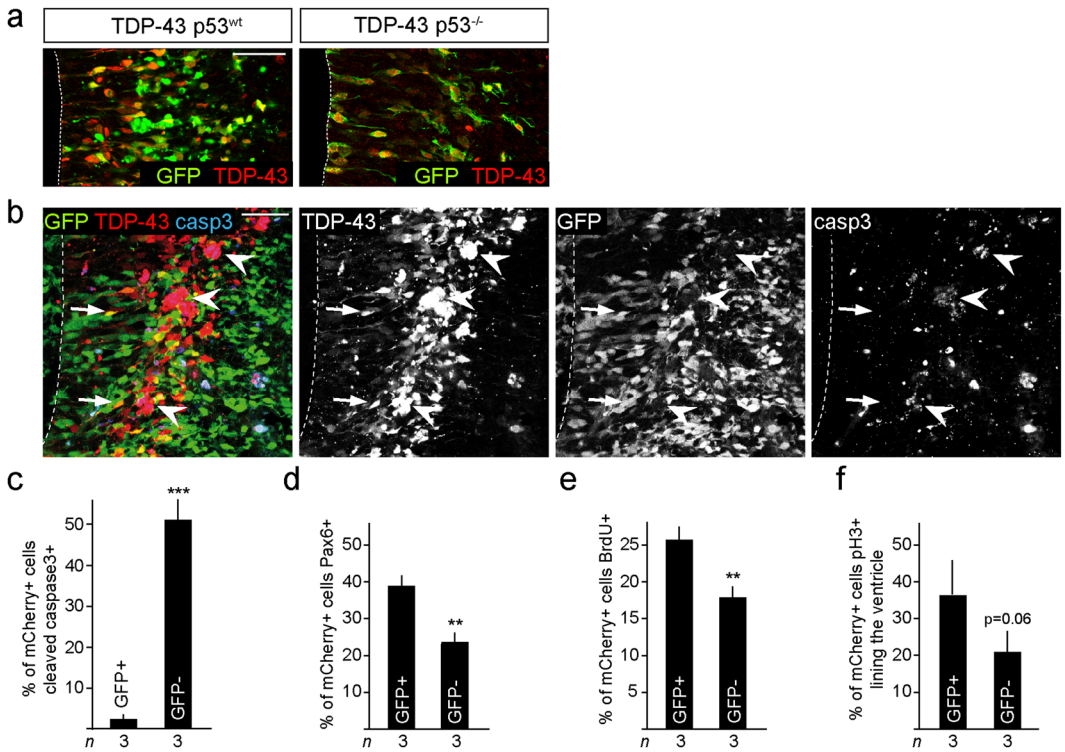


Figure 3. p53 deletion rescues TDP-43 overexpressing neural progenitors. (**a**) TDP-43 overexpression does not induce cell fragmentation or death of progenitors of $p53^{-/-}$ embryos. $p53^{-/-}$ TDP-43 expressing cells show a normal radial morphology in the ventricular zone and migrate to the subventricular zone. (**b**,**c**) Conditional deletion of floxed *Trp53* alleles by expression of Cre-recombinase (GFP^+), ($mCherry^+GFP^+p53^{-/-}$, arrows) reduces caspase-3 activation following expression of TDP-43 compared to none Cre-recombined cells ($mCherry^+GFP^-p53^{+/+}$, arrowheads). Values are shown as the proportion of TDP-43 overexpressing cells ($mCherry^+$). (**d–f**) *Trp53* ablated (GFP^+), TDP-43 overexpressing ($mCherry^+$) cells ($mCherry^+GFP^+p53^{-/-}$, arrows) show rescue of $Pax6^+$ (**d**), $BrdU^+$ (**e**) and $pH3^+$ expression (**f**) compared to none Cre-recombined *Trp53* wild type cells ($mCherry^+GFP^-p53^{+/+}$). Values are shown as the proportion of transfected cells ($mCherry^+$). Scale bar = 25 µm. Dashed line marks the telencephalic vesicle lining. tTest P = 0.06, $^{**}<0.01$, $^{***}<0.001$.

were decreased upon TDP-43 expression, but this was not rescued by inhibition of p53 (Fig. 4d). Furthermore, PFT-α treatment did not induce a change in the number of $pH3^+$ cells lining the ventricle in control (35.6 +/− 3.1 versus 39.0 +/− 1.8%) or TDP-43 overexpressing embryos (18.1 +/− 2.8% versus 23.4 +/− 3.0) (compare Figs 2g and 4d). Reduced proliferation in the ventricular zone following TDP-43 overexpression was also supported by BrdU incorporation experiments with less labeled cells in the TDP-43 overexpressing embryos even after PFT-α treatment (Fig. 4e). These findings are suggestive of two independent mechanisms of action for TDP-43 overexpression, one triggering p53-dependent apoptosis and a second perturbing cell cycle dynamics that likely does not involve p53.

RNA recognition motif 1 of TDP-43 is critical for toxicity. The underlying mechanisms through which TDP-43 induces cell death are unclear. Both full length as well as truncated and C-terminal fragments of TDP-43 accumulate in aggregates during neurodegeneration[31]. TDP-43 binds thousands of RNAs, including its own, regulating many aspects of RNA biogenesis, stability, splicing and transport[15,16,32]. The RNA recognition motif 1 (RRM1) of TDP-43 is necessary and sufficient for RNA binding[6,33]. We addressed whether RNA binding is required for TDP-43 induced toxicity in neural progenitors by expressing a mutant form where the RRM1 had been deleted ($TDP\text{-}43^{\Delta RRM1}$) (Supplementary Fig. 5a,b). Unlike TDP-43 and $TDP\text{-}43^{A315T}$, expression of $TDP\text{-}43^{\Delta RRM1}$ in neural progenitors did not result in obvious signs of cell death *in vivo* (Supplementary Fig. 5c). The number of ventricular zone progenitors ($Pax6^+$) and basal progenitors ($Tbr2^+$) were comparable to controls following $TDP\text{-}43^{\Delta RRM1}$ expression (Supplementary Fig. 5d–f). This indicated that functional RNA binding of TDP-43 is crucial in the cell death response and aberrant neurogenesis induced by TDP-43.

TDP-43 and $TDP\text{-}43^{A315T}$ induce p53 targets and proapoptotic gene expression. TDP-43 and $TDP\text{-}43^{A315T}$ induced cell death was blocked by PFT-α. TDP-43 and $TDP\text{-}43^{A315T}$ increased *Trp53* mRNA and p53 protein levels consistent with an activation of a stress response (Fig. 5a,b). In addition to the increase in total p53 levels, the phosphorylated and active form of p53 was also increased substantially following TDP-43 and $TDP\text{-}43^{A315T}$ expression (Fig. 5b). p53 regulates proapoptotic gene expression including *Bbc3* (PUMA) and *Bax* as well as cell cycle regulators including *Cdkn1a*. Consistent with an activation of p53, TDP-43 and $TDP\text{-}43^{A315T}$ induced *Bbc3* and *Bax* expression in neural progenitors but not the antiapoptotic gene *Bcl2* (Fig. 5c). In contrast,

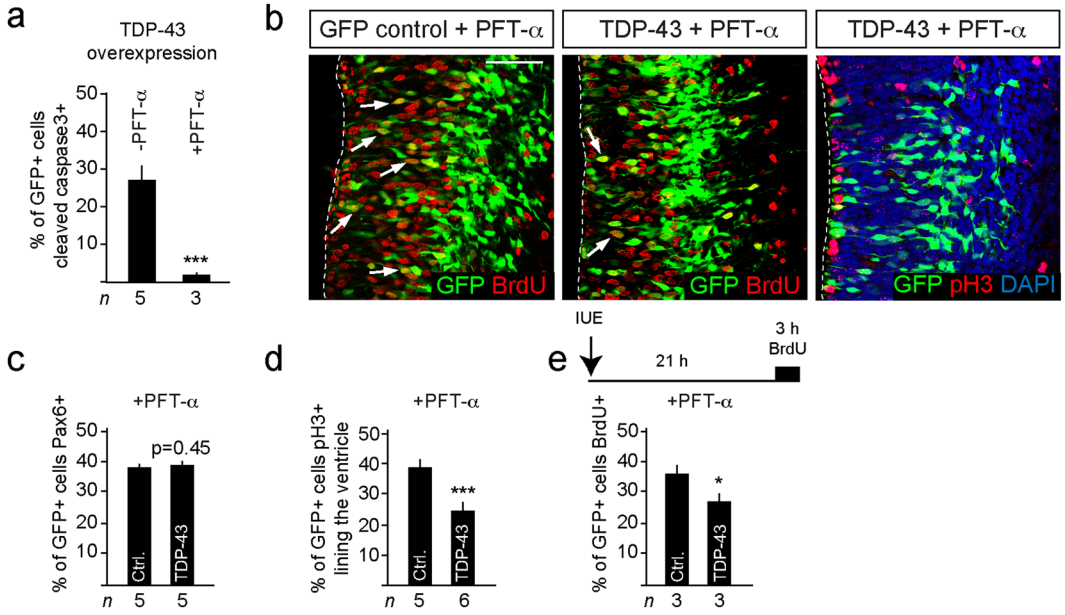


Figure 4. p53 inhibition with PFT-α rescues TDP-43 mediated apoptosis and integrity but not cell cycle defects. (**a**) TDP-43 induced caspase-3 activity and apoptosis is rescued by blocking p53 with PFT-α. (**b**) Inhibition of p53 with PFT-α *in utero* rescues TDP-43 induced apoptosis but does not reverse cell cycle defects. Arrows indicate BrdU$^+$ transfected cells (GFP$^+$). (**c**) Inhibition of p53 with PFT-α rescues TDP-43 expressing Pax6$^+$ progenitors in embryos. (**d**) The proportion of TDP-43 overexpressing cells (GFP$^+$) at the ventricular lining and expressing pH3$^+$ is reduced and not rescued by PFT-α compared to control transfected cells (Ctrl.). Values are shown as the proportion of transfected cells (GFP$^+$). (**e**) The proportion of TDP-43 overexpressing cells (GFP$^+$) in the VZ that are BrdU labeled is not rescued by PFT-α compared to control cells (Ctrl.). Values are shown as the proportion of transfected cells (GFP$^+$). Scale bars = 25 µm. Dashed line marks the telencephalic vesicle lining. tTest $^*P < 0.05$, $^{***} < 0.001$.

TDP-43$^{\Delta RRM1}$ overexpression did not affect p53 or proapoptotic gene expression (Fig. 5a,c). The balance between pro- and antiapoptotic genes regulates entry into the apoptotic pathway by stimulating mitochondrial permeability, the release of cytochrome C, and activation of caspase-9 and caspase-3[34]. Bbc3 binds Bcl2 releasing the proapoptotic BH3-only proteins Bax and Bak. Hence, Bbc3 induced by TDP-43 via p53 could bind and inhibit Bcl2, and the concomitant enhanced Bax expression would result in increased cell death.

Cdkn1a is a direct transcriptional target of p53. TDP-43 and TDP-43^{A315T} but not TDP-43$^{\Delta RRM1}$ induced *Cdkn1a* expression (Fig. 5a). PFT-α treatment did not affect the expression of the endogenous *Tardbp* but significantly inhibited TDP-43 and TDP-43^{A315T} induced *Trp53* and *Cdkn1a* expression (Fig. 5d). Hence, overexpression of TDP-43 and a mutant form, TDP-43^{A315T}, result in activation of multiple p53 target genes and this requires the RRM1 domain.

Cdkn1a mRNA associates with TDP-43. As TDP-43 reduced cell cycle, and induced *Cdkn1a* expression (Figs 2–5), we examined the cell cycle defects in more detail. Therefore, we assessed a potential TDP-43 regulation of *Trp53* and *Cdkn1a* by CLIP (cross-linked RNA immunoprecipitation). In accordance with previous findings[15,16,32,35], TDP-43 bound and decreased endogenous *Tardbp* mRNA levels (Fig. 5e and Supplementary Fig. 6a,b). In addition, *Cdkn1a* but not *Trp53* mRNA was CLIPed together with TDP-43 (Fig. 5e and Supplementary Fig. 6a) and expression of TDP-43 and TDP-43^{A315T} translated into an increase in Cdkn1a protein in N2A cells (Supplementary Fig. 6c). Hence, although PFT-α reduced *Trp53* and *Cdkn1a* mRNA expression in neural progenitors *in vitro* (Fig. 5d), and rescued cell death, Cdkn1a protein levels still increased in response to TDP-43, likely as a result of regulation of Cdkn1a at the post-transcriptional level. We did not observe an interaction between TDP-43 and p53 or Cdkn1a at the protein level (Supplementary Fig. 6d). Forced expression of Cdkn1a by IUE resulted in exit of progenitor cells from the ventricular zone and their exit from the cell cycle but, unlike TDP-43 overexpression, did not induce cell death or precocious differentiation (Supplementary Fig. 7a–d). Interestingly, we observed a reciprocal down regulation of TDP-43 mRNA and protein expression following overexpression of Cdkn1a (Supplementary Fig. 6b and Fig. 2b). Hence, TDP-43 regulates Cdkn1a expression that potentially contributes to the changes in cell cycle and explains the insensitivity of these effects to p53 inhibition.

p53 inhibition partially rescues TDP-43^{A315T} embryos. Homozygous mice expressing TDP-43^{A315T} from the *Tardbp* locus (*Tardbp*$^{315/315}$) die *in utero* (Supplementary Fig. 7a)[36]. The cause of death is unclear but post-implantation e6.5 mutant embryos were developmentally retarded, showed hemorrhage around the primitive endoderm and limited trophectodermal expansion and infiltration. At e9.5 most *Tardbp*$^{315/315}$ embryos were dead and the few remaining mutants were severely developmentally delayed and died soon thereafter

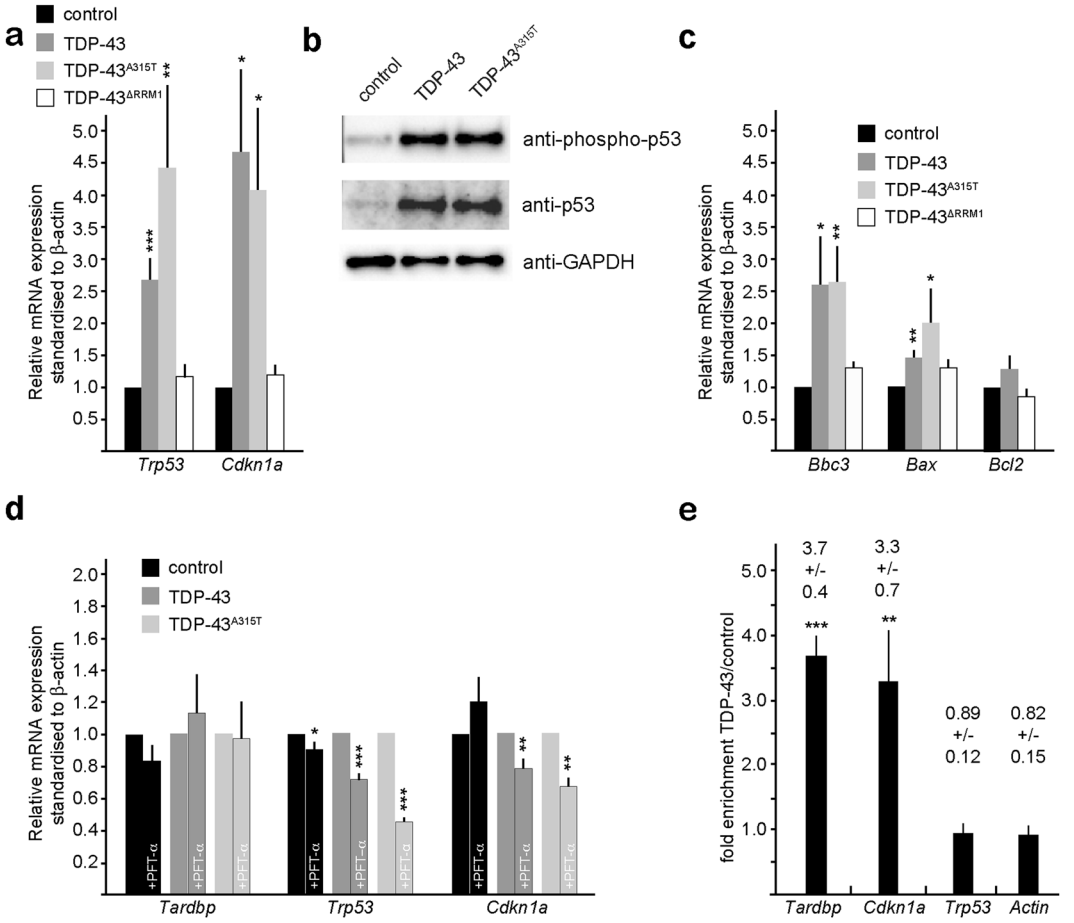


Figure 5. TDP-43 regulates proapototic BH3-only, Trp53 and Cdkn1a mRNA expression. (**a**) TDP-43 and TDP-43^{A315T} expression in neural progenitors increases *Trp53* and *Cdkn1a* mRNA levels, analyzed by quantitative RT-PCR analysis. TDP-43$^{\Delta RRM1}$expression does not affect *Trp53* and *Cdkn1a* mRNA levels. (**b**) TDP-43 and TDP-43^{A315T} expression in neural progenitors results in increased level of activated, phosphorylated p53 protein and total p53 protein. (**c**) TDP-43 and TDP-43^{A315T} expression in neural progenitors increases mRNA levels of the proapoptotic BH3-only proteins *Bbc3* (PUMA) and *Bax*, but not the antiapoptotic factor *Bcl2*. TDP-43$^{\Delta RRM1}$expression does not affect *Bbc3*, *Bax* or *Bcl-2* mRNA levels. (**d**) Pharmacological inhibition of p53 reduces TDP-43 and TDP-43^{A315T} induced *Trp53* and *Cdkn1a* mRNA levels but does not affect endogenous *Tardbp* mRNA expression, analyzed by quantitative RT-PCR analysis. Values are shown as relative to non-treated, standardized to *β-actin*. (**e**) TDP-43 binds endogenous *Tardbp* and *Cdkn1a* but not *Trp53* mRNAs. Quantitative RT-PCR analysis of *Tardbp*, *Cdkn1a*, *Trp53* and *β*-actin transcripts CLIPed together with TDP-43 from neural progenitors. Values are fold enrichment over control CLIPed (flag-GFP) transcripts. Statistical analysis of CLIPed products corrected relative to input RNA concentrations compared to flag-GFP CLIPed samples. Agarose gel analysis of the amplicons is shown in Supplementary Fig. 6a.

(Supplementary Fig. 8a). We treated pregnant *Tardbp*$^{wt/315}$ females from *Tardbp*$^{wt/315}$ inter-crosses with PFT-α every day after day 3.5 post-coitum. Inhibition of p53 *in utero* rescued *Tardbp*$^{315/315}$ embryos past e7.5 to Mendelian ratios up to e14.5 (Supplementary Fig. 8b,c). The PFT-α rescued mice looked phenotypically normal although slightly smaller than age-matched siblings. Brain morphology and development did not show obvious defects and proliferation within the brain of e14.5 *Tardbp*$^{315/315}$ embryos was indistinguishable from wild type siblings (Supplementary Fig. 8d,e). In order to address whether loss of p53 was able to rescue *Tardbp*$^{315/315}$ mice to birth, we inter-crossed *Tardbp*$^{wt/315}$ and Trp53$^{-/-}$ mice to generate *Tardbp*$^{315/315}$, p53$^{-/-}$ animals. However, after analysis of >100 offspring, we were unable to find *Tardbp*$^{315/315}$ mice in multiple litters from different parents. Hence, inhibition of p53 with PFT-α rescued early embryonic death of *Tardbp*$^{315/315}$ mutant mice but inhibition of p53 function is unable to rescue *Tardbp*$^{315/315}$ embryos to birth.

TDP-43 induced cell death of human iPS-derived cortical neurons can be rescued by blocking p53. TDP-43 and TDP-43^{A315T} expression induces p53-dependent apoptosis in the developing mouse telencephalon. In order to assess the effects of accumulated TDP-43 on human cortical cells, we differentiated human iPS cells to cortical progenitors and neurons for 39 days (Fig. 6a)[37]. Expression of wild type TDP-43 and mutant TDP-43^{A315T} at day 37 caused a loss of human iPS-derived cortical cells within 48 hours including neurons and progenitors (Fig. 6b–d). The surviving TDP-43 overexpressing cells had stunted morphologies and pyknotic nuclei (Fig. 6b).

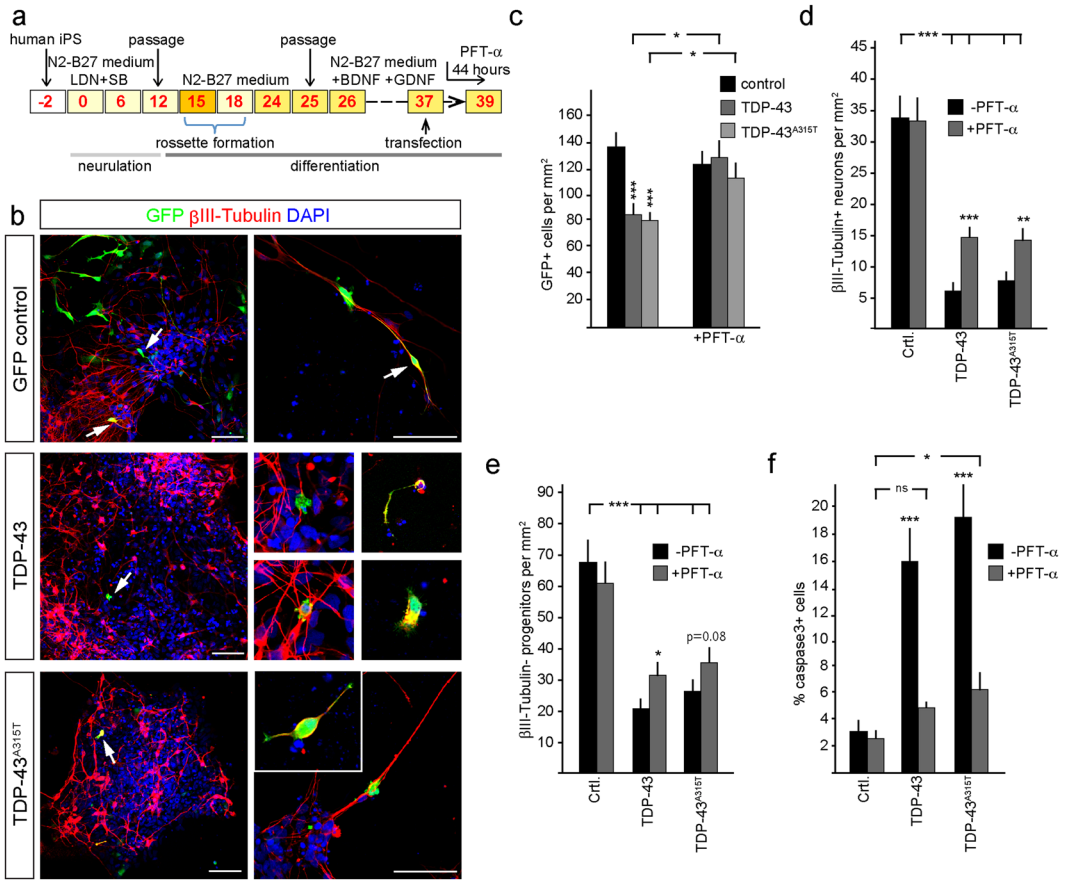


Figure 6. TDP-43 and human mutant TDP-43^{A315T} toxicity in human iPS-derived neurons is rescued by inhibition of p53. (**a**) Scheme of the human iPS cortical neuron differentiation profile and analysis of TDP-43 overexpression. Control human iPS cells were neuralized from day 0 to day 12 in N2-B27 Medium + LDN and SB and passaged on day 15 and day 25 and differentiated in N2 medium minus B27 (N2-B27) for 39 days. TDP-43 and TDP-43^{A315T} expression constructs were transfected on day 37. (**b**) Transfected GFP-expressing human iPS cells have neuronal and progenitor morphologies. TDP-43 and TDP-43^{A315T} cells are reduced and show stunted morphologies and cellular fragmentation compared to controls (arrows). (**c**) Expression of TDP-43 and TDP-43^{A315T} reduces the number of iPS-derived cells within the cultures. Treatment with PFT-α for 44 hours significantly rescues the number of transfected cells to control (GFP) levels. Together with the increased expression of activated caspase-3 these findings confirm that TDP-43 and TDP-43^{A315T} are toxic and rapidly induce apoptotic cell death that is dependent upon p53. (**d**,**e**) TDP-43 and TDP-43^{A315T} expression result in a reduction of iPS-derived neurons (βIII-Tubulin^{+}) and progenitors cells (βIII-Tubulin^{-}). PFT-α treatment of the cultures for 44 hours prior to analysis partially rescued the TDP-43 and TDP-43^{A315T} induced loss of neurons and progenitors. (**f**) TDP-43 and TDP-43^{A315T} expressing human neurons activate caspase-3 and die by apoptosis. Inhibition of p53 with PFT-α for 44 hours rescues cell death. Scale bar = 25 µm. tTest $^{*}P < 0.05$, $^{**} < 0.01$, $^{***} < 0.001$, ns not significant.

Treatment of TDP-43 and TDP-43^{A315T} expressing human iPS-derived cortical cultures with the PFT-α for the last 44 hours of culture rescued cell number and partially rescued the loss of neurons and progenitors and reduced the number of active caspase-3 expressing cells (Fig. 6c–f). Hence, TDP-43 accumulation in human cortical neurons can also induce p53-dependent neuronal cell death.

TDP-43 induces increased proapoptotic gene expression in human iPS-derived cortical cells. As TDP-43 overexpression in human iPS-derived cortical cells induced cell death, we examined whether this involved increased expression of the proapoptotic genes *BBC3*, *BAX* and the antiapoptotic gene *BCL2*. We transfected iPS-derived cortical cells at day 37 of differentiation, sorted the TDP-43 overexpressing human cortical cells (GFP^{+}) after 48 hours and examined gene expression by qRT-PCR. In agreement with our findings in the mouse telencephalon, overexpression of TDP-43 in human cells significantly increased the expression of *TRP53*, *CDKN1A*, *BBC3*, and *BAX* but also *BCL2* (Fig. 7a). We examined whether the increase in proapoptotic gene expression in response to TDP-43 overexpression was dependent upon p53 activity by treating the cells for the last 44 hours of culture with PFT-α. Blocking p53 activity prevented the increase in *TRP53*, *CDKN1A*, *BBC3*, *BAX* and *BCL2* expression above the levels seen in PFT-α treated control cultures, suggesting that also in human cortical

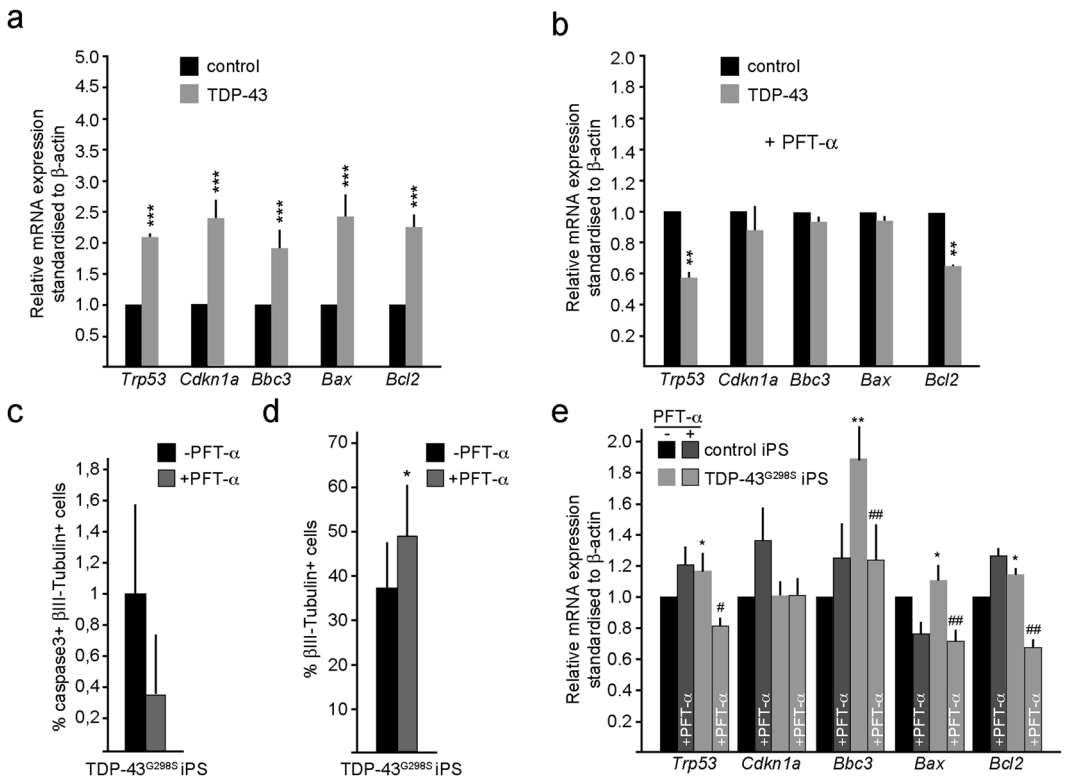


Figure 7. Proapoptotic gene expression is increased in human iPS-derived cortical cells and TDP-43^{G298S} mutant iPS derived cortical cultures. (**a**) TDP-43/GFP expression in human iPS followed by fluorescent assisted cell sorting after 48 hours. TDP-43 expression in human iPS cortical neuron increases mRNA levels of the proapoptotic genes *Trp53*, *Cdkn1a*, *Bbc3* (PUMA), *Bax* and *Bcl2*. (**b**) TDP-43/GFP expression in human iPS followed by sorting after 48 hours. p53 inhibition by PFT-α treatment for 44 hours prior to sorting rescues TDP-43 induced increase of proapoptotic genes in human iPS compared to control. (**c**) TDP-43^{G298S} iPS-derived cells show reduced numbers of activated caspase-3^{+} cells upon PFT-α treatment for 44 hours. (**d**) TDP-43^{G298S} iPS cells and control iPS cells show reduced numbers of βIII-Tubulin^{+} cells upon PFT-α treatment for 44 hours. (**e**) TDP-43^{G298S} iPS cells display increases mRNA levels of proapoptotic genes *p53*, *Bbc3* (PUMA), *Bax* and *Bcl2* upon PFT-α treatment for 44 hours. In comparison mRNA levels of the very genes in Control iPS cells are unchanged. tTest $^{*}P < 0.05$, $^{**} < 0.01$, $^{***} < 0.001$ relative to control cells. $^{\#}P < 0.05$, $^{\#\#}P < 0.01$ relative to cells not treated with PFT-α.

cells, the expression of TDP-43 induces a p53-dependent increase in proapoptotic genes that may contribute to the observed cell death (Fig. 7b).

TDP-43 mutations are associated with neurodegenerative disease including FTLD and ALS. We used patient-derived *TDP-43^{G298S}* mutant iPS cells[38] and differentiated them into cortical neurons using the same protocol used for differentiation of the control human iPS cells (Fig. 6a). These cortical cultures contained activated caspase-3^{+} neurons and treatment with PFT-α for 44 hours reduced the number of apoptotic neurons suggesting p53 activity in these cultures (Fig. 7c). In agreement with p53-mediated neuronal cell death, PFT-α treatment significantly increased the number of neurons derived from TDP-43^{G298S} patient iPS cells (Fig. 7d). Therefore, we examined the expression levels of *TRP53*, *CDKN1A*, *BBC3*, *BAX* and *BCL2* in the *TDP-43^{G298S}* mutant iPS-derived cortical cultures. *Trp53* expression was significantly higher in the *TDP-43^{G298S}*-derived cells compared to control. In support of an increased p53 activity, the expression of *BBC3*, *BAX* and *BCL2* were all elevated in the *TDP-43^{G298S}* iPS cortical neuron cultures. In addition, treatment of the iPS-derived cultures with PFT-α to block p53, significantly reduced expression of *BBC3*, *BAX* and *BCL2* in *TDP-43^{G298S}* but not by control iPS-derived cortical neurons (Fig. 7e).

Discussion

Our findings identify activation of proapototic gene expression as a mechanism through which TDP-43 accumulation and mutations can induce cell death in murine neural progenitors *in vivo* and human cortical neurons *in vitro*. This effect was rescued by blocking or depleting p53. p53 is induced by cellular stress and recent findings indicate that TDP-43 accumulation, as seen in neurodegenerative diseases including ALS and FTLD, induces cellular stress response[24,39–41]. Thus, it is possible that induced p53 in our paradigms is also a result of cellular stress caused by the accumulation of TDP-43. In addition, TDP-43 induces phosphorylation of Eif2α, which results in a general inhibition of translation[30]. Eif2α phosphorylation induces translation of ATF4 (activating transcription factor 4), which enhances expression of ATF3 and CHOP (CCAAT/enhancer-binding protein homologous protein). CHOP inhibits Bcl2 functions at the protein level and induces Bim to enhance cell death through activation

of Bax. In agreement, we found that neither TDP-43 nor TDP-43^{A315T} induce *Bcl2* expression in mouse cells, whereas *Bax* is increased. However, it is conceivable that Eif2α inhibition of antiapoptotic BH3-only proteins and the p53-induced proapoptotic gene (*Bbc3* and *Bax*) expression synergize to promote cell death in TDP-43 mutant cells. Hence, TDP-43 seems to activate a bifurcated pathway to control BH3-only protein-mediated apoptosis.

Increase of TDP-43 and TDP-43^{A315T} in cortical neurons derived from human iPS cells also resulted in an apoptosis phenotype. In concordance with the overexpression in mouse cells, we observed a strong increase in *TRP53*, *CDKN1A*, *BBC3* and *BAX* in the human cortical neural cells *in vitro*. However, *BCL2* was also increased in the human cortical neurons in contrast to mouse neural progenitor cells. This opposing result might stem from differences in pathway regulation between species. However, inhibition of p53 by PFT-α rescued the TDP-43 mediated increase in the proapoptotic genes in mouse neural progenitor cells and human cortical neurons.

Besides the TDP-43 mediated increase in p53, we also observed a strong increase in *Cdkn1a* mRNA levels. Concomitant, we observed that TDP-43 overexpression induced a block in cell cycle progression in mouse neural progenitors *in vivo*. *Cdkn1a* is activated by p53, however, we additionally show an association of TDP-43 to *Cdkn1a* mRNA, which suggests a regulation of Cdkn1a expression by TDP-43. This is supported by the observation that blocked cell cycle progression could not be rescued by p53 inhibition. Thus, TDP-43 expression not only induces apoptosis but also induces a cell cycle block, probably at G2/M phase transition through cyclin dependent kinase inhibition via Cdkn1a, which in turn could stimulate cell death. Due to the fact that we performed the CLIP experiments with PFA cross-linking, we cannot be sure that TDP-43 binds directly to *Cdkn1a* mRNA. Confirmation of direct binding of TDP-43 to *Cdk1a* mRNA will require further analysis. Previously, it has been reported that TDP-43 overexpression induces G2/M arrest in HeLa cells[42]. It was shown that the effect was partially dependent on p53 but apoptosis could not be rescued by p53-inhibition. These partially contradictory results may be explained by the differences in cells used in the experiments. HeLa cells have an altered cell cycle regulation and might not respond in the fashion as neural progenitors and neurons to inhibition of p53.

Many lines of transgenic mice expressing wild type or mutant forms of TDP-43 have been analyzed in the last years, all of which show early mortality. We employed homozygous mice expressing TDP-43^{A315T} and partially rescued the animals by p53 inhibition. Similar rescue experiments have been preformed in *hSOD1^{G82A}* transgenic mice. Interestingly, p53 deletion did not rescue the early lethality in these *hSOD1^{G82A}* animals, which indicates a p53-independent mechanism, possibly involving death receptors[34]. This is supported by findings that death receptor 6 (DR6) levels are elevated in *hSOD1^{G82A}* transgenic mice[43]. In addition, treatment with blocking antibodies against DR6 increased motor neuron survival and, therefore, provided a neuroprotective effect in *hSOD1^{G82A}* expressing animals. The partial rescue of TDP-43^{A315T} embryos indicates probable involvement of additional factors triggering lethality of these animals. However, this remains to be addressed in the future.

We also investigated how loss of TDP-43 affects cell survival and differentiation of neural stem/progenitor cells. In striking contrast to the overexpression experiments, we did not observe signs of apoptosis upon KD of TDP-43 *in vivo*. Although TDP-43 is indispensable for early development[3], it does not seem to have a crucial function in neural progenitor maintenance in the brain at the time points we analyzed. This might be due to the fact that we only achieved a 70% reduction in TDP-43 levels at the protein level. However, we did observe a significant increase in the number of proliferating cells in the *in vivo* KD experiment. *Cdkn1a* KD induces proliferation by promoting S-phase entry[44]. We hypothesize that loss of TDP-43 results in a reduction in Cdkn1a, which in turn activates proliferation, reduces cell cycle exit and delays differentiation.

In conclusion, we show that TDP-43 accumulation in mouse neural progenitors *in vivo* as well as in human cortical progenitors and neurons and in TDP-43^{G298S} mutant iPS-derived cortical cultures induces p53-dependent apoptosis. This is associated with up regulation of apoptotic genes including *Bbc3* and *Bax*. We did not observe a direct interaction between TDP-43 and p53 on the protein or mRNA levels. It therefore remains elusive how TDP-43 activates p53 expression but this is likely, at least in part, to be indirect via induction of cellular stress response[30]. Future work should address this as well as how the multiple pathways downstream of TDP-43 contribute to cell death in disease.

Methods

Animals and animal husbandry. *Trp53^{tm1Tyj}*, *Trp53^{tm1Brn}*, *TDP-43^{A315T}*, *Rosa-CAG::GFP* and *Hes5::CreER^{T2}* mice have been described elsewhere[28,36,45–47]. Mice were maintained on a 12-hour day/night cycle with adequate food and water under SPF conditions. All methods were carried out in accordance with guidelines and regulations of Max-Planck Institutional and German Federal regulations and under license numbers H-05/01, 0-06/02, G-09/18, G-09/19, G-08/26 (Ethical Commission Freiburg, Germany) and 2437 and 2438 (Veterinary commission Basel). All experimental protocols were approved by Max-Planck Institutional and German Federal regulations and under license numbers H-05/01, 0-06/02, G-09/18, G-09/19, G-08/26 (Ethical Commission Freiburg, Germany) and 2437 and 2438 (Veterinary commission Basel). The day of vaginal plug was counted as embryonic day 0.5 (e0.5).

Tamoxifen, BrdU and PFT-α administration. *Trp53^{tm1Brn}/Hes5::CreER^{T2}/Rosa-CAG::GFP* mice were given 2 mg Tamoxifen by gavage at e11.5 to induce recombination. Stock solutions of Tamoxifen (Sigma) were prepared at a concentration of 20 mg/ml in corn oil (Sigma). Bromodesoxyuridine (BrdU, Sigma) was administered to the adult animals via a single intraperitoneal injection (50 mg/kg body weight). PFT-α (stock solution 20 mM in DMSO) was injected intraperitoneal into pregnant mice (2.2 mg/kg).

***In utero* electroporation for overexpression in neural stem/progenitors *in vivo*.** Female C57BL/6 J and *Trp53^{tm1Brn}* mice were used for the *in utero* expression analyses at 13.5 days after detection of vaginal plug. DNA constructs were injected into the forebrain telencephalic vesicles of the embryonic mice using a microinjector (Pneumatic Pico Pump, WPI Rnage) and pulled Borosilicate glass capillaries (Kwick-Fil^{TM}).

The capillaries were pulled in a micropipette puller (Sutter Instrument Co.) and sharpened using a capillary sharpener (Bachofer). The capillaries were backend-loaded with 10 µl of the plasmid. Plasmid stocks were prepared under endotoxin free conditions and suspended in sterile phosphate buffered saline (PBS) at a concentration of 3 µg/µl. Fast green contrast dye (10%) was added to the plasmids to visualize the targeted area of the telencephalon. The overexpression or shRNA knockdown vectors were electroporated in a molecular ratio of 3:1 with the transfection reporter vector (*pCAGGS::eGFP*). The pregnant female mice were administered analgesic (Temgesic; 0.8 mg/kg) by intraperitoneal injection and anesthetized with 1–2% isoflurane (Baxter) in a constant flow of O_2, secured on a heated operating table. Body temperature was monitored continually. The fur was removed from the abdomen using depilation cream. Throughout the operation, the peritoneal cavity was moistened with sterile Hank's buffered saline solution (HBSS). The uterine horns containing the embryos (e13.5) were manipulated under sterile conditions by hand. A cold light source was used to illuminate the embryos. 1–2 µl of DNA solution were injected into the telencephalic vesicles of each embryo. The embryos were electroporated (Electro Square Pavator™, BTX® Harvard Apparatus) with 10 pulses of 40 V and a pulse length of 50 ms at 950 ms intervals. The anode of the electrode was oriented toward the injected side. After injection and electroporation, the uterus was returned to the abdomen, the muscle and the skin sutured and the females allowed to recover under a heating lamp with constant observation. Operated females had free postoperative access to analgesic (Temgesic; 0.8 mg/kg) in sterile agar. The animals were sacrificed after a defined time and the embryos isolated and prepared for sectioning.

Expression plasmids and constructs. Full-length coding region cDNAs for GFP, Cdkn1a, human TDP-43, human TDP-43^{A315T} and human TDP-43$^{\Delta RRM1}$ were subcloned into *pCAGGS* expression vectors with a beta-globin 3-prime untranslated and polyadenylation sequence. *pSuper-shRenilla* as well as *pSuper-shTDP-43* were cloned according to the manufacturer's instructions (Oligoengine). *p3X-FLAG-myc-CMV™-26* expression construct was obtained from Sigma and *p3X-TDP-43-flag* was cloned using Not1 and Xba1 restriction sites (Primers: fwd 5′-TTCTGCGGCCGCCACCATGTCTGAATATATTCG-3′ rev 5′-CTTTCTAGACTACATTCCCCAGCCAGAAG-3′).

***In situ* RNA hybridization.** For *in situ* RNA hybridization a digoxigenin- (DIG)-labeled antisense RNA probe was generated for mouse *Tardbp* RNA (amplified from mouse cDNA using primers: fwd 5′-ATTCCTTCCCGTCTGTGCTT-3′ rev 5′-TGCTTAGGTTCAGCATTGGA-3′) using the procedures described previously. Expression was detected by colorimetric reaction using NBT (nitroblue tetrazolium chloride) and BCIP (5-bromo-4-chloro-3-indolylphosphate p-toluidine) as reaction substrates and images taken using an Axioplan microscope (Zeiss) with an Axiocam CCD camera (Zeiss).

Quantitative real-time PCR analysis of gene expression. For RNA isolation, cells were lysed directly in Trizol (Invitrogen) reagent. RNA was prepared according to the manufacturer's instructions. 1 µg of total RNA was used for cDNA synthesis by Oligo-dT priming and BioScript (Bioline). For quantitative RT-PCR the SensiMix SYBR kit (Bioline) was used following the manufacturer's instructions (see Supplemental Materials and Methods for detailed description of primers) The reaction was run in a Rotor-Gene™ 6000 Real-time PCR machine (Corbett) and analyzed using Rotor-gene 6000 series software 1.7. β-actin and GAPDH mRNA levels were measured as endogenous controls and for quantification.

Cross-linked RNA immunoprecipitation. The detailed protocol can be found at http://www.bio-protocol.org/e398. In summary, N2A cells were transfected using Lipofectamin2000 (Invitrogen) according to manufacturer's instructions with *p3X-flag-GFP* or *p3X-flag-TDP-43* and trypsinized after 48 hours. The cells were then fixed in 3% formaldehyde in PBS for 10 minutes and lysed by sonication (10 pulses for 10 seconds). Immunoprecipitation was performed overnight at 4 °C using anti-Flag M2 Affinity Gel (Sigma-Aldrich). After collection by centrifugation at 2000g and washing 3–4 times with lysis buffer the complexes were reverse cross-linked and RNA extracted using Trizol reagent (Invitrogen) according to the manufacturers instructions. Isolated RNA was treated with RNase-free DNaseI (Roche) to remove any genomic DNA contamination. First strand cDNA was generated using BioScript (Bioline) and random hexamer primers followed by real-time PCR using SensiMix SYBR Kit (Bioline).

Immunoprecipitation and Western blotting. For immunoprecipitation, N2A cells were transfected with expression plasmids (*p3X-flag-TDP-43* or *p3X-flag-GFP* as control) and cells were harvested after 24 hours. Cells were lysed and 1/10 of the lysate was used as input control. The remaining lysate was added to pre-blocked beads coupled to anti-flag antibody (ANTI-FLAG® M2 Affinity Gel) or to un-coupled beads (Sepharose-G beads) and incubated at 4 °C. Thereafter the beads with the bound proteins were washed and the samples were incubated for 1 hour at 70 °C to release the bound proteins. The supernatant containing the immunoprecipitated proteins was then analyzed on a SDS-gel. For low molecular weight proteins like TDP-43, 15% separation gels were used according to the manufacturer's instruction. 20–40 µg proteins with sample loading buffer were denatured and loaded on the gel. Proteins were separated and after blotted onto a PVDF membrane. To detect proteins the membrane was incubated with specific antibodies after blocking with blocking buffer for 1 hour (see Supplemental Tables for detailed description of antibodies). The membrane was thereafter incubated with a secondary antibody coupled to horse reddish peroxidase and the antibody was detected using Amersham™ ECL™ Western Blotting Detection reagent.

Immunohistochemistry and immunostaining. Embryos were harvested and the brains fixed in 4% paraformaldehyde (PFA) solution in 0.1 M phosphate buffer (PB) overnight. Brains were cryoprotected in a 30% sucrose solution in 0.1 M PB for 24 hours, embedded and frozen over dry ice in OCT (TissueTEK). Horizontal sections (20 µm) were collected on Superfrost glass slides (Thermo Scientific), and stored at −20 °C until use.

For immunostaining, sections were incubated overnight at 4 °C with the primary antibody diluted in blocking solution of 2% normal donkey serum (Jackson ImmunoResearch), 0.3% Triton X-100 in 0.1 M PB. Sections were washed three times in PBS and incubated at room temperature for 3 hours with the corresponding secondary antibodies in blocking solution. When signal amplification was needed, sections were washed and incubated for 1 hour at room temperature in streptavidin–FITC (Jackson ImmunoResearch; 1:500) and counter-stained with DAPI (1 μg/ml). For BrdU detection, sections were treated with 2 M HCl at 37 °C for 15 minutes prior to primary antibody incubation. HCl-treated sections where then equilibrated in borate buffer (0.1 M, pH 8.5) for 10 minutes. For active caspase-3, Tbr1, Tbr2 and TDP-43 detection, antigens were recovered at 80 °C for 30 minutes in Sodium Citrate solution (10 mM, pH 6.0) with 0.05% Tween. Stained sections were embedded in mounting medium containing diazabicyclo-octane (DABCO; Sigma) as an anti-fading agent. Antibodies, dilutions and conditions used for immunolabeling are described in the supplementary tables. Sections were analyzed with an Axioscope (Zeiss) or confocal (Zeiss LSM510) fluorescence microscope. Images were acquired using Axiovision or Zeiss LSM 4.2 (Zeiss) and processed with ImageJ 1.33 or Photoshop CS (Adobe) software.

iPS cell culture and fluorescent assisted cell sorting. Human iPS cells (Nas2 cells)[48] were differentiated into cerebral cortex neurons for 39 days according to the protocol from Shi *et al*.[37]. Patient-derived *TDP-43*^{G298S} mutant iPS have been characterized and described previously by Alami *et al*.[38]. iPS cells were washed, filtered through a 70 μm cell strainer (Miltenyi Biotec) and sorted on a BD FACS Aria III. Live cells were discriminated by forward and side-scatter (for live cells – from the control) and gated for GFP$^-$ (non transfected) or single GFP$^+$ populations.

Quantification and statistical analysis of the data. Randomly selected, stained cells were analyzed with fixed photomultiplier settings on a Zeiss LSM510 confocal microscope (Zeiss). n numbers represent the number of animals used, and images from at least 3 sections per animal were quantified. Data are presented as average percentages of co-labeled cells and statistical comparisons were conducted by two-tailed unpaired Student's tTest. Significance was established at $P < 0.05$. In all graphs, error bars are standard deviation (SD).

References

1. Ling, S. C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**, 416–438 (2013).
2. Ou, S. H., Wu, F., Harrich, D., Garcia-Martinez, L. F. & Gaynor, R. B. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *Journal of virology* **69**, 3584–3596 (1995).
3. Sephton, C. F. *et al*. TDP-43 is a developmentally regulated protein essential for early embryonic development. *J Biol Chem* **285**, 6826–6834 (2010).
4. Wu, L. S. *et al*. TDP-43, a neuro-pathosignature factor, is essential for early mouse embryogenesis. *Genesis* **48**, 56–62 (2010).
5. Buratti, E. *et al*. Nuclear factor TDP-43 and SR proteins promote *in vitro* and *in vivo* CFTR exon 9 skipping. *EMBO J* **20**, 1774–1784 (2001).
6. Ayala, Y. M. *et al*. Human, Drosophila, and C.elegans TDP43: nucleic acid binding properties and splicing regulatory function. *J Mol Biol* **348**, 575–588 (2005).
7. Buratti, E. *et al*. TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail: an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing. *J Biol Chem* **280**, 37572–37584 (2005).
8. Wang, I. F., Reddy, N. M. & Shen, C. K. Higher order arrangement of the eukaryotic nuclear bodies. *Proc Natl Acad Sci USA* **99**, 13583–13588 (2002).
9. Igaz, L. M. *et al*. Dysregulation of the ALS-associated gene TDP-43 leads to neuronal death and degeneration in mice. *J Clin Invest* **121**, 726–738 (2011).
10. Walker, A. K. *et al*. Functional recovery in new mouse models of ALS/FTLD after clearance of pathological cytoplasmic TDP-43. *Acta neuropathologica* **130**, 643–660 (2015).
11. Arnold, E. S. *et al*. ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. *Proc Natl Acad Sci USA* **110**, E736–745 (2013).
12. Barmada, S. J. *et al*. Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J Neurosci* **30**, 639–649 (2010).
13. Wegorzewska, I., Bell, S., Cairns, N. J., Miller, T. M. & Baloh, R. H. TDP-43 mutant transgenic mice develop features of ALS and frontotemporal lobar degeneration. *Proc Natl Acad Sci USA* **106**, 18809–18814 (2009).
14. De Conti, L. *et al*. TDP-43 affects splicing profiles and isoform production of genes involved in the apoptotic and mitotic cellular pathways. *Nucleic acids research* **43**, 8990–9005 (2015).
15. Tollervey, J. R. *et al*. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci* **14**, 452–458 (2011).
16. Polymenidou, M. *et al*. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci* **14**, 459–468 (2011).
17. Lagier-Tourenne, C. *et al*. Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat Neurosci* **15**, 1488–1497 (2012).
18. Suzuki, H., Lee, K. & Matsuoka, M. TDP-43-induced death is associated with altered regulation of BIM and Bcl-xL and attenuated by caspase-mediated TDP-43 cleavage. *J Biol Chem* **286**, 13171–13183 (2011).
19. Gonzalez de Aguilar, J. L. *et al*. Alteration of the Bcl-x/Bax ratio in a transgenic mouse model of amyotrophic lateral sclerosis: evidence for the implication of the p53 signaling pathway. *Neurobiology of disease* **7**, 406–415 (2000).
20. Martin, L. J. p53 is abnormally elevated and active in the CNS of patients with amyotrophic lateral sclerosis. *Neurobiology of disease* **7**, 613–622 (2000).
21. Ranganathan, S. & Bowser, R. p53 and Cell Cycle Proteins Participate in Spinal Motor Neuron Cell Death in ALS. *Open Pathol J* **4**, 11–22 (2010).
22. Prudlo, J. *et al*. Motor neuron cell death in a mouse model of FALS is not mediated by the p53 cell survival regulator. *Brain research* **879**, 183–187 (2000).
23. Kuntz, C. T., Kinoshita, Y., Beal, M. F., Donehower, L. A. & Morrison, R. S. Absence ofp53: no effect in a transgenic mouse model of familial amyotrophic lateral sclerosis. *Experimental neurology* **165**, 184–190 (2000).
24. Bentmann, E. *et al*. Requirements for stress granule recruitment of fused in sarcoma (FUS) and TAR DNA-binding protein of 43 kDa (TDP-43). *J Biol Chem* **287**, 23079–23094 (2012).
25. Lagier-Tourenne, C. & Cleveland, D. W. Rethinking ALS: the FUS about TDP-43. *Cell* **136**, 1001–1004 (2009).

26. Guo, W. *et al*. An ALS-associated mutation affecting TDP-43 enhances protein aggregation, fibril formation and neurotoxicity. *Nat Struct Mol Biol* **18**, 822–830 (2011).
27. Culmsee, C. & Mattson, M. P. p53 in neuronal apoptosis. *Biochem Biophys Res Commun* **331**, 761–777 (2005).
28. Lugert, S. *et al*. Homeostatic neurogenesis in the adult hippocampus does not involve amplification of Ascl1(high) intermediate progenitors. *Nature communications* **3**, 670 (2012).
29. Komarov, P. G. *et al*. A chemical inhibitor of p53 that protects mice from the side effects of cancer therapy. *Science (New York, N.Y* **285**, 1733–1737 (1999).
30. Kim, H. J. *et al*. Therapeutic modulation of eIF2alpha phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. *Nature genetics* **46**, 152–160 (2014).
31. Neumann, M. *et al*. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science (New York, N.Y* **314**, 130–133 (2006).
32. Ayala, Y. M. *et al*. TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J* **30**, 277–288 (2011).
33. Buratti, E. & Baralle, F. E. Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. *J Biol Chem* **276**, 36337–36343 (2001).
34. Mc Guire, C., Beyaert, R. & van Loo, G. Death receptor signalling in central nervous system inflammation and demyelination. *Trends in neurosciences* **34**, 619–628 (2011).
35. Sephton, C. F. *et al*. Identification of neuronal RNA targets of TDP-43-containing ribonucleoprotein complexes. *J Biol Chem* **286**, 1204–1215 (2011).
36. Stribl, C. *et al*. Mitochondrial dysfunction and decrease in body weight of a transgenic knock-in mouse model for TDP-43. *J Biol Chem* **289**, 10769–10784 (2014).
37. Shi, Y., Kirwan, P. & Livesey, F. J. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nature protocols* **7**, 1836–1846 (2012).
38. Alami, N. H. *et al*. Axonal transport of TDP-43 mRNA granules is impaired by ALS-causing mutations. *Neuron* **81**, 536–543 (2014).
39. Warraich, S. T., Yang, S., Nicholson, G. A. & Blair, I. P. TDP-43: a DNA and RNA binding protein with roles in neurodegenerative diseases. *The international journal of biochemistry & cell biology* **42**, 1606–1609 (2010).
40. Aulas, A., Stabile, S. & Vande Velde, C. Endogenous TDP-43, but not FUS, contributes to stress granule assembly via G3BP. *Molecular neurodegeneration* **7**, 54 (2012).
41. Parker, S. J. *et al*. Endogenous TDP-43 localized to stress granules can subsequently form protein aggregates. *Neurochem Int* **60**, 415–424 (2012).
42. Lee, K., Suzuki, H., Aiso, S. & Matsuoka, M. Overexpression of TDP-43 causes partially p53-dependent G2/M arrest and p53-independent cell death in HeLa cells. *Neurosci Lett* **506**, 271–276 (2012).
43. Huang, G. *et al*. Death receptor 6 (DR6) antagonist antibody is neuroprotective in the mouse SOD1G93A model of amyotrophic lateral sclerosis. *Cell Death Dis* **4**, e841 (2013).
44. Spencer, S. L. *et al*. The proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. *Cell* **155**, 369–383 (2013).
45. Jonkers, J. *et al*. Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nature genetics* **29**, 418-425 (2001).
46. Tchorz, J. S. *et al*. A modified RMCE-compatible Rosa26 locus for the expression of transgenes from exogenous promoters. *PloS one* **7**, e30011 (2012).
47. Jacks, T. *et al*. Tumor spectrum analysis in p53-mutant mice. *Current biology: CB* **4**, 1–7 (1994).
48. Devine, M. J. *et al*. Parkinson's disease induced pluripotent stem cells with triplication of the alpha-synuclein locus. *Nature communications* **2**, 440 (2011).

Acknowledgements

We thank Dr. E. Buratti and F.E. Baralle for providing the TDP-43$^{\Delta RRM1}$ cDNA. We are grateful to Frank Sager for excellent technical assistance and to the members of the Taylor lab for continual discussion and for critical reading of the manuscript. This work was supported by the Forlen Stiftung, Deutsche Forschungsgemeinschaft (DFG SFB592; TA-310-1; TA-310-2), and the Max Planck Society and the University of Basel. The work of Wolfgang Wurst was supported by the German Science Foundation Collaborative Research Centre (CRC) 870 and the Munich Cluster of Systems Neurology (WXC1010 SyNergy).

Author Contributions

M.V., Z.E., P.K. and V.T. designed the protocols and the experiments; M.V., Z.E., P.K. performed the experiments and analyzed the data; M.S.H., W.W. and T.F. generated the *Tardbp*^{315} animals, performed the PFTα treatment of the embryo and isolated the embryos for analysis; T.K., K.E. and L.W. provided human iPS cell lines; A.H. and P.J.S. provided TDP-43 mutant cDNAs; M.V. and V.T. wrote the manuscript; all authors read, commented and corrected on the manuscript; V.T. supervised the project.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-26397-2.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2-3 Multigene delivery in hiPSCs and genome editing of these cells

Multigene delivery and subsequent cellular expression is a key technology for a wide range of applications in biology including synthetic biology, cellular reprogramming, functional pharmaceutical screening and cell labelling. For certain cell types including widely used HEK293 and HeLa cells, these goals can be achieved by plasmid-based transfection. However, a large number of cell lines and particularly primary cells and stem cells are usually recalcitrant to plasmid transfection, thus different approaches are required. Infection by viral vectors emerged as the dominant method of choice to deliver genes into hard-to-transfect cells including primary cells and stem cells. Common viral vectors including Lentivirus, Adenoviruses, Adeno-Associated Viruses (AAV) suffer from one or more of the following disadvantages: small cargo capacity, integration to the host DNA genome, replication in mammalian cells, inability to transduce both dividing and non-dividing cells.

In contrast baculovirus, has a large capacity for foreign DNA delivery and infect not only insect but also mammalian cells (both dividing and non-dividing cells). Transduction by baculovirus is transient without DNA integration into the target cell genome. Therefore, baculoviral vectors can be consider safe for gene and cell therapy purposes. In the last decade, baculovirus has emerged as a useful and safe technology to deliver heterologous genetic material to mammalian cell types both *in vitro* and *in vivo*.

In a project led by Dr. Philipp Berger at Paul Scherrer Institute, we could tackle the formidable challenge of multigene delivery to “hard-to-transfect” mammalian cells. In collaboration with our lab, Dr. Berger and colleagues introduce MultiPrime, a baculovirus-based tool-kit for highly efficient multigene delivery in a wide range of mammalian cells including immortalized cell lines, primary cells, and stem cells. The MultiPrime system consists of an easy-to-use assembly pipeline to put together many genes into large DNA constructions. These are then inserted into an engineered baculovirus optimized for mammalian cell transduction. They showed versatility of MultiPrime for multigene delivery in primary and stem cells by production of complex proteins (e.g., human IgGs), direct-trans differentiation of mouse embryonic fibroblasts to induced-neuron, expression of multiple fluorescently-labelled biomarkers for cell trafficking studies and also in vivo multigene delivery in Zebrafish embryos. My role in this paper (Mansouri et al, Nat. Commun. 2016) included

transduction of hiPSC with two different types of baculovirus (with or without VSV-G psedutyped baculovirus) and analysis of samples by microscopy.

Emerging technologies such as CRISPR/Cas9-based genome engineering, which enable any alteration at the genomic level of an organism, require delivery of multiple constructs in a single cell. In the second project led by Dr. Philipp Berger, MultiPrime was customized as an all-in-one viral vector for efficient HDR-based genome editing in a varity of mammalian cells including cell line, primary cells and hiPSCs. This project showed that a single baculovirus particle carrying Cas9, gRNA(s) and template can be used successfully to insert EGFP cDNA into the C-terminal of the endogenous HMGA-1 allele. My role in this paper included preparation of human induced-pluripotent stem cells, transduction of hiPSC with baculovirus harbouring Cas9-gRNA-HDR or Cas9-gRNA1-gRNA2-HDR constructs, microscopic assessment of EGFP expression from the endogenous HMGA-1 allele, genomic DNA extraction and PCR genotyping of hiPSC genome edited cells. Subsequently, I isolated the EGFP labelled iPSCs using FCAS to use in the co-culture experiment mentioned in my main paper.

CHAPTER 3
Discussion

The limited availability of human brain tissue requires alternative methods to study fundamental processes underlying brain physiology including the mechanisms controlling human cortical development. During recent years, *in vitro* recapitulation of cortical differentiation derived from human PSCs gained attention as a cell source for studying human cerebral cortex development as well as disease-specific mechanisms including Alzheimer[63,64], schizophrenia[65,66], Miller dieker syndrome[54] and microcephaly[67]. The *in vitro* culture system derived from pluripotent cells recapitulate the prominent aspects of neurogenesis *in vitro.* For instance, the differentiation process starts with a neural induction step in which the pluripotent stem cells give rise to different neural stem and progenitor cells[68]. The neural stem/progenitors generate different types of cortical neurons including deep and upper layer neurons in a sequential fashion. Astrocytes are also generated at late stages of differentiation. Interneurons are also generated using specific protocols utilizing ventralizing factors[69].

How neural stem cells/progenitors generate the great diversity of neuronal subtypes including deep and upper layer neurons and how these sequential waves of neuron generation are achieved during cortical development remains to be elucidated. Over the last decade, two different models that explain neural stem cell specification in different projection neuron cell types have been proposed. In the “common progenitor model” a single type of stem cell generates different subtypes of neurons and also glial cells. These multipotent stem cells become increasingly fate-restricted during development. In the “multiple progenitors model” different stem cells are committed to generating specific subtypes of neural neurons and glia[35].

With the aim to explore how neural stem cells/progenitors generate different neuronal subtypes during human cortical neurogenesis, we analyzed the fate of iPSCs-derived neural progenitors using 2D culture-based cortical differentiation. The 2D culture based cortical differentiation has several advantages over the current 3D (or organoids) cultures based system. First, the appearance of a homogenous neuroepithelial cell population after neural induction step rather than heterogeneous cell pool in the 3D structure. Second, in the 2D differentiation model upper layer neurons are generated in a reproducible manner when compare to 3D culture based systems.

In this study, the potency of neural progenitors over the course of differentiation was analyzed by birth-dating studies using retroviruses. With this method the NPs can be genetically labeled as the retroviruses labeled actively dividing cells. With this approach we could labeled NPs and followed their fate in culture and assess their

proliferation and specification in to neurons when they grow in their environment in contact with other cells. In order to ensure, that the retroviruses labeled only NPs, the cells were analyzed after 48 hours and the majority of GFP-labeled cells expressed PAX6 and SOX2. Although the differentiation potential of labeled NPs changes over time, we found that both early and late NP populations are multipotent and produce deep and upper layer cortical neurons. This argues against the hypothesis that late progenitors are restricted to generate upper layer neurons and NPs lose the ability to generate deep layer neurons over time[56]. The lineage-tracing studies in mice using retrovirus showed that single VZ progenitor could give rise to neurons corresponding to different cortical layers. Additionally, late progenitors mainly give rise to upper layer fate, demonstrating that the fate potential of NPs was progressively restricted[70-72].

Interestingly, our birth dating results indicate that the human NP potency does not diminish over time and they generate both deep and upper layer neurons. This could be explained by the fact that humans have notably larger cerebral cortices that contain more neurons than in rodents and some primate such as macaque. Analyzing the progeny of day 28 labeled progenitors we found the appearance of $SATB2^+$ upper layer neurons after almost 40 days of differentiation. However, $SATB2^+$ neurons derived from late-labeled progenitors are generated 20 days after labeling. This suggests that the NPs acquire an upper layer $SATB2^+$ fate progressively during differentiation. Therefore, our results support the common progenitor model in which the progenitors give rise to different types of neurons. However, we could not identify fate-restricted NP even at late time points during differentiation. One limitation of our approach is that we performed an analysis of NSCs differentiation at the population level without analyzing the progeny of single progenitors. Therefore, it is still possible that some of the single GFP labeled progenitor are giving rise to specific neuronal subtypes. Future experiments will address this point by performing clonal analysis of hiPSC-derived NSCs and by labeling with low titer of retrovirus single cell in culture. This approach will allow studying the lineage fate of single NSCs/progenitors during human cortical development. To further study molecular identity of human cortical NPs, I isolated NPs by FACS over the course *in vitro* cortical differentiation. I performed a comprehensive characterization of the FACS sorted NPs. By comparing gene expression with RNA-seq analysis we found that isolated NPs retain the expression human radial glial identity markers throughout cortical development. Moreover, we found the global transcriptomic profiles of NSCs/progenitors are dynamic over time. We hypothesize that the neuropotency of NPs is regulated by changes in gene

expression profile and stage specific transcription factor drives the differentiation process. Differential expression analysis of isolated NPs organized the progenitors into at 3 distinct clusters, which are corresponding to 3 stages during *in vitro* corticogenesis: Early progenitors, transitional progenitors, and late progenitors. Therefore, indicating that NSCs/progenitors are undergoing a transcriptional control of sequential regulators.

PCA analysis separates day 40 NP from early and progenitors suggesting this stage is transitional stage from VRG to oRG. In line with that, pairwise analysis of expression profile of consecutive time points through the differentiation protocol (i.e. from day 25 to 90 days), we found that day 40 and day 60 NSCs/progenitors exhibit the largest difference in gene expression. Moreover, these differentially expressed genes correlate with cell fate switch of VRG to oRG and we found the overexpression of molecular oRG transcriptional signatures[45] including HOPX, TNC, MOXD1 and PTPRZ1 from day 60 NPs. These markers are involved in epithelial to mesenchymal transition, stem cell maintenance and extracellular matrix production.

In human, oRG populations are responsible for the generation of the majority of cortical neurons especially upper layer neurons[19,45]. Therefore, the increase in the oRG cell population markers at late stages of differentiation correlates well with the increased generation of upper layer neurons by late NPs. One can overexpress the oRG markers in early NPs to accelerate production of upper layer neurons during differentiation process.

Our NPs transcriptomic data will allow us to further examine the molecular mechanisms underlying human cortical development. One of the most interesting open question is: Which signaling pathways orchestrate NPs fate acquisition at different stages of cortical development? We found that the expression of LIFR and STAT3 increase from early NPs to late NPs. Consistent with our data, LIF signaling together with STAT3 promote human oRG maintenance. Moreover, addition of LIF to culture medium can be used to evaluate a possible role for LIF in inducing upper layer neuron generation. Late isolated NPs (day 60-90) enriched expression of glial markers such as GFAP, S100B, NFIB, NFIX, and CD44.

One exciting finding in this study is the identification of NPs with the prominent characteristics of vRG and oRG cells through the differentiation process.

Our transcriptomic analyses indicate that NPs is coordinated by cell intrinsic mechanisms. Moreover, we could also show that environmental signals also control the fate potential of NPs. The neuropotency of day 25 and 90 NPs changed after they were exposed to a hetechronic environment. The capacity of day 90 NPs to generate $SATB2^{+}$ upper layer neurons is reduced in the presence of early cortical

cultures. However, day 25 NPs do not increase SATB2$^+$ upper layer neuron production in presence of day 90 culture. This could be due to intrinsic developmental program of NPs, and the fact that signals produced in the culture could not immediately promote the upper layer fate.

The signaling factors derived from day 25 differentiated cells inhibit differentiation of SATB2 upper layer neurons from late day 90 progenitors. This inhibitory effect might be critical for generation of deep and upper layer neurons in a sequential fashion. What could these environmental factors be that regulate NSC differentiation? Cell-cell interactions and diffusible signals could be ideal candidates to exert this function. For example, FGF, WNT and NOTCH signaling play key roles during cortical development. Consistently, we see the down regulation of WNT ligands including WNT1 and WNT8b from early to late NPs. It would be interesting to block these signaling pathways after re-plating day 25 NPs to see whether they could give rise to SATB2 neurons.

In summary, our findings provide an in-depth analysis of *in vitro* human cortical differentiation and shed light on some mechanisms that underlie neural specification. I reported here that multipotent population of NPs give rise to different neuronal subtypes during *in vitro* cortical differentiation. A fine interplay between intrinsic mechanisms and extracellular signals dictate NSC fate specification during cortical development. The data presented represent a valuable resource that clarifies how the generation of specific neuronal cell populations from human iPSCs *in vitro* is achieved.

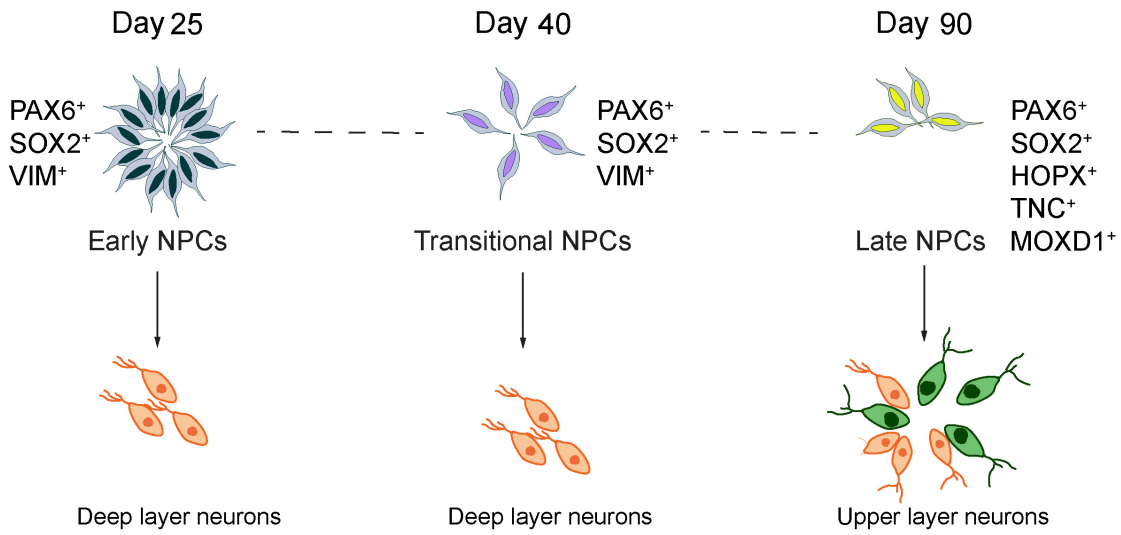


Figure 1: Three stages of *in vitro* cortical differentiation.

The early NPs express canonical radial glial markers and give rise to deep layer neurons. Day 40 NPs also differentiate to only deep layer neurons but the molecular profile of these cells are different from early and late NPs. The NPs acquire oRG identity at late stages of differentiation. Moreover, compared with early NPs the late NPs generate more upper layer neurons.

CHAPTER 4
Conclusion and Outlook

Our study applies an *in vitro* cellular model of human cortical differentiation using iPSC to investigate differentiation potential and gene expression profile of human cortical neural progenitors (NPs). During the differentiation process NPs give rise to different cortical neurons in a sequential order in a similar way of cortical development. We identified multipotent NPs generate great diversity of neurons during *in vitro* cortical differentiation. Nevertheless, early and late NPs are different in terms of differentiation potential and day 90 NPs generate significantly less $TBR1^+$ deep layer neurons compare to day 28 NPs.

Through detailed analysis of transcription profile of isolated NPs we identified 3 major type of NPs during differentiation including, early NPs which are expressing canonical radial glial markers, mid NPs which are in transitional stage from vRG to oRG and late NPs with the molecular features of oRG cells. Our result from transcriptome study of NPs and co-culture experiments indicate that extrinsic signals combined with cell-intrinsic mechanisms regulate the neuropotency of NP. The signals from day 90 culture, as the potent differentiation signals, promote increase in deep and upper layer neurons derived from day 25 NPs. Conversely, day 90 NPs exposed to the young environment reduces their differentiation and produced less $SATB2^+$ and $TBR1^+$ neurons.

The research presented in this thesis opened a number of research question that should be in address in future. Together, this study could lead not only to a better understanding of *in vitro* model of human corticogenesis but also provide insights to generate specific neuronal subtypes for drug screening and cellular therapy. The single cell resolution fate mapping of progenitors is required to find the fate-committed progenitors over the course of differentiation. Furthermore, fate potential of isolated NPs can be examined by transplantation of isolated NPs to animal models. Diverse neuronal subtypes are generated in protracted process during *in vitro* cortical differentiation. Based on our study over expression of oRG correlated markers in early NPs can accelerate differentiation of cortical neurons. I also believe that elucidating signaling pathways involved in switch from vRG cells to oRG cells is a way to identify conditions, which rapidly differentiate hPSCs into different type of cortical neurons.

CHAPTER 5
Materials and Methods

Methods

Cell line models

Human induced pluripotent stem cell (hiPSCs) lines used were; NAS2, provided by Tilo Kunath[66], and HMGA1::eGFP cell line, were made using Baculovirus based CRISPR/Cas9. Engineered line was generated as previously described[61]. See METHOD DETAILS for more details.

Maintenance, culture of hiPS

The iPSCs were cultured in using mTESR (Stem Cell Technologies, mTeSR1 Basal Medium, Cat#05850 and mTeSR1 5x supplement, Cat#05851) media on Matrigel (BD Bioscience, hESC qualified Matrigel, Cat#354277). Before plating the cells, the cell culture dishes were coated with hESC qualified Matrigel according to manufacturer's guidelines. In brief: Thaw Matrigel on ice and add 25ml DMEM-F12 according to certificate of analysis. This step has to be quick and everything should stay cold below 10°C. Add 1ml to (B3.5 plate), add 2 ml (B6 plate) and add 7ml (B10 plate). The dishes were incubated at room temperature for at least 1 hour. The coted dishes can be stored in up to 1 week at 4°C (fridge).

The growing medium was taken from cells. Cells have been washed with PBS once. Cells were incubated with Accutase for 2-5 minutes. When the cells were rounded and detach from dishes, the medium was added to dilute Accutase and then centrifuge at 1000rpm for 4 minutes to pellet the hiPSCs. The cells were resuspended in mTESR1 medium + 10µM ROCK-Inhibitor. The cells were fed daily and passaged every 4 days when the cells reached to 80-90% confluent.

Freezing of hiPSCs:

The hiPSCs were frozen from cultured iPSCs at early passage (maximum passage number 10). For freezing the cells were detached using Accutase, then harvested in a 15 ml falcon tube. The supernatant removed and cells frozen Freezing Solution (10%DMSO + 90% FBS + 10uM ROCK-Inhibitor). The cells were frozen in density of $1.5*10^6$ cells up to $4*10^6$ cells per cryovial/ml. the vials were kept in cold Mr. Frosty boxes and stored at -80°C for 24 h and then transferred into gas phase of the nitrogen tank.

Thawing, culture of hiPSCs from frozen vials

Cryovial were thawed in in a water bath at 37°C until only a small amount of ice is left. Then, the cells were transfer to a 15 ml centrifuge tube containing 4 ml of mTESR1 medium. Centrifuge at 1000rpm for 4 minutes to pellet the hiPSCs. The cells after thawing were cultured on Matrigel coated plate on the density of $2*10^6$ cells per B6 plate. The cells were incubated at 37°C, 95% humidity and 5% CO_2 and monitored daily. The growth medium should be changed after every 24 hours.

Characterization of hiPSCs

Morphological examination and immunocytochemistry for pluripotency markers were performed as indicators of pluripotency. hiPSCs were evaluated as high qualified pluripotent stem cells when they showed high nucleo-cytoplasmic ratio as well as a well-defined nucleoli with a variable number. Further analysis of pluripotency was performed by immunostaining for the pluripotency markers OCT-3/4, NANOG and SOX2. Antibodies are described in Table S1.

Figure 1. a. Phase contrast image of iPSCs shows the intact colony of hiPSCs with high nucleo-cytoplasmic ratio morphology. b. Analysis of three key pluripotency markers (Nanog, Oct3/4 and Sox2) expression by immunofluorescence shows that all of the cells express these markers.

Genome engineering of hiPSCs

Plasmids for targeting the HMGA1 locus were kindly provided by Dr.Stefan Jakobs (Ratz et al., 2015). Cas9 was excised together with gRNA1 by AflIII/NotI digestion from pX330-HMGA1-gRNA1 vector and cloned into the NcoI/NotI site of pSI-AG15 and pSI-AG16 generating the pSIAG15/16-gRNA1-Cas9 vector. The HDR construct AscI/PacI fragment for insertion of EGFP was obtained from a pGEM-HDR vector and cloned AscI/PacI into pSI-AG15/16-gRNA1-Cas9 to generate pSI-AG15/16-gRNA1-Cas9-HDR vector. The second HMGA1-gRNA (gRNA2) was isolated from pX330-HMGA1-gRNA2 as a AflIII/XhoI fragment and cloned into the BsaI/NheI digested pSI-AG15/16-gRNA1-Cas9-HDR vector to generate pSI-AG15/16-gRNA1-Cas9-HDR-gRNA2. Control plasmid pSIAG15/16-HDR was constructed by cloning of HDR through AscI/PacI to pSI-AG15/16.

Baculoviruse was generated in sf21 insect cells as previously described (Mansouri et al.,2016)[67]. For transduction of hiPSCs, the cells were plated on Matrigel coated dishes at the density of $2.5*10^5$ cells per well in six well plate one day before

transduction. Transduced cells were incubated at 37 ^{0}C for 8 hours and then the medium was replaced with fresh medium. The cells were cultured for longer time and analyzed 5 or 6 days after transduction. The iPSCs were transfected with FusionHD (Promega) according to manufacturer's recommendations and cells were analyzed 42 h after transfection.

Differentiation of human pluripotent stem cells to cortical neurons

The cells were directed differentiated to different cortical neurons with slight modification of Shi, et. al 2012. The hiPSCs cells were dissociated to single cells using accutase for 3 minutes, centrifugedat 800 rpm for 4 minutes and plated at the density of 2*10^{5}cells/cm^{2} on Matrigel coated plate in mTESR medium supplemented with Y- 27632 ROCK inhibitor (10μM). Once the cells became 100% confluent (one day after plating), the medium was changed to N2B27 differentiation medium supplemented with SB431542 (10 μM; Stemgent, Cat#040074) and LDN193189 (100 nM; Miltenyi Biotec, Cat#130-096-226). N2B27 differentiation medium consists of a 1:1 mixture of N2 medium and B27 medium, where N2 medium is: DMEM/F12 glutamax (Gibco, Cat#31331-028), N2 supplement (Gibco, Cat#17502-048), MEM non essential amino acids (Gibco, Cat# 11140-050), 2 mercaptoethanol (Gibco, Cat# 21985-023), and B27 medium is: Neurobasal (Gibco, Cat# 21103), B27 supplement (Gibco, Cat#17504-044), glutamax (Gibco, Cat#35050038).

The medium has to be changed every day until day 11 of differentiation. The cells passaged at day 12 of differentiation using dispase (BD, Cat#354235) diluted 1:50 in K/O DMEM and plated as 1:1 ratio on poly-L-Lysine/laminin coated coverslips or multiwells (0.1mg/mL poly-L-Lysine Sigma, Cat#P9155, 10 μg/mL laminin-Roche Cat#11243217001). The cells were incubated with dispase for 3-4 minutes. Then cells were washed with K/O DMEM once and detached with cell scraper. This step has to be done gently in order to preserve clumps. The sheet of cells were dissociated in the well using P1000 with a blue tip. Adjust timing of pipetting for each line (for Nas2 iPS cells we use 5 times pipetting). The cell suspension was transferred into the 15ml Falcon tube containing 0.5 ml N2B27. The walls were washed with 0.5 ml N2B27 wash the well and the remaining cells were transferred to the same tube. The suspension was mixed by gentle swirling. The tube was standed in the hood for 1 minute exactly then, the supernatant was removed (leave 100-150 μL). Before transferring the clump were resuspended with 7 times pipetting with a blue tip in 1 ml N2B27 then the cell suspension was transferred to one well of 12 well plate. The cells were fed once another day with 1 ml N2B27 (12 well plate) and these cells were formed Rosette Between days 15- 17.

Between days 20-30 of differentiation substantial neurogenesis should occur, when neurons first begin to accumulate around of the Rosettes, cells should be passaged using Accutase. At day 25 of differentiation the cells were passaged using accutase and plated on poly-L-Lysine/laminin (0.1mg/mL poly-L-Lysine 20 µg/mL laminin) coated dishes in the density of 50000 cells/cm^2. From day 25 of differentiation till 110 the cells were grown in N2B27 medium supplemented with BDNF BDNF (10 ng/mL; Miltenyi Biotec, Cat#130-096-286), GDNF (10 ng/mL; Miltenyi Biotec, Cat#130-098-449), Further, the cells passaged once more at day 31 in the density of 50000 cells/cm^2 on poly-L-Lysine/laminin (0.1mg/mL poly-L-Lysine 20 µg/mL laminin) coated coverslips or multi wells. At each stages of dissociation cells were grown in medium containing ROCK inhibitor. Half volume of medium was changed every day and the cells were cultured for cortical differentiation until day 110 according to experimental goals

The coated wells have to be prepared one day before splitting. The dishes were incubated with poly L lysine (0.1mg/ml)(300 µL for 24 well plate and 600 µL for 12 well plate) at 37°C for one hour. Then the coated wells were washed 3 times with PBS. Then laminin (20 µg/ml) (300 µl for 24 well plate and 500 µL for 12 well plate) was added to each well and incubated overnight at 37°C

Viral infection of neural progenitors

Retroviruses were generated using the packaging cell line PLAT-E (Cell Biolabs). PLAT-E cells were transfected with viral plasmid constructs and using Calphos™ mammalian transfection kit according to manufacturer's instructions. Viral supernatants were collected 24 and 48 hours after transfection and the virus was purified using Retro-X™ concentrator according to the manufacturer's instructions.

For viral infection, the cortical cultures were infected with retrovirus at day 28, 60, 80 and 90. Virus was added in a small volume of N2B27 medium for 3-4 hours and the cells afterwards cultured in adequate volume N2B27 medium. Cortical cultures were then chased until 110 days of differentiation. Cells were being fixed and analyzed at day 40, 60, 80, 90 and 110 of differentiation. Cells were then immunostained for GFP and each layer-specific marker ($CTIP2^+$, $TBR1^+$, $BRN2^+$, $SATB2^+$) and the percentages of GFP^+ cells for $CTIP2^+$/$TBR1^+$/ $BRN2^+$/ $SATB2^+$cells were quantified.

Fluorescence activated cell sorting for isolation of neural stem cells during in vitro cortical differentiation

In order to sort neural progenitors we used the protocol adapted from Yuan et al. the cortical cultures were detached using Accutase for 5 minutes with gentle resuspension to dissociate the cells to single cells. After centrifugation at 800 rpm the cell pellets were incubated with DNase (100 UNIT/mL; Roche, Cat#10104159001) for 10 minutes at room temperature. Cells were then centrifuged at 800 rpm for 4 minutes to remove the DNase and resuspended in sorting media. Sorting media is composed of N2B27 growth medium (w/o phenol red) with addition of 0.5% BSA and 5mM EDTA. Then the cell suspensions were strained through a 70µm cell strainer (BD Biosciences, Cat#352350). The cells afterwards stained with four flourochrome-conjugated antibodies for 20 minutes on ice in dark. Antibodies used for flow cytometric analysis are shown in Supplementary Table S1. After primary antibody incubation, the cells were washed at least twice and centrifuged at 800 rpm for 4 minutes. The cells were re-suspended in sorting media and filtered through a 70µm cell strainer. The cells were sorted on a fluorescence-activated cell sorter FACSAria^{TM} III (BD Bioscience) and analyzed using FACSDiva software (BD Biosciences). Data were additionally analyzed and presented using FlowJo software. For each cell preparation, a negative control (minimal number of 100.000 cells) was incubated just in the presence of the secondary antibodies. For stained cells, a minimal range of 300.000-500.000 cells was analyzed by flow cytometry. All analyses and sorts were repeated at least three times. A 85-µm ceramic nozzle (BD Biosciences), sheath pressure of 20–25 pounds per square inch (PSI), and an acquisition rate of 1.000–3.000 events per second were used. Antibodies used for flow cytometry analysis were firstly titrated.

Neural differentiation of sorted neural progenitors

The neural progenitors after sorting were plated on poly-L-Lysine/laminin (0.1mg/mL poly-L-Lysine, 20 µg/mL laminin) coated dish in the density of 50000 cells/cm^2. The cells were differentiated for 20 days in the N2B27 media supplemented with BDNF and GDNF, and the medium for first day was supplemented with Rock inhibitor.

For in vitro grafting experiment, the donor cortical cultures were derived from HMGA1::eGFP hiPSCs. After the differentiation of HMGA1::eGFP hiPSCs, the neural progenitors were sorted at day 25 and 90 of differentiation and plated on the host cortical cultures (day 25 and day 90) derived from NAS2 hiPSCs.

RNA extraction and RNA-seq

In order to isolate RNA, cells were lysed directly in Trizol (Invitrogen) reagent.

RNA was isolated according to the manufacturer's instructions.

Cells were directly sorted into 800ul of Trizol. RNA was extracted in the aqueous phase after addition of 160ul chloroform, precipitated with an equal volume of isopropanol and washed once in fresh 80% ethanol. The RNA pellet was re-suspended in 3ul water at 50C for 10 minutes and then kept on ice.

RNA-sequencing

The integrity of the RNA samples was check using the RNA 6000 Pico Complete Kit on the Agilent 2100 Bioanalyzer. Their concentration was measured using Quant-IT RiboGreen RNA Assays (Life Technologies). Illumina RNA-sequencing libraries were prepared using the Smart-seq2 protocol for reverse transcription and cDNA amplification (Picelli et al, 2014) and the Nextera XT DNA library preparation kit (Illumina) following Fluidigm's guidelines for single-cell RNA-seq on the C1 system. In brief, 2 ng of total RNA in 2.5 µl were used as input for the Smart-seq2 reverse transcription reaction and 18 PCR cycles for the cDNA amplification. The quality of the cDNAs was evaluated using the Fragment Analyzer high-sensitivity NGS kit (AATI) and the concentration quantified with the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies). For Nextera XT DNA library preparation using 25% of the standard reaction volume the cDNAs were adjusted to 0.3 ng/µl. Libraries were quality checked using the Fragment Analyzer high-sensitivity NGS kit (AATI) and pooled according to the concentration values given by the Fragment Analyzer. The pool was purified twice using 0.9x and 1.0x volume AMPure XP beads (Beckman Coulter) and sequenced SR75 on an Illumina NextSeq 500 system (75 cycles High Output v2 kit).

Imaging

Stained coverslips or stained wells in microplates were analyzed with Apotome (Zeiss observer. Z1) or Zeiss LSM510 confocal microscope (Zeiss). Images were acquired using Zen pro 2012 (Zeiss) or Zeiss LSM4.2 (Zeiss) and processed with ImageJ64 or photoshop CS5 (adobe software).

Quantification and statistical analysis of data

Randomly selected, stained cells were analyzed and Data are presented as averages of a minimum of three independent differentiation experiments. Percentages were converted by arcsine transformation. Statistical comparisons were determined by two-tailed unpaired Student's *t*-test, Kruskal-Wallis with Dunn post-

hoc test and two way ANOVA test for cross comparison of 3 and more data sets. Significance was established at $p<0.05$. In all graphs, deviance from mean was displayed as standard error of the mean, SEM.

RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
BRN2	Santa Cruz	Cat#SC-6029
CTIP2	Abcam	Cat#AB18465
MAP2	Neuromics	Cat#CH22103
SATB2	Abcam	Cat# ab51502
OCT-3/4	Santa Cruz	Cat#SC-5279
SOX2	Santa Cruz	Cat#SC-17320
TBR1	Abcam	Cat# AB31940-100
βIII-TUBULIN	Sigma	Cat#T8660
APC anti-human CD184 (CXCR4) Antibody	Biolegend	Cat#306510
Briliant violet 421 anti human CD24 Antibody	Biolegend	Cat#311122
Brilliant Violet 605™ anti-mouse/human CD44 Antibody	Biolegend	Cat#103047
PE Mouse Anti-Human CD271	BD Pharmingen	Cat# 557196
Alexa 488 donkey anti-rabbit Alexa 488 donkey anti-mouse Alexa 488 donkey anti-goat Alexa 488 donkey anti-rat Alexa 488 donkey anti-chicken	Jackson ImmunoResearch Laboratories	Cat#711-545-152 Cat#715-546-151 Cat#705-545-147 Cat#712-546-153 Cat#703-545-155
Cyanine 3 donkey anti-rabbit Cyanine 3 donkey anti-mouse Cyanine 3 donkey anti-goat Cyanine 3 donkey anti-rat	Jackson ImmunoResearch Laboratories	Cat#711-165-152 Cat#715-165-151 Cat#705-165-147 Cat#712-166-153

Cyanine 5 donkey anti-rabbit Cyanine 5 donkey anti-mouse Cyanine 5 donkey anti-goat Cyanine 5 donkey anti-rat	Jackson ImmunoResearch Laboratories	Cat#711-496-152 Cat#715-175-151 Cat#705-176-147 Cat#712-175-153
Chemicals, Peptides, and Recombinant Proteins		
mTESR1 complete kit for hESC and hiPSC maintenance	Stem Cell Technologies	Cat#5850
Neurobasal Medium	Gibco	Cat#21103
DMEM/F12, GlutaMAX	Gibco	Cat#31331-028
2-Mercaptoethanol	Gibco	Cat#21985-023
N2 supplement (100x)	Gibco	Cat#17502-048
B27 supplement (50x)	Gibco	Cat#17504-044
GlutaMAX™-I Supplement	Gibco	Cat#35050-061
Pen/Strepto	Gibco	Cat#15070063
MEM Non-Essential Amino Acids (NEAA)	Gibco	Cat#11140-050
StemPro Accutase	Gibco	Cat#A1110501
HyClone ES Cell Screened Fetal Bovine Serum	Thermo Scientific	Cat#SH30070033E
DMSO Hybri-Max	Sigma	Cat#D2650
PBS w/o Ca&Mg	Dulbecco's	Cat#H15-002
Trypan Blue Solution	Sigma	Cat#T8254
Matrigel hESC-qualified Matrix	BD Biosciences	Cat#354277
Laminin	Roche	Cat#11 243 217 001
Poly-L-Lysine hydrobromide	Sigma	Cat# P9155 (5 mg)
StemMACS Y27632	Miltenyi Biotec	Cat#130-104-169
LDN193189	Miltenyi Biotec	Cat#130-096-226
SB431542	Miltenyi Biotec	Cat#130-095-561
Purmorphamine	Miltenyi Biotec	Cat#130-104-465

BDNF	Miltenyi Biotec	Cat#130-096-286
GDNF	Miltenyi Biotec	Cat#130-098-449
Normal donkey serum	Jackson ImmunoResearch	Cat#017-000-121
Triton X-100	Fisher	Cat#BPE151-500
Albumin from bovine serum	Sigma	Cat#A3294
Paraformaldehyde	Sigma	Cat#41678-0030
Tween-20	Sigma	Cat#P1379
TRIzol Reagent	Invitrogen	Cat#VX15596026
OligodT(20) primer	Thermo Fischer Scientific	Cat#18418020
dATP	Thermo Fischer Scientific	Cat#10216-018
dCTP	Thermo Fischer Scientific	Cat#10217-016
dGTP	Thermo Fischer Scientific	Cat#10218-014
dTTP	Thermo Fischer Scientific	Cat#10219-012
Superscript III first strand kit	Invitrogen	Cat#18080051
5x FIREPol Master Mix	Solis BioDyne	Cat#04-11-00125
SensiFast Sybr HiROX kit	Bioline	Cat#Bio-92020
1kb ladder	Invitrogen	Cat#10787018
Agarose	Fischer Bioreagents	Cat#BPE1356-100
DNase I	Roche	Cat#10104159001
Resorufin β-D-glucopyranoside	Sigma	Cat#83639
RNEasy Mini Kit	Qiagen	Cat # 74104
Ion AmpliSeq RNA Library Kit		
Experimental Models: Cell Lines		

NAS2 hiPSCs		*Cell line name:* SA001 *Alternative names:* SA01, CEBe033-A *Registration:* NIH Human Embryonic Stem Cell Registry; NIHhESC-10-0085
Software and Algorithms		
Fiji (Image J)	NIH	https://imagej.nih.gov/ij/
FlowJo	TreeStar Inc.	https://www.flowjo.com/
Photoshop	Adobe	
Adobe Illustrator	Adobe	
Prism	GraphPad	
Zeiss ZEN	Zeiss	
Other		
Petri Dish 35 x 10 mm	BD	Cat#353001
6-well plate	BD	Cat#353224
12-well plate	BD	Cat#353225
24-well plate	BD	Cat#353226
96 well, black/clear, tissue culture treated plate, flat bottom with lid	BD	Cat#353219
Combitips advanced 1 mL	Eppendorf	Cat#0030089430
Combitips advanced 2.5 mL	Eppendorf	Cat#0030089448
Combitips advanced 5 mL	Eppendorf	Cat#0030089456
Combitips advanced 10 mL	Eppendorf	Cat#0030089464
Stripette serological pipettes, individual plastic wrapped, 2 mL	Carl Roth	
Stripette serological pipettes, individual plastic wrapped, 5 mL	Carl Roth	

Stripette serological pipettes, individual plastic wrapped, 10 mL	Carl Roth	
Stripette serological pipettes, individual plastic wrapped, 25 mL	Carl Roth	
Stripette serological pipettes, individual plastic wrapped, 50 mL	Carl Roth	
Falcon tube, 15 mL	BD Biosciences	
Falcon tube, 50 mL	BD	
Eppendorf tube, 1.5 mL	Sigma	
Eppendorf tube, 0.5 mL	Sigma	
Filter tips, 10 uL		
Filter tips, 20 uL		
Filter tips, 200 uL		
Filter tips, 1 mL		
500 mL rapid-flow bottle to filter, 0.2 um SFCA membrane	Thermo Fischer Scientific	Cat#291-4520
70 µm cell strainer	BD Biosciences	Cat#352350

References:

1 Herculano-Houzel, S. The human brain in numbers: a linearly scaled-up primate brain. *Front Hum Neurosci* **3**, 31, doi:10.3389/neuro.09.031.2009 (2009).

2 Azevedo, F. A. *et al.* Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *The Journal of comparative neurology* **513**, 532-541, doi:10.1002/cne.21974 (2009).

3 Eric R. Kandel, E., James H. Schwartz, Editor, Thomas M. Jessell, Editor, Steven A. Siegelbaum, Editor, A. J. Hudspeth, Editor, Sarah Mack, Art Editor.

4 Stiles, J. & Jernigan, T. L. The basics of brain development. *Neuropsychol Rev* **20**, 327-348, doi:10.1007/s11065-010-9148-4 (2010).

5 Silbereis, J. C., Pochareddy, S., Zhu, Y., Li, M. & Sestan, N. The Cellular and Molecular Landscapes of the Developing Human Central Nervous System. *Neuron* **89**, 248-268, doi:10.1016/j.neuron.2015.12.008 (2016).

6 DeFelipe, J., Alonso-Nanclares, L. & Arellano, J. I. Microstructure of the neocortex: comparative aspects. *J Neurocytol* **31**, 299-316 (2002).

7 Hansen, D. V., Lui, J. H., Parker, P. R. & Kriegstein, A. R. Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* **464**, 554-561, doi:10.1038/nature08845 (2010).

8 Fertuzinhos, S. *et al.* Selective depletion of molecularly defined cortical interneurons in human holoprosencephaly with severe striatal hypoplasia. *Cerebral cortex (New York, N.Y. : 1991)* **19**, 2196-2207, doi:10.1093/cercor/bhp009 (2009).

9 Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-956, doi:10.1016/j.cell.2005.08.020 (2005).

10 Gilbert., S. F. Developmental Biology, 6th edition.

11 Stern, C. D. Neural induction: 10 years on since the 'default model'. *Curr Opin Cell Biol* **18**, 692-697, doi:10.1016/j.ceb.2006.09.002 (2006).

12 Hemmati-Brivanlou, A. & Melton, D. A. Inhibition of activin receptor signaling promotes neuralization in Xenopus. *Cell* **77**, 273-281 (1994).

13 Sasai, Y., Lu, B., Steinbeisser, H. & De Robertis, E. M. Regulation of neural induction by the Chd and Bmp-4 antagonistic patterning signals in Xenopus. *Nature* **376**, 333-336, doi:10.1038/376333a0 (1995).

14 Levine, A. J. & Brivanlou, A. H. Proposal of a model of mammalian neural induction. *Dev Biol* **308**, 247-256, doi:10.1016/j.ydbio.2007.05.036 (2007).

15 Milet, C. & Monsoro-Burq, A. H. Neural crest induction at the neural plate border in vertebrates. *Dev Biol* **366**, 22-33, doi:10.1016/j.ydbio.2012.01.013 (2012).

16 Rallu, M., Corbin, J. G. & Fishell, G. Parsing the prosencephalon. *Nat Rev Neurosci* **3**, 943-951, doi:10.1038/nrn989 (2002).

17 Le Dreau, G. & Marti, E. Dorsal-ventral patterning of the neural tube: a tale of three signals. *Dev Neurobiol* **72**, 1471-1481, doi:10.1002/dneu.22015 (2012).

18 Kiecker, C. & Lumsden, A. The role of organizers in patterning the nervous system. *Annu Rev Neurosci* **35**, 347-367, doi:10.1146/annurev-neuro-062111-150543 (2012).

19 Lui, J. H., Hansen, D. V. & Kriegstein, A. R. Development and evolution of the human neocortex. *Cell* **146**, 18-36, doi:10.1016/j.cell.2011.06.030 (2011).

20 O'Rahilly, R. & Muller, F. The longitudinal growth of the neuromeres and the resulting brain in the human embryo. *Cells Tissues Organs* **197**, 178-195, doi:10.1159/000343170 (2013).

21 Goldman-Rakic, P. S. Architecture of the prefrontal cortex and the central executive. *Ann N Y Acad Sci* **769**, 71-83 (1995).
22 Bayer, S. A., Altman, J., Russo, R. J. & Zhang, X. Timetables of neurogenesis in the human brain based on experimentally determined patterns in the rat. *Neurotoxicology* **14**, 83-144 (1993).
23 Pinto, L. & Gotz, M. Radial glial cell heterogeneity--the source of diverse progeny in the CNS. *Prog Neurobiol* **83**, 2-23, doi:10.1016/j.pneurobio.2007.02.010 (2007).
24 Weissman, T., Noctor, S. C., Clinton, B. K., Honig, L. S. & Kriegstein, A. R. Neurogenic radial glial cells in reptile, rodent and human: from mitosis to migration. *Cerebral cortex (New York, N.Y. : 1991)* **13**, 550-559 (2003).
25 Clancy, B., Darlington, R. B. & Finlay, B. L. Translating developmental time across mammalian species. *Neuroscience* **105**, 7-17 (2001).
26 Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S. & Kriegstein, A. R. Neurons derived from radial glial cells establish radial units in neocortex. *Nature* **409**, 714-720, doi:10.1038/35055553 (2001).
27 Parnavelas, J. G., Alifragis, P. & Nadarajah, B. The origin and migration of cortical neurons. *Prog Brain Res* **136**, 73-80 (2002).
28 Corbin, J. G., Nery, S. & Fishell, G. Telencephalic cells take a tangent: non-radial migration in the mammalian forebrain. *Nat Neurosci* **4 Suppl**, 1177-1182, doi:10.1038/nn749 (2001).
29 Anderson, S. A., Marin, O., Horn, C., Jennings, K. & Rubenstein, J. L. Distinct cortical migrations from the medial and lateral ganglionic eminences. *Development* **128**, 353-363 (2001).
30 Huang, Z. Molecular regulation of neuronal migration during neocortical development. *Mol Cell Neurosci* **42**, 11-22, doi:10.1016/j.mcn.2009.06.003 (2009).
31 Valiente, M. & Marin, O. Neuronal migration mechanisms in development and disease. *Curr Opin Neurobiol* **20**, 68-78, doi:10.1016/j.conb.2009.12.003 (2010).
32 Molyneaux, B. J., Arlotta, P., Menezes, J. R. & Macklis, J. D. Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci* **8**, 427-437, doi:10.1038/nrn2151 (2007).
33 Gaspard, N. & Vanderhaeghen, P. Laminar fate specification in the cerebral cortex. *F1000 Biol Rep* **3**, 6, doi:10.3410/B3-6 (2011).
34 Tomassy, G. S., Lodato, S., Trayes-Gibson, Z. & Arlotta, P. Development and regeneration of projection neuron subtypes of the cerebral cortex. *Sci Prog* **93**, 151-169 (2010).
35 Franco, S. J. & Muller, U. Shaping our minds: stem and progenitor cell diversity in the mammalian neocortex. *Neuron* **77**, 19-34, doi:10.1016/j.neuron.2012.12.022 (2013).
36 Qian, X., Goderie, S. K., Shen, Q., Stern, J. H. & Temple, S. Intrinsic programs of patterned cell lineages in isolated vertebrate CNS ventricular zone cells. *Development* **125**, 3143-3152 (1998).
37 McConnell, S. K. & Kaznowski, C. E. Cell cycle dependence of laminar determination in developing neocortex. *Science* **254**, 282-285 (1991).
38 Frantz, G. D. & McConnell, S. K. Restriction of late cerebral cortical progenitors to an upper-layer fate. *Neuron* **17**, 55-61 (1996).
39 Desai, A. R. & McConnell, S. K. Progressive restriction in fate potential by neural progenitors during cerebral cortical development. *Development* **127**, 2863-2872 (2000).
40 Franco, S. J. *et al.* Fate-restricted neural progenitors in the mammalian cerebral cortex. *Science (New York, N.Y.)* **337**, 746-749, doi:10.1126/science.1223616 (2012).

41 Zecevic, N., Chen, Y. & Filipovic, R. Contributions of cortical subventricular zone to the development of the human cerebral cortex. *The Journal of comparative neurology* **491**, 109-122, doi:10.1002/cne.20714 (2005).
42 Johnson, M. B. *et al.* Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat Neurosci* **18**, 637-646, doi:10.1038/nn.3980 (2015).
43 Lui, J. H. *et al.* Radial glia require PDGFD-PDGFRbeta signalling in human but not mouse neocortex. *Nature* **515**, 264-268, doi:10.1038/nature13973 (2014).
44 Florio, M. *et al.* Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science* **347**, 1465-1470, doi:10.1126/science.aaa1975 (2015).
45 Pollen, A. A. *et al.* Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55-67, doi:10.1016/j.cell.2015.09.004 (2015).
46 Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199-206, doi:10.1038/nature13185 (2014).
47 Espuny-Camacho, I. *et al.* Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo. *Neuron* **77**, 440-456, doi:10.1016/j.neuron.2012.12.011 (2013).
48 Gaspard, N. *et al.* Generation of cortical neurons from mouse embryonic stem cells. *Nat Protoc* **4**, 1454-1463, doi:10.1038/nprot.2009.157 (2009).
49 Gaspard, N. *et al.* An intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature* **455**, 351-357, doi:10.1038/nature07287 (2008).
50 van de Leemput, J. *et al.* CORTECON: a temporal transcriptome analysis of in vitro human cerebral cortex development from human embryonic stem cells. *Neuron* **83**, 51-68, doi:10.1016/j.neuron.2014.05.013 (2014).
51 Edri, R. *et al.* Analysing human neural stem cell ontogeny by consecutive isolation of Notch active neural progenitors. *Nat Commun* **6**, 6500, doi:10.1038/ncomms7500 (2015).
52 Yao, Z. *et al.* A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development. *Cell Stem Cell* **20**, 120-134, doi:10.1016/j.stem.2016.09.011 (2017).
53 Shi, Y., Kirwan, P., Smith, J., Robinson, H. P. & Livesey, F. J. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nat Neurosci* **15**, 477-486, S471, doi:10.1038/nn.3041 (2012).
54 Bershteyn, M. *et al.* Human iPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia. *Cell Stem Cell* **20**, 435-449 e434, doi:10.1016/j.stem.2016.12.007 (2017).
55 Chambers, S. M. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* **27**, 275-280, doi:10.1038/nbt.1529 (2009).
56 Qian, X. *et al.* Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. *Neuron* **28**, 69-80, doi:10.1016/S0896-6273(00)00086-6 (2000).
57 Shen, Q. *et al.* The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nat Neurosci* **9**, 743-751, doi:10.1038/nn1694 (2006).
58 Yuan, S. H. *et al.* Cell-surface marker signatures for the isolation of neural stem cells, glia and neurons derived from human pluripotent stem cells. *PLoS One* **6**, e17540, doi:10.1371/journal.pone.0017540 (2011).
59 Lukaszewicz, A. *et al.* G1 phase regulation, area-specific cell cycle control, and cytoarchitectonics in the primate cortex. *Neuron* **47**, 353-364, doi:10.1016/j.neuron.2005.06.032 (2005).

60 Smart, I. H., Dehay, C., Giroud, P., Berland, M. & Kennedy, H. Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. *Cerebral cortex (New York, N.Y. : 1991)* **12**, 37-53 (2002).
61 Mansouri, M., Ehsaei, Z., Taylor, V. & Berger, P. Baculovirus-based genome editing in primary cells. *Plasmid* **90**, 5-9, doi:10.1016/j.plasmid.2017.01.003 (2017).
62 Rakic, P. Neurons in rhesus monkey visual cortex: systematic relation between time of origin and eventual disposition. *Science* **183**, 425-427 (1974).
63 Yang, J., Li, S., He, X. B., Cheng, C. & Le, W. Induced pluripotent stem cells in Alzheimer's disease: applications for disease modeling and cell-replacement therapy. *Mol Neurodegener* **11**, 39, doi:10.1186/s13024-016-0106-3 (2016).
64 Israel, M. A. *et al.* Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature* **482**, 216-220, doi:10.1038/nature10821 (2012).
65 Hoffman, G. E. *et al.* Transcriptional signatures of schizophrenia in hiPSC-derived NPCs and neurons are concordant with post-mortem adult brains. *Nat Commun* **8**, 2225, doi:10.1038/s41467-017-02330-5 (2017).
66 Brennand, K. J. *et al.* Modelling schizophrenia using human induced pluripotent stem cells. *Nature* **473**, 221-225, doi:10.1038/nature09915 (2011).
67 Lancaster, M. A. & Knoblich, J. A. Organogenesis in a dish: modeling development and disease using organoid technologies. *Science* **345**, 1247125, doi:10.1126/science.1247125 (2014).
68 Aaboud, M. *et al.* Measurements of the production cross section of a [Formula: see text] boson in association with jets in pp collisions at [Formula: see text] TeV with the ATLAS detector. *Eur Phys J C Part Fields* **77**, 361, doi:10.1140/epjc/s10052-017-4900-z (2017).
69 Hansen, D. V., Rubenstein, J. L. & Kriegstein, A. R. Deriving excitatory neurons of the neocortex from pluripotent stem cells. *Neuron* **70**, 645-660, doi:10.1016/j.neuron.2011.05.006 (2011).
70 Luskin, M. B., Pearlman, A. L. & Sanes, J. R. Cell lineage in the cerebral cortex of the mouse studied in vivo and in vitro with a recombinant retrovirus. *Neuron* **1**, 635-647 (1988).
71 Price, J. & Thurlow, L. Cell lineage in the rat cerebral cortex: a study using retroviral-mediated gene transfer. *Development* **104**, 473-482 (1988).
72 Walsh, C. & Cepko, C. L. Clonally related cortical cells show several migration patterns. *Science* **241**, 1342-1345 (1988).
73 Devine, M. J. *et al.* Parkinson's disease induced pluripotent stem cells with triplication of the alpha-synuclein locus. *Nat Commun* **2**, 440, doi:10.1038/ncomms1453 (2011).
74 Mansouri, M. *et al.* Highly efficient baculovirus-mediated multigene delivery in primary cells. *Nat Commun* **7**, 11529, doi:10.1038/ncomms11529 (2016).

CHAPTER 7
Appendix

Plasmid 90 (2017) 5–9

Contents lists available at ScienceDirect

Plasmid

journal homepage: www.elsevier.com/locate/yplas

Baculovirus-based genome editing in primary cells

Maysam Mansouri [a,b], Zahra Ehsaei [c], Verdon Taylor [c], Philipp Berger [a,*]

[a] Paul Scherrer Institute, Biomolecular Research, Molecular Cell Biology, CH-5232 Villigen, Switzerland
[b] ETH Zürich, Department of Biology, CH-8093 Zürich, Switzerland
[c] University of Basel, Department of Biomedicine, CH-4058 Basel, Switzerland

ARTICLE INFO

Article history:
Received 18 November 2016
Received in revised form 19 January 2017
Accepted 19 January 2017
Available online 22 January 2017

Keywords:
Genome editing
CRISPR/Cas9
Baculovirus
Primary cells

ABSTRACT

Genome editing in eukaryotes became easier in the last years with the development of nucleases that induce double strand breaks in DNA at user-defined sites. CRISPR/Cas9-based genome editing is currently one of the most powerful strategies. In the easiest case, a nuclease (e.g. Cas9) and a target defining guide RNA (gRNA) are transferred into a target cell. Non-homologous end joining (NHEJ) repair of the DNA break following Cas9 cleavage can lead to inactivation of the target gene. Specific repair or insertion of DNA with Homology Directed Repair (HDR) needs the simultaneous delivery of a repair template. Recombinant Lentivirus or Adenovirus genomes have enough capacity for a nuclease coding sequence and the gRNA but are usually too small to also carry large targeting constructs. We recently showed that a baculovirus-based multigene expression system (MultiPrime) can be used for genome editing in primary cells since it possesses the necessary capacity to carry the nuclease and gRNA expression constructs and the HDR targeting sequences. Here we present new Acceptor plasmids for MultiPrime that allow simplified cloning of baculoviruses for genome editing and we show their functionality in primary cells with limited life span and induced pluripotent stem cells (iPS).

1. Introduction

Technologies for genome editing are important developments in molecular biology to study gene function. Genome editing allows modifications including insertion, deletion or point mutations to be made in a targeted gene (Hsu et al., 2014). Therefore, this technology holds much promise for addressing many fundamental as well as applied questions (Barrangou and Doudna, 2016) including, the generation of human disease models (Dow, 2015; Platt et al., 2014), ex vivo gene therapy (Benjamin et al., 2015; Savić and Schwank, 2016), and recombinant protein production (Peng et al., 2015). Four main classes of site specific DNA nucleases including meganuclease, zinc finger nuclease (ZFN), TALEN, and Cas9 have been developed to facilitate targeted genome editing in eukaryotic cells (Gaj et al., 2013; Nelson and Gersbach, 2016). The CRISPR/Cas9 system is currently the most popular tool due to its simplicity, high efficiency and versatility. Clustered regularly interspaced short palindromic repeats (CRISPRs), together with CRISPR-associated (Cas) proteins, provide adaptive immunity against viruses and plasmids in many bacterial species and most archaea (Jiang and Marraffini, 2015). The two crucial components in CRISPR/Cas9-based genome editing are, Cas9 as a nuclease and a gRNA (Jinek et al., 2013; Ran et al., 2013). Together, Cas9 and the gRNA are able to induce a double strand break (DSB) at specific sites in DNA of host cells. Subsequently, the DSB DNA damage is repaired by the cellular DNA repair machinery through either NHEJ or HDR pathways (Hsu et al., 2014; Jiang and Marraffini, 2015; Jinek et al., 2013). NHEJ is an error-prone process which fills the gaps with small insertions or deletions (indel) which can be taken advantage of to disrupt the coding region of a gene of interest (knockout). In contrast to NHEJ, the HDR pathway is suitable for precise genome editing by inserting (knockin) fragments of exogenous DNA at a precise position in the genome (Sander and Joung, 2014). However, HDR is an inefficient process that requires an exogenous repair DNA template with sequence homology to regions that flank the DSB (van Erp et al., 2015; Hsu et al., 2014; Sander and Joung, 2014; Wang et al., 2015). The recombinant viruses currently used for gene transfer, do not have sufficient capacity to carry Cas9 coding sequences, the gRNAs, and the construct for HDR in one vector.

One possibility for delivery of the necessary elements for CRISPR/Cas9-HDR mediated knock-in is co-transfection or co-infection of Cas9, gRNA(s) and HDR constructs (Mali et al., 2013; Nelson and Gersbach, 2016). Co-transfection/co-infection is an easy and straightforward way to co-deliver multiple constructs, but it can cause low targeting efficiency in cells due to heterogeneity of vector delivery. In addition, some cells such as primary cells, are resistant to DNA transfection (Mansouri and Berger, 2014).

We recently showed that MultiPrime, a baculovirus-based system that was developed for multigene expression in primary cells, is suitable

* Corresponding author at: Paul Scherrer Institute, Molecular Cell Biology, CH-5232 Villigen, Switzerland.
E-mail address: Philipp.Berger@psi.ch (P. Berger).

http://dx.doi.org/10.1016/j.plasmid.2017.01.003
0147-619X/

for genome editing because it has sufficient capacity to carry the CRISPR machinery and HDR targeting vectors (Mansouri et al., 2016). Here, we report new plasmids that allow simplified cloning of these constructs. Available elements from existing, commercial vectors like Cas9 or gRNAs are compatible with our cloning sites. As proof of concept, we tagged the *HMGA1* locus with EGFP in HEK293 cells, an easy to transfect mammalian cell line. In addition, we transduced HUVEC and hiPS cells that are difficult to transfect. We demonstrate that our new plasmids support the rapid production of baculoviruses for genome editing in primary cells.

2. Materials and methods

2.1. Vector design

We designed an Acceptor plasmid for MultiPrime with unique cloning sites that allow the insertion of a nuclease, gRNAs, and a HDR construct, preferentially using fragments from CRISPR plasmids that are widely used in the genome engineering community (e.g. pX derivatives). The pSI-AG10 vector with additional I-*Sce*I and I-*Ceu*I sites was used as the backbone for the constructs. This part of the plasmid contains all of the necessary elements for MultiPrime baculovirus generation. pX derived vectors (including pX330, pX458, pX459, pX460, pX461, and pX462), contain a gRNA cassette composed of an U6 promoter-driven crRNA-tracrRNA that can be removed by *Afl*III/*Xba*I restriction digestion. We included two sites in our synthetic construct for insertion of two gRNAs (*Nco*I → *Xba*I and *Bsa*I → *Nhe*I). Cas9 is flanked by *Age*I/*Not*I sites in the pX vectors. We have included the same restriction enzyme sites in our vector. Therefore, Cas9, or each variation of it including Cas9n or dCas9, can easily be placed at this position. The CMV promoter that drives Cas9 is localized between the *Xba*I/*Age*I sites (in pSI-AG15/16 as well as in the pX plasmids). Thus, the CVM promoter sequences can be replaced by cell-type specific or regulatable elements for controlled expression of Cas9 if desired. In addition, an *Afl*III/*Not*I fragment from the pX vector containing the gRNA and Cas9 expression cassette can be cloned directly into the *Nco*I/*Not*I digested pSI-AG15/16 vector. The repair construct can be cloned into the *Asc*I/*Pac*I sites in the vector. Alternatively, the repair construct can be cloned in parallel in a MultiPrime Donor vector and then inserted by a Cre/LoxP-mediated recombination (Fig. 1a).

2.2. Cloning of acceptor plasmids

Plasmid pSI-AGZ10 was used as a template to construct the pSI-AG15 vector (Mansouri et al., 2016). The *Age*I site was removed from pSI-AG10 by digesting the plasmid with *Age*I followed by refilling with Phusion HF polymerase (NEB, USA) and relegation with T4 ligase (NEB, USA). The backbone of this plasmid was amplified using primers: AG15/16-forward-w-loxP 5′-CTA GGC TAG CGG TAT TAA CTA TAA CGG TCC TAA GGT AGC GAA ACC CAT CTA ATT GGA ACC AGA TAA G-3′ and primer AG15/16-back-w-lox P 5′-CTT ACC ATG GGG ACC ATT ACC CTG TTA TCC CTA GTT ATA GTT ACA GAT AAC TTC-3′ as forward and reverse

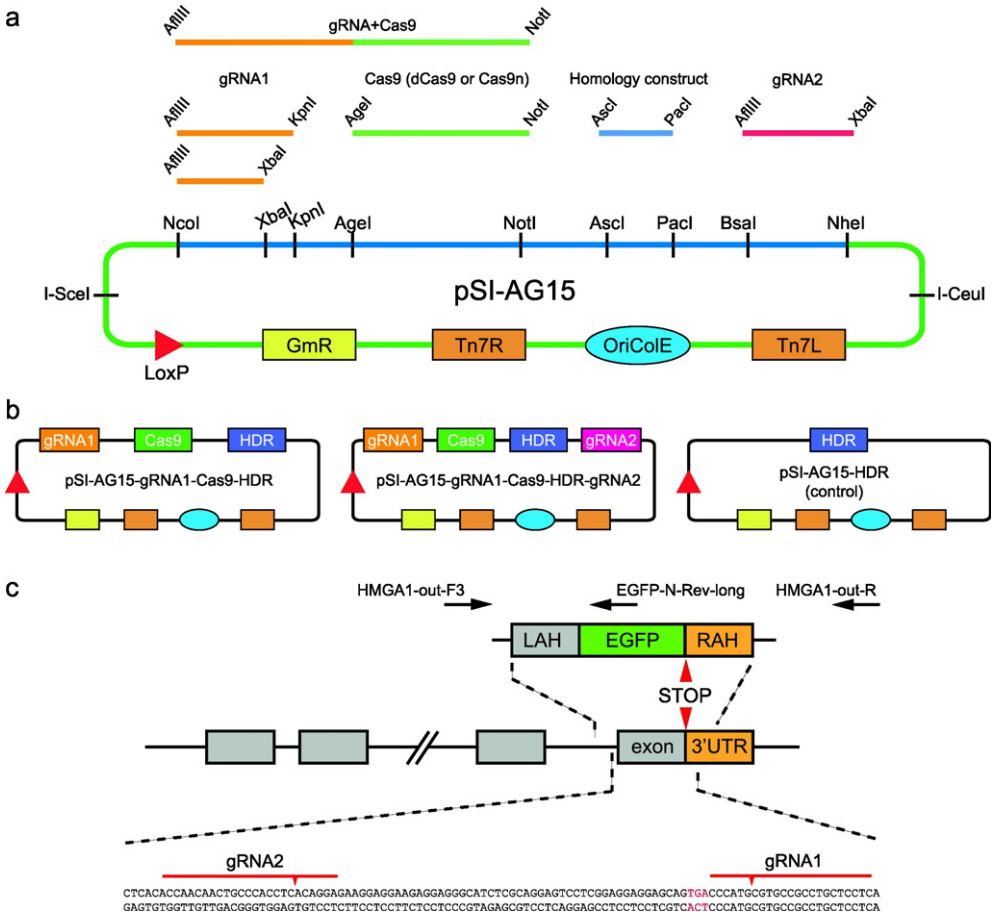


Fig. 1. Design of Acceptor plasmids. (a) Relevant elements of pSI-AG15 and cloning strategy. The backbone of pSI-AG15 (green) contains a LoxP site, a gentamycin resistance gene (GmR), ColE1 as origin of replication, and elements for transposition of the plasmid into the baculovirus genome (Tn7R/L). pSI-AG16 does not contain the LoxP site, but is otherwise identical. The blue part contains a multiple cloning site for cloning of CRISPR/Cas9 components from pX-derived vectors to pSI-AG15 and pSI-AG16. Possible strategies to insert first gRNA, Cas9, gRNA + Cas9, homology construct (HDR) and a second gRNA into pSI-AG15/16 are indicated with different colored lines. (b) Plasmids used in this study for validation. We inserted the previously established gRNAs and homology construct that allow insertion of EGFP in the human *HMGA1* locus. pSI-AG15-HDR is the negative control construct. (c) Graphical overview of the relevant part of the *HMGA1* locus. Primers for genotyping are given on top, binding and cleavage sites of gRNAs are given at the bottom.

primers, respectively. This resulted in a fragment with *Nco*I and *Nhe*I sites that was fused with a synthetic sequence (Genewiz, USA) containing unique restriction sites for the insert of all relevant elements for genome editing (Fig. 1a). In addition, we generated an Acceptor plasmid without LoxP site (pSI-AG16). The LoxP site was eliminated by site-directed mutagenesis using PCR with primer AG15/16-forward-wo-loxP 5′-CTA GGC TAG CGG TAT TAA CTA TAA CGG TCC TAA GGT AGC GAA ACC CAT CTA ATT GGA ACC AGA TAA G-3′ and primer AG15/16-back-wo-lox P5′-GGC GCC ATG GAC TTC GCA TTA CCC TGT TAT CCC TAT GCT GCC CAA GGT TGC CGG GTG AC-3′ as forward and reverse primers, respectively.

2.3. Cloning procedure

Plasmids for targeting the *HMGA1* locus were kindly provided by Dr. Stefan Jakobs (Ratz et al., 2015). Cas9 was excised together with gRNA1 by *Afl*III/*Not*I digestion from pX330-HMGA1-gRNA1 vector and cloned into the *Nco*I/*Not*I site of pSI-AG15 and pSI-AG16 generating the pSI-AG15/16-gRNA1-Cas9 vector. The HDR construct *Asc*I/*Pac*I fragment for insertion of EGFP was obtained from a pGEM-HDR vector and cloned *Asc*I/*Pac*I into pSI-AG15/16-gRNA1-Cas9 to generate pSI-AG15/16-gRNA1-Cas9-HDR vector. The second HMGA1-gRNA (gRNA2) was isolated from pX330-HMGA1-gRNA2 as a *Afl*III/*Xho*I fragment and cloned into the *Bsa*I/*Nhe*I digested pSI-AG15/16-gRNA1-Cas9-HDR vector to generate pSI-AG15/16-gRNA1-Cas9-HDR-gRNA2. Control plasmid pSI-AG15/16-HDR was constructed by cloning of HDR through *Asc*I/*Pac*I to pSI-AG15/16 (Fig. 1b).

2.4. Cell culture

HEK293 cells were cultured in DMEM (Amimed) containing 10% FBS (Life Technologies) and 100 units/ml Penicillin and 100 μg/ml Streptomycin (Life Technologies) in a humified incubator with 5% CO_2 at 37 °C. HUVECs (Life Technologies) were maintained in M-200 medium (Life Technologies). Human iPS cells (NAS2, kind gift from Tilo Kunath, University of Edinburgh) were cultured in mTESR1 medium (Stem Cell Technologies). Baculovirus generation in sf21 insect cells and transduction of mammalian cells was described previously (Mansouri et al., 2016). Transfection was performed with Lipofectamin3000 (Life Technologies) according to manufacturer's recommendations. Cells were analyzed 5 or 6 days after transfection or transduction. The quantifications described in the result part are from single transduction or transfection experiment and represent mean ± SEM from ten independent areas of three coverslips.

2.5. Analysis of cells

HEK293 and HUVECs were plated on glass slides coated with poly-L-lysine (Sigma P4707) or 0.1% gelatin (Sigma G1393). iPS cells were plated on hESC qualified Matrigel (BD Bioscience). Cells were fixed with 4% formaldehyde/PBS. iPS cells were immunostained with an Oct4 antibody (sc-5279, 1:500, Santa Cruz Technology) and nuclei were counter-stained with DAPI (1 μg/ml). Samples were mounted on glass slides with Gelvatol. Images were taken on a ZEISS Apotome.2 fluorescent microscope. Genomic DNA of cells was isolated by QiaAmp DNA mini Kit (Qiagen) and used as template in PCR genotyping reaction. PCR was performed in 1× (98 °C for 5 min), 30× (98 °C for 30 s, 72 °C for 30 s and 72 °C for 180 s) and 1× (72 °C for 10 min). Primer HMGA1-out-F3 5′-TTG TCC TTC CTA TGA GCC TCT GCA G-3′ was used as forward primer and either HMGA1-out-R 5′-CAA TGA CGG ATG TCG AAG AAT GGA ACA TTG AAC-3′ or EGFP-N-Rev-long 5′-GTC AGC TTG CCG TAG GTG GCA TC-3′ were used as reverse primers for amplification of wildtype or mutated alleles, respectively.

3. Results

3.1. Designing of pSI-AG15 and pSI-AG16

We showed previously that baculoviruses is an elegant tool for genome editing in mammalian cells and that it has sufficient capacity to harbor all relevant elements including a large homology construct (Mansouri et al., 2016). MultiPrime uses cre/LoxP technology to assemble multiple cassettes. This technology is extremely flexible but not essential for standardized applications. We therefore designed new acceptors for MultiPrime that allow straightforward cloning making use of commonly available CRISPR vectors (Fig. 1). To validate our vectors, we used constructs that were previously used to modify the human *HMGA1* locus (Ratz et al., 2015). HMGA1 is a nuclear DNA-binding non-histone high mobility group protein. Targeting with these constructs leads to the expression of a HMGA1-EGFP fusion protein. Genome edited cells can easily be recognized by a green nucleus. The use of an established system allows us to compare our strategy with previously used strategies.

3.2. Targeting of HMGA1 in HEK293 cells

We first used HEK293 cells to compare the efficiencies of transfection and baculovirus transduction procedures. HEK293 cells are easy to transfect and we therefore do not expect a major difference in efficiency between the two methods. Image analysis by fluorescence microscopy revealed cells with green nuclei in transfected cells (data not shown) as well as in transduced cells (Fig. 2a). We observed approximately 4% edited cells following transfection which was similar to the efficiency observed by Ratz et al. for co-transfection of the targeting plasmids (Ratz et al., 2015). Transduction was more efficient and led to approximately 6% of targeted cells (Fig. 2b). Genotyping PCR and sequencing of the resulting PCR products confirmed correct integration of the HDR template. We only obtain a product for the mutated allele in cells that were transduced with a baculovirus that contains Cas9, gRNA, and repair construct (Fig. 2c). Sequence analysis of the mutant and the wildtype product show the correctly integrated and targeted sequence in the cell pool (Fig. 2d).

3.3. Targeting of the HMGA1 gene in iPS cells and HUVECs

A large number of cells and particularly primary cells are difficult to transfect with current reagents. We used iPS cells that are derived from primary human fibroblasts and primary HUVECs. iPS cells can go through clonal expansion whereas HUVEC have a limited life span and stop dividing after 7–10 passages. Again we compared transfection vs transduction for introducing the targeting construct in these two cell types. Only very few cells with green nuclei were observed in HUVEC and iPS cells following transfection with plasmids containing all three Cas9, gRNA and HDR targeting vector components (pSI-AG15_gRNA1-Cas9-HDR or pSI-AG15_gRNA1-Cas9-HDR-gRNA2, data not shown). In contrast, transduction with baculoviruses containing the constructs led to more GFP-positive cells (Fig. 3a, d). Quantification revealed that approximately 4.5% (HUVEC) and 3% (iPS) of the cells had been genome modified (11 and 4 times better than for transfection, respectively) (Fig. 3b, e). Genotyping with PCR only resulted in a band for the mutated allele when all three components were transduced (Fig. 3c, f).

4. Discussion

A major challenge in the CRISPR/Cas9 system is safe and effective delivery of all components at the same time to host cells. This can be done by transfection or infection. We now developed Acceptor plasmids for the generation of baculovirus for genome editing. Baculovirus is superior to other viruses since it has enough capacity to harbor all elements that are necessary for genome editing with HDR. In addition,

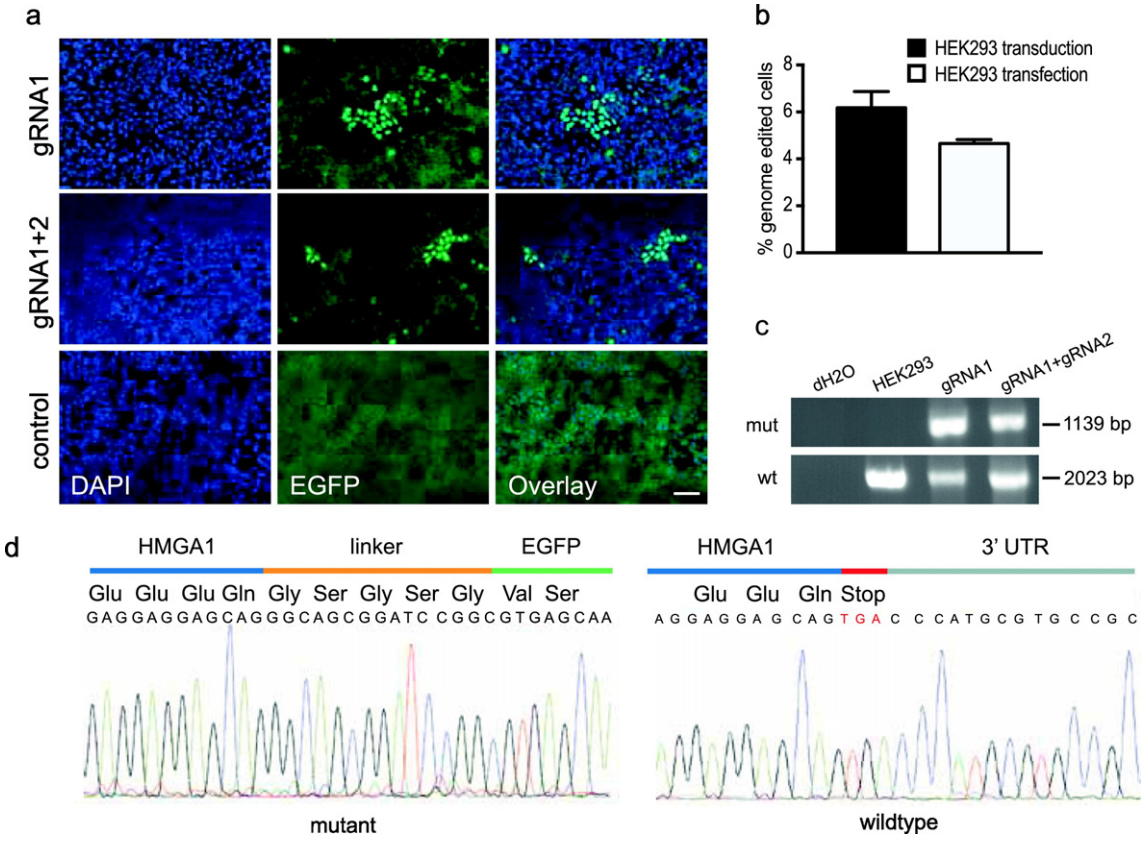


Fig. 2. Genome editing of the *HMGA1* locus in HEK293 cells. (a) HEK293 cells were transduced with baculovirus harboring the constructs described in Fig. 1b. Constructs with gRNAs and Cas9 (top: pSI-AG15-gRNA1-Cas9-HDR; middle: pSI-AG15-gRNA1-Cas9-HDR-gRNA2) lead to cells that express a HGMA1-EGFP fusion protein (green nuclei). Transduction with the control construct did not lead to green nuclei (bottom panel). Scale bar, 50 μm. (b) Efficiency of genome editing in HEK293 cells when constructs were delivered by baculovirus (transduction) or by transfection with the equivalent plasmid. Error bars indicate SEM (n = 10). (c) PCR genotyping using genomic DNA extracted from HEK293 cells. The mutant allele results in a fragment with 1139 bp (top), whereas the wild-type allele yielded a fragment of 2023 bp. (d) Sequencing analysis of the PCR products for HMGA1 knock-in (mutant) and wildtype locus.

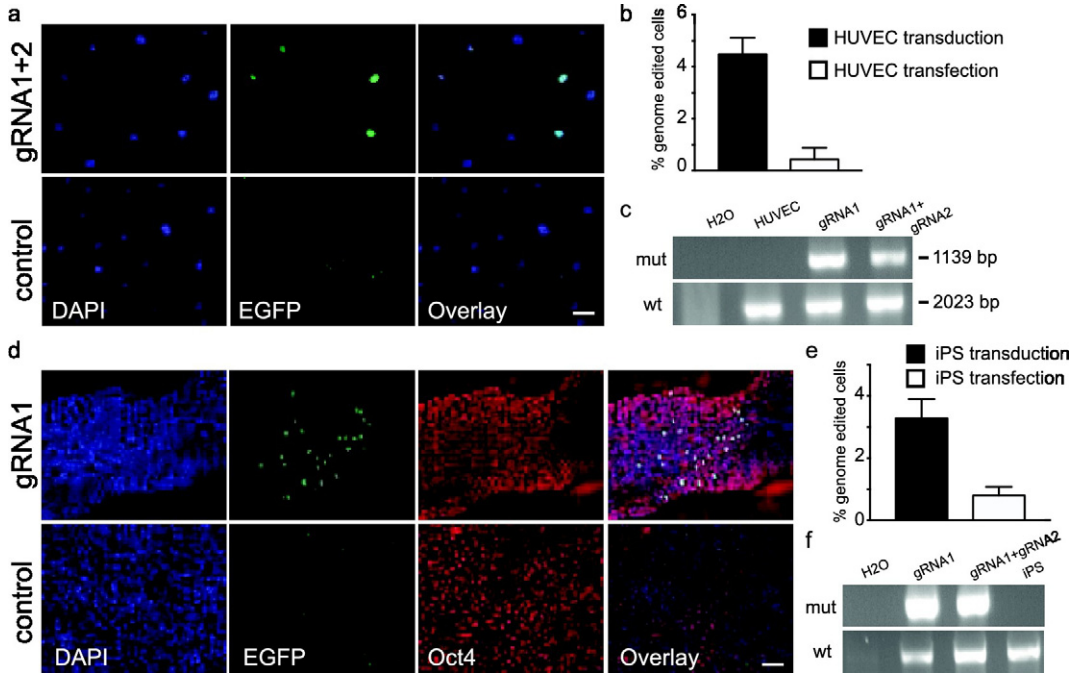


Fig. 3. Genome editing in primary cells. (a–c) EGFP knock-in in the *HMGA1* locus of HUVEC. Transduction with pSI-AG15-gRNA1-Cas9-HDR-gRNA2 or pSI-AG15-HDR (control). Green nuclei indicate successful genome modification. (b) Quantification of efficiency of genome engineering of these samples. Error bars indicate SEM (n = 10). (c) PCR genotyping of these samples. (d–f) EGFP knock-in in the *HMGA1* locus of iPS cells. Transduction with pSI-AG15-gRNA1-Cas9-HDR-gRNA2 or pSI-AG15-HDR (control). Green nuclei indicate successful genome modification. Oct-4 was used as a marker of pluripotent stem cells. (e) Quantification of efficiency of genome engineering. Error bars indicate SEM (n = 10). (f) PCR genotyping of these samples. Scale bar: 50 μm.

baculovirus is a gentle tool to transfer DNA to cells. We showed before that it has little toxicity and the functionality of transduced cells is maintained (Mansouri et al., 2016). Adenovirus has 8–10 kb foreign DNA capacity and is therefore also a potential candidate for homology-based genome editing but it was to our knowledge not yet used for this purpose (Chen and Gonçalves, 2016; Kennedy and Parks, 2009; Luo et al., 2007).

Our results show that a simple transient transfection in easy-to-transfect cells (e.g., HEK293) results in 4% genome edited cells (Fig. 2). Transduction with a baculovirus led to better targeting but the increased efficiency may not be worth the investment to establish baculovirus-mediated targeting in these cell types. The situation changes in primary cells that are difficult to transfect. We observed very few genome targeted cells (<1%) when we transfected HUVECs and iPS cells with our plasmids. This low efficiency is in many cases not sufficient to generate enough cells for analysis if targeted cells cannot be expanded extensively. In the case of iPS cells, clonal expansion to obtain 100% positive cell populations is possible but would require the parallel insertion of a selection marker which might be undesirable. HUVECs on the other hand undergo a limited number of doublings in vitro and divide only 5–7 times in culture. Clonal expansion is therefore not possible and the maximum possible modification rate must be reached with the targeted cells.

We never reached >6% of modified cells in all cell types tested even though we routinely transduce 60–80% of all cells. This indicates that CRISPR mediated HDR is an inefficient process and that transfer of CRISPR elements does not guarantee successful modification. Future developments have to identify elements that enhance efficiency. MultiPrime has the advantage that it can also deliver these elements, e.g. proteins that enhance recombination efficiency or reduce off-target effects, as soon as they are identified. In addition, transient suppression of the NHEJ pathway using inhibitors of Ku70, Ku80 and DNA ligase IV might also improve HDR efficacy after transduction with baculovirus (Chu et al., 2015; Maruyama et al., 2015),

In summary, we constructed all-in-one plasmids encoding Cas9, gRNAs and HDR on a single plasmid. We increased efficiency of genome editing in primary cells and stem cells using baculovirus. We anticipate many applications in genome engineering (such as activation, repression, deletion or insertion of single or many genes) will benefit from our system.

Acknowledgement

This work was supported by the Swiss National Science Foundation (31003A_146975 to PB). We would like to thank Dr. Stefan Jakobs for plasmids. This work was also supported by the Swiss National Science Foundation through SystemX.ch under the RTD project NeuroStemX to VT.

References

Barrangou, R., Doudna, J.A., 2016. Applications of CRISPR technologies in research and beyond. Nat. Biotechnol. 34, 933–941.
Benjamin, D., Cox, T., Platt, R.J., Zhang, F., 2015. Therapeutic genome editing: prospects and challenges. Nat. Med. 21.
Chen, X., Gonçalves, M.A.F.V., 2016. Engineered viruses as genome editing devices. Mol. Ther. 24, 447–457.
Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., Kühn, R., 2015. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat. Biotechnol. 33, 543–548.
Dow, L.E., 2015. Modeling disease in vivo with CRISPR/Cas9. Trends Mol. Med. 21, 609–621.
van Erp, P.B., Bloomer, G., Wilkinson, R., Wiedenheft, B., 2015. The History and Market Impact of CRISPR RNA-Guided Nucleases.
Gaj, T., Gersbach, C.a., Barbas, C.F., 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 31, 397–405.
Hsu, P.D., Lander, E.S., Zhang, F., Barrangou, R., van der Oost, J., Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., et al., 2014. Development and applications of CRISPR-Cas9 for genome engineering. Cell 157, 1262–1278.
Jiang, W., Marraffini, L.A., 2015. CRISPR-Cas: new tools for genetic manipulations from bacterial immunity systems. Annu. Rev. Microbiol. 69, 209–228.
Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., Doudna, J., 2013. RNA-programmed genome editing in human cells. eLife 2, e00471.
Kennedy, M.A., Parks, R.J., 2009. Adenovirus virion stability and the viral genome: size matters. Mol. Ther. 17, 1664–1666.
Luo, J., Deng, Z.-L., Luo, X., Tang, N., Song, W.-X., Chen, J., Sharff, K.A., Luu, H.H., Haydon, R.C., Kinzler, K.W., et al., 2007. A protocol for rapid generation of recombinant adenoviruses using the AdEasy system. Nat. Protoc. 2, 1236–1247.
Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., 2013. RNA-guided human genome engineering via Cas9. Science 339, 823–826.
Mansouri, M., Berger, P., 2014. Strategies for multigene expression in eukaryotic cells. Plasmid 75, 12–17.
Mansouri, M., Bellon-Echeverria, I., Rizk, A., Ehsaei, Z., Cianciolo Cosentino, C., Silva, C.S., Xie, Y., Boyce, F.M., Davis, M.W., Neuhauss, S.C.F., et al., 2016. Highly efficient baculovirus-mediated multigene delivery in primary cells. Nat. Commun. 7, 11529.
Maruyama, T., Dougan, S.K., Truttmann, M.C., Bilate, A.M., Ingram, J.R., Ploegh, H.L., 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat. Biotechnol. 33, 538–542.
Nelson, C.E., Gersbach, C.A., 2016. Engineering delivery vehicles for genome editing. Annu. Rev. Chem. Biomol. Eng. 7, 637–662.
Peng, J., Wang, Y., Jiang, J., Zhou, X., Song, L., Wang, L., Ding, C., Qin, J., Liu, L., Wang, W., et al., 2015. Production of human albumin in pigs through CRISPR/Cas9-mediated knockin of human cDNA into swine albumin locus in the zygotes. Sci. Rep. 5, 16705.
Platt, R.J., Chen, S., Zhou, Y., Yim, M.J., Swiech, L., Kempton, H.R., Dahlman, J.E., Parnas, O., Eisenhaure, T.M., Jovanovic, M., et al., 2014. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell 159, 440–455.
Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8, 2281–2308.
Ratz, M., Testa, I., Hell, S.W., Jakobs, S., 2015. CRISPR/Cas9-mediated endogenous protein tagging for RESOLFT super-resolution microscopy of living human cells. Sci. Rep. 5, 9592.
Sander, J.D., Joung, J.K., 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 32, 347–355.
Savić, N., Schwank, G., 2016. Advances in therapeutic CRISPR/Cas9 genome editing. Transl. Res. 168, 15–21.
Wang, B., Li, K., Wang, A., Reiser, M., Saunders, T., Lockey, R.F., Wang, J.W., 2015. Highly efficient CRISPR/HDR-mediated knock-in for mouse embryonic stem cells and zygotes. Biotechniques 59 (4), 201–202, 204, 206–208.

nature COMMUNICATIONS

ARTICLE

Received 17 Feb 2016 | Accepted 4 Apr 2016 | Published 4 May 2016

DOI: 10.1038/ncomms11529 OPEN

Highly efficient baculovirus-mediated multigene delivery in primary cells

Maysam Mansouri[1], Itxaso Bellon-Echeverria[2], Aurélien Rizk[1], Zahra Ehsaei[3], Chiara Cianciolo Cosentino[4], Catarina S. Silva[2], Ye Xie[1], Frederick M. Boyce[5], M. Wayne Davis[6], Stephan C.F. Neuhauss[4], Verdon Taylor[3], Kurt Ballmer-Hofer[1], Imre Berger[2,7] & Philipp Berger[1]

Multigene delivery and subsequent cellular expression is emerging as a key technology required in diverse research fields including, synthetic and structural biology, cellular reprogramming and functional pharmaceutical screening. Current viral delivery systems such as retro- and adenoviruses suffer from limited DNA cargo capacity, thus impeding unrestricted multigene expression. We developed MultiPrime, a modular, non-cytotoxic, non-integrating, baculovirus-based vector system expediting highly efficient transient multigene expression from a variety of promoters. MultiPrime viruses efficiently transduce a wide range of cell types, including non-dividing primary neurons and induced-pluripotent stem cells (iPS). We show that MultiPrime can be used for reprogramming, and for genome editing and engineering by CRISPR/Cas9. Moreover, we implemented dual-host-specific cassettes enabling multiprotein expression in insect and mammalian cells using a single reagent. Our experiments establish MultiPrime as a powerful and highly efficient tool, to deliver multiple genes for a wide range of applications in primary and established mammalian cells.

[1] Biomolecular Research, Molecular Cell Biology, Paul Scherrer Institute, CH-5232 Villigen, Switzerland. [2] European Molecular Biology Laboratory (EMBL), Grenoble Outstation, B.P. 181, 38042 Grenoble Cedex 9, France. [3] Department of Biomedicine, University of Basel, CH-4058 Basel, Switzerland. [4] Institute of Molecular Life Sciences, University of Zürich, CH-8057 Zürich, Switzerland. [5] Department of Neurology, Massachusetts General Hospital, Cambridge, Massachusetts 02139, USA. [6] Department of Biology and Howard Hughes Medical Institute, University of Utah, Salt Lake City, Utah 84112-0840, USA. [7] School of Biochemistry, University of Bristol, Bristol BS8 1TD, UK. Correspondence and requests for materials should be addressed to I.B. (email: imre.berger@bristol.ac.uk) or to P.B. (email: Philipp.Berger@psi.ch).

Multigene delivery into cultured cells or tissues is emerging as an indispensable tool for many applications in biological research and development. Examples include simultaneous labelling of living cells with various fluorescently-tagged sensors for monitoring changes in cellular architecture or metabolism, lineage tracing during morphogenesis to follow regenerative tissue processes, visualization of multi-component molecular pathways for high-content screening in pharmacological applications or the construction of recombinant adeno-associated viruses for gene therapy[1–5]. Multigene delivery systems also allow reprogramming of somatic cells to stem cells[6] or to specifically differentiated cell lines[7]. The construction of complex multigene circuits in mammalian cells is a core concept in synthetic biology requiring the flexible generation of modular multigene expression systems[8,9]. Moreover, structural and biophysical characterization of multiprotein complexes relies on co-expression of an ensemble of genes that may include ancillary factors, such as chaperones or protein modifying enzymes[10]. All applications share in common that they require versatile tool-kits to flexibly engineer and to simultaneously, efficiently and reproducibly deliver multiple genes into target host cells.

Several strategies for multigene expression in mammalian cells exist, each with its own merits[11]. All of these applications require specific boundary conditions. For instance, it is essential that all transfected cells in a population express all heterologous genes at the same defined level, on an equal time frame. Other applications require that the proteins of interest retain native N- or C termini. Furthermore, long-term stable expression versus transient expression is a crucial parameter to be considered. Ideally, an efficient multigene-delivery system would provide the means to afford many or all of these requirements.

We have developed systems for the delivery of multigene constructs in prokaryotic and eukaryotic hosts[12–14]. A central feature of these technologies is the assembly of multiple gene expression cassettes by recombineering[15], from custom designed plasmids encoding specific genes, into a single multicomponent DNA construct for gene delivery. This approach was shown to overcome the limitations hampering classical co-transfection or co-infection techniques, which for statistical reasons, are inherently unbalanced[16,17]. More recently, we introduced MultiLabel[14] and demonstrated that homogenous mammalian cell populations could be achieved by transient introduction of single recombineering-based multigene expression plasmids by classical transfection methods. This method performs well with cell lines that are readily transfected, such as HEK293 or HeLa cells. However, a large number of cell lines and particularly primary cells are markedly recalcitrant to plasmid transfection, thus requiring a different approach. Primary cells are a central focus of contemporary biological research efforts, and efficient multigene delivery in primary cells is thus highly desirable.

Infection by viral vectors emerged as the dominant method of choice to deliver genes into primary cells[18]. An ideal viral vector for multigene delivery should have virtually unlimited foreign DNA cargo capacity allowing for integration of a multitude of independent expression cassettes, functionalities and regulatory elements. Moreover, such an optimal viral vector should exhibit low cytotoxicity in mammalian cells and should enable transduction of dividing and non-dividing mammalian cells alike. Currently used lenti- and other retroviruses, as well as adeno- and adeno-associated viruses have a limitation on DNA cargo size due to spatial constraints imposed by the tight geometry of their capsids.

Baculoviral vectors, in contrast, can accommodate very large DNA cargo insertions[19]. The *Autographa californica* multiple nuclear polyhedrovirus (AcMNPV), is a baculovirus with a large (134 kb) double-stranded circular DNA genome that normally infects specific moth larvae[19]. Transgene capacity of AcMNPV is very large, extending probably beyond 100 kbp. Replication of AcMNPV is highly insect-cell specific; however, AcMNPV is capable of efficiently transducing not only insect but also mammalian cells. Transduction is usually transient without DNA integration into the target cell genome and such viruses are replication deficient[20,21]. In baculoviruses used for mammalian cell transduction (BacMam)[22,23] heterologous genes are placed under the control of mammalian promoters and inserted into the baculoviral genome, and viral stocks are produced in insect cells. Once the baculovirus enters mammalian cells, these genes are actively transcribed within 9 h and the cells produce the heterologous gene product. In the last decade, baculovirus has emerged as a useful and safe technology to deliver heterologous genetic material to mammalian cell types both *in vitro* and *in vivo*[24].

Here we introduce MultiPrime, a novel tool-kit specifically designed for efficient delivery of large multigene constructs into primary and established mammalian cells. MultiPrime enables simultaneous expression of multiple, independent cassettes in mammalian cells. This system combines the ease and flexibility of our recombination-mediated multigene DNA assembly technology with the superior performance of baculovirus as a viral vector for transducing mammalian cells. We transduced with MultiPrime a large variety of cell types including difficult to transfect stem cells and primary neurons. Moreover, we utilized MultiPrime for reprogramming mouse embryo fibroblasts (MEFs) into neurons. Further, we show that our system is not limited to mammalian transduction but can also be used to transduce zebrafish embryos. We applied our system to create synthetic multicomponent intracellular biosensor systems, such as Rab GTPases regulating vesicular membrane traffic in cells, phosphoinositide binding signalling proteins or fluorescently-labelled cytoskeletal markers. These biosensors were simultaneously delivered and expressed in mammalian cells allowing to quantitatively monitor a large variety of intracellular parameters. A wide range of promoters are available in MultiPrime, to regulate and fine-tune individual heterologous target gene expression.

With the objective to provide a means to concomitantly exploit with one single expression system the advantages of high-level protein production in baculovirus-infected insect cells and protein expression in a 'native' mammalian environment, we incorporated dual-host-specific promoters in MultiPrime that are functional in both insect and mammalian cells.

Our MultiPrime system is not limited to the assembly of RNA polymerase II-based expression cassettes. It can likewise harbour U6-driven expression cassettes and homology constructs that are required for CRISPR/Cas9-mediated genome engineering. We demonstrate the aptitude of our system by applying MultiPrime-CRISPR/Cas9 to express a HMGA1-EGFP fusion protein in primary cells from the native genomic HMGA locus.

Results

MultiPrime system design. We developed MultiPrime specifically to overcome the limitations of transfecting mammalian cells for multigene transfer applications. We redesigned our previous pFL plasmid[25], which contains the Tn7R and Tn7L DNA elements required for Tn7-transposase-mediated integration into a baculovirus genome containing a Tn7-attachment site[26]. Expression cassettes with promoters active in mammalian cells, or, alternatively, with activity in both mammalian and insect cells, were inserted into this MultiPrime acceptor plasmid. This acceptor plasmid is poised to receive further multigene expression cassettes by incorporating donor plasmids (Supplementary

Figs 1–3). All donors from our previous plasmid-based MultiLabel system can be used for incorporation into this new acceptor to generate acceptor-donor fusions by recombineering. Moreover, expression cassettes can be freely exchanged between donors and acceptors due to the modular design[14].

Acceptors or acceptor–donor fusions containing multigene expression cassettes are integrated into baculoviral genomes by means of Tn7 transposition[25]. We utilized two different baculoviral genomes in our experiments. In addition to our EMBacY baculoviral genome[25] we generated in this study a new baculoviral genome, MultiBacMam, which expresses a vesicular stomatitis virus glycoprotein (VSV-G) and the fluorescent protein mCherry in insect cells during virus production. Both genomes were generated from the original MultiBac virus by integrating genes encoding EYFP (EMBacY) or mCherry and VSV-G (MultiBacMam) by Cre recombination into the LoxP site present on the MultiBac virus[12,25]. The expression of mCherry results in a characteristic purple colour of MultiBacMam infected cell cultures (Supplementary Fig. 1), thereby simplifying the tracking of virus amplification by eye. Moreover, the MultiBacMam virus gives rise to baculovirions displaying VSV-G on their surface. The presence of VSV-G in the baculoviral envelope has been shown to increase the efficacy of mammalian cell transduction[27]. All composite baculoviruses prepared in this study were produced in Sf21 cells. Virus was obtained with comparable efficiency and resulting in similar viral titers to what we had observed in previous multigene expressions in insect cells with recombinant MultiBac or EMBacY viruses[25]. Our multigene viruses are characterized by multiple use of regulatory elements such as the CMV promoter, which in theory could introduce genomic instability during repeated rounds of viral amplicfication[14,25]. We had developed previously efficient protocols to amplify baculoviral constructs containing multiple copies of late viral promoters (polh and p10) by stringently applying a low multiplicity of infection (MOI) regimen and few, ideally not more than two amplification rounds thus preventing accumulation of non-productive viruses containing genomic deletions[28]. Strictly adhering to this strategy for amplifying baculoviruses containing multiple copies of CMV promoter-driven expression cassettes again resulted in stable virus producing all proteins of choice in each transduced mammalian cell in homogenous cultures, while failure to adhere to the protocol resulted in heterogeneous cell populations where expression of individual heterologous genes had been lost, presumably due to accumulation of deletion virus species (Supplementary Fig. 4). Occasionally, we observed cell fusion in tissue culture plates of insect cells infected with MultiBacMam viruses, which, however, was found not to be detrimental to transduction experiments with the viral supernatant. In the following, we use the EMBacY baculoviral genome to prepare composite MultiPrime baculoviruses, unless indicated otherwise.

Transduction and viability. As proof-of-concept, we generated a composite MultiPrime baculovirus expressing five fluorescently-tagged proteins (EBFP2-nucleus, mTFP1-FYVE (PI-3-P-binding, early endosomes), EYFP-tubulin, Mito-dsRED (mitochondria) and Plum-PLCδ-PH (PI-4,5-P_2-binding, plasma membrane) localizing to different subcellular compartments. We used this baculovirus to test a variety of different cell types for their propensity to be transduced efficiently. We initially transduced well-established cell lines including HEK293, PAE, COS7, HeLa, SK-MEL-28, CCL39 and Swiss 3T3. All cell lines that were tested proved to be highly transduceable by MultiPrime and efficiently expressed all heterologous proteins (Supplementary Fig. 5 and data not shown). Typically, transduction efficiencies ranged between 20 and 100%. Transduction frequencies up to 100% were observed even in cell lines, such as PAE that are traditionally considered to be difficult to transfect.

We then asked whether we can use MultiPrime to transfect primary cells. Primary cells are an intense focus of contemporary research efforts for many reasons, and it is well documented that they are typically extremely difficult to transfect. For this experiment, we utilized human umbilical vein endothelial cells (HUVEC), rat embryo fibroblasts (REF), rat cortical neurons and human-induced pluripotent stem cells (iPS). With our MultiPrime virus, all of these primary cell types could be transduced efficiently and again expressed all heterologous genes of interest, compellingly underscoring the utility of our approach (Fig. 1 and Supplementary Fig. 5).

To analyze viability and functionality, we proceeded to express multiple intracellular sensors simultaneously from a MultiPrime baculovirus. COS7 cells expressing epidermal growth factor receptor (EGFR) endogenously were used to monitor trafficking of EGFR. Genes encoding fluorescently-tagged markers for early (RAB5A), recycling (RAB11A) and late (RAB7A) endosomes were expressed from a single MultiPrime virus and the cells were stimulated 40 h post-transduction with Cy5-labelled epidermal growth factor (EGF). As expected, EGF was found in early (RAB5A+) and late (RAB7A+) endosomes, but not in recycling (RAB11A+) endosomes after 30 min (Fig. 2a). A time-resolved quantitative analysis confirmed that EGF was transferred to late endosomes but not to RAB11-positive, recycling vesicles (Supplementary Movie 1). Next, we used PAE cells stably expressing VEGFR2 and neuropilin-1 (NRP1). A composite MultiPrime baculovirus was used to express fluorescently-tagged Rab4, Rab5 and Rab7 in these cells. Cells were then stimulated with Nt647-labelled VEGF-A165a for 3 h. VEGF was found in Rab5 and Rab7 vesicles clearly indicating that it was properly internalized, in accordance with previous reports[3]. Using the same cell line transduced with the RAB5A-RAB11A-RAB7A virus in a time-resolved study, we could show that VEGFR2 behaves differently from EGFR. VEGFR2 is, in contrast to EGFR, partly recycled through the RAB11 compartment (Supplementary Movie 2).

Next, we tested the functionality of baculovirus-transduced HUVEC in migration and angiogenesis assays. HUVEC transduced with a MultiPrime baculovirus encoding mTFP1-actin, EYFP-tubulin and Mito-DsRed were plated on matrigel and incubated for 14 h. HUVEC are known to establish a characteristic vascular network within this time frame. Transduced cells efficiently integrated in this network, clearly demonstrating that MultiPrime-transduced HUVEC show identical behaviour compared to untransduced cells (Fig. 2c). Moreover, in a migration assay, we could show that transduced cells migrate with similar efficiency as untransduced cells (Fig. 2d).

We quantitatively compared baculovirus-mediated transduction with the classical plasmid transfection approach. In addition to EMBacY, we used our MultiBacMam baculovirus displaying VSV-G on its surface in this experiment. As expected, we observed increased transduction rates with MultiBacMam compared to EMBacY, albeit the gain in efficacy in our hands turned out to be modest in many cases. Clearly, transduction with both EMBacY and MultiBacMam outperformed the classical transfection approach (Fig. 3a). We routinely obtained transduction efficiencies higher than 50% when using the MultiPrime baculovirus compared to transfection efficiencies well below 20% with the corresponding plasmid (13,305 bp) that had been used to generate the composite baculovirus.

Baculovirus displaying VSV-G on its surface was superior to virus lacking VSV-G at all tested MOI. Saturation was usually obtained at a MOI of 500 (Fig. 3b). The relative expression levels

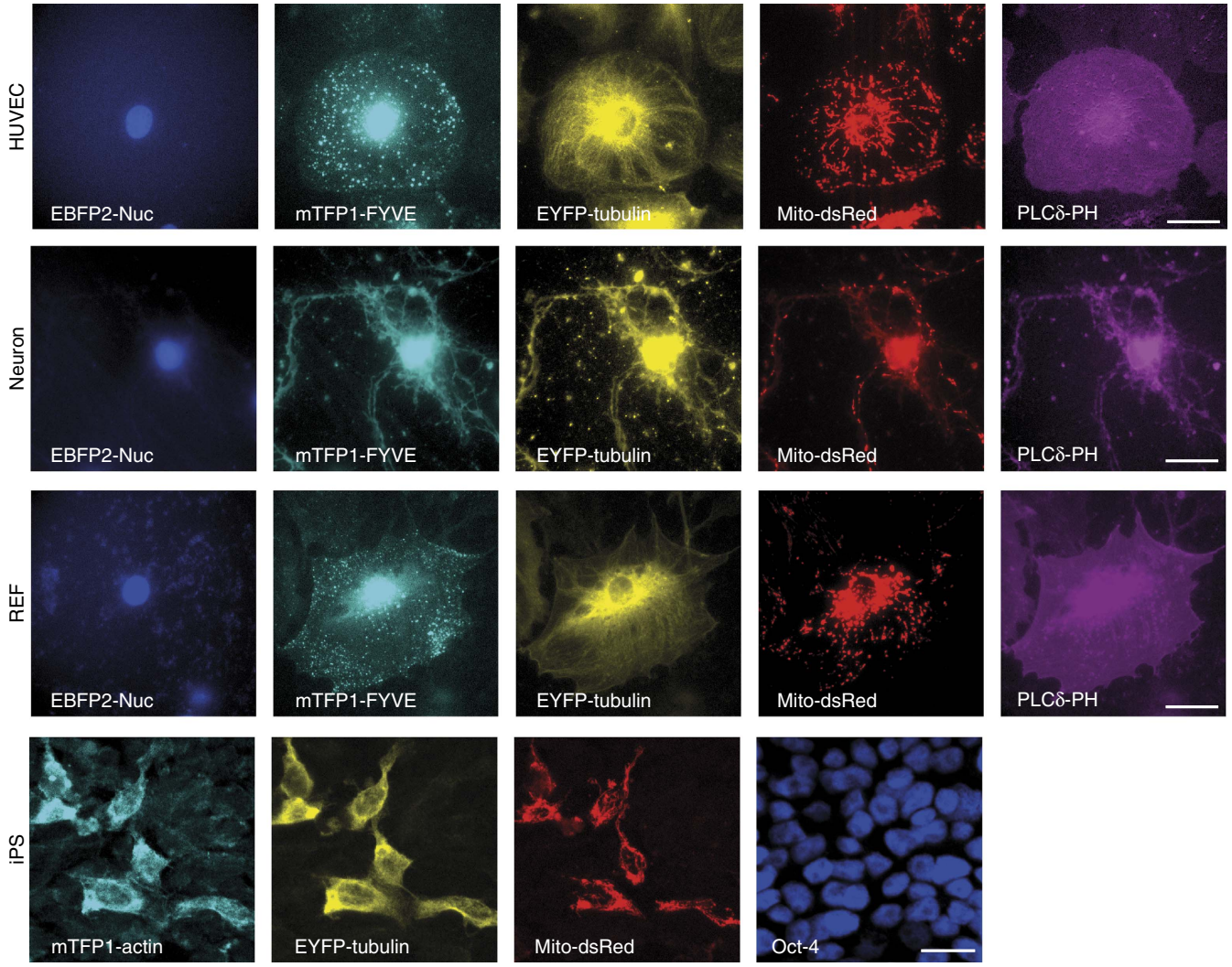


Figure 1 | Multigene expression in primary cells by MultiPrime. HUVEC, REF cells and rat cortical neurons were infected with a MultiPrime baculovirus encoding EBFP2-Nuc (labelling the nucleus), mTFP1-FYVE (PI-3-P containing endosomes), tubulin-EYFP (cytoskeleton), Mito-dsRed (mitochondria) and PLCδ-PH (PI-4,5-P2; plasma membrane). iPS cells were transduced with a virus encoding mTFP1-actin, EYFP-tubulin and Mito-dsRed. Oct4 was used as a marker for pluripotent stem cells. All infected cells express all heterologous proteins. Scale bar, 20 μm.

between cells appear to be similar (Fig. 3c). This is in contrast to transfected cells that typically show a wide variety of expression levels (data not shown). Since we use relatively high MOI, the toxicity of the virus could conceivably be an issue. We tested baculovirus toxicity at MOI 500 compared to plasmid transfection with Fugene HD, which is considered to be a mild transfection reagent. Both EMBacY- and MultiBacMam-derived viruses exhibit negligible toxicity similar to plasmid-based transfection (Fig. 3d).

Baculovirus transduction of mammalian cells is transient in nature as the foreign DNA does not integrate into the host genome. We therefore tested the persistence of recombinant expression following transduction with a MultiPrime baculovirus by immunofluorescence and western blotting. In our experiments, the percentage of positive cells decreased to ~20% after 20 days and 5% after 30 days (Fig. 3e).

Modulation of expression levels in mammalian cells. The hCMV-IE1 promoter we used in our experiments is considered to be the strongest promoter available for heterologous expression in most mammalian cells. It may be desirable to have alternative promoters that are characterized by lower levels of expression. We expanded our tool-box by incorporating the SV40, PGK and UBC promoters in alternative expression cassettes in our MultiPrime system (Supplementary Fig. 6). We determined expression levels from these alternative promoters by expressing EYFP-tubulin, and simultaneously expressing citrine from a CMV promoter as a bench-mark to normalize expression levels. All three alternative promoters show distinctly lower expression levels in HEK293 and PAE cells as well as in primary REF compared to CMV promoter-driven expression (Fig. 4a,b). Furthermore, we included a tetracycline-inducible promoter in our system (Supplementary Fig. 6). Tetracycline-inducible promoters are dependent on a transactivator, for example tTA, to initiate expression[29]. We observed approximately four times higher expression levels in the absence of doxycycline in HeLa cells stably producing tTA, which were transduced with a MultiPrime baculovirus containing a tetracycline-inducible expression cassette, in good agreement with reports involving tetracycline-inducible promoters on plasmids (Fig. 4c,d).

Bifunctional dual-host promoters. Expression plasmids that could be used for heterologous protein production in insect as well as in mammalian cells have not found wide-spread

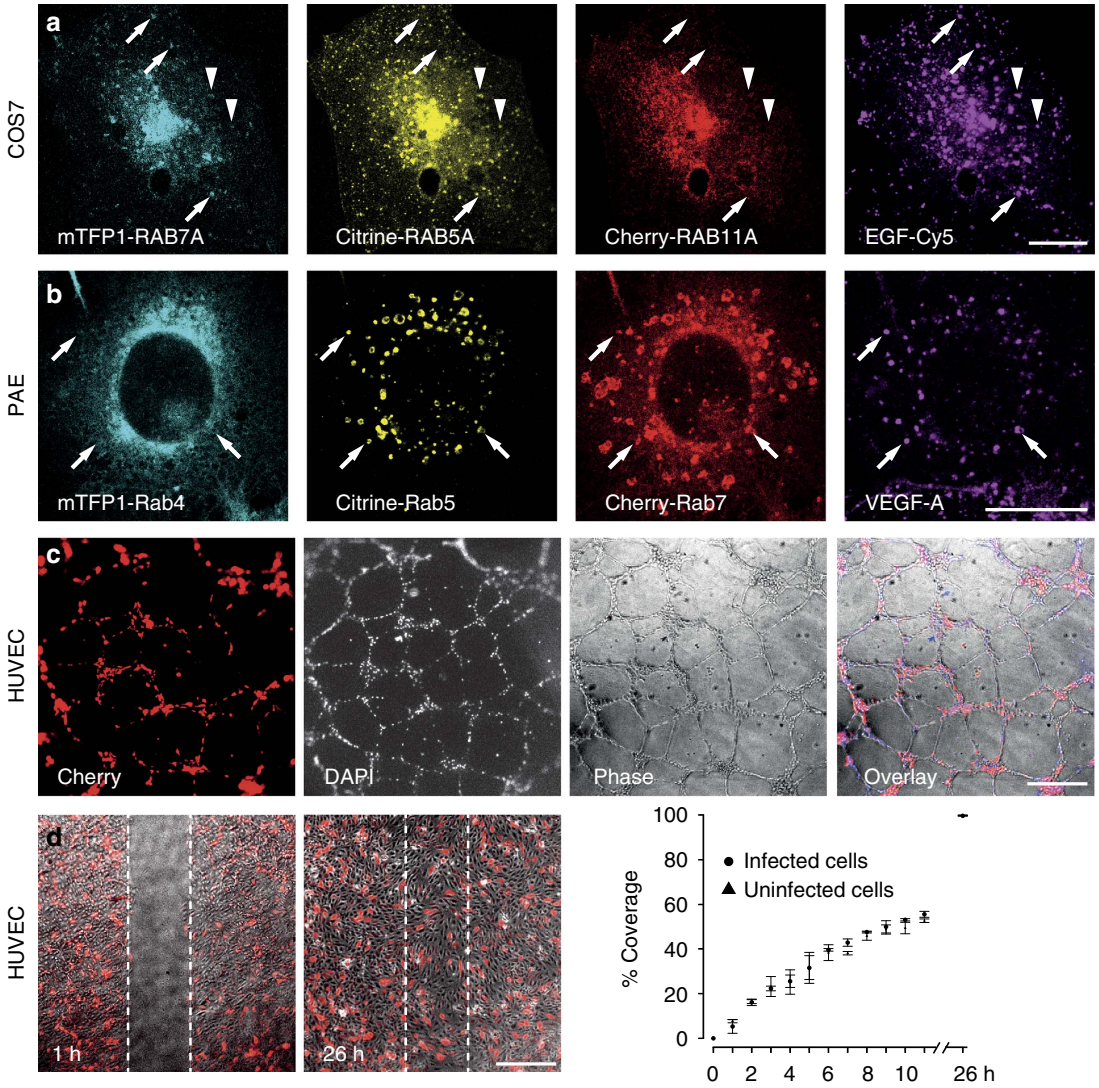


Figure 2 | Infected cells retain functionality. (**a**) COS7 cells were infected with MultiPrime baculoviruses expressing the indicated fluorescently-tagged RAB GTPases. Cells were stimulated for 3 h with Cy5-labelled EGF. As expected, EGF was found in RAB7A vesicles (arrows) and RAB5A vesicles (arrowheads) but not in RAB11 vesicles. A quantitative time-resolved analysis is provided (Supplementary Movie 1). (**b**) PAE cells stably expressing VEGFR2 and neuropilin-1 were infected with baculoviruses expressing the indicated fluorescently-tagged RAB GTPases. Cells were then stimulated with Nt647-labelled VEGF. VEGF can be found in RAB5 and Rab7 vesicles (arrows). (**c**) Tube formation: HUVEC were infected with a baculovirus expressing mTFP1-actin, EYFP-tubulin and Mito-dsRed (only red channel is shown). DAPI was used to counterstain nuclei of all cells. Approximately 30% of cells were infected. Infected and uninfected cells contribute to tubes. (**d**) Migration: HUVEC were infected with a baculovirus expressing mTFP1-actin, EYFP-tubulin and Mito-dsRed (only red channel is shown). Approximately 30% of cells were infected. Infected and uninfected cells migrate at same rates. Scale bar, 20 μm (**a**,**b**); 500 μm (**c**,**d**).

application so far, possibly because comprehensive comparative data which would have encouraged their use is currently lacking. We addressed this issue by creating, validating and incorporating dual-host promoters as a choice in our MultiPrime system. Our objective was to provide a single expression reagent, which is the composite MultiPrime baculovirus containing the genes of choice controlled by this validated dual-host promoter, for example to produce a protein or protein complex of choice efficiently in insect cells for structural studies and in mammalian cell lines for functional validation. We used two promoters, the first one (denoted CMVP10) is a fusion of the CMV promoter and the baculoviral very late promoter p10, the second (denoted CMVintP10) contains the p10 promoter in an intron of the CMV transcription unit (Fig. 5a and Supplementary Fig. 7). These two dual-host promoters were validated in mammalian cells by expressing EYFP-tubulin from a MultiPrime baculovirus, which also expressed citrine driven by a CMV promoter for normalization purposes. In HEK293 and PAE cells, the dual function promoters expressed at comparable levels to the original mammalian-only CMV promoter. In REF cells, the intron-less CMVP10 promoter resulted in lower expression (Fig. 5b,c). We quantified expression from these MultiPrime baculoviruses in insect cells and found them entirely satisfactory (Fig. 5d,e). Furthermore, we tested MultiPrime constructs expressing human transcription factors, which we had produced before for structural studies in insect cells with our MultiBac insect-cell expression system (Supplementary Fig. 8)[30,31]. We observed virtually indistinguishable levels of expression for complexes formed by these human TATA-box associated factors (TAFs) from dual function promoters as compared to the MultiBac expressed complexes. Transduction of HeLa cells with the TAF producing MultiPrime baculoviruses resulted

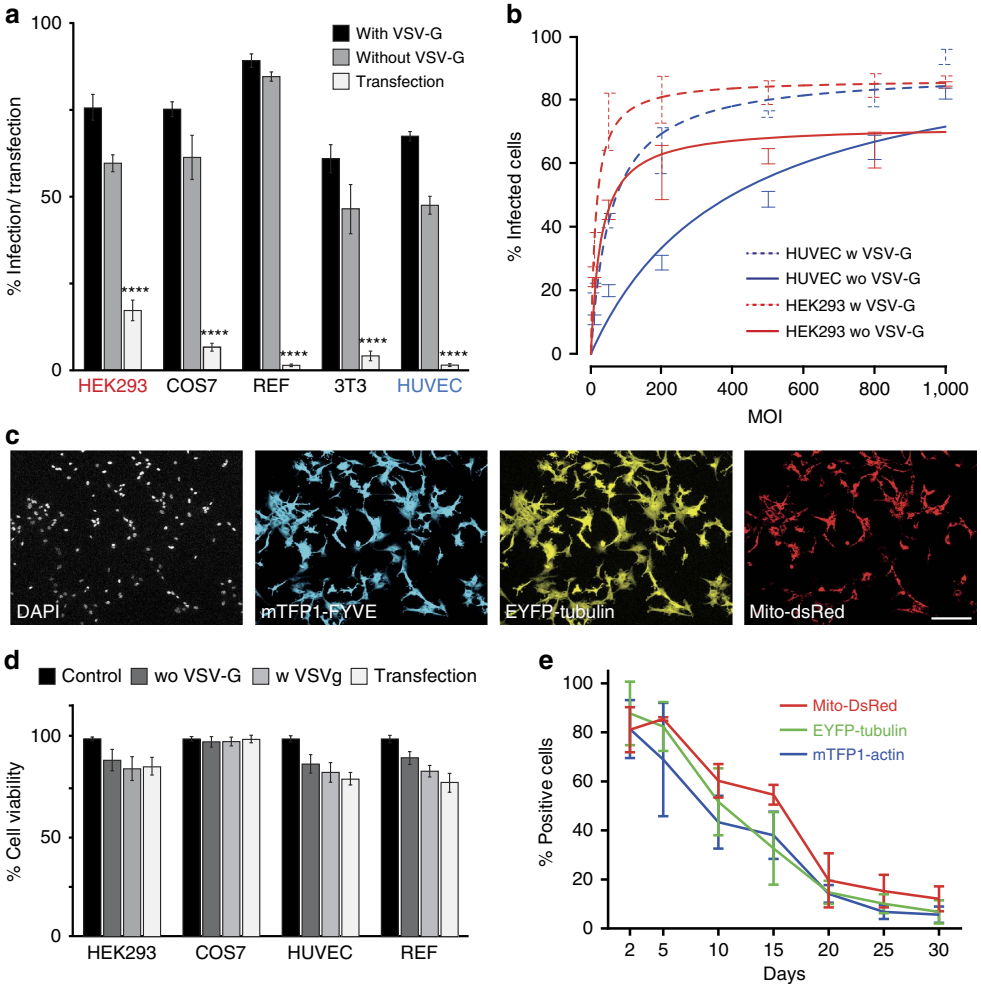


Figure 3 | Transduction efficacy. (**a**) Transduction by MultiPrime baculovirus was compared to plasmid-based transfection. MultiBacMam bacoluvirus expressing VSV-G and EMBacY baculovirus devoid of VSV-G were used. For plasmid-based transfection, the plasmid (13 kb) used originally inserted into the recombinant baculoviruses was utilized. MultiPrime-mediated transduction is markedly superior in all cell types tested. Data shows mean value ± s.d.; $n=3$; ****$P<0.0001$ determined by comparing transfection with transduction with or without VSV-G using one way analysis of variance followed by the Dunnet's *post hoc* test. (**b**) Effects of the MOI are shown. (**c**) PAE cells were infected with a MultiPrime baculovirus expressing three proteins at a MOI of 500. DAPI was used to counterstain nuclei of all cells. Virtually all cells are infected and express all heterologous proteins. Scale bar, 100 µm. (**d**) Toxicity of transduction was measured by means of a MTT assay. MultiPrime transduction exhibits comparable, low toxicity as plasmid transfection. (**e**) The persistence of heterologous expression was quantified by fluorescence in REF cells. The percentage of positive cells for each individual protein is shown over time.

in close to complete transduction rates (Supplementary Fig. 8).

Genome engineering by CRISPR/Cas9. CRISPR/Cas9-mediated genome engineering requires the expression of Cas9, the concomitant expression of a U6-driven guide RNA (gRNA) and the provision of a DNA construct for homologous recombination. Currently used viral systems can harbour Cas9 and the gRNA but are unable to include a homology construct due to limited cargo capacity[32]. We assembled DNAs for the expression of a HMGA1-EGFP fusion protein from its endogenous locus in a MultiPrime virus[33] (see Supplementary Fig. 9 for details). Transduction of HEK293 and HUVEC led to expression of HMGA1-EGFP in the nucleus in ~1% of the cells. Successful homologous integration of the DNA construct was verified by PCR (Fig. 6a and Supplementary Fig. 9).

Reprogramming by MultiPrime. We next investigated whether MultiPrime viruses are suitable for reprogramming of cells. Currently, this is mainly carried out with lentivirus, which is a retrovirus that stably integrates into the genome of cells. We assembled a MultiPrime virus expressing the transcription factors Asc1, Brn2 and Myt1L, which were shown to convert MEFs into neurons[7]. Transduction of MEFs with this MultiPrime virus resulted in cells with neuron-like morphology, which expressed the neuronal markers MAP2 and β-tubulin III 20 days after transduction, indistinguishable from co-infection with three lentiviruses each expressing one of the transcription factors (Fig. 6b). Our results provide compelling evidence that reprogramming can be successfully achieved with a transient expression system such as MultiPrime.

Functional antibody production. Our MultiPrime approach can not only induce morphological changes in cells but also potentially interfere with it. We addressed this by using MultiPrime to express functional antibodies in primary cells. Our previously described single-chain antibodies targeting VEGF (SZH9) and VEGFR2 (ADH9), and a control single-chain

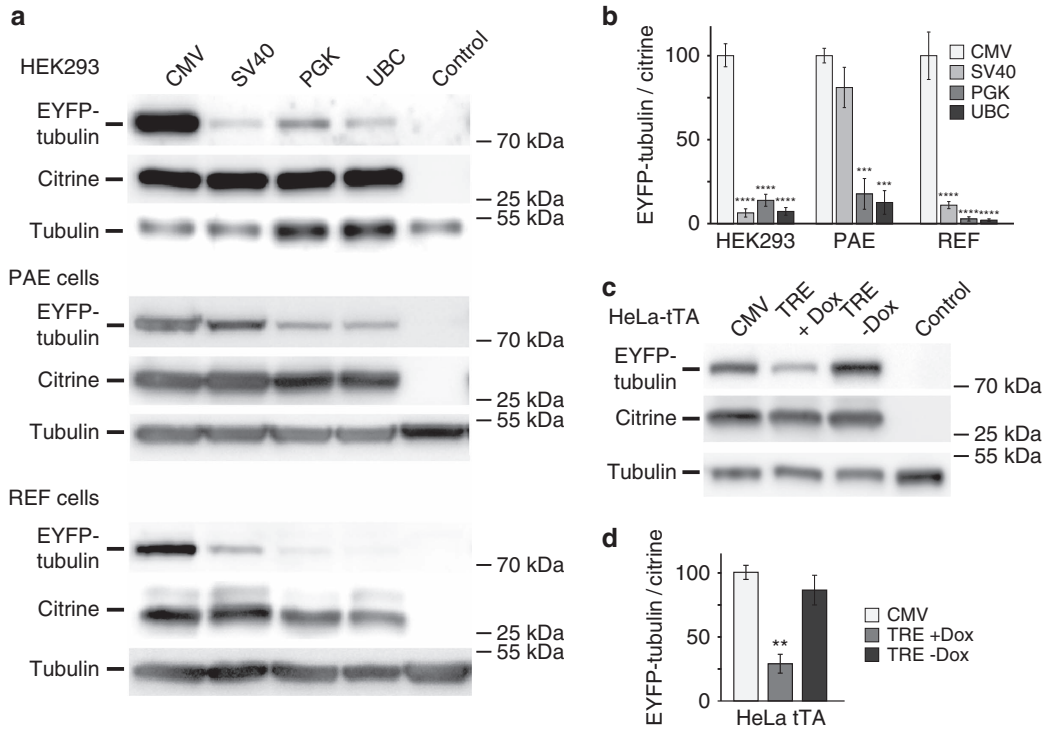


Figure 4 | Modulation of expression levels with alternative promoters. (**a**) Transduction of HEK, PAE and REF cells with baculoviruses expressing EYFP-tubulin under the control of the indicated promoters and citrine under the control of the CMV promoter is shown. (**b**) Quantification of the blots. EYFP-tubulin/citrine ratio was used as a measure for promoter strength. Endogenous tubulin was used as loading control. Data shows mean value ± s.d.; $n=3$; ****$P<0.0001$, ***$P<0.001$ determined by comparing CMV promoter with alternative promoters using one-way analysis of variance followed by the Dunnet's *post hoc* test. The CMV promoter is the strongest promoter in all tested cell lines. (**c**) HeLa-tTA cells were infected with a baculovirus expressing EYFP-tubulin under the control of a tetracyclin inducible element and citrine under the control of the CMV promoter. (**d**) Quantification of the blots shown above. Data shows mean value ± s.d.; $n=3$; **$P<0.01$ determined by comparing induced versus non-induced tet promoter using one-way analysis of variance followed by the Tukey *post hoc* test. No significant difference between CMV promoter and induced tet promoter.

antibody (A1) were converted into a full length IgG consisting of light and heavy chains[34,35]. The dual-host promoter CMV-CMVintP10 was utilized to drive recombinant IgG expression. The resulting MultiPrime baculovirus was successfully tested for expression in HEK293 and insect cells (Fig. 6c). All IgG antibodies tested were expressed at comparable levels. The same virus was then used to transduce HUVEC that were then placed into a tube formation assay. Only the function-blocking anti-VEGF antibody SZH9 was able to interfere with tube formation. All other antibodies, including the VEGFR2-binding but not function-blocking antibody ADH9, did not interfere with tube formation (Fig. 6c).

Zebrafish transduction. It was previously shown that mammalian promoters can be used for heterologous expression in zebrafish embryos[36]. We set out to establish whether MultiPrime viruses are restricted to mammalian and insect cells, or whether they can also be used to transduce zebrafish. A MultiPrime virus encoding mTFP1-actin, EYFP-tubulin and Mito-dsRed under control of mammalian CMV promoters was injected into intercellular spaces in the brain region of zebrafish embryos at 24 h post fertilization. Injection of this virus showed heterologous expression of all genes in zebrafish embryos. Expression was restricted to the site of injection and could be detected for at least 5 days (Fig. 6d).

Discussion

In the three decades since their inception, baculovirus-based expression systems have become well-established and widely used for recombinant protein production in insect cells. Later, it was discovered that baculoviruses not only infect insect cells but can also drive heterologous protein expression in mammalian cells if appropriate mammalian regulatory elements are provided in the recombinant baculovirus genome[22,23,37]. This so-called 'BacMam' method has been applied to produce heterologous proteins in academic and industrial research and development, notably for pharmacological screening[37]. Today, it is becoming increasingly evident that most physiological activities are mediated by multiple proteins forming complex assemblies. Therefore, a powerful tool that exploits recombinant baculovirus to deliver multiple genes simultaneously and reproducibly into a range of mammalian cell types and notably primary cells is highly desirable to study health and disease states, and to analyze molecular mechanisms of cell fate. Notwithstanding, such a tool has been lacking so far. Therefore, we developed MultiPrime, a versatile, flexible and fully modular system for efficient multigene delivery and expression in any mammalian cell type, primary and established. MultiPrime relies on a set of customized DNA plasmid modules, called acceptors and donors that provide the means to combine a theoretically unlimited number of genes of interest with different promoters, terminators and other control elements in multiple expression cassettes to generate multigene-delivery constructs, which are then inserted into engineered baculoviral genomes. Moreover, they can comprise all the elements necessary for genome engineering including editing functions and the sequences required for homologous recombination. We provide and validate a range of mammalian promoters that can be introduced into our MultiPrime system in this way. In addition,

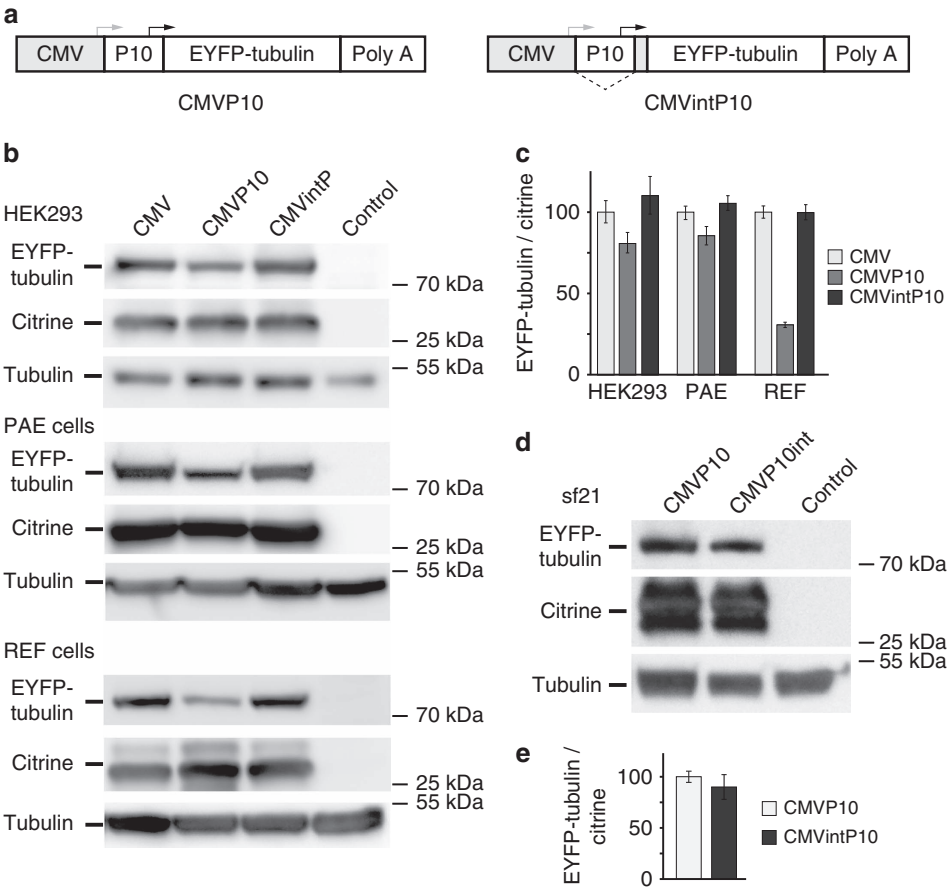


Figure 5 | Promoters active in mammalian and insect cells. (**a**) Structure of the tested dual promoters is shown schematically. In CMVP10, the baculoviral very late promoter p10 was inserted downstream of the CMV promoter. In CMVintP10, the p10 promoter was placed within an intron and is spliced out from the transcript of the CMV promoter. Grey arrow: transcription initiation of the CMV promoter; black arrow: transcription initiation of the p10 promoter. (**b**) HEK293, PAE and REF cells were infected with a MultiPrime baculovirus expressing EYFP-tubulin under the control of the above promoters and citrine under the control of the CMV promoter. The lysates of the cells were analysed by western blotting. (**c**) Quantification of blots shown in **b**. EYFP-tubulin/citrine ratio was used as a measure for promoter strength. Endogenous tubulin was used as loading control. The CMVintP10 promoter expresses at a similar level as the original CMV promoter, while the CMVP10 promoter expresses at a lower level in mammalian cells. (**d**,**e**) The same baculoviruses were used to infect insect (Sf21) cells. Citrine was expressed by a P10 driven expression cassette in the backbone of the baculovirus. Data shows mean value ± s.d.; $n = 3$; there is no significant difference ($P > 0.05$) by comparing the two CMVP10 promoter variants using the Student's *t*-test.

we provide dual-host promoters to drive heterologous multiprotein production in both insect and mammalian cells. This highly versatile and flexible tool-box allows users to conveniently introduce many different proteins simultaneously into mammalian and insect cells. Corroborating previous observations, we found negligible toxicity and sustained viability when infecting a range of mammalian cells with recombinant MultiPrime reagents. Importantly, we demonstrate here that MultiPrime infected cells are competent to divide and migrate normally and are capable of adequately responding to external stimuli as, for example, growth factors.

In this study, we utilized two baculovirus types, EMBacY and MultiBacMam. These engineered baculoviral genomes are characterized by reduced proteolysis and delayed cell lysis during virus amplification in insect cells, resulting in high quality, high titre virus[25]. The EMBacY and MultiBacMam viruses express either EYFP or mCherry fluorescent marker genes, to signal late replication cycle entry. Expression of fluorescent marker proteins during virus production is a convenient tool to simplify and standardize monitoring the production of baculovirus, in particular for non-specialist users. MultiBacMam virus generated in this study expresses also VSV-G during virus production in insect cells. Decorating baculovirus with VSV-G has been shown to improve mammalian transduction efficiencies. Consequently, we observed superior transduction efficiencies with our MultiBacMam-derived viruses that display VSV-G on their capsids. We note here that, at least in Switzerland where these experiments were performed, MultiBacMam-derived reagents expressing VSV-G have to be handled at biosafety level 2, which requires specific laboratory infrastructure. To circumvent this complication the EMBacY virus variant can be utilized, which is devoid of VSV-G but still resulted in satisfactory transduction rates in our experiments. Nonetheless, for experiments which may rely on maximum transduction efficiencies, the VSV-G containing MultiBacMam virus is recommended.

All viral genomes we utilized contain a site-specific integration site in the backbone distal from the Tn7-attachment site. This LoxP site allows introduction of additional genes by Cre-LoxP mediated fusion *in vivo*[12]. This enables a range of options to modify and tailor the baculovirus genomes for specific applications. For example, a baculovirus called SweetBac was developed to achieve mammalian-type glycosylation of recombinant secreted proteins such as antibodies[38,39].

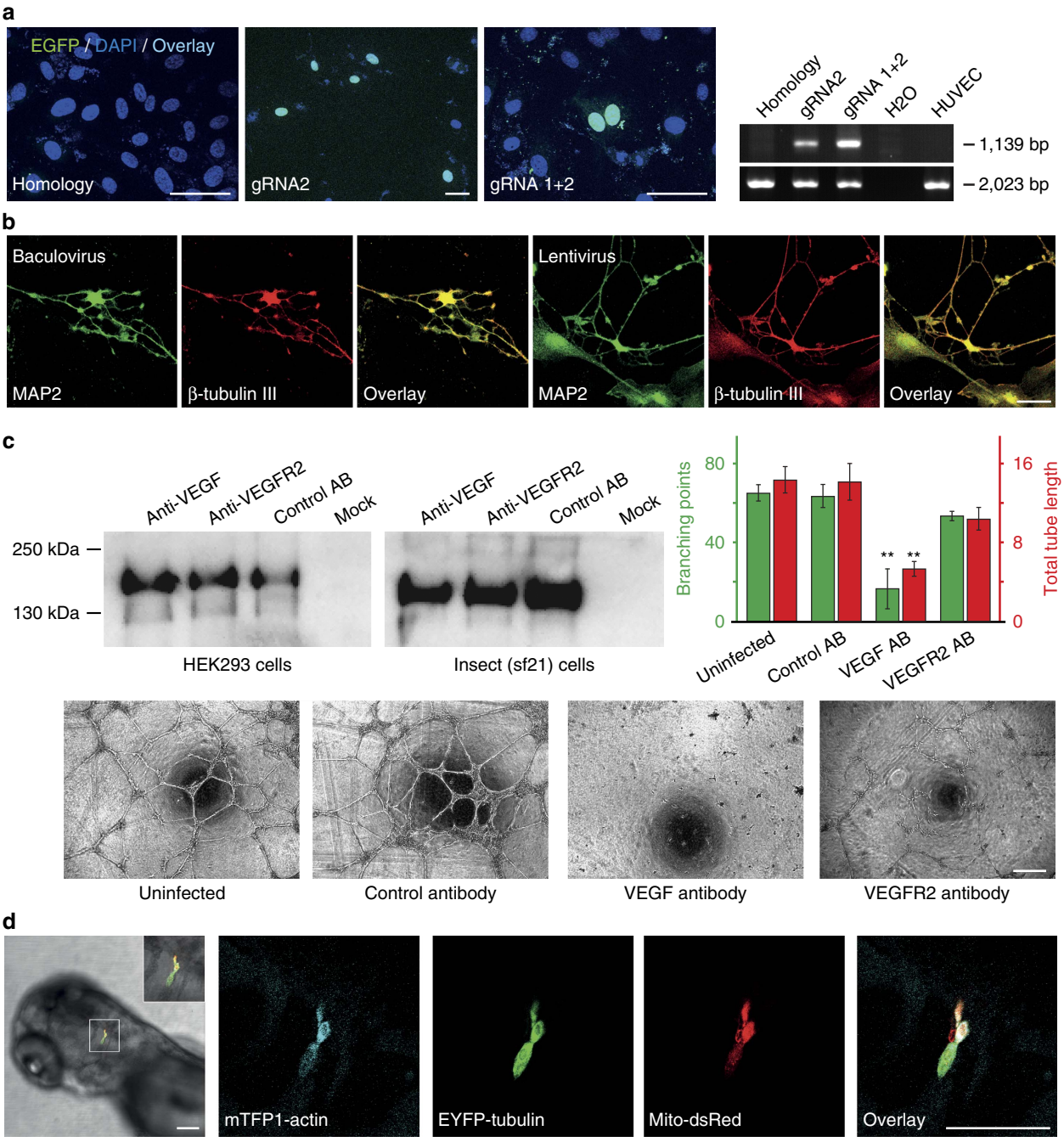


Figure 6 | MultiPrime applications. (**a**) Genome engineering. Infection of HUVEC with a baculovirus containing a HMGA1-EGFP homology construct did not lead to cells with HMGA1-EGFP expression (left). Co-expression with Cas9 and HMGA1-gRNA1 (2nd panel) or co-expression with Cas9, HMGA1-gRNA1 and HMGA1-gRNA2 (3rd panel) led to cells with HMGA1-EGFP expression in the nucleus. Correct integration of the homology construct was verified with PCR. The wild-type allele yielded a fragment of 2,023 bp, whereas the mutant allele results in a fragment with 1,139 bp (right). Scale bar, 50 μm. (**b**) Reprogramming of cells. MEF cells were infected with a Multiprime virus expressing Ascl1, Brn2 and Myt1L or co-infected with three lentiviruses individually expressing the same transcription factors. Both strategies led to cells with neuron-like morphology that express the neuronal markers MAP2 and β-tubulin III. Scale bar, 50 μm. (**c**) Functional antibody expression. MultiPrime viruses, encoding light and heavy chains of three different IgGs (anti-VEGF, anti-VEGFR2, and unspecific) were used to express antibodies in HEK293 (left) and also in insect cells (middle). HUVEC cells were infected with these viruses and the cells were used in a Matrigel-based angiogenesis assay. As expected, only the anti-VEGF antibody is capable of blocking tube formation Scale bar, 500 μm. Data shows mean value ± s.d.; $n=4$; $^{**}P<0.01$ when comparing VEGF function-blocking antibody versus control antibody. There is no significant difference ($P>0.05$) when comparing the non-function-blocking VEGFR2 antibody with the control antibody. Both P values are determined using one way analysis of variance followed by the Dunnet's *post hoc* test. (**d**) Baculovirus-mediated gene expression in zebrafish. Dorsal view of the head of a 3-day-old zebrafish larva after injection of MultiPrime baculoviruses expressing mTFP1-actin, EYFP-tubulin and Mito-DsRed into the hindbrain region at 24 h post fertilization. All infected cells express all heterologous proteins. Scale bar, 100 μm.

Currently, these functionalizations are limited to applications in insect cells. We anticipate that a wide range of functions to modify, enhance and regulate multiprotein production in mammalian cells will be exploited by modifying the baculoviral genome accordingly, providing appropriate expression cassettes active in mammalian cells in the LoxP locus of these vectors.

Multigene expression systems are rapidly gaining prominence for producing protein complexes for structural and functional studies. Often, several expression systems must be tested to obtain functional complexes in sufficient quantity and quality. This typically requires recloning of genes into different sets of expression plasmids given that the regulatory elements in each system, here mammalian and insect cells, are optimized for a particular host, and are typically not compatible between the different species. The incorporation of dual-host promoters into MultiPrime allows simultaneous testing of expression constructs in insect and mammalian cells by using the same reagent. This feature can be conveniently exploited if high-level production of a complex protein of interest is carried out in insect cells, while functional analysis of the same complex is performed in mammalian cells, which is increasingly the case in current structural biology. The possibility to use the same reagent for both host systems will also benefit analysis of structure–function correlations requiring multiple mutational analysis. MultiPrime affords the means to carry out such elaborate studies, notably also of complexes controlling cell fate, which can be mechanistically dissected by infecting primary cells.

Baculovirus constitutes an attractive tool for gene therapy for a number of reasons. Due to its flexible envelope structure, very large heterologous DNA cargo can be incorporated into the baculoviral genome. Moreover, baculovirus is replication incompetent in mammalian cells, and virtually no viral protein expression occurs on transfection in a mammalian host. Initial *in vivo* experiments had limited success since injected baculoviruses are rapidly inactivated and cleared by the immune complement system. Strategies were developed to overcome this impediment and many successful *in vivo* applications were published since then (reviewed in ref. 24). For example, expression of VEGF-D-induced vascularization in rabbit skeletal muscle suggesting that baculovirus-driven VEGF-D expression might be an option to cure lymphatic disorders[40]. Nevertheless as a non-integrative virus it is *a priori* limited to transient expression, which can be an advantage or a disadvantage depending on the application. Transient expression may be desirable, for example, for vaccination or to promote changes in cell fate. Of note, altering cell fate is a particularly interesting application for multigene expression systems, as it relies on the simultaneous and temporally restricted expression of several transcription factors. Induced-pluripotent stem cells have been generated before with a BacMam virus *in vitro* using a fusion protein construct[41]. Four transcription factors were expressed as a fusion protein from a single open reading frame (ORF) via self-cleaving 2A peptides. Here, we converted MEFs into neurons using independent expression cassettes, which offer advantages especially when different protein combinations need to be tested in a combinatorial fashion.

MultiPrime is not restricted to the delivery of expression constructs. We anticipate that genome engineering will be an important future application, owed to the very large cargo capacity of baculoviruses. Other viruses such as lentiviruses or adeno-associated viruses cannot accommodate all DNA elements needed to produce Cas9, a gRNA and a construct for homologous recombination. With MultiPrime, we were able to modify the HMGA1 locus of HUVEC, which are human primary cells that show restricted replication potential.

Our results compellingly validate MultiPrime as a powerful vehicle for multigene delivery, protein expression and genome engineering, relevant for a large number of applications, *in vitro* and *in vivo*, and underscore the enormous potential of our baculoviral system to deliver large multigene DNA constructs into a wide range of mammalian cells, notably including primary cells. A multitude of genes and regulatory elements can be delivered due to the very large heterologous DNA cargo capacity of the system, offering novel exciting possibilities for biological research. Entire signalling cascades, gene regulatory systems or metabolic pathways and multiple mutants thereof, can be efficiently engineered with MultiPrime. We anticipate that many applications will benefit from MultiPrime, notably when efficient transfer of multiple genes or efficient engineering of genomes is required.

Methods

Molecular biology. DNA construction in MultiPrime follows the high-throughput compatible logic of our ACEMBL concept to prepare multicomponent DNA constructs from acceptor and donor plasmid DNA modules that are conjoined by the Cre-LoxP fusion reaction[13,14]. Plasmid pSI-AGR10 is the common acceptor in MultiPrime, and has been developed from our previous pFL plasmid[25]. The ampicillin resistance gene and an internal SapI site were removed and the expression cassettes for insect-cell expression replaced by the CMV-based expression cassette from plasmid pSI-AKR1 by standard cloning methods (Supplementary Fig. 2a)[14]. In addition, pSI-AGR10 contains a gentamycin resistance marker, a LoxP site, and the DNA elements (Tn7R, Tn7L) required for transposition into the baculovirus genome by Tn7 transposase.

All donor plasmids of the original MultiLabel system are compatible with this acceptor[14]. Donors are fused to pSI-AGR10 by Cre-LoxP recombination concomitantly or in a sequential manner. Acceptor–donor assembly was performed as described and electrocompetent DH10β or $CaCl_2$ competent XL1-blue cells were used for transformation[42]. Sequences were assembled *in silico* using the 'Multi-Cre Recombination Tool' in the plasmid editor software Ape (http://biologylabs.utah.edu/jorgensen/wayned/ape/) or, alternatively, with software Cre-ACEMBLER[43]. Integrity of all fusion plasmids was confirmed by restriction mapping. Alternative mammalian promoters and dual-host promoters active in both mammalian and insect cells were synthesized by Genewiz (South Plainfield, USA) on the basis of sequences provided in the Supplementary Materials and inserted as AscI–HindIII fragments into parent Acceptor plasmid pSI-AGR10.

Recombinant baculoviral genomes. Two baculoviral genomes were used in this study, our previously described EMBacY genome and the novel MultiBacMam genome, which we constructed in this study. Both baculoviral genomes are present as bacterial artificial chromosomes (BAC) in *E. coli* cells (DH10EMBacY and DH10MultiBacMam, respectively). EMBacY produces yellow fluorescent protein (YFP) as a marker in infected insect cells as a means to track virus amplification and performance by monitoring the fluorescence signal[25].

Display of a VSV-G on the baculovirion was reported to enhance mammalian transduction efficiency by baculovirus[27]. We therefore constructed a novel MultiBacMam baculovirus by modifying our original MultiBac baculoviral genome, retaining its advantageous features including reduced proteolysis and delayed cell lysis[12]. A synthetic gene (Genscript, Psicataway, NJ) encoding for VSV-G was inserted into a modified pUCDM donor plasmid[25] by using BamHI and XbaI restriction sites to yield plasmid pLox-VSV-G. Subsequently, a second cassette containing a synthetic gene for mCherry (Genscript) was inserted by using the multiplication module as described[13]. The resulting pLox-VSV-G-mCherry Donor plasmid was incorporated into the MultiBac virus by transforming DH10MultiBac^{Cre} cells harbouring the MultiBac baculoviral genome as a BAC and Cre recombinase expressed from a pBADZHisCre helper plasmid on arabinose induction[12]. Positive integrands were selected by antibiotic screening. Successful Cre-mediated integration was further verified by PCR analysis as described[44]. Competent DH10MultiBacMam cells were prepared following standard protocols and contain in addition to the MultiBacMam baculovirus also a helper plasmid expressing Tn7 transposase on induction with isopropyl β-D-1-thiogalactopyranoside. Expression of mCherry from this baculovirus in infected insect cells during virion production results in the cell culture adopting a characteristic purple color, allowing tracking of successful viral infection and production easily by eye.

Generation of composite MultiPrime baculovirus. MultiPrime acceptors or acceptor–donor fusions were transformed into electrocompetent DH10EMBacY or DH10MultiBacMam cells, respectively. Composite baculovirus generation occurred by Tn7 tranposition mediated by Tn7 transposase expressed from a helper plasmid. Transformants were selected and composite baculoviral genomes prepared as described[44] Sf21 insect cells were transfected with Cellfectin II (Life Technologies) at a density of 0.5×10^6 cells ml^{-1} according to manufacturer's recommendations.

We took particular care during virus amplification to prevent accumulation of defective virus, which would not express all heterologous genes. We applied a protocol we had developed previously for successful amplification of composite baculovirus containing multiple copies of viral late promoters (polh, p10), preserving the integrity of the viral genome[28]. Briefly, primary baculovirus stock (V0, 2 ml) was harvested 50–60 h after transfection and 0.5 ml was used to infect 4 ml new Sf21 cells for 75 h yielding V1 stock. Overall 3 ml of this V1 baculovirus stock was then used for a further round of virus amplification for 60 h (V2, 100 ml). The amplification of the virus was followed in this phase by monitoring EYFP (EMBacY) or mCherry (MultibacMam) expression from the viral backbone (Supplementary Fig. 1). Less than 1% of cells were positive when harvesting V0. When harvesting V1, 20–30% of cells were positive and after V2, 80–90% of cells were positive. Incubation times must not be extended during amplification, otherwise over-amplification of the virus can occur, resulting in loss of heterologous insert (Supplementary Fig. 4). The V2 virus stocks were stored either at 4 °C or after addition of 5% FBS at − 80 °C. For sensitive cells (for example, iPS) or zebrafish, virus was concentrated by ultracentrifugation. For this purpose, virus supernatant was placed on a sucrose cushion (25% sucrose/ 5 mM NaCl/ 10 mM EDTA) and then centrifuged for 90 min at 80,000*g*. The pellet containing the virus was resuspended in PBS pH 6.2 (ref. 45). The titre of baculovirus stocks was determined using end-point dilution assay[46]. Viruses displaying VSV-G were handled as biosafety level 2 agents in Switzerland.

Cell culture. Insect cells (Sf21, Sf9; Life Technologies) were cultured in SF-4 BaculoExpress ICM medium (Amimed) containing 1% FBS at 27 °C. Mammalian cells were incubated at 37 °C in a humidified atmosphere containing 5% CO_2. HEK293, COS7, REF, Swiss 3T3 and HeLa cells were cultured in DMEM (Amimed) containing 10% FBS (Life Technologies) and 100 units ml^{-1} penicillin and 100 µg ml^{-1} Streptomycin (Life Technologies). PAE and HUVEC (Life Technologies) cells were maintained in Ham-F12 media supplemented with 10% FBS and penicillin/ streptomycin and M-200 medium (Life Technologies), respectively. Primary rat cortex neurons (Life Technologies) were cultured in Neurobasal medium supplemented with 1% B27 and 1% Glutamax (Life Technologies). Human iPS cells (NAS2) were obtained from Tilo Kunth (University of Edinburgh) and cultured in feeder-free maintenance medium for human ES/ iPS cells (mTESR1 medium; Stem Cell Technology)[47]. iPS cells were immunostained with an Oct4 antibody (Santa Cruz Biotechnology, sc-5279, 1:500 in 10% NDS/ 0.2% Triton X100/ 1% BST/ PBS), transfections were performed with FusionHD (Promega) according to manufacturer's recommendations and cells were analyzed 42 h after transfection. For endothelial tube formation assays, MultiPrime-transduced HUVEC cells were seeded on a ibiTreat µ-slide angiogenesis plate (ibidi GmbH, Germany) at a density of 5,000 cells/slide in EGM-2 medium (Life Technologies) and analyzed 16 h later. Migration assays were performed by seeding 21000 MultiPrime-transduced HUVEC cells in ibidi Culture-Insert plate (ibidi GmbH, Germany). The culture-insert was removed after 14 h and migration into the gap was monitored every 2 h for 24 h.

Transduction of mammalian cells. Mammalian cells were plated at a density of 2.5×10^5 cells per well in six-well plates 1 day before transduction. Baculovirus was added at a MOI between 100 and 500 in 80% insect medium/ 20% DMEM without any FBS or antibiotics. Transduced cells were incubated at 37 °C for 8 h, and the medium was then replaced with fresh mammalian cell culture medium. Plates were cultured for one or two additional days. Cytotoxicity of baculovirus transduction was monitored by MTT assay in various mammalian cells lines. Overall 10^4 cells were plated in a 96-well plate and transduced with baculovirus at a MOI 500. After 24 h, the medium was replaced with culture medium containing 20 µM resazurin and the cells were incubated for 2–4 h. The number of viable cells was obtained by monitoring resazurin fluorescence with a microplate spectrofluorometer (Tecan Ltd).

For CRISPR/Cas9-mediated genome engineering, HEK293 cells and HUVEC were transduced with a Multiprime virus that expresses CMV-driven Cas9, U6-driven gRNAs and a homology construct as described in Supplementary Fig. 9. Immunostaining for Cas9 after 40 h revealed 80% transduction. Cells were fixed after 4 or 6 days and analyzed for nuclear EGFP expression by microscopy or DNA was extracted with QIAamp DNA mini kit (Qiagen). Correct integration was verified by PCR using primers described in Supplementary Fig. 9.

MEF cells used for reprogramming to neurons were obtained from Amsbio, and were used at passage 3. Cells were transduced with concentrated bacuolovirus expressing Ascl1, Brn2 and Myt1L in MEF medium for 8 h. In parallel, MEFs were infected with lentiviruses containing expression constructs Tet-o-FUW-Ascl1, Tet-o-FUW-Brn2 and Tet-o-FUW-Myt1L (all from Addgene) in presence of polybrene (8 µg ml^{-1}). Cells were cultured in N3 medium (DMEM/F12, B27, N2 (all (Life Technologies), 25 µg ml^{-1} Insulin (Sigma-Aldrich)). Doxycycline (2 µg ml^{-1}) was added to lentivirus-transduced cells[7,48]. Cells were fixed after 6, 12 and 20 days. Immunostaining was performed with chicken anti MAP2 (Neuromics, CH22103, 1:5,000 in 10% NDS/ 0.2% Triton X100/ 1% BST/ PBS) and mouse anti β-tubulin III (Sigma, T8578, 1:600 in 10% NDS/ 0.2% Triton X100/ 1% BST/ PBS) antibodies.

Microscopy. Cells for microscopic analysis were plated on glass coverslips. Untreated coverslips were used for COS7, REF, Swiss 3T3 and PAE cells. Poly-L-lysine (Sigma P4707) treated coverslips were used for HEK293 cells and 0.1% gelatin (Sigma G1393) treated coverslips for HUVEC. Poly-D-lysine hydrobromide coated coverslips (Sigma P7280) were used for primary rat cortex neurons and iPS were plated on hESC qualified Matrigel (BD Bioscence). Analysis of cells was performed 27, 42 and 48 h after transduction. Cells were fixed with 4% formaldehyde in PBS and mounted with Gelvatol. Imaging was performed on a Leica SP5 laser scanning confocal microscope or on an Olympus IX81 equipped with an Andor iXonEM camera. On Leica SP5, EBFP2 was excited with the 405 nm laser line and the emission was collected from 430 to 450 nm (405/ 430-450). The other fluorescent proteins were analyzed as follows: mTFP1 (458/ 485-510), mCitrine (514/ 525-545), mCherry (543/ 585-620) and mPLUM (633/ 640-800). In addition, the spectral mode (xyλ) of the microscope was used to verify the presence of all fluorescent proteins (data not shown). Standard excitation and emission filters were used on the Olympus IX81. Quantification was performed with Squassh[49].

Western blotting. Mammalian cells were lysed 42 h after transduction with lysis buffer (0.5% Triton X100, 50 mM Tris–HCl, 100 mM NaCl, pH 7.5). The supernatant was used for western blotting after sonification and centrifugation. Rabbit anti-GFP (Abcam ab137827; diluted 1:2500 in 3% BSA/TBST) and mouse anti-tubulin (Sigma T5168; diluted 1:2,500 in 3% BSA/TBST) were used as primary antibodies. As secondary antibodies, alkaline phosphatase-coupled goat anti-rabbit and anti-mouse as well as donkey anti-human IgGs (Southern Biotech, diluted 1:10,000 in TBST) were used, followed by chemiluminescence detection. Quantification was performed with ImageJ. Original western blots are shown in Supplementary Fig. 10.

Transduction of zebrafish embryos. All experiments were performed in accordance with the animal welfare guidelines of the Federal Veterinary Office of Switzerland. Zebrafish (*Danio rerio*) were maintained as described[50]. Embryos of the wild-type strain WIK were raised at 28°C in E3 medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$ and 0.33 mM $MgSO_4$), and pigment development was inhibited by phenylthiourea (1-phenyl-2-thiourea; Sigma-Aldrich) as described in Westerfield[50]. For injections, individual dechorionated embryos at 24 h post fertilization were anesthetized in 200 mg ml^{-1} 3-aminobenzoic acid methyl ester (MESAB, Sigma-Aldrich) and 4.6 mM $NaHCO_3$ and embedded in 1% low melting agarose (Lonza). Concentrated viruses were injected in intercellular spaces using borosilicate glass microcapillary injection needles (1 mm outside diameter × 0.78 mm inside diameter, Science Products GmbH, Hofheim, Germany) and a PV820 Pneumatic PicoPump (World Precision Instruments, Sarasota, Florida, USA). After injection, the infected embryos were returned to E3 medium. For in vivo imaging, the larvae were anesthetized and embedded in 1% low melting agarose in a 35 mm imaging dish with a glass bottom (Ibidi) and imaged using a CLSM SP5 Mid UV–vis Leica inverted microscope.

Bifunctional CMV/ P10 dual-host promoters. Subcomplexes of human general transcription factor TFIID were produced in insect cells from the CMV-CMVintP10 promoter (Supplementary Fig. 7). A complex formed by human TBP associated factor (TAF) 8 and TAF10, and a complex formed by TAF5, TAF6 and TAF9, were expressed. Production levels of these complexes on the basis of the dual-host promoter was compared to previously established production levels on the basis of the baculoviral polyhedrin (polh) promoter. A polyprotein strategy was utilized[51], which we had developed for high-level expression of complexes[30,31]. Briefly, the genes encoding for the TAFs were placed in a single ORF flanked by genes encoding for tobacco etch virus NIa (TEV) protease at 5′ and a cyan fluorescent protein at the 3′-end[51]. The ORFs give rise to self-processing polyproteins, which are cleaved by TEV protease at high specific TEV protease cleavage sites in between the constituent proteins. The polyproteins are shown schematically in Supplementary Fig. 8.

Polyproteins were expressed from the CMVintP10 and the polh promoter, respectively, by using EMBacY, and purified as described[30,31]. Production levels of the polyproteins in insect cells were indistinguishable notwithstanding the promoter used (Supplementary Fig. 8). Next, the MultiBacMam virus (see above) was used in conjunction with the CMVintP10 promoter, again leading to indistinguishable expression levels (Supplementary Fig. 8). HeLa cell cultures were transfected with MultiBacMam-derived baculoviruses that had been used to express the proteins in insect cells. Complete DMEM containing 10% foetal calf serum and 8 mM L-glutamine was utilized. Best results were achieved by supplementing the media with 3 mM sodium butyrate for cell recovery after aspirating the virus. Nearly all HeLa cells were transduced as revealed by measuring the specific fluorescence of the cyan fluorescent protein marker encoded by the polyprotein (Supplementary Fig. 8).

References

1. Kretzschmar, K. & Watt, F. M. Lineage tracing. *Cell* **148,** 33–45 (2012).
2. Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450,** 56–62 (2007).

3. Ballmer-Hofer, K., Andersson, A. E., Ratcliffe, L. E. & Berger, P. Neuropilin-1 promotes VEGFR-2 trafficking through Rab11 vesicles thereby specifying signal output. *Blood* **118,** 816–826 (2011).
4. Trowitzsch, S., Klumpp, M., Thoma, R., Carralot, J. P. & Berger, I. Light it up: highly efficient multigene delivery in mammalian cells. *Bioessays* **33,** 946–955 (2011).
5. Marsic, D. *et al.* Vector design Tour de Force: integrating combinatorial and rational approaches to derive novel adeno-associated virus variants. *Mol. Ther* **22,** 1900–1909 (2014).
6. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126,** 663–676 (2006).
7. Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463,** 1035–1041 (2010).
8. Church, G. M., Elowitz, M. B., Smolke, C. D., Voigt, C. A. & Weiss, R. Realizing the potential of synthetic biology. *Nat. Rev. Mol. Cell. Biol.* **15,** 289–294 (2014).
9. Guye, P., Li, Y., Wroblewska, L., Duportet, X. & Weiss, R. Rapid, modular and reliable construction of complex mammalian gene circuits. *Nucleic Acids Res.* **41,** e156 (2013).
10. Barford, D., Takagi, Y., Schultz, P. & Berger, I. Baculovirus expression: tackling the complexity challenge. *Curr. Opin. Struct. Biol.* **23,** 357–364 (2013).
11. Mansouri, M. & Berger, P. Strategies for multigene expression in eukaryotic cells. *Plasmid* **75,** 12–17 (2014).
12. Berger, I., Fitzgerald, D. J. & Richmond, T. J. Baculovirus expression system for heterologous multiprotein complexes. *Nat. Biotechnol.* **22,** 1583–1587 (2004).
13. Bieniossek, C. *et al.* Automated unrestricted multigene recombineering for multiprotein complex production. *Nat. Methods* **6,** 447–450 (2009).
14. Kriz, A. *et al.* A plasmid-based multigene expression system for mammalian cells. *Nat. Commun.* **1,** 120 (2010).
15. Haffke, M., Viola, C., Nie, Y. & Berger, I. Tandem recombineering by SLIC cloning and Cre-LoxP fusion to generate multigene expression constructs for protein complex research. *Methods Mol. Biol.* **1073,** 131–140 (2013).
16. Vijayachandran, L. S. *et al.* Robots, pipelines, polyproteins: enabling multiprotein expression in prokaryotic and eukaryotic cells. *J. Struct. Biol.* **175,** 198–208 (2011).
17. Sokolenko, S. *et al.* Co-expression vs co-infection using baculovirus expression vectors in insect cell culture: Benefits and drawbacks. *Biotechnol. Adv.* **30,** 766–781 (2012).
18. Anliker, B. *et al.* Specific gene transfer to neurons, endothelial cells and hematopoietic progenitors with lentiviral vectors. *Nat. Methods* **7,** 929–935 (2010).
19. Kost, T. A., Condreay, J. P. & Jarvis, D. L. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat. Biotechnol.* **23,** 567–575 (2005).
20. Ames, R. S., Kost, T. A. & Condreay, J. P. BacMam technology and its application to drug discovery. *Expert Opin. Drug Discov.* **2,** 1669–1681 (2007).
21. Kost, T. A., Condreay, J. P. & Ames, R. S. Baculovirus gene delivery: a flexible assay development tool. *Curr. Gene. Ther.* **10,** 168–173 (2010).
22. Hofmann, C. *et al.* Efficient gene transfer into human hepatocytes by baculovirus vectors. *Proc. Natl Acad. Sci. USA* **92,** 10099–10103 (1995).
23. Boyce, F. M. & Bucher, N. L. Baculovirus-mediated gene transfer into mammalian cells. *Proc. Natl Acad. Sci. USA* **93,** 2348–2352 (1996).
24. Airenne, K. J. *et al.* Baculovirus: an insect-derived vector for diverse gene transfer applications. *Mol. Ther.* **21,** 739–749 (2013).
25. Fitzgerald, D. J. *et al.* Protein complex expression by using multigene baculoviral vectors. *Nat. Methods* **3,** 1021–1032 (2006).
26. Luckow, V. A., Lee, S. C., Barry, G. F. & Olins, P. O. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in *Escherichia coli*. *J. Virol.* **67,** 4566–4579 (1993).
27. Barsoum, J., Brown, R., McKee, M. & Boyce, F. M. Efficient transduction of mammalian cells by a recombinant baculovirus having the vesicular stomatitis virus G glycoprotein. *Hum. Gene. Ther.* **8,** 2011–2018 (1997).
28. Berger, I. *et al.* The multiBac protein complex production platform at the EMBL. *J. Vis. Exp.* **77,** e50159 (2013).
29. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl Acad. Sci. USA* **89,** 5547–5551 (1992).
30. Bieniossek, C. *et al.* The architecture of human general transcription factor TFIID core complex. *Nature* **493,** 699–702 (2013).
31. Trowitzsch, S. *et al.* Cytoplasmic TAF2-TAF8–TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules. *Nat. Commun.* **6,** 6011 (2015).
32. Marx, V. Cell biology: delivering tough cargo into cells. *Nat. Methods* **13,** 37–40 (2016).
33. Ratz, M., Testa, I., Hell, S. W. & Jakobs, S. CRISPR/Cas9-mediated endogenous protein tagging for RESOLFT super-resolution microscopy of living human cells. *Sci. Rep.* **5,** 9592 (2015).
34. Rubio Demirovic, A. *et al.* Targeting human cancer cells with VEGF receptor-2-directed liposomes. *Oncol. Rep.* **13,** 319–324 (2005).
35. Zeisberger, S. M. *et al.* Clodronate-liposome-mediated depletion of tumour-associated macrophages: a new and highly effective antiangiogenic therapy approach. *Br. J. Cancer* **95,** 272–281 (2006).
36. Wagle, M. & Jesuthasan, S. Baculovirus-mediated gene expression in zebrafish. *Mar. Biotechnol.* **5,** 58–63 (2003).
37. van Oers, M. M., Pijlman, G. P. & Vlak, J. M. Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology. *J. Gen. Virol.* **96,** 6–23 (2015).
38. Palmberger, D., Wilson, I. B., Berger, I., Grabherr, R. & Rendic, D. SweetBac: a new approach for the production of mammalianised glycoproteins in insect cells. *PLoS ONE* **7,** e34226 (2012).
39. Palmberger, D., Klausberger, M., Berger, I. & Grabherr, R. MultiBac turns sweet. *Bioengineered* **4,** 78–83 (2013).
40. Heikura, T. *et al.* Baculovirus-mediated vascular endothelial growth factor-D(DeltaNDeltaC) gene transfer induces angiogenesis in rabbit skeletal muscle. *J. Gene Med.* **14,** 35–43 (2012).
41. Takata, Y. *et al.* Generation of iPS cells using a BacMam multigene expression system. *Cell Struct. Funct.* **36,** 209–222 (2011).
42. Kriz, A., Schmid, K., Ballmer-Hofer, K. & Berger, P. Integration of multiple expression cassettes into mammalian genomes in a single step. *Protocol Exchange*, doi:10.1038/protex.2011.249 (2011).
43. Becke, C., Haffke, M. & Berger, I. Ce-ACEMBLER Software User Manual. 10.13140/2.1.1068.11278; https://github.com/christianbecke/Cre-ACEMBLER (2012).
44. Bieniossek, C., Richmond, T. J. & Berger, I. MultiBac: multigene baculovirus-based eukaryotic protein complex production. *Curr. Protoc. Protein Sci.* Chapter **5,** Unit 5 20 (2008).
45. O'Reilly, D. R., Miller, L. K. & Luckov, V. A. *Baculovirus Expression Vectors: A laboratory Manual* (Oxford University Press, 1994).
46. Dee, K. U. & Shuler, M. L. Optimization of an assay for baculovirus titer and design of regimens for the synchronous infection of insect cells. *Biotechnol. Prog.* **13,** 14–24 (1997).
47. Devine, M. J. *et al.* Parkinson's disease induced pluripotent stem cells with triplication of the alpha-synuclein locus. *Nat. Commun.* **2,** 440 (2011).
48. Chanda, S. *et al.* Generation of induced neuronal cells by the single reprogramming factor ASCL1. *Stem cell Rep.* **3,** 282–296 (2014).
49. Rizk, A. *et al.* Segmentation and quantification of subcellular structures in fluorescence microscopy images using Squassh. *Nat. Protoc.* **9,** 586–596 (2014).
50. Westerfield, M. *The Zebrafish Book, A guide for the laboratory use of zebrafish (Danio rerio)*. 5th edn (University of Oregon Press, 2007).
51. Nie, Y., Bellon-Echeverria, I., Trowitzsch, S., Bieniossek, C. & Berger, I. Multiprotein complex production in insect cells by using polyproteins. *Methods Mol. Biol.* **1091,** 131–141 (2014).

Acknowledgements

This work was supported by the Swiss National Science Foundation (31003A-146975 and 3100A0-118351 to P.B.), the European Commission (EC) Framework programme (FP) seven projects ComplexINC, SynSignal and Biostruct-X (contracts nr. 279039, 613879 and 283570 to I.B.) and a fellowship of the Iranian Ministry of Science (to M.M.). We acknowledge the assistance and support of the Center for Microscopy and Image Analysis, University of Zürich for performing zebrafish *in vivo* imaging. We would like to thank Mayanka Asthana for help with VEGF labelling and Dr Stefan Jakobs for providing gRNA plasmids.

Author contributions

I.B. and P.B. conceived the study. M.M., I.B.E., Z.E., C.C.C., C.S.S., Y.X. and P.B. performed experiments; A.R. and M.W.D. performed data analysis or developed software; M.M., F.M.B., S.C.F.N., V.T., K.B-H., I.B. and P.B. designed experiments; M.M., K.B-H., I.B. and P.B. wrote the paper with input from all authors.

Additional information

Accession codes: Nucleotide sequences have been deposited in GenBank under the following accession codes: pSI-AGR10: KX001770; pSI-AGZ10: KX001771; pSI-DSZ2cx-TRE: KX001772; pSI-DSZ2cx-DI: KX001773; pSI-DSZ2x-DM: KX001774; pSI-DSZ2cx-PGK: KX001775; pSI-DSZ2cx-UBC: KX001776; pSI-DSZcx-SV40: KX001777; pSI-AGL10: KX001778; pSI-AGL10-DM: KX001779; pSI-AGL10-DI: KX001780; pSI-DAZ-DM: KX001781; pSI-DAZ-DI: KX001782.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare competing financial interest. I.B. is inventor on patents that comprise components of the systems here described (EP 1945773, EP 2403940).

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Mansouri, M. *et al.* Highly efficient baculovirus-mediated multigene delivery in primary cells. *Nat. Commun.* 7:11529 doi: 10.1038/ncomms11529 (2016).

CHAPTER 8
Acknowledgements

I would like to express my gratitude to my supervisor Prof. Verdon Taylor for giving me the opportunity to carry out my Ph.D in his lab. Thank you Verdon, I will never forget all of your supports, patience, and motivation during my Ph.D. Also, I will never forget meetings with you where I learned how to think like a scientist especially at the beginning of my Ph.D.

I am happy to be a member of our iPSC team including Dr. Ginetta Collo and Dr. Stefania Fedelle. I really appreciate Ginetta's help during my Ph.D in regards to planning of my research and experiments. I also like to thank Stefania. We worked together for 4 years and we had a lot of memories together. She was one of the best friends and colleague, who helped me a lot.

I also like to have great thank to Dr. Chiara Rolando for her continuous help, and motivation. Thank you, Chiara not only for great time we had during long day sorting but also for your help during the writing of the thesis and articles and preparation of my presentations. I learned a lot from you and it is my pleasure to meet you here. I also like to thank current and former lab members for all exciting time we had together especially Dr. Elena Parmigiani and Frank Sager. I also say thanks to Dr. Dirk junghans. When I was just naive Ph.D student he always encouraged me and he also helped me with correction of my introduction.

I would like to extend my appreciation to my scientific committee member, Prof. Raphael Guzman and Dr. Jan Pruszak for their helpful comments and discussion during my annual committee meeting.

I also would like to thank my family. My mom for her support and all of her kindness. I really thank my father for all freedom he gave me for my entire life. My parents taught me to live as an independent girl and decide for each step of my life. I also deeply appreciate my brothers, Sobhan and Sajjad as well as my sisters, Saeedeh and Salimeh.

Additionally, I would like to thank my parents in law for their support. I also take this chance to thank my grandfather, who always encourages me during my study since I was a bachelor student. Unfortunately, he is not anymore with us.

Last but not least I like to say thank to my love, Maysam Mansouri. He was always with me and encouraged me. We start our journey since we were master student and I wish great years ahead. Maysam I don't forget your support especially in difficult moments. I know that it was difficult to plan holidays with my experiments but you always support me.