

Neuronal identity and maturation

Citation for published version (APA):

van den Hurk, M. (2016). *Neuronal identity and maturation: insights from the single-cell transcriptome*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20161213mh>

Document status and date:

Published: 01/01/2016

DOI:

[10.26481/dis.20161213mh](https://doi.org/10.26481/dis.20161213mh)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

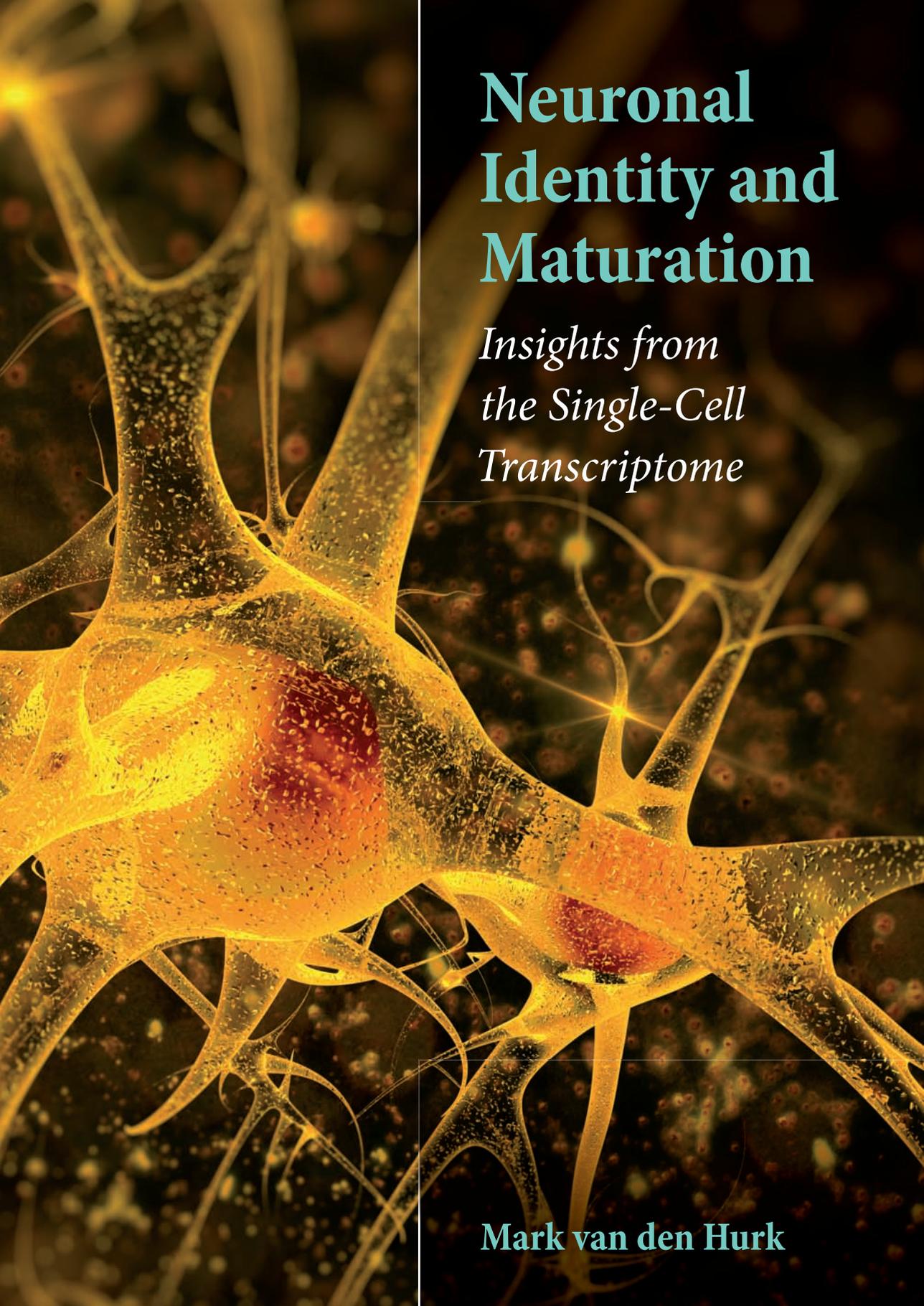
www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

A microscopic image of neurons, likely from a mouse brain, showing their complex branching structure. The neurons are stained with a yellow/gold dye, and several cell bodies and processes are highlighted with a red/orange glow. The background is dark with some out-of-focus light spots.

Neuronal Identity and Maturation

*Insights from
the Single-Cell
Transcriptome*

Mark van den Hurk

**Neuronal Identity and Maturation:
Insights from the Single-Cell Transcriptome**

Mark van den Hurk

© 2016 by Mark van den Hurk, Maastricht, the Netherlands

Neuronal Identity and Maturation: Insights from the Single-Cell Transcriptome.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, without permission in writing from the copyright holder.

ISBN 978-94-6169-989-3

Layout and printing: Optima Grafische Communicatie, Rotterdam, the Netherlands

**Neuronal Identity and Maturation:
Insights from the Single-Cell Transcriptome**

DISSERTATION

To obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus, Prof. Dr. R.M. Letschert,
in accordance with the decision of the Board of Deans,
to be defended in public
on Tuesday December 13th, 2016 at 14:00 hours

by

Mark van den Hurk

Supervisors

Prof. Dr. H.W.M. Steinbusch

Prof. Dr. B.P.F. Rutten

Co-supervisor

Dr. G. Kenis

Dr. C. Bardy (Salk Institute for Biological Studies, United States)

Assessment Committee

Prof. Dr. J.C.S. Kleinjans (Chairman)

Prof. Dr. C.T.A. Evelo

Prof. Dr. K-P. Lesch (University of Würzburg, Germany)

Dr. L. Nguyen (University of Liège, Belgium)

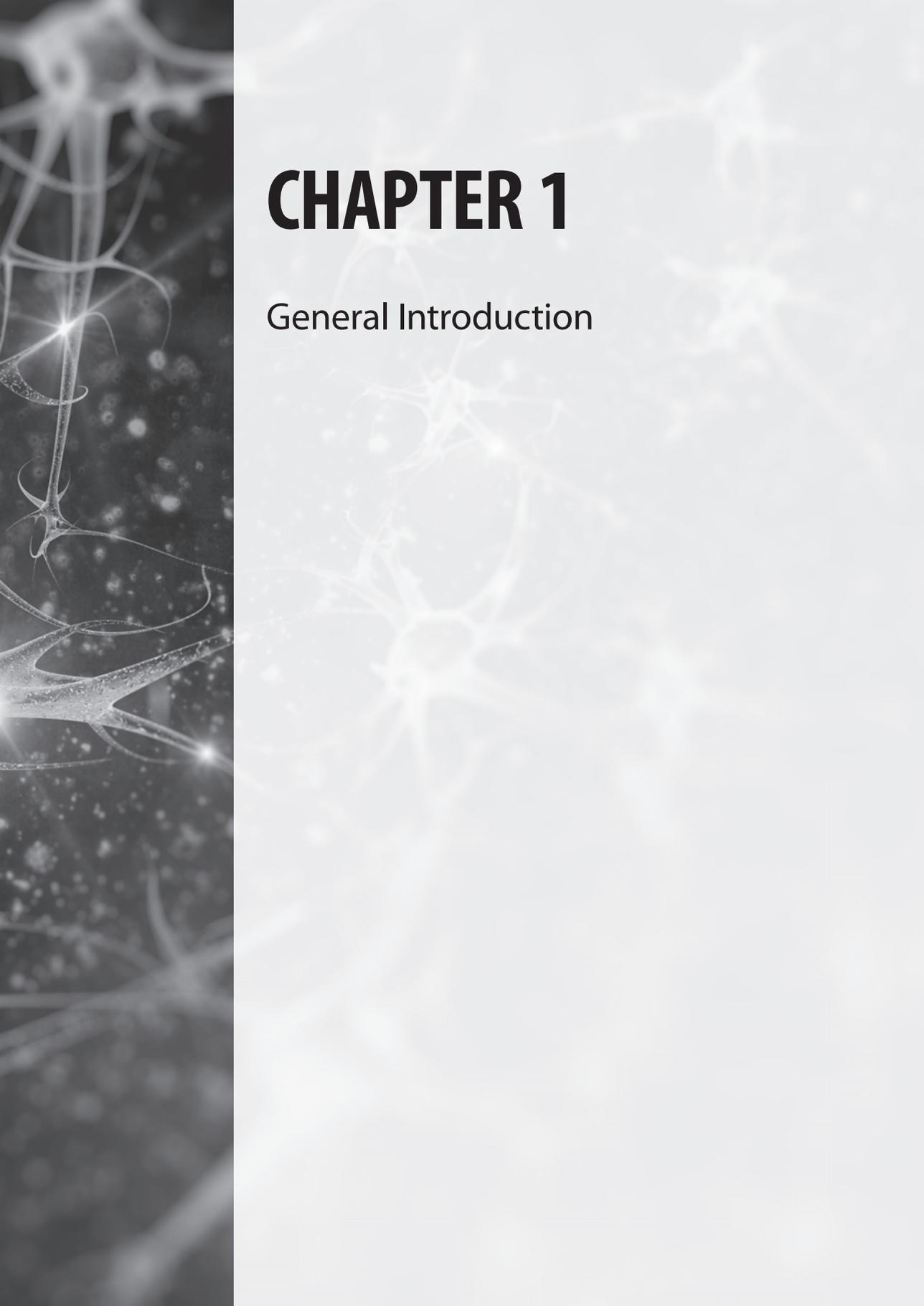
Prof. Dr. H.J.M. Smeets

“I AM THE MASTER OF MY FATE. I AM THE CAPTAIN OF MY SOUL.”
- WILLIAM ERNEST HENLEY

TABLE OF CONTENTS

CHAPTER 1	General Introduction	9
CHAPTER 2	Transcriptional and epigenetic mechanisms of cellular reprogramming to induced pluripotency	19
CHAPTER 3A	Dulbecco's Modified Eagle's Medium impairs spontaneous calcium activity in a human stem cell-derived <i>in vitro</i> model of neuronal networks	51
CHAPTER 3B	Neuronal medium that supports basic synaptic functions and activity of human neurons <i>in vitro</i>	75
CHAPTER 4A	Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology	123
CHAPTER 4B	Establishment of a pre-sequencing quality control pipeline for single-neuron Patch-seq experiments	179
CHAPTER 5	Identification of a co-expression gene network and transcriptional regulators for functional maturity based on electrophysiological, morphological and transcriptomic profiling of iPSC-derived single human neurons	193
CHAPTER 6	General Discussion and Conclusion	221
	Summary	235
	Valorization	241
	Acknowledgments	251
	Curriculum Vitae	257
	List of Publications	263
	Thesis Defenses from MHeNS	267





CHAPTER 1

General Introduction

Innovative research on the cellular and molecular mechanisms underlying neurological and psychiatric disorders is strongly needed to discover and validate new therapeutic targets and treatment options. Although human trials would be the most accurate and desired approach, studying the precise physiologic dysfunction of a neurologic or psychiatric disorder in the brains of living patients is both technically and ethically restricted. Whereas neuroimaging represents an attractive method to study disease-related alterations in brain structure, connectivity and functional activity in specific neuronal areas, it cannot resolve which specific neuronal subtypes are principally affected in the disease [1]. In addition, long-term pharmacological treatments may induce structural and/or functional brain alterations, and could therefore potentially confound neuroimaging studies of patients [1]. Whereas much of our current knowledge on the biological mechanisms of brain disorders has been derived from *in vivo* studies with experimental animals, it has repeatedly been proven difficult to translate promising findings in animal models into new effective treatments for humans due to factors such as species differences and human brain complexity [2]. Now, with the advent of human cell reprogramming [3-6] and direct cell fate conversion technologies [7-9], unlimited quantities of live human neurons can be generated from a variety of readily accessible somatic cell sources, such as skin fibroblasts. Representing a promising and practical intermediate between animal studies and human clinical trials, such cultures of human neurons reprogrammed from patient's somatic tissue offer great opportunities for regenerative medicine, and for studying the complex aetiology and pathophysiology of neurological and psychiatric disorders at a cellular and molecular biological level (Figure 1).

NEUROPSYCHIATRIC DISEASE MODELING WITH PATIENT-DERIVED NEURONS AND CELLULAR HETEROGENEITY

Cellular reprogramming for the purpose of disease modeling and drug screening entails the conversion of patient-derived differentiated cells into so-called induced pluripotent stem cells (iPSCs), a process that is achieved through enforced ectopic expression of essential pluripotency transcription factors (most commonly Oct4, Sox2, Klf4 and c-Myc) [3, 10]. The remarkable developmental capacity of iPSCs enables their *in vitro* differentiation into a variety of cell types, including those from the nervous system, and the resulting neuronal or glial cells represent valuable models for studying the basic principles of nervous system function and disease. The idea behind disease modeling is to recapitulate and study at a molecular and cellular level various features of neurological and psychiatric disorders, and efficiently test and validate potential new treatments “in a laboratory dish” [11]. In recent years, various features of several genetic and sporadic disorders of the nervous system, including Alzheimer's disease [12, 13] and Parkinson's

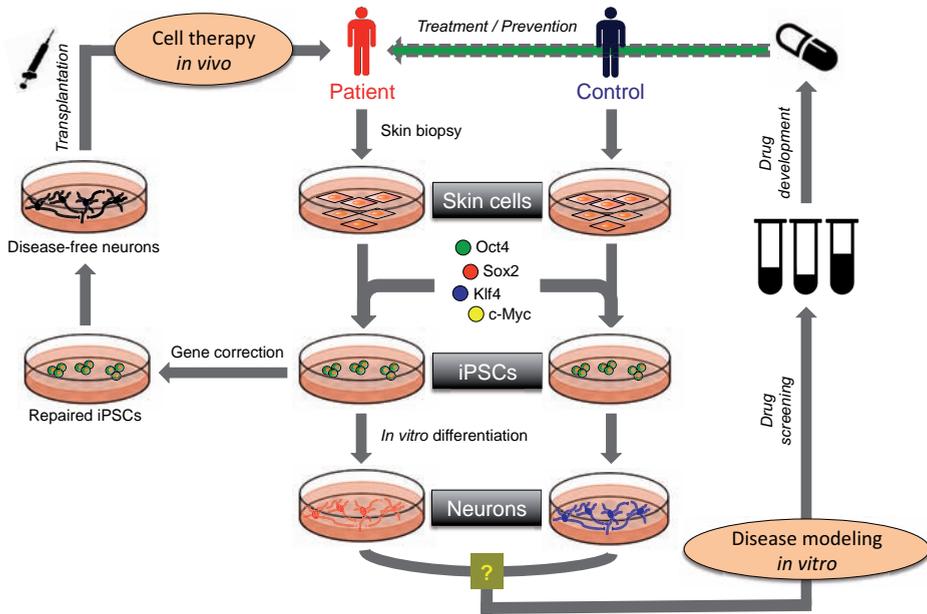


Figure 1. The promise of induced pluripotent stem cell technology in research and regenerative medicine. Disease-specific induced pluripotent stem cells (iPSCs) can be generated through reprogramming of patient-derived somatic cells (most commonly skin fibroblasts). The remarkable developmental potential of iPSCs enables their *in vitro* differentiation into a variety of cell types, such as neurons, for the study of various cellular and molecular aspects of disease (i.e., disease modeling), and the testing of potential therapeutics. Cell lines derived from healthy individuals are commonly included as controls. Since the disease-specific iPSCs possess the genetic background of the patient, attempts can also be made to correct the genetic defects by genome editing technology in order to generate disease-free cells for use in cell transplantation therapies. However, the use of iPSCs for generation of transplantable tissue is not free of risks and its applicability in clinical setting warrants further study.

disease [14-18] as well as complex psychiatric disorders such as schizophrenia [19-21], have successfully been modeled and studied using human patient-derived iPSCs.

While unlimited access to tissue with the genetic background of patients opens unprecedented avenues for medical discoveries, great care should be taken in employing stem cell-based approaches for the modeling and study of human neuropsychiatric illnesses. One major challenge to accurate *in vitro* disease modeling is the large degree of variability inherent to stem cell culture and differentiation paradigms. Subtle differences in derivation protocols and micro-environmental culturing conditions can generate observable variation among pluripotent cell lines [22, 23]. Another source of variation in stem cell-derived neuronal models is the generation of heterogeneous cultures comprised of different types of neurons and varying states of functional maturity [24]. How exactly neurons induced from stem cells change and mature over time *in vitro* is still unclear

and remains a topic of intense research. Yet, if one wishes to compare neuronal cultures in the context of a disease modeling experiment, one should (i) ensure that the individual neurons in these cultures differentiate and mature at similar rates, resulting in the production of neuronal populations of equivalent functional maturity, or (ii) isolate and enrich functionally comparable types of neurons from the heterogeneous neuronal cultures. A failure to thoroughly characterize individual pluripotent stem cell lines and their differentiated derivatives may obscure the identification of disease-related phenotypes in a disease modeling study, or worse, may lead to false findings and conclusions [24].

UNDERSTANDING CELLULAR HETEROGENEITY WITH SINGLE-CELL ANALYSIS

Cellular heterogeneity is a common feature of any cell population, and even seemingly identical types of cells exhibit a high degree of variability in behavior and function. While whole-cell patch-clamp electrophysiology is the prevailing technique used to assess in detail the functionality of individual neurons, understanding how these cells function differently at the molecular level is a fundamental goal of biology key to discovering the basic mechanisms of brain function. While essentially every cell of a multicellular organism has an identical set of genes (i.e., the genome), it is the variation in gene expression (i.e., the transcriptome) between individual cells that contributes to their phenotypic, behavioral and functional differences. Recent advances in next-generation sequencing technologies have enabled the identification and quantification of complete transcriptomes in virtually any population of cells or tissue. In the last couple of years, transcriptome profiling of single cells by RNA-sequencing (scRNA-seq) has emerged as a new frontier in cell biology that disregards previous assumptions on cell phenotype and captures in molecular code a remarkably detailed, unbiased picture of cellular state and function [25].

With the rapidly evolving fields of single-cell transcriptomics, bioinformatics and big data analytics, it should be possible in the near future to profile the transcriptome of almost all cell types of an organism. An accurate and unbiased classification of cellular types and states will aid in the understanding of cellular functions and behaviors in adult as well as in developing and pathological tissues. Single-cell analysis has the power to identify and characterize different subtypes of cells, something that is not possible with bulk analysis of cells where *en masse* pooling obscures subpopulations and results in an average gene expression profile [25-27]. Neuropsychiatric disease-specific abnormalities may only be manifested in a particular subtype of neurons, or at a specific maturational stage, and single-cell expression profiling provides the resolution required for identifying rare or clinically important cell types [28, 29].

AIM AND OUTLINE OF THIS THESIS

Although neuronal identity has traditionally been defined as a combination of morphological and neurophysiological traits, investigations are only beginning to explore how cell-molecular profiles determine such features [30]. In this thesis, we stratified the functional heterogeneity of human stem cell-derived neurons *in vitro* by combined single-neuron measurements of electrophysiological activity, morphology and transcriptomic profiles.

To acquire a better understanding of the molecular regulation of cell identity and the generation of iPSCs from somatic cells, I first reviewed recent research on the transcriptional and epigenetic mechanisms of pluripotency induction in somatic cells (CHAPTER 2). In the lab, we used iPSCs derived from fibroblasts of four healthy donors, as well as human embryonic stem cell (ESC) lines, to establish midbrain-like human neuronal circuits *in vitro*. Characterization of neural activity in these circuits using several electrophysiological techniques that included live calcium imaging, patch-clamp recording and multi-electrode array measurements revealed that classically used cell culture media (i.e., DMEM/F12 basal and Neurobasal), as well as serum, profoundly impair essential neurophysiological functions (CHAPTER 3A and CHAPTER 3B). To overcome this problem, we developed a serum-free neuronal basal medium (BrainPhys™) that, when combined with the appropriate supplements, (i) adequately sustains neuronal functions and survival *in vitro*, (ii) more closely mimics the *in vivo* brain microenvironment, and (iii) is also applicable to the culturing of rodent primary neurons (CHAPTER 3B). The improvements made in this new neuromedium constitute an important step toward improving *in vitro* human neuronal models and may ultimately result in a more successful translation of findings from the laboratory to the clinic.

The rate at which human iPSC/ESC-derived neural progenitors develop into electrophysiologically mature neurons is highly heterogeneous and poses another significant challenge to *in vitro* studies of neurological disorders. By means of single-neuron multidimensional analysis of electrophysiological function, morphology and transcriptome (a method nowadays referred to as “Patch-seq” [31]), we profiled the *in vitro* functional heterogeneity of a large sample of human neurons (CHAPTER 4A). Specifically, we defined a neurodevelopmental continuum of five electrophysiological types of neurons and correlated their functional properties with morphological features and molecular profiles. Using machine learning, we identified new molecular determinants predictive of the neurophysiological state of the neurons. Importantly, based on the molecular information, we were able to isolate and effectively enrich highly functional, live human neurons from heterogeneous cultures using fluorescent markers. The ability of such an approach to compensate for the wide heterogeneity of functional neuronal types among

cell lines constitutes an important step toward strengthening comparative analyses in stem cell-based models of neurological disease [24].

Sample quality control (QC) is an important part of any expression profiling experiment, and the pre-sequencing QC pipeline that we established to identify and filter out potentially spurious single-neuron captures in advance of deep sequencing is explained in detail in CHAPTER 4B. The methodology described in this chapter is likely to reduce the risk of erroneous inclusion of low-quality or compromised samples in the sequencing and downstream bioinformatics analysis pipelines. As it is rather difficult to discern technical artifacts from intrinsic biological variability in scRNA-seq data, such an improved methodological QC may increase the power with which altered gene expression is detected between experimental samples in single-cell transcriptome analyses.

To obtain further mechanistic insight into the biology underlying the different electrophysiological neuron types, we conducted additional systems-level analysis of the single-cell transcriptome data by Weighted Gene Co-expression Network Analysis (WGCNA), and identified a cluster of highly co-expressed genes, mostly related to adenosine triphosphate (ATP) metabolism and mitochondrial energetics, that strongly correlated with the neurophysiological features of highly functional neurons (CHAPTER 5). Gene Set Enrichment Analysis (GSEA) suggested that a substantial portion of these genes play a general role in neuronal maturity, confirming the notion that our functional neuronal types classification reflects different stages of neurophysiological maturation. In addition, we identified a set of 21 putative transcription factors that may regulate, or drive, the highly functional (mature) neuronal state.

Altogether, the studies conducted as part of this thesis open up new avenues for improving and homogenizing models of neurological disorders *in vitro*. The main findings and implications of the research are summarized and discussed in the concluding chapter of this thesis (CHAPTER 6).

REFERENCES

1. Brennand KJ, Gage FH. Concise review: the promise of human induced pluripotent stem cell-based studies of schizophrenia. *Stem Cells* 29(12), 1915-1922 (2011).
2. Dragunow M. The adult human brain in preclinical drug development. *Nat. Rev. Drug Discov.* 7(8), 659-666 (2008).
3. 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).
4. Yu J, Vodyanik MA, Smuga-Otto K *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318(5858), 1917-1920 (2007).
5. Park IH, Zhao R, West JA *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 451(7175), 141-146 (2008).
6. Lowry WE, Richter L, Yachechko R *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.* 105(8), 2883-2888 (2008).
7. Pfisterer U, Kirkeby A, Torper O *et al.* Direct conversion of human fibroblasts to dopaminergic neurons. *Proc. Natl. Acad. Sci. U. S. A.* 108(25), 10343-10348 (2011).
8. Caiazzo M, Dell'anno MT, Dvoretzkova E *et al.* Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature* 476(7359), 224-227 (2011).
9. Son EY, Ichida JK, Wainger BJ *et al.* Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell* 9(3), 205-218 (2011).
10. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126(4), 663-676 (2006).
11. Dolmetsch R, Geschwind DH. The human brain in a dish: the promise of iPSC-derived neurons. *Cell* 145(6), 831-834 (2011).
12. 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).
13. Yagi T, Ito D, Okada Y *et al.* Modeling familial Alzheimer's disease with induced pluripotent stem cells. *Hum. Mol. Genet.* 20(23), 4530-4539 (2011).
14. 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).
15. Imaizumi Y, Okada Y, Akamatsu W *et al.* Mitochondrial dysfunction associated with increased oxidative stress and alpha-synuclein accumulation in PARK2 iPSC-derived neurons and postmortem brain tissue. *Mol. Brain* 5 35 (2012).
16. Byers B, Lee HL, Reijo Pera R. Modeling Parkinson's disease using induced pluripotent stem cells. *Curr. Neurol. Neurosci. Rep.* 12(3), 237-242 (2012).
17. Byers B, Cord B, Nguyen HN *et al.* SNCA triplication Parkinson's patient's iPSC-derived DA neurons accumulate alpha-synuclein and are susceptible to oxidative stress. *PLoS One* 6(11), e26159 (2011).
18. Nguyen HN, Byers B, Cord B *et al.* LRRK2 mutant iPSC-derived DA neurons demonstrate increased susceptibility to oxidative stress. *Cell Stem Cell* 8(3), 267-280 (2011).
19. Brennand KJ, Simone A, Jou J *et al.* Modelling schizophrenia using human induced pluripotent stem cells. *Nature* 473(7346), 221-225 (2011).
20. Pedrosa E, Sandler V, Shah A *et al.* Development of patient-specific neurons in schizophrenia using induced pluripotent stem cells. *J. Neurogenet.* 25(3), 88-103 (2011).
21. Paulsen Bda S, De Moraes Maciel R, Galina A *et al.* Altered oxygen metabolism associated to neurogenesis of induced pluripotent stem cells derived from a schizophrenic patient. *Cell Transplant.* 21(7), 1547-1559 (2012).

22. Tomoda K, Takahashi K, Leung K *et al.* Derivation conditions impact X-inactivation status in female human induced pluripotent stem cells. *Cell Stem Cell* 11(1), 91-99 (2012).
23. Newman AM, Cooper JB. Lab-specific gene expression signatures in pluripotent stem cells. *Cell Stem Cell* 7(2), 258-262 (2010).
24. Sandoe J, Eggan K. Opportunities and challenges of pluripotent stem cell neurodegenerative disease models. *Nat. Neurosci.* 16(7), 780-789 (2013).
25. Sandberg R. Entering the era of single-cell transcriptomics in biology and medicine. *Nat. Methods* 11(1), 22-24 (2014).
26. Patel AP, Tirosh I, Trombetta JJ *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344(6190), 1396-1401 (2014).
27. Toriello NM, Douglas ES, Thaitrong N *et al.* Integrated microfluidic bioprocessor for single-cell gene expression analysis. *Proc. Natl. Acad. Sci. U. S. A.* 105(51), 20173-20178 (2008).
28. Eberwine J, Sul JY, Bartfai T, Kim J. The promise of single-cell sequencing. *Nat. Methods* 11(1), 25-27 (2014).
29. Stahlberg A, Kubista M, Aman P. Single-cell gene-expression profiling and its potential diagnostic applications. *Expert Rev. Mol. Diagn.* 11(7), 735-740 (2011).
30. Fuzik J, Zeisel A, Mate Z *et al.* Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nat. Biotechnol.* 34(2), 175-183 (2016).
31. Cadwell CR, Palasantza A, Jiang X *et al.* Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. *Nat. Biotechnol.* 34(2), 199-203 (2016).



CHAPTER 2

Transcriptional and epigenetic mechanisms of cellular reprogramming to induced pluripotency

Mark van den Hurk^{1,2,3}, Gunter Kenis^{1,2,*}, Cedric Bardy^{3,*}, Daniel L. van den Hove^{1,2,4}, Fred H. Gage³, Harry W.M. Steinbusch^{1,2}, Bart P.F. Rutten^{1,2}

¹ Department of Psychiatry and Neuropsychology, Division of Translational Neuroscience, Maastricht University, Maastricht, 6200 MD, The Netherlands

² European Graduate School of Neuroscience (EURON), Maastricht University, Maastricht, 6200 MD, The Netherlands

³ Laboratory of Genetics, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

⁴ Laboratory of Translational Neuroscience, Department of Psychiatry, Psychosomatics and Psychotherapy, University of Wuerzburg, 97080 Wuerzburg, Germany

* These authors contributed equally

Epigenomics 2016; 8(8):1131-49

doi: 10.2217/epi-2016-0032

ABSTRACT

Enforced ectopic expression of a cocktail of pluripotency-associated genes such as *Oct4*, *Sox2*, *Klf4* and *c-Myc* can reprogram somatic cells into induced pluripotent stem cells (iPSCs). The remarkable proliferation ability of iPSCs and their aptitude to re-differentiate into any cell lineage makes these cells a promising tool for generating a variety of human tissue *in vitro*. Yet, pluripotency induction is an inefficient process, as cells undergoing reprogramming need to overcome developmentally imposed epigenetic barriers. Recent work has shed new light on the molecular mechanisms that drive the reprogramming of somatic cells to iPSCs. Here, we present current knowledge on the transcriptional and epigenetic regulation of pluripotency induction and discuss how variability in epigenetic states impacts iPSCs' inherent biological properties.

Keywords: pluripotency, induced pluripotent stem cell (iPSC), reprogramming, epigenetics, chromatin, DNA methylation, histone modifications

INTRODUCTION

The transition from stem cells to mature, functionally specialized cells has long been considered a finite phenomenon. Evidence for the reversibility of cell fate commitment came from observations that an adult somatic cell could be returned to an immature state upon transplantation of its nucleus to an enucleated oocyte (somatic cell nuclear transfer – SCNT) [1-3] or by fusion with an embryonic stem cell (ESC) [4, 5]. In 2006, innovative work by Yamanaka and colleagues showed that forced ectopic expression of the pluripotency-associated transcription factors Oct4, Sox2, Klf4 and c-Myc (collectively termed OSKM) could reprogram differentiated cells, such as skin fibroblasts, to an ESC-like state, yielding induced pluripotent stem cells (iPSCs) [6]. The technology of induced pluripotency has since received considerable interest in the scientific community because it holds great promise for *in vitro* human disease modeling, drug screening, and cell-based therapies in regenerative medicine.

iPSCs are very similar to ESCs in their morphology, proliferative capacity and developmental potential, and both pluripotent cell types exhibit highly comparable global gene expression and epigenetic profiles [7-10]. The chromatin landscapes and DNA methylomes of iPSCs are, however, markedly different from those of lineage-committed cells, indicating that reprogramming requires widespread remodeling of somatic cell epigenetic profiles to reflect the ESC epigenome [11-14]. In generating iPSCs from somatic cells, epigenome-modifying molecules may be used to improve the efficiency of the overall process and even limit the number of exogenous pluripotency-inducing factors [15-18], confirming the general notion that epigenetics plays a key role in the regulation of cell fate specification.

Somatic cell reprogramming is a relatively slow and inefficient process, with only a minority of transduced somatic cells becoming fully reprogrammed to *bona fide* iPSCs after several weeks [19-21]. Observations that stem and progenitor cells reprogram with higher efficiency and kinetics than terminally differentiated cells [22-24] suggest that epigenetic barriers established during embryonic differentiation hinder efficient reprogramming to the pluripotent state (for excellent reviews, see refs. [25-27]). Somatic cell types that are developmentally closer to ESCs supposedly require less epigenetic remodeling, potentially facilitating their reprogramming into iPSCs.

Despite major advances in the methods for deriving and culturing iPSCs, the precise molecular mechanisms that drive cells to overcome developmentally imposed epigenetic barriers are only beginning to be elucidated. Most of our current information about the transcriptional and epigenetic events regulating pluripotency and reprogramming has come from studies using murine cells. Yet, strong cross-species conservation of fundamental genetic and epigenetic mechanisms controlling stem cell self-renewal and differentiation has enabled the translation of numerous experimental procedures

and insights from mouse to human (**Box 1**). In this review, we summarize the current knowledge of the transcriptional and epigenetic regulation of pluripotency induction, and discuss the sources and functional biological consequences of epigenetic variability in iPSCs. Though this review mainly focuses on murine somatic cell reprogramming, a greater understanding of the molecular events governing pluripotency induction in mouse provides important insights to improve human cell reprogramming methods and guide safe and large-scale iPSC production for therapeutic use in human [28].

Box 1. Conservation and divergence in human and murine (induced) pluripotency.

Mammalian pluripotency is conferred by a unique and highly conserved network of pluripotency transcription factors, of which Oct4, Sox2 and Nanog constitute key regulators [29-31]. Comparisons of mouse and human ESCs have, however, revealed important interspecies differences in the target genes controlled by these pluripotency regulators [30] and specific molecular signaling pathways activated [32]. For instance, while mouse ESCs require leukemia inhibitor factor (LIF)-Stat3 signaling for self-renewal and maintenance of pluripotency, human ESCs are insensitive to LIF and show elevated expression of SOCS-1, an inhibitor of STAT3 signaling [32, 33]. Despite these differences, and differences in cell culture requirements, expression of cell-surface antigens (mouse: SSEA1; human: SSEA3, SSEA4, TRA-1-60 and TRA-1-81 [34]) and developmental potential (e.g., the inability of mouse ESCs to differentiate to trophoblasts [35]), there is also a substantial overlap in gene expression and pathway activation between both species [32]. The high evolutionary conservation of core pluripotency transcriptional and epigenetic mechanisms has thus enabled many insights from studies conducted in mice to be translated to the human situation. Ectopic expression of the same set of pluripotency-associated transcription factors (Oct4, Sox2, Klf4 and c-Myc), for example, induces pluripotency in somatic cells of mouse and human origin [6, 36-38]. Likewise, a highly conserved microRNA cluster (miR-302/367) can efficiently reprogram mouse and human somatic cells to iPSCs, even in the complete absence of exogenous pluripotent factors [39]. The miR-302/367 cluster is specifically expressed in human and mouse ESCs [40], and has been identified as a direct target of the Oct4 and Sox2 pluripotency transcription factors [41], thus providing evidence for a conserved function of this specific miRNA cluster in the regulation and maintenance of the undifferentiated stem cell state. All in all, we can conclude that core members of the pluripotency regulatory network appear to be well conserved between mice and humans, enabling us to use the murine system to study human cell reprogramming mechanisms. Their downstream targets and signaling pathways, however, seem more evolutionarily divergent, which suggests the need for caution in using mouse pluripotent stem cells as a model for the differentiation of human pluripotent stem cells in the context of therapeutic applications [42].

MOLECULAR AND TRANSCRIPTIONAL ROUTES TO INDUCED PLURIPOTENCY

Reprogramming of somatic cells into iPSCs requires silencing of the somatic cell transcriptional profile and the induction of an ESC-like expression program. This change in global gene expression has been demonstrated to occur gradually and through a defined sequence of molecular events (**Figure 1**). Upon reprogramming factor expression, mouse fibroblasts initially suppress somatic cell-specific genes such as the fibroblast marker Thy1 before upregulating the early embryonic markers alkaline phosphatase (AP) and stage-specific embryonic antigen 1 (SSEA1), and ultimately activating core pluripotency factors

such as Nanog, Sall4, Oct4 and Sox2 [43-45]. Fluorescence-activated cell sorting (FACS) on the basis of those markers enables the isolation of distinct intermediate populations of iPSC generation from the highly heterogeneous reprogramming culture, thereby facilitating the study of the molecular mechanisms governing successful pluripotency induction [43-45]. In this way, Polo and colleagues showed that, compared to SSEA1+ cells, which over time progressed toward a *bona fide* pluripotent state (Oct4-GFP), cells that failed to downregulate Thy1 shortly after OSKM induction became refractory to the reprogramming process [45]. Gene expression profiling by microarray analysis further demonstrated that refractory Thy1+ cells, in contrast to progressing SSEA1+ intermediates, did not undergo proper mesenchymal-to-epithelial transition (MET) [45], a developmental process by which cells acquire epithelial characteristics and lose mesenchymal features. MET represents a key initiating event during iPSC derivation that depends on

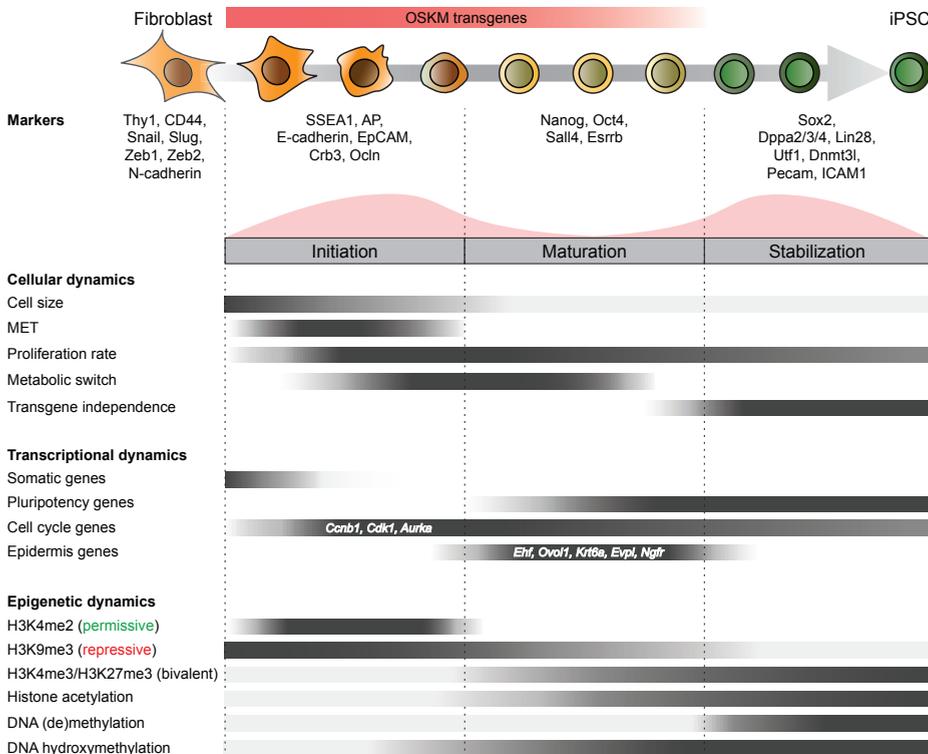


Figure 1. Cellular, transcriptional and epigenetic changes during mouse somatic cell reprogramming to induced pluripotency. Markers upregulated before and during the different phases of reprogramming are shown at the top. The two major waves of gene activity that occur during pluripotency induction are depicted in pink. Gray intensities of the bars denote the approximate magnitude of the indicated process over the time course of reprogramming.

AP: alkaline phosphatase; iPSC: induced pluripotent stem cell; MET: mesenchymal-to-epithelial transition; OSKM: Oct4, Sox2, Klf4, c-Myc.

intrinsic bone morphogenetic protein (BMP) signaling, and coincides with increased proliferation and cell size reduction [9, 46-48] (**Figure 1**).

Following the establishment of MET during the initiation phase of mouse pluripotency induction, cells gradually transit into a maturation phase, which is characterized by the activation of a subset of pluripotency-associated genes, including *Nanog*, *Sall4*, *Esrrb*, and endogenous *Oct4* [47, 49]. Yet, induction of the complete pluripotency transcriptional network with expression of *Utf1*, *Lin28*, *Dnmt3l*, *Dppa2/3/4*, *Pecam*, and endogenous *Sox2*, does not occur until the later stabilization phase [47, 49]. Studies that used doxycycline (Dox)-inducible expression systems to delineate the temporal order of molecular events in mouse iPSC formation showed that the initiation phase was not self-sustainable and that transition to the intermediate maturation phase was dependent upon continuous expression of the OSKM transgenes [43, 44, 47]. However, sustained expression of reprogramming factors suppressed the upregulation of stabilization-phase pluripotency markers late in maturation, indicating that transgene silencing was required for cells to transition from maturation to stabilization and become fully pluripotent [49]. Through comparative gene expression profiling of clonally cultured cells that had gained competency to progress to the stabilization phase versus those that had not, Golipour *et al.* identified a distinctive transcriptomic signature associated with successful acquisition of the pluripotent state [49]. Notably, genes governing maturation-to-stabilization transition were distinct from those regulating pluripotency, and bioinformatics analyses suggested their involvement in signaling pathways, cell cycle, cytoskeletal reorganization, chromosome segregation, and chromosome stability. Hence, it has been suggested that late-phase progression to transgene-independent *bona fide* iPSCs and maintenance of the pluripotent state are under the control of distinct molecular regulatory pathways [49].

Paralleling the occurrence of the initiation, maturation and stabilization phases, during reprogramming, cells destined to become iPSCs showed two major waves of gene activation that were separated by an intermediary period of less profound transcriptional alterations [45] (**Figure 1**, *pink shade*). The first transcriptional wave was initiated in the majority of cells and occurred between days 0 to 3, reflecting a common early response to ectopic OSKM expression. The second wave was activated late in reprogramming (after day 9) and only in cell populations progressing to *bona fide* pluripotency. During the first wave, genes related to proliferation, metabolism and cytoskeletal organization showed activation, whereas developmental genes were downregulated [45]. This finding is consistent with observations that most murine fibroblasts repress the somatic cell transcriptional program immediately following reprogramming factor transduction, increase their proliferation rate, and reduce their size [9, 48]. Of the OSKM reprogramming factors, c-Myc has been implicated as a major inducer of these initial reprogramming events, in part because of its potentiating effects on cell cycle progression [45, 50]. c-Myc targets, including the cell cycle regulatory genes *Ccnb1*, *Cdk1* and *Aurka*, are

robustly upregulated within the first few days after OSKM induction and only modestly change expression until the end of reprogramming (day 12; see also **Figure 1**) [45]. The combined activity of Oct4 and Sox2 seems to result in the gradual upregulation of core pluripotency genes such as *Nanog* that hallmark the second transcriptional wave and reinforce the transcriptional program to establish a stable pluripotent state [45]. Klf4, the other reprogramming factor, presumably serves a binary supporting role in iPSC generation by repressing somatic genes and activating pluripotency genes during the first and second transcriptional waves, respectively [45].

A shortcoming of the Thy1/SSEA1/Oct4-GFP route map to study the molecular mechanisms of iPSC generation is that the SSEA1 surface antigen is a poor predictor of complete reprogramming. O'Malley and colleagues found that Nanog-eGFP expression emerged in both SSEA1+ and SSEA1- cell populations after eight days of Dox-induced reprogramming, and SSEA1 expression was highly heterogeneous in established iPSCs [51]. The authors identified a new murine fibroblast-to-iPSC reprogramming route map contingent on the loss of CD44 expression and a subsequent rapid upregulation of ICAM1, events that closely paralleled the gain of Nanog-eGFP. ICAM1 was expressed, though to variable extent, in fibroblasts and early reprogramming intermediates, but most cells became ICAM1- by day 6 of Dox-induced OSKM expression. Two days later, CD44 was downregulated in the vast majority of cells and, towards the end of reprogramming (day 12), the cells appeared enriched for an ESC-/iPSC-like phenotype (CD44-/ICAM1+/Nanog-eGFP+). Nanog-eGFP was, however, already expressed in some cells as early as day 6, prior to the disappearance of CD44 and re-upregulation of ICAM1. Yet, Nanog-eGFP+ cells had a higher probability of transitioning to fully reprogrammed iPSCs than Nanog-eGFP- cells [51], underscoring the important role of Nanog in driving pluripotency acquisition.

RNA-sequencing analysis of reprogramming intermediates purified by CD44/ICAM1/Nanog-eGFP FACS status subsequently indicated that pluripotency gene activation occurred in two transcriptional waves. Several pluripotency genes, including endogenous *Oct4* and *Sall4*, already showed marked upregulation very early in the reprogramming process [51]. Other pluripotency gene markers, such as endogenous *Sox2*, *Esrrb* and *Dppa2*, demonstrated more gradual activation at later reprogramming stages. This finding was corroborated by an earlier study that used single-cell expression profiling to characterize various stages and transitions on the road to pluripotency [52]. Notably, in that study, expression of endogenous *Oct4* occurred very early in the reprogramming process and could be detected in some partially reprogrammed colonies, suggesting that *Oct4* activation is a poor predictor of successful pluripotency induction. By contrast, endogenous *Sox2* expression appeared late, occurred in intermediate cells to a very low extent, and was predicted as an upstream regulator of activation of many genes involved in pluripotency [52].

In addition to the two waves of pluripotency gene expression, O'Malley and colleagues identified a group of epidermis-related genes that were transiently upregulated in intermediate reprogramming stages [51]. Interestingly, as determined by single-cell PCR, upregulation of these epidermis genes concurred with the first wave of early pluripotency gene activation, whereas their downregulation paralleled the activation of the late-stage pluripotency genes of the second wave. Given the occurrence of a similar transient epidermal gene expression pattern in three previously published datasets [45, 47, 50], transient ectodermal gene activation has been pointed out as a common hallmark of reprogramming [51]. Future research aimed at thoroughly investigating this phenomenon of temporary epidermal gene expression may provide important novel mechanistic insights into the molecular mechanisms of cellular reprogramming.

EPIGENETIC CONTROL OF CELLULAR REPROGRAMMING

The inefficiency and slow kinetics by which pluripotency is established in somatic cells after ectopically enforced OSKM expression suggests a critical role for additional regulators, including epigenetic modulators, in the induction and/or maintenance of the pluripotent state. Close interrelationships are known to exist between the transcriptional regulatory circuitry controlling pluripotency and modifiers of chromatin structure [53, 54]. The pluripotency factors Oct4, Sox2 and Nanog have been found to physically associate with, as well as regulate the expression of, various regulators of chromatin state [29, 55-57]. Furthermore, in iPSC generation from mouse and human somatic cells, exogenously applied epigenome-modifying molecules can be used to functionally replace or complement one or more of the OSKM reprogramming factors [15-18]. In this section, we discuss in detail the various epigenetic regulators and their associated marks involved in OSKM-induced iPSC generation.

Histone Modifications in Reprogramming

The global chromatin state of pluripotent and differentiated cells is markedly different. Unlike differentiated cells, which are characterized by highly condensed domains of heterochromatin and the accumulation of repressive histone modifications including methylation of histone H3 at lysine 9 (H3K9) and 27 (H3K27) [11, 58], pluripotent stem cells are distinguished by globally decompacted, euchromatic regions of chromatin associated with activating histone marks such as methylation on histone H3 lysine 4 (H3K4) and hyperacetylation of histones H3 and H4 (H3/H4ac) [59, 60]. In ESCs, architectural chromatin proteins such as heterochromatin protein 1 (HP1) bind loosely and in a hyperdynamic fashion to chromatin whereas immobilization of such proteins occurs upon lineage commitment [60]. Chromatin reorganization during differentiation serves

to silence lineage-inappropriate genes and thus constitutes an important mechanism by which cells establish and maintain a stable cell fate-specific gene expression program [61]. It follows that transcription factor-induced reprogramming requires global remodeling of the somatic cell chromatin back to a more dispersed conformation typical of ESCs, a process associated with substantial alterations in histone modification patterns and DNA methylation status [11-14, 62].

OSKM-Resistant Chromatin

Understanding at the molecular level exactly how the reprogramming factors change the somatic cell epigenome to reset cell identity represents an important goal for the reprogramming field. In iPSC generation, a major challenge for the reprogramming factors is to gain access to and reactivate the endogenous pluripotency genes buried in densely compacted heterochromatin. A study investigating global gene expression and epigenetic changes early in mouse somatic reprogramming identified acquisition of euchromatic H3K4 dimethylation (H3K4me₂) at numerous pluripotency-associated loci as an immediate primary response to ectopic reprogramming factor induction [13] (see also **Figure 1**). Notably, gain of H3K4 methylation at these sites did not lead to changes in gene transcription. In fact, genes that were upregulated early in reprogramming generally exhibited activating H3K4 trimethylation (H3K4me₃) at their promoters in fibroblasts (i.e., prior to reprogramming factor induction) [13], suggesting that somatic cell chromatin status may limit reprogramming factor action to readily accessible promoter DNA. Supporting this idea, Soufi and colleagues identified megabase-long chromatin domains in human fibroblasts, called “differentially bound regions” (DBRs), that blocked initial access of the OSKM reprogramming factors to the genome [63]. These OSKM refractory regions contained many genes necessary for establishing pluripotency, including *SOX2* and *NANOG*, and were specifically enriched for heterochromatic H3K9 trimethylation (H3K9me₃). Transient knockdown of the H3K9 methyltransferases *SUV39H1*, *SUV39H2* or *SETDB1* by small interfering RNA (siRNA) technology enabled Oct4 and Sox2 reprogramming factors to access the DBRs and resulted in more efficient and accelerated pluripotency induction [63]. Consistent with the notion that H3K9 methylation constitutes a critical epigenetic impediment to cellular reprogramming (**Figure 2A**), Chen and colleagues demonstrated that its specific removal could fully convert mouse partially reprogrammed cells to *bona fide* iPSCs expressing core pluripotency markers [64]. Specifically, knock-down of the H3K9 methyltransferase *Setdb1* or overexpression of the H3K9 demethylase *Kdm4b* promoted the further reprogramming of pre-iPSC intermediates into Oct4-GFP-expressing colonies under reprogramming-inhibiting culture conditions. The authors also found that vitamin C-mediated enhancement of iPSC induction was profoundly inhibited upon treatment with siRNAs targeting various members of the Kdm family of H3K9 demethylases [64]. These data indicate that specific constituents of the cell culture

environment, including the small-molecule supplement vitamin C, can affect chromatin structure by regulating H3K9 methylation levels and thereby influence the efficiency and kinetics of somatic cell reprogramming.

Immunofluorescence analysis of mouse embryonic fibroblasts that underwent reprogramming revealed that the redistribution of the heterochromatin-associated histone mark H3K9me3 and its binding protein HP1, from confined foci with well-defined boundaries to a more diffuse organization with more prominent nucleoplasmic staining, is a rather early event during iPSC generation [14]. This global change in heterochromatin organization could already be observed in small colonies of OSKM-transduced cells by day 6 of the reprogramming process, prior to the expression of *Nanog* and the establishment of euchromatic chromatin features such as histone acetylation and H3K4 trimethylation (**Figure 1**). Euchromatic histone marking paralleled the expression of *Nanog*; hence establishment of an open chromatin configuration typical of ESCs does not seem to occur until relatively late in reprogramming [14]. Supporting this notion, conversion of mouse partial iPSCs into full iPSCs by dual chemical inhibition of MEK and GSK3 signaling (known as “2i” culture [65]) resulted in the decondensation of highly compacted heterochromatin into dispersed, open 10-nm fibers concurrently with transgene silencing and pluripotent gene activation [62], both of which are late events in cellular reprogramming.

Role of Polycomb and Trithorax Group Proteins

The Polycomb group (PcG) and Trithorax group (trxG) protein complexes, which catalyze methylation of different lysines of histone H3, play an important role in controlling pluripotency-related gene expression. Whereas PcG proteins assist in chromatin compaction and thus gene silencing by mediating repressive trimethylation of H3 at lysine 27 (H3K27me3), trxG proteins favor a transcriptionally permissive chromatin structure through establishment of H3K4 trimethylation (H3K4me3) marks at gene promoters [66].

In ESCs, genes required for pluripotency and self-renewal, such as *Oct4*, *Sox2* and *Nanog*, are actively transcribed, and their promoters are typically associated with activating H3K4me3 histone marks [67-70]. WD repeat domain 5 (*Wdr5*), a key component of the trxG protein complex in mammals, co-occupies numerous pluripotency genes together with *Oct4* in mouse ESCs to positively regulate their expression [71]. Reduction in H3K4me3 levels in these ESCs as a result of short-hairpin RNA (shRNA)-induced knockdown of *Wdr5* causes downregulation of pluripotency-associated genes including *Oct4*, *Sox2*, *Klf4*, *Nanog* and *Esrrb*, and induced differentiation [71]. Interestingly, *Wdr5* expression is upregulated during iPSC generation and loss of *Wdr5* function markedly reduces reprogramming efficiency [71]. These findings highlight the importance of H3K4me3 marking in controlling pluripotent cell identity and suggest a requirement for

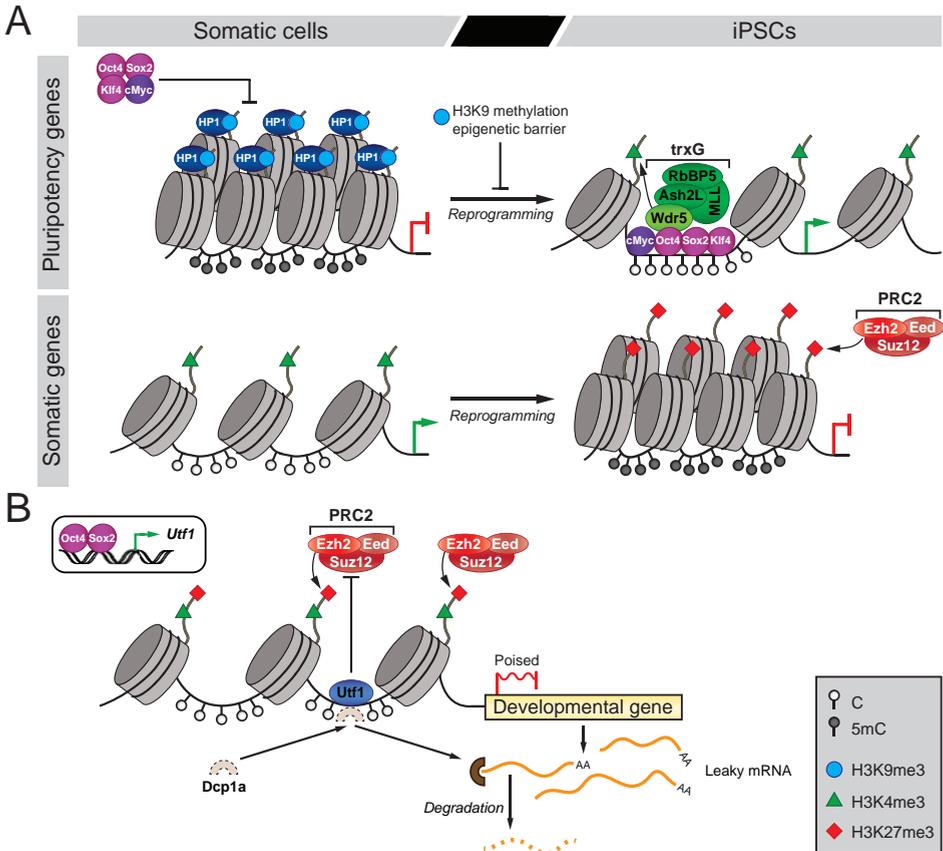


Figure 2. Roles of histone-modifying proteins and DNA methylation in pluripotency and reprogramming. (A) In somatic cells, pluripotency gene promoters are methylated (5mC) and enriched for H3K9me₃, a heterochromatin-associated histone mark that blocks access of the OSKM factors to the DNA, hindering reprogramming. Reprogramming of somatic cells to iPSCs is associated with DNA demethylation and acquisition of H3K4me₃ at promoters of pluripotency genes. The establishment of this permissive H3K4me₃ chromatin state is brought about by Trithorax group (trxG) proteins, whose core component Wdr5 was found in ESCs to be directed to pluripotency loci by Oct4. The occupancy of Oct4 (and other pluripotency factors) with trxG drives strong activation of the pluripotency transcriptional network. In apparent contrast, somatic genes become silenced during reprogramming due to *de novo* promoter DNA methylation and the action of Polycomb repressive complex 2 (PRC2) that imposes a transcriptionally silent state of chromatin characterized by H3K27me₃ modification. (B) Bivalent H3K4me₃/H3K27me₃ chromatin domains at promoter regions of developmental genes constitute a hallmark of ESC and iPSC pluripotency. The bivalent chromatin-associated protein Utf1, a target of Oct4 and Sox2, enforces a “poised” state of gene expression in ESCs by (1) preventing excessive PRC2 binding and H3K27me₃, and (2) facilitating the loading of decapping protein 1a (Dcp1a) to mRNAs generated from leaky transcription for cytoplasmic degradation. These contrasting functions of Utf1 prevent both the oversilencing and insufficient repression of bivalent genes, thereby properly coordinating developmental processes.

5mC: 5-methylcytosine; C: (non-methylated) cytosine; HP1: heterochromatin protein 1.

trxG activity by the OSKM reprogramming factors to reset the epigenome during iPSC generation (**Figure 2A**).

In addition to the trxG proteins, PcG protein complexes play an indispensable role in specifying and maintaining pluripotent cell identity, as no ESC lines can be established from mouse blastocysts lacking the Polycomb repressive complex 2 (PRC2) subunit *Ezh2* [72], and ESCs deficient in the PRC2 component *Eed* differentiate spontaneously [73]. PcG proteins exert their preventive effect on ESC differentiation by H3K27me₃-mediated silencing of key developmental regulators [73, 74]. Accordingly, ESCs lacking the PRC2 component *Suz12* have globally reduced H3K27me₃ levels and show increased expression of differentiation-specific markers [75]. PRC2 complex function is critical for transcription factor-induced reprogramming, as functional ablation of any of the three core PRC2 components, *Eed*, *Ezh2* or *Suz12*, impairs iPSC generation significantly [52, 76, 77]. Overexpression of *Ezh2*, by contrast, enhances efficiency of iPSC formation, as determined by an increased number of Nanog-GFP⁺ cells relative to control after 7 days of Dox induction of the OSKM reprogramming factors [52].

Many developmental genes in ESCs are enriched for both H3K27me₃ and H3K4me₃ modifications in their promoter regions. This concurrence of both repressive and activating histone marks at lineage-specific gene promoters establishes a bivalent (“poised”) state of chromatin activation in response to differentiation-inducing signals [69, 78]. Upon lineage commitment, most bivalent chromatin domains become resolved to the monovalent state of either H3K4me₃ or H3K27me₃. Genes activated upon differentiation acquire H3K4me₃ and lose H3K27me₃, whereas non-activated genes lose H3K4me₃ but retain H3K27me₃ [68, 69]. For successful iPSC generation to occur, H3K4 and H3K27 methylation patterns thus need to be reset to the ESC-like bivalent state. Indeed, bivalent histone marks are typically re-established at developmental gene promoters during mouse cellular reprogramming [9, 12]. The chromatin-associated transcription factor *Utf1*, a direct downstream target of *Oct4* and *Sox2*, is strongly enriched on bivalent genes in ESCs, where it tightly controls their poised state by preventing excessive PRC2 binding and promoting the tagging of mRNAs generated from “leaky” transcription for subsequent cytoplasmic degradation (a process referred to as mRNA pruning) [79] (**Figure 2B**). *Utf1* ensures rapid proliferation of ESCs by mRNA pruning-mediated inhibition of *Arf* expression [79], which might explain why *Utf1* markedly enhances the efficiency of iPSC generation even in the absence of *c-Myc* [80]. Collectively, these results highlight the complex interplay of the OSKM reprogramming factors with regulators of chromatin structure in the establishment and maintenance of a ground pluripotent state.

DNA Methylation in Reprogramming

DNA methylation is the covalent attachment of a methyl (CH₃) group to the carbon-5 position of a cytosine pyrimidine ring, forming 5-methylcytosine (5mC). This reversible

epigenetic modification is carried out by specific enzymes, called DNA methyltransferases (DNMTs), and occurs predominantly on cytosine residues at cytosine-guanine dinucleotide (CpG) sites in the DNA. There is, in general, a strong correlation between CpG density at gene promoter regions, the methylation status of those CpGs, and gene function. Promoter regions with high-CpG content are mostly hypomethylated in ESCs and relate to both ubiquitous housekeeping genes and genes subject to complex regulation during development [70, 81]. Whereas the CpG-rich promoters of housekeeping genes are actively expressed and associated with monovalent H3K4me3 marks in ESCs, those of developmental regulators are generally bivalently marked (H3K4me3/H3K27me3) and transcriptionally silent [68]. Notably, H3K4me3 marks are much less prominent at low-CpG promoters, which tend to be hypermethylated and are associated with genes with tissue-specific functions [68, 70]. Characterization of promoter DNA methylation profiles in ESCs has shown that, while methylated genes are typically silenced and involved in differentiation, pluripotency genes including *Oct4*, *Sox2* and *Nanog* are generally active and distinguished by promoter hypomethylation [82, 83]. Upon lineage commitment, many pluripotency gene promoters acquire methylation, contributing to their transcriptional silencing [84]. During the generation of iPSCs from murine fibroblasts, key pluripotency gene promoters become demethylated whereas genes associated with differentiation (e.g., *HoxA10* and *Gja8*) acquire *de novo* promoter methylation; these changes in DNA methylation patterns have been observed to take place primarily toward the end of reprogramming (**Figure 1**) [45].

In ESCs, but not in differentiated cells, the genes encoding the *Dnmt3a* and *Dnmt3b* methyltransferases are highly expressed, suggesting an important link between *de novo* DNMT function and pluripotency [85]. Yet, in the absence of DNMT activity, mouse ESCs preserve stem cell properties such as *Oct4* expression and self-renewal but demonstrate little or no ability to differentiate [86, 87], which suggests that DNA methylation is essential for differentiation but not important for ground state pluripotency. Consistent with this notion, *de novo* DNA methylation is dispensable for transcription factor-mediated reprogramming, as iPSCs can be derived from mouse embryonic fibroblasts with conditionally deleted *Dnmt3a* and *Dnmt3b* [88].

A substantial degree (up to ~25%) of all DNA methylation in mammalian ESCs occurs at non-CpG sites (CpH, where H = A, C or T) [89, 90]. As of today, little is known about the importance of this type of DNA methylation. Methylation in non-CpG contexts is much less abundant in differentiated cells compared to ESCs, despite similar levels of CpG methylation in both cell types [89, 90]. Interestingly, CpH methylation is lost upon differentiation of human ESCs and is re-established during reprogramming into iPSCs [89, 91, 92], suggesting that it is specific to a stem cell state. Ziller and colleagues, however, showed that knockdown of *DNMT3A* and *DNMT3B* in human ESCs globally reduced non-CpG methylation whereas pluripotency remained unaffected [92]. Additional sup-

port for the notion that CpH methylation might not be directly related to the pluripotent state is provided by Guo *et al.*'s recent observation that CpH methylation abundantly occurs in the mouse and human brain, where it may function in the silencing of neuronal gene expression [93]. Notably, CpH but not CpG methylation levels were found to gradually increase during postnatal neuronal maturation and to be actively preserved by DNMT3A. Further studies investigating the specific role of non-CpG methylation in the regulation of pluripotency and differentiation are likely to offer new insights into the mechanisms by which cell fate is controlled and may contribute to a better molecular understanding of the reprogramming process.

DNA Hydroxymethylation in Reprogramming

The ten-eleven translocation (TET) family of oxygenases plays a particularly important and active role in DNA demethylation by hydroxylating 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) [94, 95]. 5hmC represents a first and principal intermediate of the complete demethylation reaction (**Box 2**) and can be further oxidized by TET enzymes to subsequently generate 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), whose exact importance as functional epigenetic marks requires further investigation [96]. In mammals, the TET family of 5mC hydroxylases consists of three proteins of which two members, Tet1 and Tet2, are specifically abundant in ESCs [97, 98]. In accordance, levels of 5hmC in ESCs are high but decline following induction of differentiation [94, 99, 100]. Interestingly, reprogramming somatic cells to iPSCs induces *Tet1* and *Tet2* expression and enriches 5hmC to levels comparable to those in ESCs [100, 101]. Together, these results suggest an essential role for TET proteins and 5hmC in epigenetic regulation of the pluripotent state and differentiation. However, whereas some studies have shown that knockdown of *Tet1* causes loss of ESC identity [97, 102], others have not found a detrimental effect of Tet1 depletion on maintenance of the ESC state [101, 103, 104]. Yet, deficiency of Tet1 in ESCs is accompanied by a biased differentiation propensity, in particular toward the extraembryonic trophoblast lineage [97, 101, 103-105]. Several studies that profiled Tet1 binding and 5hmC distribution in mouse ESCs on a genome-wide scale have provided insight into the mechanisms with which TET proteins and 5hmC could regulate lineage differentiation potential (**Figure 3**). In particular, these studies revealed enrichment of Tet1 and 5hmC at transcription start sites of genes whose promoters were CpG-rich and were associated with bivalent chromatin states [106-109]. Therefore, it seems plausible that Tet1 and/or 5hmC have a function in controlling the pluripotent state and preventing lineage-specific gene expression by regulating bivalent chromatin marking. Supporting this notion, it was found that Tet1 bound to promoters of developmental regulators and recruited PRC2 to deposit H3K27me3, thereby inducing bivalency-associated transcriptional silencing [109] (**Figure 3A**). Interestingly, Mbd3, a methyl-CpG-binding domain protein that specifically binds 5hmC- but not 5mC-marked

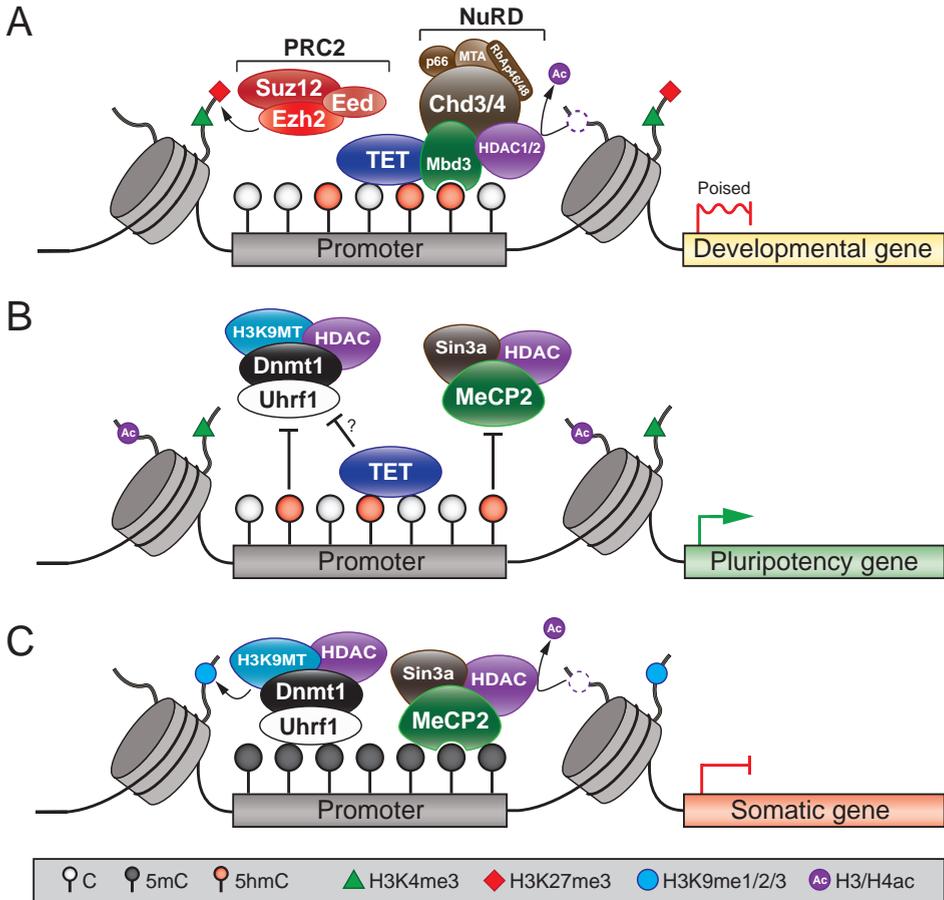


Figure 3. Roles of TET enzymes and hydroxymethylation in the control of pluripotency. (A) In ESCs, ten-eleven translocation (TET) enzymes induce a poised state of chromatin at promoters of lineage-specific genes by recruiting Polycomb repressive complex 2 (PRC2), whose Ezh2 subunit catalyzes the methylation of nearby nucleosomes at H3K27. Methyl-CpG-binding domain protein 3 (Mbd3), a transcriptional repressor of the nucleosome remodeling and deacetylation (NuRD) complex that specifically binds to hydroxymethylated cytosines (5hmC), reinforces the repressed chromatin state via associated histone deacetylases (HDACs) and chromatin remodelers. (B) At pluripotency gene promoters, 5hmC and possibly TET itself may hinder Uhrf1-mediated recruitment of the maintenance DNA methyltransferase Dnmt1 and associated transcriptional repressors, thereby keeping inhibitory DNA methylation (5mC) at a low level and safeguarding a transcriptionally permissive state of chromatin. Furthermore, methyl-CpG-binding protein 2 (MeCP2), a transcriptional repressor associated with Sin3a and HDAC-containing co-repressor complexes, fails to recognize 5hmC. (C) The unhindered binding of Dnmt1 and MeCP2 at heavily methylated promoter regions of somatic genes contributes to gene repression by recruitment of chromatin-modifying transcriptional repressors.

H3K9MT: H3K9 methyltransferase; Uhrf1: ubiquitin-like, containing PHD and RING finger domains protein 1.

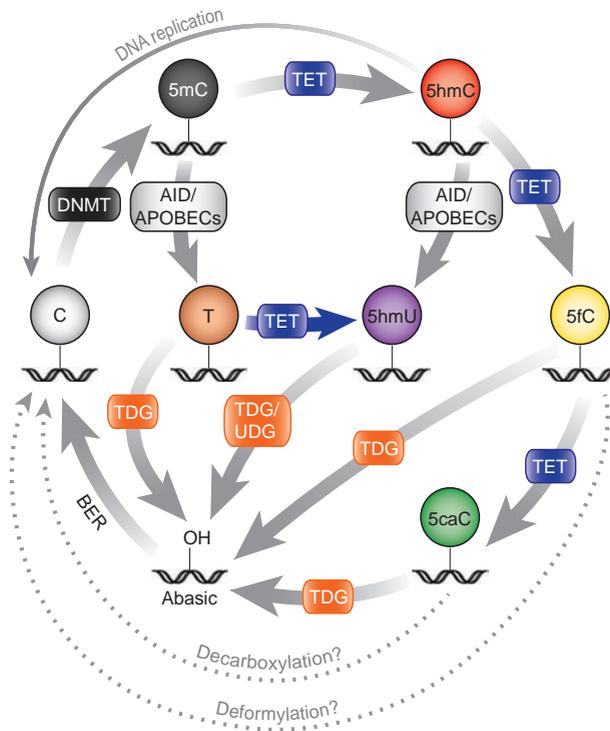
DNA, was also found to be enriched at high-CpG promoters of PcG-targeted bivalent genes and was physically associated with Tet1 in electrophoretic mobility shift assay experiments [118]. Mbd3, being a component of the NuRD complex, recruits several histone deacetylases (HDACs), chromatin remodeling complexes and transcriptional co-repressors, resulting in the imposition of a transcriptionally repressive chromatin structure [119, 120]; thus, 5hmC might play a role in the transcriptional silencing of developmentally regulated genes via its interaction with Mbd3.

Genome-wide and functional analyses have also demonstrated a role for TET proteins and 5hmC in promoting active gene expression in ESCs (**Figure 3B**). In particular, Tet1 and 5hmC are also represented at CpG-rich promoters of transcriptionally active genes, and knockdown of *Tet1* alone or in combination with *Tet2* can result in hypermethylation and downregulation of several pluripotency-related genes in conjunction with a biased differentiation potential [105, 109]. Thus, in ESCs, TET-induced conversion of 5mC to 5hmC at pluripotency genes possibly contributes to their transcriptional activation via abrogation of the repressive effect of 5mC. Consistent with this notion, the methyl-CpG-binding transcriptional repressor MeCP2, which associates with co-repressor complexes containing Sin3a and HDACs to induce a condensed, transcriptionally incompetent chromatin state at methylated gene promoters [121], fails to recognize 5hmC [122]. Furthermore, 5hmC, and potentially TET itself, may hinder Uhrf1-mediated recruitment of Dnmt1 to the DNA [110-112], resulting in the progressive loss of inhibitory methylation with each DNA replication cycle. This may explain how these rapidly dividing stem cells can maintain low levels of DNA methylation at promoters of constitutively active pluripotency genes. In addition, as Dnmt1 itself associates with transcriptional repressors, such as HDACs [123] and H3K9 methyltransferases [124, 125], a failure to recruit Dnmt1 to hydroxymethylated DNA may further prevent the establishment of a repressed chromatin state at these loci. By contrast, the unrestricted binding of Dnmt1 to the heavily methylated promoter regions of somatic genes in ESCs may ensure that these genes remain transcriptionally silent through cell division (**Figure 3C**).

Together, these results suggest that TET enzymes and/or 5hmC have a crucial function in finely controlling pluripotency gene expression and preventing differentiation. Accordingly, Tet1 facilitates reprogramming of mouse embryonic fibroblasts toward iPSCs and can even substitute for Oct4 in the original OSKM reprogramming factor cocktail [126]. Tet2 has also been found to be essential for establishing a transcriptionally permissive chromatin environment at stem cell loci during reprogramming of somatic cells [127]. It should be pointed out, however, that TET-induced 5hmC deposition might also recruit Mbd3 along with transcriptional repression machinery, potentially resulting in repression of pluripotency gene expression. Accordingly, Mbd3 has been identified as a prime impediment to the reprogramming process, and ablating Mbd3 greatly improved reprogramming efficiency [128].

Box 2. Possible pathways of DNA demethylation.

Patterns of DNA methylation (5-methylcytosine, 5mC) are established and maintained by the DNA methyltransferase (DNMT) family of enzymes. 5mC can be oxidized in a stepwise manner to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) by the action of ten-eleven translocation (TET) oxygenases. 5hmC may passively become diluted out during cell multiplication as hemi-hydroxylated DNA is not efficiently recognized by the maintenance DNA methyltransferase Dnmt1 [110-112]. Both the 5fC and 5caC oxidation derivatives are subject to removal by thymine DNA glycosylase (TDG), leaving an abasic site that can be repaired back to unmodified cytosine (C) by the base excision DNA repair (BER) machinery [95, 113, 114]. 5fC and 5caC may also be directly converted back to C by the action of putative deformylases or decarboxylases (dashed arrows) [115, 116]. In an alternative pathway, 5mC and 5hmC can be deaminated by the AID (activation-induced deaminase) or APOBEC (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like) family of cytidine deaminases to generate thymine (T) and 5-hydroxymethyluracil (5hmU), respectively. T and 5hmU are then excised by TDG or uracil DNA glycosylase (UDG) and finally repaired by BER machinery [95]. A novel interaction between the deamination and oxidation pathways is highlighted by the finding that T can serve as a substrate for TET-induced oxidation to 5hmU (dark blue arrow) [117].



TRANSCRIPTIONAL AND EPIGENETIC VARIABILITY IN CELLULAR REPROGRAMMING

Successful reprogramming to iPSCs is largely dependent on faithful remodeling of the cell's epigenetic landscape to shut down somatic gene expression and activate a

transcriptional program characteristic of pluripotent cells. Incomplete reprogramming, however, may result in improper epigenetic states that could contribute to variability in gene expression and biological function among iPSC lines, as well as between iPSC and ESC lines. Various studies have indicated distinguishable epigenetic differences between iPSCs and ESCs [91, 129-138], many of which have been extensively reviewed elsewhere [27]. In summary, dissimilarities between ESCs and iPSCs have particularly been observed at the level of DNA methylation and are generally due to a failure to reset methylation signatures from the somatic cell of origin or the acquisition of *de novo* aberrant methylation during the reprogramming process. The observation that inhibiting DNMTs by 5-aza-cytidine could efficiently and rapidly convert partially reprogrammed cells to full iPSCs implies that persistent DNA methylation or insufficient DNA demethylation, in particular at pluripotency-related genes, constitutes a barrier to full reprogramming [9]. Incomplete promoter DNA methylation represents another mechanism underlying the failure to correctly reset the DNA methylome during iPSC formation and has been shown to contribute to residual expression of somatic genes in human iPSCs [91, 132]. Importantly, retention of an epigenetic memory of the somatic tissue of origin has been shown to affect the *in vitro* differentiation potential of resulting iPSCs, favoring differentiation along donor cell-related lineages while restricting commitment towards other lineages [134-136].

Comparing Genetically Matched ESCs and iPSCs

To more reliably assess whether or not iPSCs are identical to ESCs on the molecular and functional level, one needs to rule out genetic background effects and compare genetically matched iPSC and ESC lines. Stadtfeld and colleagues used ESCs containing Dox-inducible OKSM transgenes to engineer “reprogrammable” mice from which various somatic cell types were collected for the generation of genetically identical iPSCs [139]. Genetically matched iPSCs and ESCs were found to be transcriptionally very similar and could only be distinguished by a minimal set of two transcripts (*Gtl2* and *Rian*), suggesting that earlier observations of differential gene expression between ESCs and iPSCs [140-142] may be based on differences in genetic background or on other hard-to-control variables inherent to the reprogramming process. Despite the overall similarity in transcriptional profiles, however, iPSCs generated using this transgenic reprogramming system were unable to give rise to viable and germline-competent “all-iPSC” mice (i.e., mice produced exclusively from iPSCs) in the tetraploid complementation assay [139], the golden standard for testing pluripotency. This developmental deficiency of iPSCs was attributed to aberrant epigenetic repression of the *Dlk1-Dio3* imprinted region on chromosome 12qF1, a finding that corroborates another study revealing a positive correlation between stem cell pluripotency level and the degree of activation of this particular gene cluster [143]. Bisulfite pyrosequencing and chromatin immunoprecipitation (ChIP) analyses indicated

that the abnormal silencing of normally maternally expressed genes of this locus is caused by DNA hypermethylation as well as reduction of histone modifications associated with active transcription (H3ac, H4ac, and H3K4me) [139]. Interestingly, the *Dlk1-Dio3* gene cluster was found to be silenced in a predominant number of iPSC clones derived from various somatic cell types at different differentiation stages, leading the authors to propose that loss of imprinting at this locus and incomplete pluripotency induction are common consequences of transcription factor-mediated reprogramming. In apparent contrast to those findings, Carey *et al.* [21] demonstrated that most iPSCs generated using a highly similar transgenic reprogramming system preserved normal *Dlk1-Dio3* imprinting and were able to produce all-iPSC mice via tetraploid blastocyst complementation. The discordance in findings was attributed to subtle differences in design of the polycistronic expression vectors between the two studies, most notably a switch in the order of reprogramming factors that resulted in altered levels and stoichiometry of reprogramming factor proteins. iPSCs supporting the development of all-iPSC mice were characterized by 15- and 5-fold higher Klf4 and Oct4, and 2-fold lower Sox2 and c-Myc protein levels than those with impaired developmental potential [21]. Interestingly, increased ectopic expression of Klf4 and Oct4, or treatment with the HDAC inhibitor valproic acid, could restore functional developmental competence of the poor-quality iPSCs [21, 139].

Epigenetics and Alternative Pluripotent States

Further supporting the notion that the expression level of the four reprogramming transcription factors greatly influences the epigenetic state and biological functionality of generated iPSCs, Tonge and colleagues [144-148] described, in a collection of five manuscripts, a novel stable state of induced pluripotency that is distinct from the ESC-like state and requires constitutively high OSKM transgene expression. These so-called F-class cells did not form colonies with typical ESC-like morphology but instead appeared “fuzzy” due to their low intercellular adhesion. Yet, F-class cells displayed high proliferation, expressed *Nanog*, endogenous *Oct4* and several other pluripotency-related genes at ESC levels, and were able to give rise to teratomas containing tissues representing all three germ layers, thereby effectively fulfilling some commonly used criteria for defining pluripotency [144]. However, the F-class cells were incapable of contributing to the formation of chimaeric mice, indicating impaired embryonic developmental potential of these cells. By global gene expression profiling of mouse embryonic fibroblasts, traditional (ESC-like) iPSCs and F-class cells, the authors revealed a transcriptional signature unique to the F-class state that persisted in long-term culture and exhibited low variability between independent sub-clones. Thus, F-class cells do not seem to represent an intermediate reprogramming state but rather a pluripotent alternative to iPSCs. In another paper, Hussein and colleagues [145] provide the results of an extensive multi-omic analysis that profiled the transcriptome, proteome and epigenome of the F-class and ESC-like states

to better understand, at the molecular level, their emergence. They found that a subset of genes typically expressed in ESCs and iPSCs are repressed in F cells. Although high expression levels of the reprogramming factors were initially required for a genome-wide loss of repressive H3K27me3 and concomitant general opening of somatic cell chromatin, maintenance of high transgene expression induced an F-class state characterized by reacquisition of this repressive histone modification [145]. By contrast, lower transgene levels resulted in the acquisition of activating H3K4me3 and endowed reprogramming cells with an ESC-like phenotype. In addition, whereas F-class cells failed to lose some parental fibroblast-inherited DNA methylation patterns, ESC-like cells showed proper demethylation of somatic cell-originating signatures. Taken together, these data suggest that active epigenetic mechanisms play a crucial role in directing cell identity towards either the iPSC or F-class state during cellular reprogramming. Further corroborating this notion, Tonge *et al.* [144] showed that treatment with an inhibitor of HDACs converted F-class cells to a transgene-independent, ESC-like state with embryonic developmental potential.

Lee and colleagues [146] investigated the dynamics of DNA (de)methylation throughout reprogramming in more detail and showed that the promoter methylation status of genes at the start of the process controlled their regulation by histone modification alterations. The authors indicated that DNA methylation represented a key epigenetic barrier *en route* to ESC-like induced pluripotency and served as an important determinant of pluripotent outcome (that is, whether cells transitioned towards the ESC-like or F-class state). Yet, epigenetic mechanisms other than DNA methylation are also involved in regulating ESC-like versus F-class identity. Notably, Clancy *et al.* [147] revealed the existence of a novel group of differentially expressed microRNAs that specifically control pluripotency in the F-class cells. Epigenetic changes such as these are likely to contribute to the establishment of cell state-specific patterns of protein expression. In support of this idea, Benevento and colleagues [148] showed, using mass spectrometry-based proteomics, that the somatic cell proteome was reorganized in two distinct waves during pluripotency induction and that the F-class proteome was different from the iPSC proteome. As these studies have demonstrated, understanding how the epigenome changes during reprogramming to induced pluripotency will enhance our molecular understanding of the mechanisms of cell fate change and of somatic cell reprogramming in particular.

CONCLUSION AND FUTURE PERSPECTIVES

Reprogramming of somatic cells to induced pluripotency using defined transcription factors is predominantly an epigenetic phenomenon. Epigenetic as well as (single-cell) transcriptomic analyses on isolated reprogramming intermediates *en route* to induced

pluripotency have generated invaluable mechanistic insights into the process of cell fate specification and transition. It has become evident that faithful reprogramming is very dependent on various technological parameters that can influence cellular epigenetic status, such as the levels and stoichiometric ratio of reprogramming factors and the composition of the culture environment. In the next few years, parallel profiling of the methylome and transcriptome of single cells (e.g., scM&T-seq [149]) should further advance our understanding of how the OSKM reprogramming factors, and other factors in the reprogramming milieu, impact the epigenome, gene expression pattern and functionality of reprogramming cells, and especially those of human origin. This new information will bring us one step closer to the development of more efficient reprogramming methods and, ultimately, high-fidelity generation of therapeutic-grade iPSCs with limited heterogeneity for use in clinical applications [28].

EXECUTIVE SUMMARY

Molecular and Transcriptional Routes to Induced Pluripotency

- Gene expression profiling analyses have identified three phases of the murine somatic cell reprogramming process. During the initiation phase, transcription factor-induced cells suppress somatic cell-specific genes and undergo mesenchymal-to-epithelial transition (MET). Cells then gradually transition into maturation, in which a subset of pluripotency-associated genes (e.g., *Nanog*, *Sall4*, *Oct4*) becomes activated. Ultimately, cells transition to the stabilization phase, characterized by induction of the complete pluripotency transcriptional network with expression of endogenous *Sox2*.
- Two major waves of gene activation occur during reprogramming. The first transcriptional wave occurs early and in the majority of cells, and genes related to proliferation, metabolism and cytoskeletal organization are activated. The second wave occurs late and is characterized by upregulation of core pluripotency genes in cells progressing to *bona fide* pluripotency. Transient activation of epidermis-related genes occurs at intermediate reprogramming stages, and may represent a common hallmark of reprogramming.

Epigenetic Control of Cellular Reprogramming

- Reprogramming requires global remodeling of somatic cell chromatin from a highly condensed state (heterochromatin) to a more dispersed conformation (euchromatin). Whereas heterochromatin reorganization occurs early in reprogramming, the establishment of euchromatin features constitutes a relatively late event. Heterochromatin-associated H3K9 methylation constitutes an important epigenetic barrier to cellular reprogramming.

- Trithorax (trxG) and Polycomb group (PcG) proteins play an indispensable role in specifying and maintaining pluripotent cell identity. During reprogramming, mono-valent histone methylation marks at lineage-specific gene promoters are reset to the bivalent state (H3K27me3/H3K4me3) characteristic of stem cells.
- Changes in DNA methylation take place primarily toward the end of reprogramming. Whereas the promoters of differentiation-associated genes acquire *de novo* methylation, pluripotency gene promoters become demethylated. The functional importance of the re-establishment of DNA methylation in non-CpG contexts in iPSCs requires further investigation.
- iPSC generation induces the expression of Tet1 and Tet2, two members of the ten-eleven translocation (TET) family of enzymes that hydroxylate 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) to initiate a process of DNA demethylation. These TET proteins, along with the 5hmC mark itself, control the pluripotent state and prevent lineage-specific differentiation through their association with both active (pluripotency) and bivalent (developmental) gene promoters.

Transcriptional and Epigenetic Variability in Cellular Reprogramming

- Incomplete reprogramming may result in improper epigenetic states that contribute to variability in gene expression and biological function among iPSC lines. Epigenetic aberrations are mostly observed at the level of DNA methylation and include remnant methylation signatures originating from the somatic cell as well as reprogramming-induced *de novo* aberrant methylation patterns.
- Faithful reprogramming towards high-quality iPSCs with full developmental competence is very dependent on various technological parameters that influence epigenetic status, including the levels and stoichiometry of reprogramming factor proteins. Suboptimal reprogramming conditions can cause the emergence of an alternative pluripotent state (F-class cells) with unique transcriptional and epigenetic profiles and altered developmental potential.

Future Perspectives

- Parallel profiling of single-cell methylomes and transcriptomes will enhance our understanding of the reprogramming process and of the mechanisms contributing to epigenetic, transcriptional, and functional heterogeneity of iPSCs

ACKNOWLEDGMENTS

The authors would like to thank Mary Lynn Gage for editing the manuscript.

REFERENCES

Papers of special note have been highlighted as:

- of interest
 - of considerable interest
1. Gurdon JB. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *J. Embryol. Exp. Morphol.* 10 622-640 (1962).
 2. Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH. Viable offspring derived from fetal and adult mammalian cells. *Nature* 385(6619), 810-813 (1997).
 3. Markoulaki S, Meissner A, Jaenisch R. Somatic cell nuclear transfer and derivation of embryonic stem cells in the mouse. *Methods* 45(2), 101-114 (2008).
 4. Cowan CA, Atienza J, Melton DA, Eggan K. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* 309(5739), 1369-1373 (2005).
 5. Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr. Biol.* 11(19), 1553-1558 (2001).
 6. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126(4), 663-676 (2006).
 7. Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature* 448(7151), 313-317 (2007).
 8. Wernig M, Meissner A, Foreman R *et al.* In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448(7151), 318-324 (2007).
 9. Mikkelsen TS, Hanna J, Zhang X *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454(7200), 49-55 (2008).
 10. Guenther MG, Frampton GM, Soldner F *et al.* Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell* 7(2), 249-257 (2010).
 11. Hawkins RD, Hon GC, Lee LK *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* 6(5), 479-491 (2010).
 12. Maherali N, Sridharan R, Xie W *et al.* Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* 1(1), 55-70 (2007).
 13. Koche RP, Smith ZD, Adli M *et al.* Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell* 8(1), 96-105 (2011).
 14. Mattout A, Biran A, Meshorer E. Global epigenetic changes during somatic cell reprogramming to iPSC cells. *J. Mol. Cell. Biol.* 3(6), 341-350 (2011).
 15. Huangfu D, Maehr R, Guo W *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* 26(7), 795-797 (2008).
 16. Huangfu D, Osafune K, Maehr R *et al.* Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat. Biotechnol.* 26(11), 1269-1275 (2008).
 17. Shi Y, Despons C, Do JT, Hahm HS, Scholer HR, Ding S. Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 3(5), 568-574 (2008).
 18. Hou P, Li Y, Zhang X *et al.* Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* 341(6146), 651-654 (2013).
 19. Hanna J, Saha K, Pando B *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* 462(7273), 595-601 (2009).

20. Papapetrou EP, Tomishima MJ, Chambers SM *et al.* Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 106(31), 12759-12764 (2009).
21. Carey BW, Markoulaki S, Hanna JH *et al.* Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell* 9(6), 588-598 (2011).
- **Indicates that the levels and stoichiometry of reprogramming factor proteins during reprogramming can critically impact epigenetic status and developmental potential of iPSCs.**
22. Eminli S, Foudi A, Stadtfeld M *et al.* Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nat. Genet.* 41(9), 968-976 (2009).
23. Kim JB, Zaehres H, Wu G *et al.* Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* 454(7204), 646-650 (2008).
24. Bar-Nur O, Brumbaugh J, Verheul C *et al.* Small molecules facilitate rapid and synchronous iPSC generation. *Nat. Methods* 11(11), 1170-1176 (2014).
25. Hochedlinger K, Plath K. Epigenetic reprogramming and induced pluripotency. *Development* 136(4), 509-523 (2009).
26. Papp B, Plath K. Reprogramming to pluripotency: stepwise resetting of the epigenetic landscape. *Cell Res.* 21(3), 486-501 (2011).
27. Liang G, Zhang Y. Embryonic stem cell and induced pluripotent stem cell: an epigenetic perspective. *Cell Res.* 23(1), 49-69 (2013).
28. Bhutani K, Nazor KL, Williams R *et al.* Whole-genome mutational burden analysis of three pluripotency induction methods. *Nat. Commun.* 7 10536 (2016).
29. Boyer LA, Lee TI, Cole MF *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122(6), 947-956 (2005).
30. Loh YH, Wu Q, Chew JL *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* 38(4), 431-440 (2006).
31. Babaie Y, Herwig R, Greber B *et al.* Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. *Stem Cells* 25(2), 500-510 (2007).
32. Sato N, Sanjuan IM, Heke M, Uchida M, Naef F, Brivanlou AH. Molecular signature of human embryonic stem cells and its comparison with the mouse. *Dev. Biol.* 260(2), 404-413 (2003).
33. Daheron L, Opitz SL, Zaehres H *et al.* LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells. *Stem Cells* 22(5), 770-778 (2004).
34. Thomson JA, Itskovitz-Eldor J, Shapiro SS *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* 282(5391), 1145-1147 (1998).
35. Xu RH, Chen X, Li DS *et al.* BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat. Biotechnol.* 20(12), 1261-1264 (2002).
36. 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).
37. Lowry WE, Richter L, Yachechko R *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.* 105(8), 2883-2888 (2008).
38. Park IH, Zhao R, West JA *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 451(7175), 141-146 (2008).
39. Anokye-Danso F, Trivedi CM, Juhr D *et al.* Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* 8(4), 376-388 (2011).
40. Suh MR, Lee Y, Kim JY *et al.* Human embryonic stem cells express a unique set of microRNAs. *Dev. Biol.* 270(2), 488-498 (2004).

41. Card DA, Hebbar PB, Li L *et al.* Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. *Mol. Cell. Biol.* 28(20), 6426-6438 (2008).
42. Schnerch A, Cerdan C, Bhatia M. Distinguishing between mouse and human pluripotent stem cell regulation: the best laid plans of mice and men. *Stem Cells* 28(3), 419-430 (2010).
43. Brambrink T, Foreman R, Welstead GG *et al.* Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* 2(2), 151-159 (2008).
44. Stadtfeld M, Maherali N, Breault DT, Hochedlinger K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2(3), 230-240 (2008).
45. Polo JM, Anderssen E, Walsh RM *et al.* A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* 151(7), 1617-1632 (2012).
- **Genome-wide profiling of the transcriptome, chromatin status and DNA methylation in reprogramming intermediates en route to induced pluripotency. Reveals a c-Myc-driven early and Oct4/Sox2-driven late wave of transcriptional activation, and identifies DNA methylation changes as a late event in pluripotency induction.**
46. Li R, Liang J, Ni S *et al.* A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* 7(1), 51-63 (2010).
47. Samavarchi-Tehrani P, Golipour A, David L *et al.* Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* 7(1), 64-77 (2010).
48. Smith ZD, Nachman I, Regev A, Meissner A. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nat. Biotechnol.* 28(5), 521-526 (2010).
49. Golipour A, David L, Liu Y *et al.* A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. *Cell Stem Cell* 11(6), 769-782 (2012).
- **Indicates that silencing of OSKM transgenes is required for cells to become fully pluripotent. A transcriptomic signature associated with successful acquisition of the pluripotent state was identified that was distinct from the network of genes regulating maintenance of pluripotency.**
50. Sridharan R, Tchieu J, Mason MJ *et al.* Role of the murine reprogramming factors in the induction of pluripotency. *Cell* 136(2), 364-377 (2009).
51. O'malley J, Skylaki S, Iwabuchi KA *et al.* High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. *Nature* 499(7456), 88-91 (2013).
- **Identifies a reprogramming route map contingent on the loss and upregulation of, respectively, CD44 and ICAM1 cell surface markers. RNA-sequencing analysis of intermediate reprogramming populations demonstrated upregulation of pluripotency genes in two distinct waves, as well as transient activation of epidermis-related genes.**
52. Buganim Y, Faddah DA, Cheng AW *et al.* Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 150(6), 1209-1222 (2012).
- **Single-cell gene expression profiling to characterize various stages and transitions during reprogramming. Reveals substantial variation in gene expression early in reprogramming and identifies Sox2 as an upstream regulator of activation of many genes involved in pluripotency.**
53. Orkin SH, Hochedlinger K. Chromatin connections to pluripotency and cellular reprogramming. *Cell* 145(6), 835-850 (2011).
54. Apostolou E, Hochedlinger K. Chromatin dynamics during cellular reprogramming. *Nature* 502(7472), 462-471 (2013).
55. Ho L, Ronan JL, Wu J *et al.* An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc. Natl. Acad. Sci. U. S. A.* 106(13), 5181-5186 (2009).

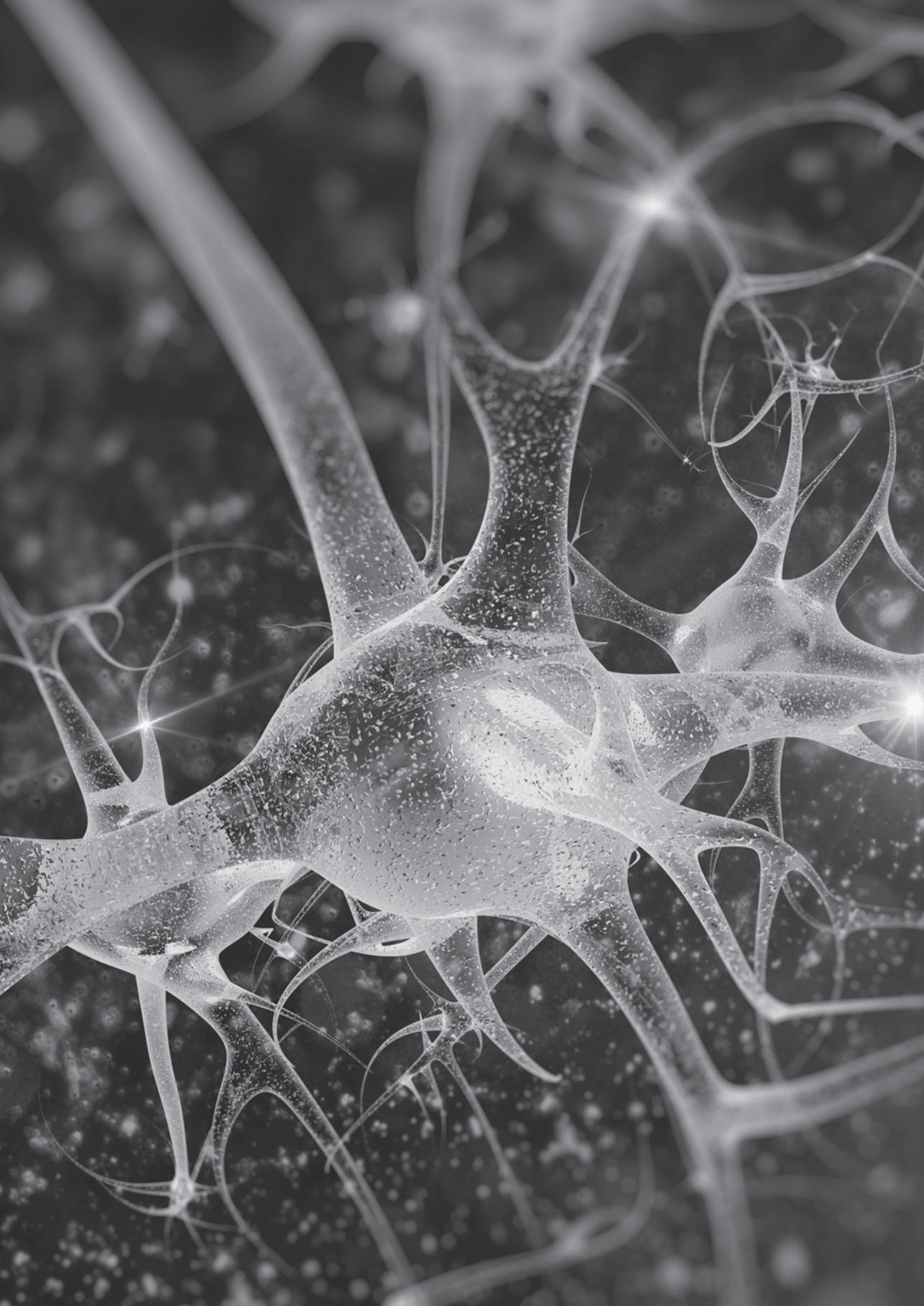
56. Liang J, Wan M, Zhang Y *et al.* Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nat. Cell Biol.* 10(6), 731-739 (2008).
57. Wang J, Rao S, Chu J *et al.* A protein interaction network for pluripotency of embryonic stem cells. *Nature* 444(7117), 364-368 (2006).
58. Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat. Genet.* 41(2), 246-250 (2009).
59. Krejci J, Uhlirova R, Galiova G, Kozubek S, Smigova J, Bartova E. Genome-wide reduction in H3K9 acetylation during human embryonic stem cell differentiation. *J. Cell. Physiol.* 219(3), 677-687 (2009).
60. Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell* 10(1), 105-116 (2006).
61. Gaetz J, Clift KL, Fernandes CJ *et al.* Evidence for a critical role of gene occlusion in cell fate restriction. *Cell Res.* 22(5), 848-858 (2012).
62. Fussner E, Djuric U, Strauss M *et al.* Constitutive heterochromatin reorganization during somatic cell reprogramming. *EMBO J.* 30(9), 1778-1789 (2011).
63. Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* 151(5), 994-1004 (2012).
- **Identifies megabase-long chromatin domains in fibroblasts, enriched for heterochromatic H3K9me3 and including many genes required for establishing pluripotency, that block binding of the OSKM factors and hinder reprogramming.**
64. Chen J, Liu H, Liu J *et al.* H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat. Genet.* 45(1), 34-42 (2013).
65. Silva J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, Smith A. Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol.* 6(10), e253 (2008).
66. Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. *Cell* 128(4), 735-745 (2007).
67. Pan G, Tian S, Nie J *et al.* Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* 1(3), 299-312 (2007).
68. Mikkelsen TS, Ku M, Jaffe DB *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448(7153), 553-560 (2007).
69. Bernstein BE, Mikkelsen TS, Xie X *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125(2), 315-326 (2006).
70. Meissner A, Mikkelsen TS, Gu H *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454(7205), 766-770 (2008).
71. Ang YS, Tsai SY, Lee DF *et al.* Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* 145(2), 183-197 (2011).
72. O'carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T. The polycomb-group gene Ezh2 is required for early mouse development. *Mol. Cell. Biol.* 21(13), 4330-4336 (2001).
73. Boyer LA, Plath K, Zeitlinger J *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441(7091), 349-353 (2006).
74. Lee TI, Jenner RG, Boyer LA *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125(2), 301-313 (2006).
75. Pasini D, Bracken AP, Hansen JB, Capillo M, Helin K. The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol. Cell. Biol.* 27(10), 3769-3779 (2007).
76. Ding X, Wang X, Sontag S *et al.* The polycomb protein Ezh2 impacts on induced pluripotent stem cell generation. *Stem Cells Dev.* 23(9), 931-940 (2014).

77. Onder TT, Kara N, Cherry A *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature* 483(7391), 598-602 (2012).
78. Azuara V, Perry P, Sauer S *et al.* Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* 8(5), 532-538 (2006).
79. Jia J, Zheng X, Hu G *et al.* Regulation of pluripotency and self-renewal of ESCs through epigenetic-threshold modulation and mRNA pruning. *Cell* 151(3), 576-589 (2012).
80. Zhao Y, Yin X, Qin H *et al.* Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell Stem Cell* 3(5), 475-479 (2008).
81. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci. U. S. A.* 103(5), 1412-1417 (2006).
82. Fouse SD, Shen Y, Pellegrini M *et al.* Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* 2(2), 160-169 (2008).
83. Imamura M, Miura K, Iwabuchi K *et al.* Transcriptional repression and DNA hypermethylation of a small set of ES cell marker genes in male germline stem cells. *BMC Dev. Biol.* 6 34 (2006).
84. Mohn F, Weber M, Rebhan M *et al.* Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* 30(6), 755-766 (2008).
85. Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* 19(3), 219-220 (1998).
86. Jackson M, Krassowska A, Gilbert N *et al.* Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol. Cell Biol.* 24(20), 8862-8871 (2004).
87. Tsumura A, Hayakawa T, Kumaki Y *et al.* Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* 11(7), 805-814 (2006).
88. Pawlak M, Jaenisch R. De novo DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state. *Genes Dev.* 25(10), 1035-1040 (2011).
89. Lister R, Pelizzola M, Dowen RH *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462(7271), 315-322 (2009).
90. Ramsahoye BH, Biniszkiwicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. U. S. A.* 97(10), 5237-5242 (2000).
91. Lister R, Pelizzola M, Kida YS *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* 471(7336), 68-73 (2011).
92. Ziller MJ, Muller F, Liao J *et al.* Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet.* 7(12), e1002389 (2011).
93. Guo JU, Su Y, Shin JH *et al.* Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat. Neurosci.* 17(2), 215-222 (2014).
94. Tahiliani M, Koh KP, Shen Y *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324(5929), 930-935 (2009).
95. Guo JU, Su Y, Zhong C, Ming GL, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* 145(3), 423-434 (2011).
96. Ito S, Shen L, Dai Q *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333(6047), 1300-1303 (2011).
97. Ito S, D'alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466(7310), 1129-1133 (2010).

98. Huang Y, Chavez L, Chang X *et al.* Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 111(4), 1361-1366 (2014).
99. Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.* 38(19), e181 (2010).
100. Ruzov A, Tsenkina Y, Serio A *et al.* Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. *Cell Res.* 21(9), 1332-1342 (2011).
101. Koh KP, Yabuuchi A, Rao S *et al.* Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* 8(2), 200-213 (2011).
102. Freudenberg JM, Ghosh S, Lackford BL *et al.* Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. *Nucleic Acids Res.* 40(8), 3364-3377 (2012).
103. Dawlaty MM, Ganz K, Powell BE *et al.* Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* 9(2), 166-175 (2011).
104. Dawlaty MM, Breiling A, Le T *et al.* Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev. Cell* 24(3), 310-323 (2013).
105. Ficiz G, Branco MR, Seisenberger S *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* 473(7347), 398-402 (2011).
106. Williams K, Christensen J, Pedersen MT *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* 473(7347), 343-348 (2011).
107. Pastor WA, Pape UJ, Huang Y *et al.* Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* 473(7347), 394-397 (2011).
108. Xu Y, Wu F, Tan L *et al.* Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol. Cell* 42(4), 451-464 (2011).
109. Wu H, D'Alessio AC, Ito S *et al.* Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* 473(7347), 389-393 (2011).
110. Valinluck V, Sowers LC. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res.* 67(3), 946-950 (2007).
111. Hashimoto H, Liu Y, Upadhyay AK *et al.* Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.* 40(11), 4841-4849 (2012).
112. Otani J, Kimura H, Sharif J *et al.* Cell cycle-dependent turnover of 5-hydroxymethyl cytosine in mouse embryonic stem cells. *PLoS One* 8(12), e82961 (2013).
113. He YF, Li BZ, Li Z *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333(6047), 1303-1307 (2011).
114. Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J. Biol. Chem.* 286(41), 35334-35338 (2011).
115. Xu S, Li W, Zhu J *et al.* Crystal structures of isoorotate decarboxylases reveal a novel catalytic mechanism of 5-carboxyl-uracil decarboxylation and shed light on the search for DNA decarboxylase. *Cell Res.* 23(11), 1296-1309 (2013).
116. Schiesser S, Hackner B, Pfaffeneder T *et al.* Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. *Angew. Chem. Int. Ed. Engl.* 51(26), 6516-6520 (2012).
117. Pfaffeneder T, Spada F, Wagner M *et al.* Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat. Chem. Biol.* 10(7), 574-581 (2014).
118. Yildirim O, Li R, Hung JH *et al.* Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* 147(7), 1498-1510 (2011).

119. Morey L, Brenner C, Fazi F *et al.* MBD3, a component of the NuRD complex, facilitates chromatin alteration and deposition of epigenetic marks. *Mol. Cell. Biol.* 28(19), 5912-5923 (2008).
120. Zhang Y, Ng HH, Erdjument-Bromage H, Tempst P, Bird A, Reinberg D. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* 13(15), 1924-1935 (1999).
121. Nan X, Ng HH, Johnson CA *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393(6683), 386-389 (1998).
122. Valinluck V, Tsai HH, Rogstad DK, Burdzy A, Bird A, Sowers LC. Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res.* 32(14), 4100-4108 (2004).
123. Fuks F, Burgers WA, Brehm A, Hughes-Davies L, Kouzarides T. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat. Genet.* 24(1), 88-91 (2000).
124. Fuks F, Hurd PJ, Deplus R, Kouzarides T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.* 31(9), 2305-2312 (2003).
125. Esteve PO, Chin HG, Smallwood A *et al.* Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev.* 20(22), 3089-3103 (2006).
126. Gao Y, Chen J, Li K *et al.* Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* 12(4), 453-469 (2013).
127. Doege CA, Inoue K, Yamashita T *et al.* Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* 488(7413), 652-655 (2012).
128. Rais Y, Zviran A, Geula S *et al.* Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* 502(7469), 65-70 (2013).
129. Bock C, Kiskinis E, Verstappen G *et al.* Reference Maps of human ES and iPSC cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* 144(3), 439-452 (2011).
130. Deng J, Shoemaker R, Xie B *et al.* Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nat. Biotechnol.* 27(4), 353-360 (2009).
131. Doi A, Park IH, Wen B *et al.* Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat. Genet.* 41(12), 1350-1353 (2009).
132. Ohi Y, Qin H, Hong C *et al.* Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPSCs. *Nat. Cell Biol.* 13(5), 541-549 (2011).
133. Bar-Nur O, Russ HA, Efrat S, Benvenisty N. Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. *Cell Stem Cell* 9(1), 17-23 (2011).
134. Polo JM, Liu S, Figueroa ME *et al.* Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat. Biotechnol.* 28(8), 848-855 (2010).
135. Kim K, Doi A, Wen B *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* 467(7313), 285-290 (2010).
136. Kim K, Zhao R, Doi A *et al.* Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat. Biotechnol.* 29(12), 1117-1119 (2011).
137. Ruiz S, Diep D, Gore A *et al.* Identification of a specific reprogramming-associated epigenetic signature in human induced pluripotent stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 109(40), 16196-16201 (2012).
138. Nishino K, Toyoda M, Yamazaki-Inoue M *et al.* DNA methylation dynamics in human induced pluripotent stem cells over time. *PLoS Genet.* 7(5), e1002085 (2011).
139. Stadtfeld M, Apostolou E, Akutsu H *et al.* Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* 465(7295), 175-181 (2010).

140. Chin MH, Mason MJ, Xie W *et al.* Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5(1), 111-123 (2009).
141. Marchetto MC, Yeo GW, Kainohana O, Marsala M, Gage FH, Muotri AR. Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PLoS One* 4(9), e7076 (2009).
142. Ghosh Z, Wilson KD, Wu Y, Hu S, Quertermous T, Wu JC. Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. *PLoS One* 5(2), e8975 (2010).
143. Liu L, Luo GZ, Yang W *et al.* Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. *J. Biol. Chem.* 285(25), 19483-19490 (2010).
144. Tonge PD, Corso AJ, Monetti C *et al.* Divergent reprogramming routes lead to alternative stem-cell states. *Nature* 516(7530), 192-197 (2014).
- **Identifies epigenetic modifications that actively direct reprogramming cells toward an alternative stable state of pluripotency (F-class cells) that is characterized by a unique transcriptional signature and is dependent on high reprogramming factor expression.**
145. Hussein SM, Puri MC, Tonge PD *et al.* Genome-wide characterization of the routes to pluripotency. *Nature* 516(7530), 198-206 (2014).
146. Lee DS, Shin JY, Tonge PD *et al.* An epigenomic roadmap to induced pluripotency reveals DNA methylation as a reprogramming modulator. *Nat. Commun.* 5 5619 (2014).
147. Clancy JL, Patel HR, Hussein SM *et al.* Small RNA changes en route to distinct cellular states of induced pluripotency. *Nat. Commun.* 5 5522 (2014).
148. Benevento M, Tonge PD, Puri MC *et al.* Proteome adaptation in cell reprogramming proceeds via distinct transcriptional networks. *Nat. Commun.* 5 5613 (2014).
149. Angermueller C, Clark SJ, Lee HJ *et al.* Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* doi:10.1038/nmeth.3728 (2016).



CHAPTER 3A

Dulbecco's Modified Eagle's Medium impairs spontaneous calcium activity in a human stem cell-derived *in vitro* model of neuronal networks

Mark van den Hurk^{1,2,3}, Tameji Eames³, Ruben V. Hernandez³, Mariko Kellogg³, Vanessa Palomares³, Leah Boyer³, Suzanne Simon³, Gunter Kenis^{1,2}, Bart P.F. Rutten^{1,2}, Harry W.M. Steinbusch^{1,2}, Fred H. Gage³, Cedric Bardy³

¹ Department of Psychiatry and Neuropsychology, Division of Neuroscience, Maastricht University, Maastricht, 6200 MD, The Netherlands

² School for Mental Health and Neuroscience (MHeNS), European Graduate School of Neuroscience (EURON), Maastricht University, Maastricht, 6200 MD, The Netherlands

³ Laboratory of Genetics, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

ABSTRACT

In vitro cultures of neurons are being employed in research laboratories all over the world as a tool to better understand nervous system function and disease. *In vivo*, correlated neural electrical activity is essential for the proper formation, maturation and survival of developing neuronal circuits. In this study, we characterized spontaneous and electrical stimulation-induced activity in a human stem cell-derived *in vitro* model of neuronal network development using calcium imaging. We discovered that spontaneous calcium activity is low in standard Dulbecco's Modified Eagle's Medium (DMEM) compared to artificial cerebrospinal fluid (ACSF, a commonly used solution for assessing the functionality of neurons) – even after matching for ionic composition, pH and osmolarity. Interestingly, removal of amino acids, vitamins or extra components from the complete DMEM rescued the low spontaneous calcium spike activity of the cells, indicating that DMEM basal medium contains various components that (acutely) interfere with neurophysiological function. Future research efforts should be directed toward designing specific neuronal media that adequately support neuronal activity, so as to create more realistic conditions for *in vitro* neuronal disease models and improve their rate of translational success.

INTRODUCTION

In vitro cultures of neurons derived from human stem cells are an attractive model to study the formation and maturation of human neuronal networks. The human brain has numerous unique features that are difficult to model in non-human preclinical models [1], and human stem cell-derived neuronal cultures may represent the most practical means to study human neuronal function and development in a live environment [2].

Multiple approaches exist to differentiate stem cells into functional neurons [3]. The most common strategy involves the initial differentiation of stem cells into neural progenitor cells (NPCs) through the formation of so-called embryoid bodies [4]. These NPCs can then be expanded and maintained, or differentiated into any of the three neural lineages: neurons, astrocytes, or oligodendrocytes. Differentiation of NPCs into specific types of neurons has become a widely standardized procedure, involving the use of specific combinations of growth factors and chemically defined media under standard cell culture conditions.

Spontaneous correlated activity in neural circuits is known to be essential for proper maturation and refinement of developing networks [5-7]. As the brain develops and neural networks mature, coordinated spontaneous activity becomes increasingly complex, involving more neuronal assemblies [8, 9]. Since the *in vitro* differentiation of stem cells is believed to resemble the complex mechanisms of brain development *in vivo* [10, 11], we aimed to examine spontaneous activity in a human cell culture model of neural network maturation. To that end, we performed calcium imaging to monitor intracellular calcium oscillations in human embryonic stem cell (ESC)-derived neuronal populations grown in standard cell culture conditions.

The most commonly used buffer solution for calcium imaging and electrophysiological recordings of neuronal activity is artificial cerebrospinal fluid (ACSF). ACSF is an isotonic buffer with neurophysiological concentrations of important inorganic salts, energy substrates and phosphate buffers. Although ACSF can sustain neurophysiological function for a couple of hours, it does not support long-term survival or growth of neurons *in vitro*, even when growth factors and other cell culture supplements are added (unpublished observation). Consequently, neurons are generally differentiated and cultured *in vitro* in standard basal media such as Dulbecco's Modified Eagle's Medium (DMEM) with added supplements (e.g., serum, growth factors, hormones, antioxidants and antibiotics). When examining the calcium activity of human neurons in the medium in which they had been cultured (DMEM/F12+supplements), we surprisingly found very few neurons to be spontaneously active – in contrast to recordings performed in ACSF, in which many more cells showed spontaneous activity. Even after omitting all supplements, and matching the inorganic salt concentration, pH and osmolarity of our ACSF to those in DMEM, we established that cells exhibit low calcium activity in DMEM basal

medium compared to ACSF solution. This impairment in neurophysiological activity could be rescued by removal of amino acids, vitamins or extra components from the complete DMEM solution, indicating that DMEM basal contains various components that (acutely) interfere with neuronal function. Further characterization of the calcium activity in ACSF revealed that both the generation and network propagation of calcium signals are independent of the activation of ionotropic glutamate and GABA receptors, suggesting that these processes are mediated by neurotransmitters other than glutamate and GABA, or that they rely on mechanisms other than synaptic transmission. All in all, the research presented here constitutes a first step toward optimizing media for neuronal cell culture so as to adequately support neurophysiological function and improve successful translation from the laboratory to the clinic.

MATERIALS AND METHODS

Cell Culture

Culturing of Human Embryonic Stem Cell-Derived Neural Progenitor Cells. Human ESC-derived NPCs were cultured on Matrigel (BD Biosciences, Cat No. 356231)-coated wells of 6-well culture plates (CytoOne, Cat No. CC7682-7506) in neural progenitor medium (NPM). NPM was composed of DMEM/F12 + GlutaMAX™ basal medium (Gibco, Cat No. 10565-018) supplemented with 1x N2 (Gibco, Cat No. 17502-048), 1x B27 (Gibco, Cat No. 17504-044), Fibroblast Growth Factor 8 (FGF8b, 100 ng/ml; Peprotech, Cat No. 100-25), Sonic Hedgehog (SHH, 200 ng/ml; R&D Systems, Cat No. 1314SH) and laminin (1 µg/ml; Invitrogen, Cat No. 23017-015). Cells were fed every other day with new, 37 °C pre-warmed. About once a week, or when grown confluent, NPCs were passaged and re-plated onto Matrigel-coated plates.

Differentiation of Neural Progenitor Cells into Neurons. NPCs maintained in NPM were enzymatically dissociated using TrypLE Express (Invitrogen, Cat No. 12604-013). Subsequently, the cells were pelleted through centrifugation and resuspended in neural differentiation medium (NDM) to induce differentiation. NDM was composed of DMEM/F12 + GlutaMAX™ basal medium (Gibco, Cat No. 10565-018) supplemented with 1x N2 (Gibco, Cat No. 17502-048), 1x B27 (Gibco, Cat No. 17504-044), Brain-Derived Neurotrophic Factor (BDNF, 20 ng/ml; Peprotech, Cat No. 450-02), Glial-Derived Neurotrophic Factor (GDNF, 20 ng/ml; Peprotech, Cat No. 450-10), Ascorbic Acid (AA, 200 nM; Sigma, Cat No. A0278), dibutyryl cyclic AMP (cAMP, 1 mM; Sigma, Cat No. D0627) and laminin (1 µg/ml; Invitrogen, Cat No. 23017-015). Following trituration to dissociate NPC clumps, cells were plated on 12-mm round glass coverslips (Fisher Scientific, Cat. No. 12-545-80) in 24-well culture plates (CytoOne, Cat No. CC7682-7524) at a density of $\sim 10^4$ cells per cm^2 . The coverslips used were coated with 10 µg/ml poly-L-

ornithine (Sigma, Cat. No. P-3655) and 5 µg/ml laminin (Invitrogen, Cat. No. 23017-015) to promote adherence of the differentiating NPCs. Following induction of neuronal differentiation, cells were given a 4-day rest period during which they were left undisturbed in a humidified incubator at 37 °C and 5% CO₂, allowing them to adhere. Neurons were subsequently fed every other day by replacement of half (~0.5 ml) of the media with new, 37 °C pre-warmed NDM, minimizing fluid perturbations. Cells received double the amount of media (~1 ml) over the weekend. Cultures were regularly checked under the microscope for their neuronal characteristics as well as signs of possible contamination.

Calcium Imaging

Preparation of Fluo-4 AM Loading Stock Solution. A 1 mM loading stock solution of the calcium-sensitive fluorescent dye Fluo-4 AM was prepared by dissolving 50 µg of Fluo-4 AM (Invitrogen, Cat No. F14201) with 45.6 µl of 20% Pluronic F-127 solution in DMSO (Invitrogen, Cat No. P-3000MP).

Fluo-4 AM Pretreatment (Loading) of Cells. For each calcium imaging experiment, an individual coverslip containing neurons was transferred to a new 24-well culture plate. ~0.5 ml of NDM covering the cells was aspirated and transferred to the new plate after which the coverslip was grasped and transferred into this medium using a fine, sterile forceps rinsed in sterile H₂O. 2 µl of Fluo-4 AM loading stock solution was added to the media containing the transferred coverslip to reach a 4 µM final concentration of the calcium indicator dye. To achieve an even distribution of Fluo-4 AM in the media surrounding the cells, the dye was gently but well resuspended in the media. Cells were incubated with Fluo-4 AM for ~20 minutes under standard tissue culture conditions (37 °C, 5% CO₂, and 95% humidity).

Imaging of Spontaneous Calcium Activity. For live imaging of the neuronal cultures, coverslips containing cells loaded with Fluo-4 AM were individually bathed in a custom-designed plastic perfusion chamber (Warner Instruments). The use of a perfusion pump (Bio-Rad, Econo Pump) allowed for a constant fresh supply of perfusate and provided an easy means of subjecting the cells to different media and neuropharmacological antagonist solutions for the purpose of real-time recording of calcium activity in different conditions. All perfusates were continuously gassed with carbogen (95% O₂ and 5% CO₂) and pre-heated to 35 °C using a dual channel temperature controller with in-line solution heater (Warner Instruments, Model No. TC-344B) before entering the chamber bath. To wash out excess Fluo-4 AM dye, the cultures were perfused with perfusate for at least 15 minutes prior to movie acquisition with the multi-photon microscope system (Olympus, Fluoview FV1000MPE). After movie recording in first (control) perfusate, cells were subjected to a different medium or drug solution, after which the perfusion solution was typically changed back again to the control perfusate in an attempt to recover from the medium or drug effect; perfusate flow rate was maintained at 0.50 or 1.00 ml/minute

during all experiments. Two consecutive 5-minute movies were recorded in each condition (control, medium/drug, and recovery), and one to three fields of view (FOVs) were analyzed per coverslip. Movies were acquired using a 40X (Olympus, LUMPLFLN NA 0.80) or 25X (Olympus, XLPLN NA 1.05) objective, and the imaging parameters applied were typically as follows: FOV size = 506.94 x 506.94 μm , frame rate = 1.8 Hz, scan-matrix = 256 x 256 pixels, scan rate = 4.0 $\mu\text{s}/\text{Pixel}$. For each FOV, a high-resolution image was also acquired at 2048 x 2048 pixels.

Imaging of Stimulation-Evoked Calcium Activity. A similar imaging setup was used for imaging of calcium activity evoked by electrical stimulation, except that a glass pipette stimulation electrode was positioned close to the neurons and approximately in the middle of the FOV; a reference electrode was placed in the chamber bath. The first 60 seconds consisted of recording of spontaneous activity after which two stimulations (amplitude: 200 μA ; pulse duration: 100 μs) were applied 60 seconds apart, at time points 60 and 120 seconds. Different antagonist solutions were perfused to evaluate their effects on induced calcium activity.

Solutions. In our initial experiments showing reduced spontaneous activity of cells in neural differentiation medium (NDM), control/recovery and experimental perfusates were standard ACSF (124 mM NaCl, 3 mM KCl, 1.3 mM MgSO_4 , 26 mM NaHCO_3 , 1.25 mM NaH_2PO_4 , 20 mM glucose, 2 mM CaCl_2 ; pH adjusted to 7.4 with 95% O_2 and 5% CO_2) and NDM, respectively, or vice versa. To exclude the possibility that the low spontaneous activity in NDM versus standard ACSF was due to a difference in ionic composition between the two solutions, we repeated our experiments with a matched ACSF and DMEM (i.e., NDM without N2/B27 supplements and growth factors) solution having the exact same concentration of inorganic salts, pH and osmolarity. In an attempt to pinpoint the component or components of the complete DMEM that are specifically responsible for the low spontaneous activity of cells in this medium, we performed calcium imaging experiments with custom-made DMEM media devoid of either amino acids, vitamins or extra components (see **Supplementary Table 1** for a detailed specification of DMEM composition/components). In another series of experiments, different neuropharmacological antagonists were consecutively added to control ACSF perfusate to explore in more detail the mechanisms that mediate the spontaneous firing as well as the propagation of stimulation-induced calcium signals within the neuronal cultures: kynurenic acid (KYN, 5.0 mM; Sigma, Cat No. K3375-5G) to non-specifically block ionotropic glutamate receptors, gabazine (SR95531, 10 μM ; Sigma, Cat No. S106) to selectively and competitively block GABA_A receptors, tetrodotoxin (TTX, 1.0 μM ; Tocris, Cat No. 1078) to block voltage-gated sodium channels, and nifedipine (NIFE, 10 μM ; Tocris, Cat No. 1075) to selectively block L-type calcium channels. All solutions were stored at 4 $^\circ\text{C}$, but were warmed to ~ 25 $^\circ\text{C}$ prior to initiation of recordings. **Table 1** summarizes the different perfusates used in each experimental paradigm.

Table 1. Summary of perfusates used in the different calcium imaging experiments.

Experiment	Control perfusate	Experimental perfusate #1	Experimental perfusate #2	Experimental perfusate #3	Recovery perfusate
1 Spont activity ACSF vs NDM <i>standard, non-matched</i>	ACSF-std	NDM-std			ACSF-std
2 Spont activity ACSF vs DMEM <i>custom-made, matched</i>	ACSF*	DMEM*			ACSF*
	ACSF*	DMEM no AA*			ACSF*
	ACSF*	DMEM no vit*			ACSF*
	ACSF*	DMEM no ext*			ACSF*
	DMEM*	ACSF*			DMEM*
	DMEM*	DMEM no AA*			DMEM*
	DMEM*	DMEM no vit*			DMEM*
3 Spont activity Effect of antagonists	ACSF-std	+KYN +SR95531	+KYN +SR95531 +TTX	+KYN +SR95531 +TTX +NIFE	No recovery
4 Induced activity Effect of antagonists	ACSF-std	+KYN +SR95531	+KYN +SR95531 +TTX	+KYN +SR95531 +TTX +NIFE	No recovery

*: Solutions having matched inorganic salt concentration, pH and osmolarity.

ACSF: Artificial cerebrospinal fluid; DMEM: Dulbecco's Modified Eagle's Medium; KYN: Kynurenic acid (5.0 mM); NDM: Neural differentiation medium; NIFE: Nifedipine (10 μ M); no AA: without amino acids; no ext: without extra components; no vit: without vitamins; SR95531: Gabazine (10 μ M); std: standard; TTX: Tetrodotoxin (1.0 μ M).

Data Analysis

Quantification and Characterization of Spontaneously Active Cells. Cells showing a positive change in Fluo-4 AM fluorescence intensity from baseline were selected as regions of interest (ROIs) for time series analyses with the Olympus Fluoview FV1000 software. For paired analysis between conditions, ROIs selected in one particular movie were copied to all other movies corresponding to the same FOV, regardless of the recording condition. This way, any given ROI represents a cell that is active in at least one of the different conditions. When unpaired analyses were performed, we treated the data as if the different recordings were independent and thus ROIs were selected for any of the conditions independently. The number of active ROIs was determined for each FOV and then averaged across FOVs of the same condition to obtain a quantitative measure of the activity effect of the different perfusates. At the same time, all cells were characterized according to their type(s) of spontaneous calcium signals (see **RESULTS** for description). Except for our initial experiments on spontaneous activity in NDM versus standard, non-matched ACSF, all analysis was performed blinded to the recording condition with the code being broken only after quantification and characterization of active cells was complete.

Dynamics and Kinetics Analysis of Individual Calcium Transients. The events of cells with fast spontaneous calcium spikes (i.e., “unitary” responses) were further analyzed with MiniAnalysis software (Synaptosoft, Version 6.0.3) for calculation of spike frequency, average amplitude, and rise and decay kinetics (10-90% rise time, 100-37% decay time,

area, halfwidth and 10-90% slope). We applied the following settings for peak detection: threshold = 10 pA; period to search a local maximum = 30,000,000 μ s; period to search a decay time = 30,000,000 μ s; fraction of peak to find a decay time = 0.37; period to average a baseline = 1,000,000 μ s; area threshold = 200,000; number of points to average for peak = 3; direction of peak = positive. Time before a peak for baseline was usually set to 20,000,000 μ s, but was allowed to vary between 0-30,000,000 μ s if that improved baseline measurement. Rise and decay kinetics were not determined for partial events occurring at the immediate beginning or very end of a 5-minute recording. IGOR Pro (Wavemetrics, Version 6) was used for graphical presentation of the calcium traces following merging and baseline correction of the Fluoview time series data. For import of the calcium traces into MiniAnalysis, individual IGOR wave files were converted to axon binary file (ABF) formats using the ABF Utility program (Synptosoft). For paired analysis, we compared the calcium transients of all cells with unitary responses in the control condition to their corresponding events in the other conditions. Non-paired analysis was performed if the number of cells with unitary events in the control condition was too small for a proper paired comparison.

Statistics

Data in all figures represent group means \pm S.E.M. values, and unless otherwise stated, *n* designates the number of FOVs analyzed per condition. Non-parametric Wilcoxon tests and Mann-Whitney tests were used, respectively, for comparison of paired and non-paired group data, with differences between groups being considered significant for *P* values < 0.05. GraphPad Prism (GraphPad Software Inc, Version 5.0) was used for all statistical analysis and graph presentation.

RESULTS

Cells Exhibit Low Spontaneous Calcium Activity in Standard DMEM compared to Matched ACSF

Initial calcium imaging experiments from our group revealed that the number of spontaneously active neurons is significantly lower in recordings obtained in neural differentiation medium (NDM) compared to standard, non-matched ACSF (data not shown). This difference in spontaneous activity is not due to differences in ionic composition between the two solutions as similar results were obtained with a custom-made ACSF and DMEM solution having the exact same (matched) concentration of inorganic salts, pH and osmolarity. Whereas the number of active cells was lower in calcium imaging recordings obtained in complete DMEM compared to matched control ACSF, spontaneous activity appropriately recovered after reperfusion with ACSF perfusate (**Figure 1**, same as Figure 1B of CHAPTER 3B).

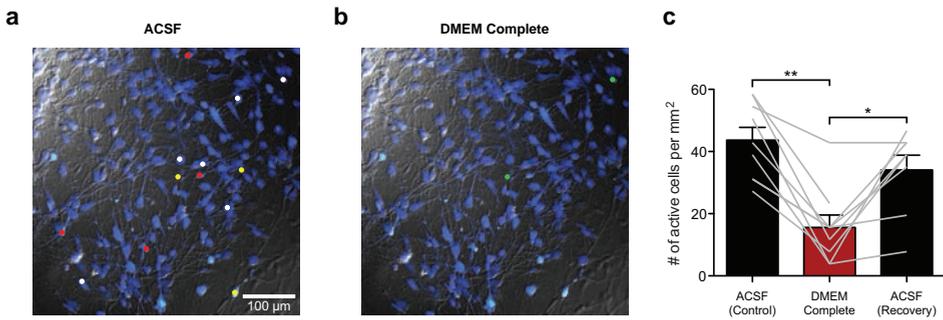


Figure 1. Low spontaneous calcium activity in standard DMEM compared to matched ACSF. (a,b) Representative images of calcium activities recorded in the different perfusates. The number of spontaneously active cells, indicated by filled contours, decreased in recordings obtained in complete DMEM (**b**) compared to control ACSF (**a**), and showed appropriate recovery in ACSF perfusate (**a**): yellow filled contours indicate cells active in control ACSF ($n = 3$), white filled contours indicate cells active in recovery ACSF ($n = 6$), and red filled contours represent cells active in both control and recovery ACSF ($n = 4$); green-filled contours are active cells in complete DMEM ($n = 2$). (**c**) The number of spontaneously active cells was significantly lower in complete DMEM compared to both ACSF control and recovery. Bars represent means \pm S.E.M.; gray lines show results from individual FOVs. * $P < 0.05$, ** $P < 0.01$; Wilcoxon matched pairs test.

Removal of Amino Acids, Vitamins or Extra Components from Complete DMEM Rescues Low Calcium Activity

To investigate which solution component or components are responsible for the low spontaneous calcium activity in complete DMEM, we conducted calcium imaging experiments with various custom-made DMEM media depleted of either amino acids, vitamins or extra components (collectively abbreviated DMEM without*). Whereas complete DMEM reduced the number of spontaneously active cells in ACSF by $64.4 \pm 9.2\%$ ($n = 9$ FOVs) (**Figure 2a**), the three DMEM without* solutions did not ($n = 2$ FOVs per solution) (**Figure 2b**). In fact, the number of spontaneously active cells tended to be higher in the three DMEM without* media compared to ACSF ($P = 0.0625$; **Figure 2b**; see **Supplementary Figure 1a** for the effect of the individual media).

Relative to complete DMEM, spontaneous calcium activity increased by $156.1 \pm 46.3\%$ for the three DMEM without* media together, and recovered back to control levels after washout with complete DMEM ($n = 12$ total FOVs, $n = 4$ FOVs per custom-made solution) (**Figure 2c**; see **Supplementary Figure 1b** for the effect of the individual media). Cells showing one or more positive changes in Fluo-4 AM fluorescence intensity over time were identified as spontaneously active cells (**Figure 2d**, orange contours). As an example, spontaneous calcium activity was much more apparent in DMEM depleted of amino acids ($n = 21$ active cells) compared to complete DMEM ($n = 6$ and 9 active cells for control and recovery conditions, respectively) (**Figure 2e**). However, custom-made DMEM solution with the addition of standard supplements but without amino acids failed to sustain neuronal survival and maturation for more than one week (see Supple-

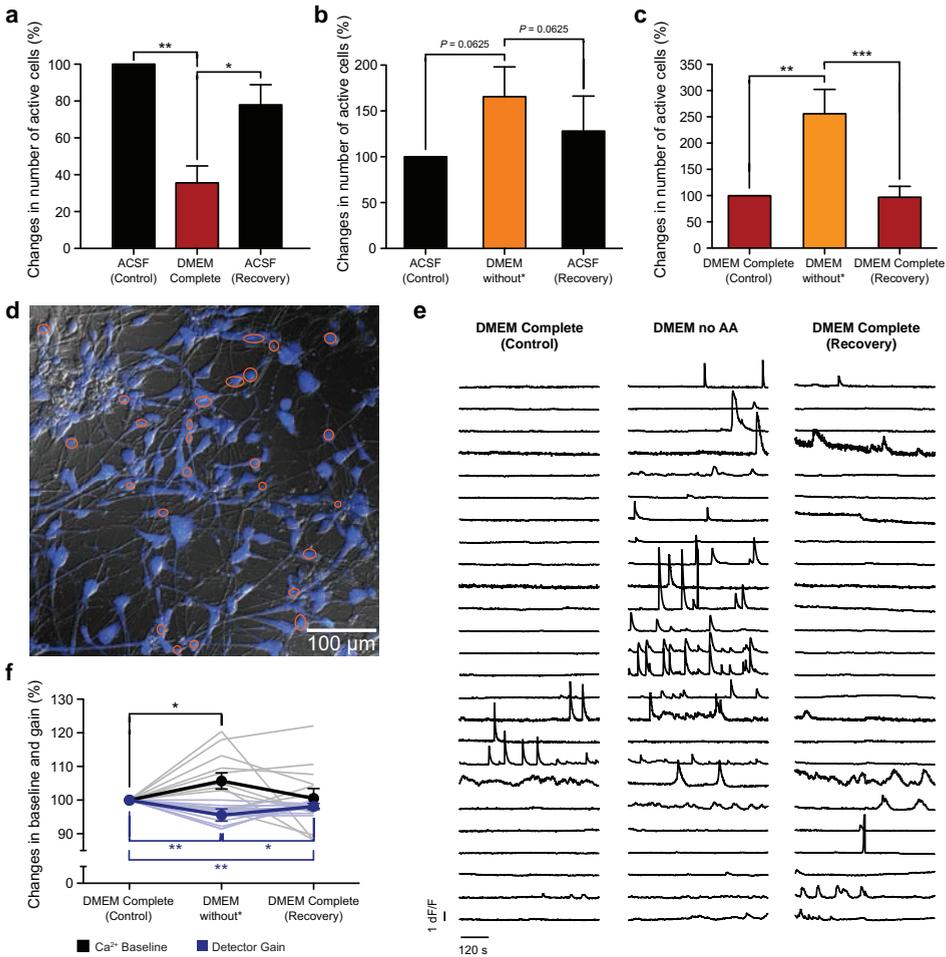


Figure 2. Removal of amino acids, vitamins or extra components rescues the particularly low calcium activity in complete DMEM. (a) Relative to ACSF, spontaneous calcium activity decreased more than two-fold in complete DMEM ($n = 9$ FOVs), and recovered after switching perfusion back to ACSF ($n = 8$ FOVs). (b) This reduction in spontaneous activity was not apparent in custom-made DMEM media depleted of amino acids, vitamins or extra components (DMEM without*, $n = 6$ FOVs); in contrast, there was a tendency ($P = 0.06$) towards an increase in spontaneous activity compared to ACSF. (c) Relative to complete DMEM, spontaneous calcium activity more than doubled in DMEM without* media ($n = 12$ FOVs). (d) Fluo-4 AM fluorescence and DIC overlay image of an example neuronal population subjected to complete DMEM and DMEM without amino acids. (e) Calcium traces from the 25 spontaneously active cells indicated in (d) by orange contours, measured as relative changes in fluorescence (dF/F) over time. (f) Changes in calcium baseline and photomultiplier (PMT) detector gain in DMEM without* versus complete DMEM conditions. To maintain constant baseline fluorescence, PMT voltage had to be significantly reduced upon perfusate change from complete DMEM to DMEM without* ($n = 12$ FOVs). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; Wilcoxon matched pairs tests.

mentary Figure 2C in CHAPTER 3B). Taken together, these findings indicate that amino acids in DMEM, as well as various other basal medium constituents, acutely interfere with neurophysiological function.

Resting Level of Intracellular Calcium is Low in Complete DMEM

We then asked whether the resting level of intracellular calcium was different between solutions. To investigate this, we analyzed changes in calcium baseline and photomultiplier (PMT) detector gain between recordings in complete DMEM and DMEM without* media. Significant downregulation of the PMT voltage was necessary to prevent the calcium baseline from increasing by more than 10% when switching perfusate from complete DMEM to DMEM without* ($n = 12$ FOVs) (Figure 2f). This finding, that amplification of fluorescence had to be significantly reduced to maintain a constant level of baseline fluorescence between conditions, suggests that cells have substantially lower levels of intracellular calcium in complete DMEM compared to DMEM without* medium.

Cells With Different Types of Spontaneous Calcium Events Become Inactive in Complete DMEM

Different types of spontaneous calcium signals could be identified through time-series analysis on selected ROIs. Two main categories of spontaneous calcium transients were distinguished, termed “calcium spikes” (fast rising phase events) and “calcium waves” (slow rising phases), and could be further subcategorized as follows (Figure 3a). Clean, single-peaked calcium transients having a fast rise and a complementary, relatively fast decay were characterized as “unitary” spikes, whereas events showing multiple fast and clean peaks before returning back to baseline level were referred to as “multipeak” responses. “Bursts” of spikes, typically recognized as multiple consecutive events having very short (commonly < 10 ms) inter-event intervals [12], were found to be highly irregular. We classified multipeak and burst activities together as “complex” events given the repetitive nature of both these types of responses. Among the calcium waves, “slow” events were typically identified as transients with a slow onset and complementary slow decay. Transients of cells with both complex and slow events were characterized as “mixed”, and calcium waves that could not be recognized as belonging to one of the abovementioned groups were classified as “slow undefined”; this in contrast with “inactive” calcium traces that did not show any rise over the course of two 5-minute movies.

Using this method of characterization of calcium signals, we studied the changes in distribution of cells with different types of spontaneous calcium transients between ACSF and complete DMEM perfusate (Figure 3b). Because there was no clear difference between ACSF and complete DMEM in the overall distribution of cells with different calcium event types (data from $n = 9$ FOVs) (Figure 3b₁), we excluded all cells with slow and undefined calcium waves since these types of calcium signals are likely to be

generated by astrocytes rather than by neurons [13-15]. Limiting our selection to cells with calcium spikes and mixed responses in ACSF, we found that $85.2 \pm 4.8\%$ of these cells became inactive in complete DMEM (**Figure 3b₂**). No clear changes were found in the distribution of cell types between complete DMEM and the three DMEM without* media ($n = 12$ FOVs) (**Figure 3c**).

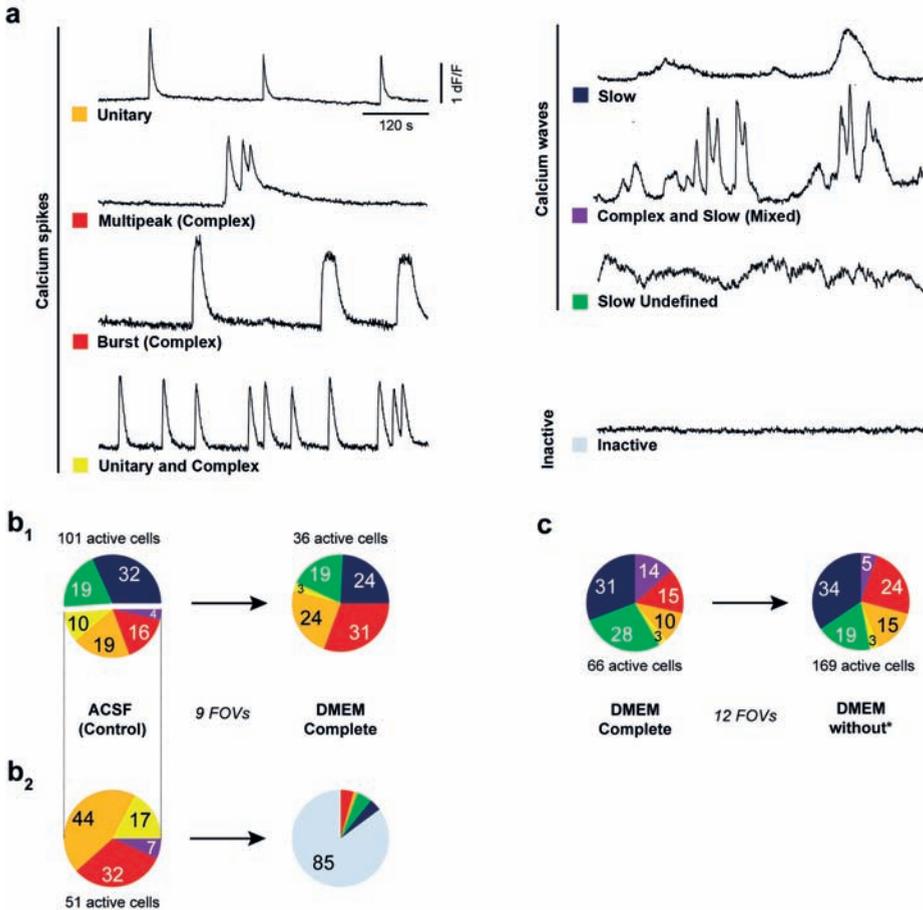


Figure 3. Types of spontaneous calcium signals recorded in the different perfusates. (a) Characterization of different types of spontaneous calcium transients (see **Results** for complete description). **(b,c)** Distribution of cells with different types of spontaneous calcium signals in the different solutions. **(b₁)** There was no clear difference between ACSF (*left*) and complete DMEM (*right*) in overall distribution of cells with different calcium transient types. **(b₂)** Excluding all cells with slow and undefined responses (which typically represent astroglial calcium signals) in ACSF (*left*), we found that the vast majority of cells classified as unitary, complex and mixed, became inactive in complete DMEM (*right*). **(c)** Distribution of cell types was not different between the various DMEM without* media (*right*) and complete DMEM (*left*). Values in pies represent average percentages rounded to the nearest whole percent.

Blockage of Ionotropic Glutamate and GABA Receptors Does Not Lower the Frequency of Unitary Calcium Events

To investigate in more detail the mechanisms mediating spontaneous firing in the neuronal cultures, we attempted to modulate spontaneous calcium activity by subjecting the cells to different neuropharmacological antagonists. Whereas consecutive addition of the glutamate receptor antagonist kynurenic acid (KYN), the GABA_A receptor antagonist gabazine (SR95531), the voltage-gated sodium channel blocker tetrodotoxin (TTX) and the L-type calcium channel blocker nifedipine (NIFE) did not affect the total number of spontaneously active cells ($n = 4$ FOVs) (Figure 4a), we were able to observe a clear reduction in activity upon exclusion of all cells with slow and undefined calcium waves. Of the cells with calcium spikes and mixed responses in ACSF, an increasing percentage became inactive with sequential addition of the drugs (Figure 4b and Supplementary Table 2). In addition, cells with slow and undefined calcium signals tended to constitute a greater percentage of total active cells after addition of TTX (Figure 4a, insets; Supplementary Table 2).

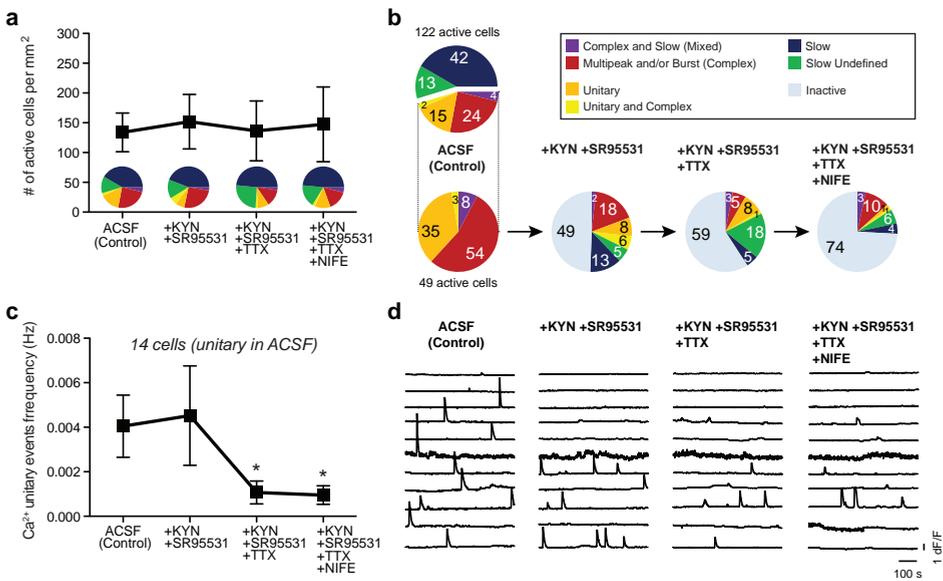


Figure 4. Influence of various neuropharmacological antagonists on spontaneous calcium activity.

(a) Consecutive addition of kynurenic acid (KYN), gabazine (SR95531), tetrodotoxin (TTX) and nifedipine (NIFE) to the ACSF perfusate did not reduce the total number of spontaneously active cells ($n = 4$ FOVs). However, after addition of TTX, cells with slow and undefined calcium signals tended to constitute a greater portion of total active cells (*insets*). (b) Of the cells with responses classified as unitary, multipeak, burst and mixed in ACSF, an increasing number became inactive with consecutive addition of antagonists (*bottom*). Values in pies represent average percentages rounded to the nearest whole percent. (c) The frequency of events in cells with unitary responses in ACSF became significantly lower after addition of TTX with or without NIFE. * $P < 0.05$; Wilcoxon matched pairs tests. (d) Spontaneous calcium traces in different antagonist conditions, measured as relative changes in Fluo-4 AM fluorescence over time.

Because fast fluorescence calcium dynamics presumably reflect neuronal calcium events [14, 15], we further investigated the influence of the antagonists on the frequency of events in cells ($n = 14$) with unitary calcium spikes in ACSF. Whereas addition of TTX with or without nifedipine significantly reduced the unitary events frequency in these cells, application of KYN and SR95531 alone had no effect (**Figure 4c,d** and **Supplementary Table 2**), suggesting that the generation of spontaneous neuronal activity is independent of the activation of glutamatergic and GABAergic receptors, but relies on intact transmembrane flux of sodium.

Neuropharmacological Antagonists Diminish the Magnitude and Amplitude of Stimulation-Induced Calcium Signals

To evaluate the effect of KYN, SR95531, TTX and NIFE on evoked calcium activity, as opposed to spontaneous activity, we positioned a glass pipette stimulation electrode in the middle of the FOV (**Figure 5a**) to apply two consecutive stimulations (amplitude: 200 μ A, pulse duration: 100 μ sec, inter-stimulus interval: 60 sec) to the neuronal cultures. Whereas electrical stimulation readily induced calcium responses in control ACSF perfusate, consecutive addition of KYN plus SR95531 and TTX markedly affected both the total number of cells responding to stimulation as well as the amplitude of induced responses (**Figure 5b-d** and **Supplementary Table 2**). Although perfusion with KYN plus SR95531 had only a small effect on the average evoked calcium response (bold black line in **Figure 5b**), virtually no calcium transients (thin gray lines in **Figure 5b**; $n = 20$ total cells) could be induced in presence of TTX. Clearly, we observed a marked decrease in both the number of responding cells (**Figure 5c**, $n = 30$ total cells) and the amplitude of evoked calcium transients (**Figure 5d**) with consecutive blockage of ionotropic glutamate receptors, GABA_A receptors, and sodium channels (data from $n = 2$ FOVs).

Blockage of Voltage-Gated Sodium Currents by TTX Impairs Functional Connectivity within Neuronal Networks

In order to characterize the mechanisms of propagation of calcium signals within the cultures, we determined how the aforementioned neuropharmacological drugs affect the calcium response at various distances from the stimulation electrode. Interestingly, ACSF containing KYN plus SR95531 and TTX, but not KYN plus SR95531 alone, completely blocked calcium transients in cells located at distances farther than 60 μ m from the electrode tip ($n = 2$ FOVs) (**Figure 5e,f** and **Supplementary Table 2**). Because we observed no change in the spread of the calcium response across the network with KYN plus SR95531 in the absence of TTX, we can conclude that voltage-gated sodium currents are specifically involved in propagating neuronal activity across the network.

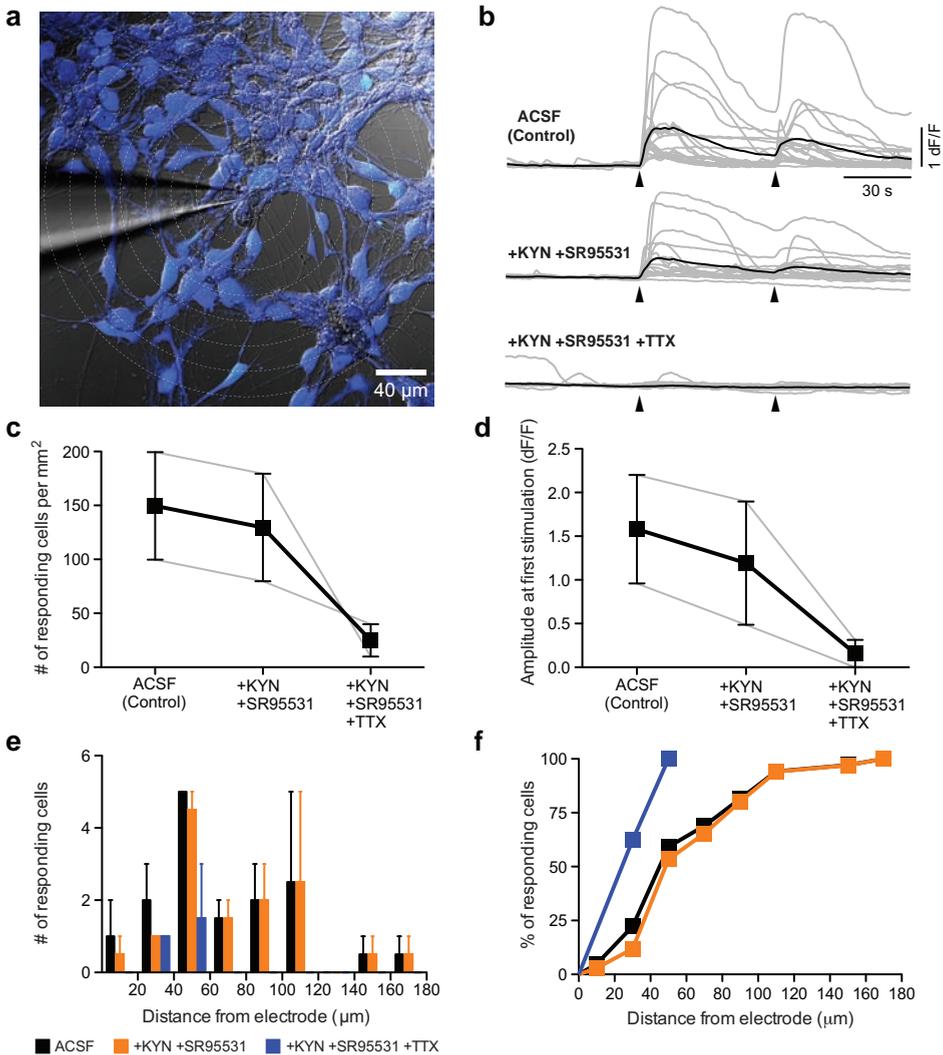


Figure 5. Influence of various neuropharmacological antagonists on stimulation-evoked calcium activity. (a) Fluorescence and DIC overlay image of an example neuronal population subjected to electrical stimulation. Drawn concentric circles are 20 µm apart. (b) Calcium responses recorded in the different antagonist solutions (thin gray lines indicate responses from individual cells [$n = 20$]; bold black lines indicate average responses). Two stimulations (amplitude: 200 µA, pulse duration: 100 µs; arrowheads) were applied 60 seconds apart. Induced responses decreased both in number and amplitude in perfusate containing KYN plus SR95531, and almost no calcium activity could be evoked in presence of TTX. (c,d) The number of responding cells and the amplitude of evoked responses at first stimulation was only slightly reduced by KYN plus SR95531, but was markedly lower when TTX was added. Gray lines show results from individual FOVs. (e,f) ACSF containing KYN plus SR95531 and TTX, but not KYN plus SR95531 alone, completely blocked calcium responses in cells located at distances >60 µm from the stimulation electrode.

DISCUSSION

In the present study, we examined spontaneous and induced calcium activity patterns in an *in vitro* model of neuronal networks derived from human ESCs. Our main finding is that DMEM, a basal medium widely used for the growing of human neuronal cultures, acutely impairs spontaneous neural activity. Removal of amino acids, vitamins or extra components from the complete DMEM solution rescued the low calcium activity of the cells; yet, in the absence of amino acids, long-term survival and maturation of the neurons *in vitro* could not be sustained. *In vivo*, spontaneous electrical activity is essential for the formation and maturation of neuronal networks [5-7] as well as for survival of developing neurons [16]. Therefore, the development of a specific neuronal medium that adequately supports neuronal activity while sustaining long-term maturation (see CHAPTER 3B) will result in a better mimicking of brain physiological conditions and thereby improve the clinical relevance of *in vitro* experimental models. Especially in the context of *in vitro* disease modeling using patient-specific neurons derived from induced pluripotent stem cells (iPSCs), neuronal models that more closely resemble the living brain will be more likely to recapitulate the dysfunctions occurring in patients' brains, enabling the discovery of new effective therapeutic targets and drugs for the treatment of neurological and psychiatric disease.

Which Mechanisms Mediate The Generation and Cell-to-Cell Propagation of Neuronal Calcium Activity?

To gain further insight into the mechanisms underlying the spontaneous firing in the neuronal cultures, we subjected the cells to neuropharmacological modulation in an attempt to alter spontaneous activity. Here, we show that ACSF perfusate containing KYN, SR95531 and TTX reduced the frequency of spontaneous calcium events in cells with unitary responses. That the unitary events frequency did not change with addition of KYN and SR95531 alone suggests that the generation of this type of spontaneous activity is independent of the activation of glutamatergic and GABAergic receptors. The significant reduction in the unitary events frequency with TTX however indicates that voltage-gated sodium currents are specifically involved in the generation of the spontaneous calcium activity. It has become apparent that many neurons in the central nervous system have intrinsic firing properties, enabling them to fire spontaneous action potentials in the absence of synaptic excitation [17-20]. These so-called autorhythmic or pacemaker-like neurons have been proposed to trigger the synchronous depolarizing activity in developing neural circuits [21]. The capability of neurons to generate intrinsic spontaneous activity is dependent, at least partly, on voltage-gated inward sodium currents [22]. Further supporting this notion, a persistent, subthreshold and TTX-sensitive sodium current has been shown to drive the autorhythmic firing of neurons [23, 24]. Taken together, these

findings raise the possibility that the spontaneous calcium oscillations that we monitored in our imaging experiments are of intrinsic origin.

We also examined the effects of KYN, SR95531 and TTX on electrical stimulation-induced activity and the propagation of evoked calcium responses across the neural network. Here we not only demonstrate that virtually no calcium activity could be evoked in presence of TTX, but we also show that addition of KYN and SR95531 alone (i.e., in the absence of TTX) did not impair the propagation of induced calcium signals. This finding suggests that the calcium responses in the neuronal cultures spread from cell to cell mainly through mechanisms other than synaptic transmission. One such mechanism might be the electrical (gap-junctional) coupling of neuronal clusters that allows for a direct transfer of ions between neurons [25-27]. Future experiments with blockers of gap junction transmission should reveal to what extent electrical coupling mediates communication between neurons in the network.

Do Spontaneous Neuronal Calcium Signals Differ From Those Generated By Astrocytes?

We characterized different types of spontaneous calcium signals in the *in vitro* neuronal cultures. These calcium transients differ in complexity as well as the time course of activity. In comparing between perfusates the distribution of cells with different types of spontaneous calcium transients, we excluded from analysis all cells with slow and undefined responses since these types of events are unlikely to represent neuronal activity. Whereas neuronal calcium transients can be characterized by fast, spike-like changes in relative fluorescence (dF/F), responses with a slow time course are more typical for astrocytes [13-15]. Although differentiation of the human ESCs was directed towards the neuronal lineage, the neuronal cultures also contained a small number of astrocytes as revealed by immunohistochemical staining for the astrocyte-specific marker glial fibrillary acid protein (GFAP) (data not shown). In an attempt to discriminate the astroglial from neuronal calcium dynamics, we combined Fluo-4 AM calcium indicator loading with sulforhodamine 101 (SR101) labeling [14, 28], but staining with this fluorescent marker was not specific to astrocytes in the cultures under study (data not shown). The possibility of the presence of some astrocytes in the cultures warrants further investigation given that we only observed a clear effect of the neuropharmacological antagonists on spontaneous activity after exclusion of all cells with slow and undefined calcium responses. Aguado *et al.* reported that TTX significantly impaired spontaneous activity of neurons, but not astrocytes, in mouse brain hippocampal slices [29]. This specific effect of TTX on neuronal spontaneous activity might explain our observation that cells with slow and undefined calcium responses tended to constitute a greater portion of the total number of active cells after adding TTX. To establish what types of spontaneous calcium activities are truly neuronal, future experiments should be directed at monitoring spontaneous

calcium signals specifically in neurons, for example by combining calcium imaging with electrophysiological analysis of single-neuron activity (patch-clamp recording) and/or by using a genetically encoded calcium indicator (e.g., GCaMP) expressed under a neuron-specific promoter [30, 31].

ACKNOWLEDGMENTS

This work was supported by a travel fellowship from the International Foundation Alzheimer's Research (ISAO) to M.v.d.H.

REFERENCES

1. Clowry G, Molnar Z, Rakic P. Renewed focus on the developing human neocortex. *J. Anat.* 217(4), 276-288 (2010).
2. Dolmetsch R, Geschwind DH. The human brain in a dish: the promise of iPSC-derived neurons. *Cell* 145(6), 831-834 (2011).
3. Mahmood A, Napoli C, Aldahmash A. In vitro differentiation and maturation of human embryonic stem cell into multipotent cells. *Stem Cells Int.* 2011 735420 (2011).
4. Itskovitz-Eldor J, Schuldiner M, Karsenti D *et al.* Differentiation of human embryonic stem cells into embryoid bodies comprising the three embryonic germ layers. *Mol. Med.* 6(2), 88-95 (2000).
5. Katz LC, Shatz CJ. Synaptic activity and the construction of cortical circuits. *Science* 274(5290), 1133-1138 (1996).
6. Ben-Ari Y. Developing networks play a similar melody. *Trends Neurosci.* 24(6), 353-360 (2001).
7. Zhang LI, Poo MM. Electrical activity and development of neural circuits. *Nat. Neurosci.* 4 Suppl 1207-1214 (2001).
8. Crepel V, Aronov D, Jorquera I, Represa A, Ben-Ari Y, Cossart R. A parturition-associated nonsynaptic coherent activity pattern in the developing hippocampus. *Neuron* 54(1), 105-120 (2007).
9. Allene C, Cattani A, Ackman JB *et al.* Sequential generation of two distinct synapse-driven network patterns in developing neocortex. *J. Neurosci.* 28(48), 12851-12863 (2008).
10. Abranches E, Silva M, Pradier L *et al.* Neural differentiation of embryonic stem cells in vitro: a road map to neurogenesis in the embryo. *PLoS One* 4(7), e6286 (2009).
11. Cazillis M, Rasika S, Mani S, Gressens P, Lelievre V. In vitro induction of neural differentiation of embryonic stem (ES) cells closely mimics molecular mechanisms of embryonic brain development. *Pediatr. Res.* 59(4 Pt 2), 48R-53R (2006).
12. Debanne D. The nodal origin of intrinsic bursting. *Neuron* 71(4), 569-570 (2011).
13. Hirase H, Qian L, Bartho P, Buzsaki G. Calcium dynamics of cortical astrocytic networks in vivo. *PLoS Biol.* 2(4), E96 (2004).
14. Nimmerjahn A, Kirchhoff F, Kerr JN, Helmchen F. Sulforhodamine 101 as a specific marker of astroglia in the neocortex in vivo. *Nat. Methods* 1(1), 31-37 (2004).
15. Gobel W, Helmchen F. In vivo calcium imaging of neural network function. *Physiology (Bethesda)* 22 358-365 (2007).
16. Golbs A, Nimmervoll B, Sun JJ, Sava IE, Luhmann HJ. Control of programmed cell death by distinct electrical activity patterns. *Cereb. Cortex* 21(5), 1192-1202 (2011).
17. Llinas RR. The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science* 242(4886), 1654-1664 (1988).
18. Raman IM, Bean BP. Ionic currents underlying spontaneous action potentials in isolated cerebellar Purkinje neurons. *J. Neurosci.* 19(5), 1663-1674 (1999).
19. Raman IM, Gustafson AE, Padgett D. Ionic currents and spontaneous firing in neurons isolated from the cerebellar nuclei. *J. Neurosci.* 20(24), 9004-9016 (2000).
20. Puopolo M, Bean BP, Raviola E. Spontaneous activity of isolated dopaminergic periglomerular cells of the main olfactory bulb. *J. Neurophysiol.* 94(5), 3618-3627 (2005).
21. Blankenship AG, Feller MB. Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nat. Rev. Neurosci.* 11(1), 18-29 (2010).
22. Cruz JS, Silva DF, Ribeiro LA *et al.* Resurgent Na⁺ current: a new avenue to neuronal excitability control. *Life Sci.* 89(15-16), 564-569 (2011).

23. Taddese A, Bean BP. Subthreshold sodium current from rapidly inactivating sodium channels drives spontaneous firing of tuberomammillary neurons. *Neuron* 33(4), 587-600 (2002).
24. Do MT, Bean BP. Subthreshold sodium currents and pacemaking of subthalamic neurons: modulation by slow inactivation. *Neuron* 39(1), 109-120 (2003).
25. Maex R, De Schutter E. Mechanism of spontaneous and self-sustained oscillations in networks connected through axo-axonal gap junctions. *Eur. J. Neurosci.* 25(11), 3347-3358 (2007).
26. Sohl G, Maxeiner S, Willecke K. Expression and functions of neuronal gap junctions. *Nat. Rev. Neurosci.* 6(3), 191-200 (2005).
27. Bennett MV, Zukin RS. Electrical coupling and neuronal synchronization in the Mammalian brain. *Neuron* 41(4), 495-511 (2004).
28. Nimmerjahn A, Helmchen F. In vivo labeling of cortical astrocytes with sulforhodamine 101 (SR101). *Cold Spring Harb. Protoc.* 2012(3), 326-334 (2012).
29. Aguado F, Espinosa-Parrilla JF, Carmona MA, Soriano E. Neuronal activity regulates correlated network properties of spontaneous calcium transients in astrocytes in situ. *J. Neurosci.* 22(21), 9430-9444 (2002).
30. Palmer AE, Tsien RY. Measuring calcium signaling using genetically targetable fluorescent indicators. *Nat. Protoc.* 1(3), 1057-1065 (2006).
31. Pologruto TA, Yasuda R, Svoboda K. Monitoring neural activity and [Ca²⁺] with genetically encoded Ca²⁺ indicators. *J. Neurosci.* 24(43), 9572-9579 (2004).

SUPPLEMENTARY INFORMATION

Supplementary Table 1. Formulation of DMEM/F12 + GlutaMAX™ basal culture medium.

Components	Molecular Weight	Concentration (mg/L)	Molarity (mM)
Inorganic Salts			
Calcium Chloride (CaCl ₂) (anhyd.)	111.0	116.6	1.0504504
Cupric Sulfate (CuSO ₄ · 5H ₂ O)	250.0	0.0013	5.2E-6
Ferric Nitrate (Fe(NO ₃) ₃ · 9H ₂ O)	404.0	0.05	1.2376238E-4
Ferric Sulfate (FeSO ₄ · 7H ₂ O)	278.0	0.417	0.0015
Magnesium Chloride (MgCl ₂) (anhyd.)	95.0	28.64	0.30147368
Magnesium Sulfate (MgSO ₄) (anhyd.)	120.0	48.84	0.407
Potassium Chloride (KCl)	75.0	311.8	4.1573334
Sodium Bicarbonate (NaHCO ₃)	84.0	2438.0	29.02381
Sodium Chloride (NaCl)	58.0	6999.5	120.68104
Sodium Phosphate Dibasic (Na ₂ HPO ₄) (anhyd.)	142.0	71.02	0.50014085
Sodium Phosphate Monobasic (NaH ₂ PO ₄ · H ₂ O)	138.0	62.5	0.45289856
Zinc Sulfate (ZnSO ₄ · 7H ₂ O)	288.0	0.432	0.0015
Amino Acids			
Glycine	75.0	18.75	0.25
L-Alanine	89.0	4.45	0.049999997
L-Alanyl-L-Glutamine	217.0	542.0	2.497696
L-Arginine Hydrochloride	211.0	147.5	0.69905216
L-Asparagine-H ₂ O	150.0	7.5	0.05
L-Aspartic Acid	133.0	6.65	0.05
L-Cysteine Hydrochloride-H ₂ O	176.0	17.56	0.09977272
L-Cystine 2HCl	313.0	31.29	0.09996805
L-Glutamic Acid	147.0	7.35	0.05
L-Histidine Hydrochloride-H ₂ O	210.0	31.48	0.14990476
L-Isoleucine	131.0	54.47	0.41580153
L-Leucine	131.0	59.05	0.45076334
L-Lysine Hydrochloride	183.0	91.25	0.4986339
L-Methionine	149.0	17.24	0.11570469
L-Phenylalanine	165.0	35.48	0.2150303
L-Proline	115.0	17.25	0.15
L-Serine	105.0	26.25	0.25
L-Threonine	119.0	53.45	0.44915968
L-Tryptophan	204.0	9.02	0.04421569
L-Tyrosine Disodium Salt Dihydrate	261.0	55.79	0.21375479

Supplementary Table 1. Formulation of DMEM/F12 + GlutaMAX™ basal culture medium. (continued)

Components	Molecular Weight	Concentration (mg/L)	Molarity (mM)
L-Valine	117.0	52.85	0.4517094
Vitamins			
Biotin	244.0	0.0035	1.4344263E-5
Choline Chloride	140.0	8.98	0.06414285
D-Calcium Pantothenate	477.0	2.24	0.0046960167
Folic Acid	441.0	2.65	0.0060090707
Niacinamide	122.0	2.02	0.016557377
Pyridoxine Hydrochloride	206.0	2.031	0.009859223
Riboflavin	376.0	0.219	5.824468E-4
Thiamine Hydrochloride	337.0	2.17	0.0064391694
Vitamin B12	1355.0	0.68	5.0184503E-4
i-Inositol	180.0	12.6	0.07
Extra Components			
D-Glucose (Dextrose)	180.0	3151.0	17.505556
Hypoxanthine Na	159.0	2.39	0.015031448
Linoleic Acid	280.0	0.042	1.4999999E-4
Lipoic Acid	206.0	0.105	5.097087E-4
Phenol Red	376.4	8.1	0.021519661
Putrescine 2HCl	161.0	0.081	5.031056E-4
Sodium Pyruvate	110.0	55.0	0.5
Thymidine	242.0	0.365	0.0015082645

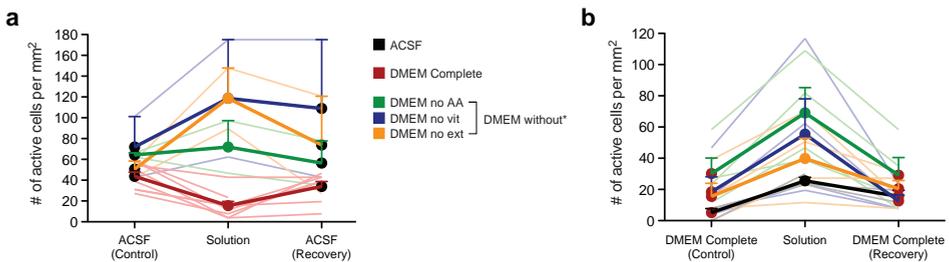
Supplementary Table 2. Quantitative effects of the different neuropharmacological antagonists on spontaneous and induced calcium activity.

	ACSF (Control)	+KYN +SR95531	+KYN +SR95531 +TTX	+KYN +SR95531 +TTX +NIFE
Spontaneous Activity				
# of active cells per mm ²	134 ± 33 (4)	152 ± 46 (4)	136 ± 50 (4)	147 ± 63 (4)
% of slow and undefined events	54.7 ± 10.6 (4)	58.9 ± 11.1 (4)	74.0 ± 12.1 (4)	66.2 ± 11.8% (4)
% of active cells that became inactive*		49.4 ± 18.5% (4)	59.2 ± 13.8% (4)	74.2 ± 11.1% (4)
Unitary events frequency (Hz)†	0.0040 ± 0.0014	0.0045 ± 0.0022	0.0011 ± 0.00051	0.00095 ± 0.00042
Induced Activity				
# of responding cells per mm ²	150 ± 50 (2)	130 ± 50 (2)	25 ± 15 (2)	<i>Not tested</i>
Amplitude at first stimulation (dF/F)	1.58 ± 0.62 (2)	1.19 ± 0.71 (2)	0.15 ± 0.16 (2)	<i>Not tested</i>
% of responding cells per distance				
0-20 μm	5.0 ± 5.0 (2)	2.8 ± 2.8 (2)	0.0 ± 0.0 (2)	<i>Not tested</i>
20-40 μm	17.5 ± 12.5 (2)	9.0 ± 3.5 (2)	62.5 ± 37.5 (2)	<i>Not tested</i>
40-60 μm	37.5 ± 12.5 (2)	42.4 ± 20.1 (2)	37.5 ± 37.5 (2)	<i>Not tested</i>
60-80 μm	10.0 ± 0.0 (2)	11.8 ± 0.7 (2)	0.0 ± 0.0 (2)	<i>Not tested</i>
80-100 μm	12.5 ± 2.5 (2)	14.6 ± 2.1 (2)	0.0 ± 0.0 (2)	<i>Not tested</i>
100-120 μm	12.5 ± 12.5 (2)	13.9 ± 13.9 (2)	0.0 ± 0.0 (2)	<i>Not tested</i>
120-140 μm	0.0 ± 0.0 (2)	0.0 ± 0.0 (2)	0.0 ± 0.0 (2)	<i>Not tested</i>
140-160 μm	2.5 ± 2.5 (2)	2.8 ± 2.8 (2)	0.0 ± 0.0 (2)	<i>Not tested</i>
160-180 μm	2.5 ± 2.5 (2)	2.8 ± 2.8 (2)	0.0 ± 0.0 (2)	<i>Not tested</i>

Number of analyzed FOVs in brackets

*: Cells with responses classified as unitary, multippeak, burst and mixed in ACSF, *n* = 49

†: *n* = 14 cells in ACSF



Supplementary Figure 1. Effect of ACSF and the various custom-made DMEM media on spontaneous electrical activity. Custom-made DMEM media were tested for their ability to support spontaneous activity relative to either ACSF (a) or complete DMEM (b) as control/recovery condition.



CHAPTER 3B

Neuronal medium that supports basic synaptic functions and activity of human neurons *in vitro*

Cedric Bardy¹, Mark van den Hurk^{1,2,*}, Tameji Eames^{1,*}, Cynthia Marchand^{1,3,*}, Ruben V. Hernandez^{1,*}, Mariko Kellogg^{1,*}, Mark Gorris¹, Ben Galet¹, Vanessa Palomares¹, Joshua Brown⁴, Anne G. Bang⁴, Jerome Mertens¹, Lena Böhnke¹, Leah Boyer¹, Suzanne Simon¹, Fred H. Gage¹

¹ Salk Institute for Biological Studies, Sanford Consortium for Regenerative Medicine, La Jolla, CA 92037

² Department of Psychiatry and Neuropsychology, Division of Translational Neuroscience, Maastricht University, 6200 MD, Maastricht, The Netherlands

³ School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

⁴ Conrad Prebys Center for Chemical Genomics, Sanford-Burnham Medical Research Institute, La Jolla, CA 92037

* M.v.d.H., T.E., C.M., R.V.H. and M.K. contributed equally to this work.

Proceedings of the National Academy of Sciences of the United States of America
2015; 112(20):E2725-34

doi: 10.1073/pnas.1504393112

ABSTRACT

Human cell reprogramming technologies offer access to live human neurons from patients and provide a new alternative for modeling neurological disorders *in vitro*. Neural electrical activity is the essence of nervous system function *in vivo*. Therefore, we examined neuronal activity in media widely used to culture neurons. We found that classic basal media, as well as serum, impair action potential generation and synaptic communication. To overcome this problem, we designed a new neuronal medium (BrainPhys basal + serum-free supplements) in which we adjusted the concentrations of inorganic salts, neuroactive amino acids and energetic substrates. We then tested that this new medium adequately supports neuronal activity and survival of human neurons in culture. Long-term exposure to this physiological medium also improved the proportion of neurons that were synaptically active. The medium was designed to culture human neurons but also proved adequate for rodent neurons. The improvement in BrainPhys basal medium to support neurophysiological activity is an important step toward reducing the gap between brain physiological conditions *in vivo* and neuronal models *in vitro*.

Keywords: BrainPhys, tissue culture milieu, Neurobasal, DMEM, neuromedium, induced pluripotent stem cells

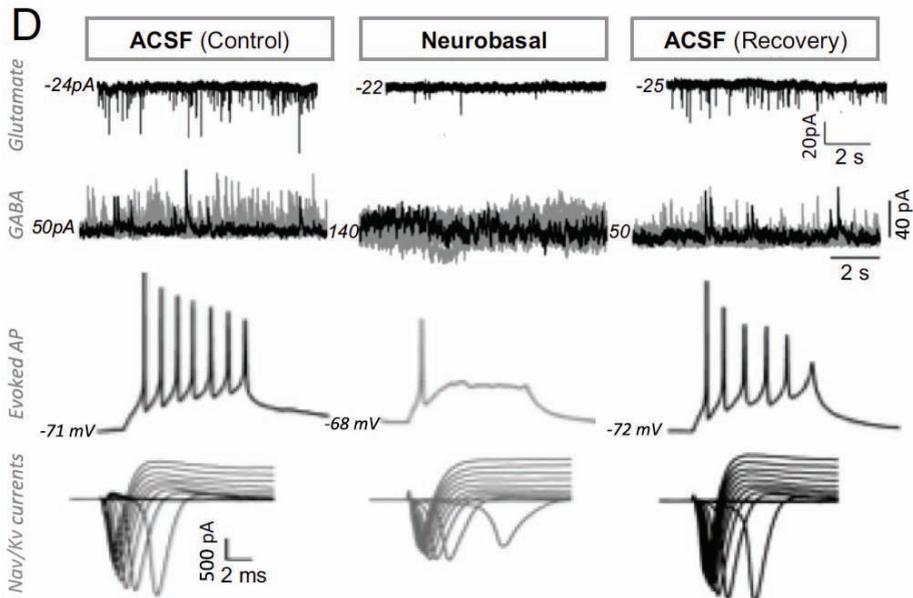
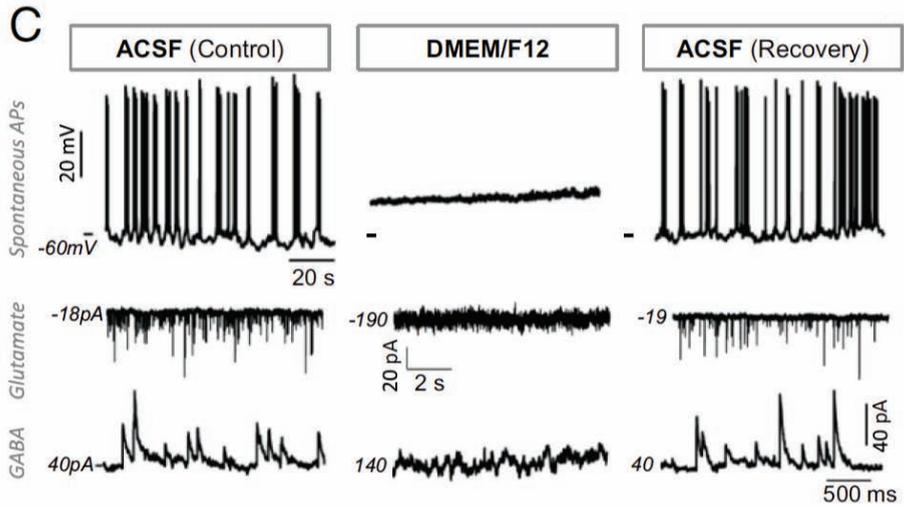
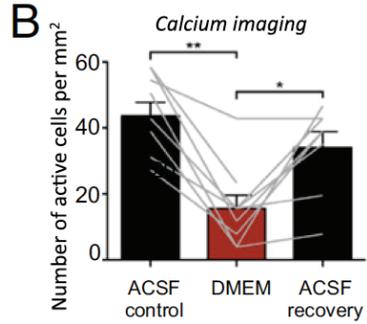
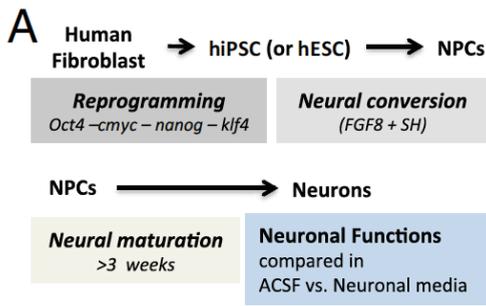
SIGNIFICANCE

Neuronal cultures are very valuable to learn about basic principles of the nervous system. *In vivo*, neural electrical activity is the essence of nervous system function, controlling emotion, memory, sensory modalities and behavior. In this study, we discovered that many crucial neurophysiological properties were strongly altered in classic culture media that are widely used by the research community. To overcome this problem, we designed and tested a new tissue culture neuromedium that adequately supports *in vitro* neuronal activity. The improvements made in this medium reduce the gap between *in vivo* brain physiological conditions and neuronal models *in vitro*. Improving physiological conditions *in vitro* may lead to more successful translation from bench to clinics.

INTRODUCTION

Induced pluripotent stem cell (iPSC) technology is currently being employed to model human diseases *in vitro* and may contribute to the discovery and validation of new pharmacological treatments [1-3]. In particular, neuroscientists have seized the opportunity to culture neurons from patients with neurological and psychiatric disorders and have demonstrated that phenotypes associated with particular disorders can be recapitulated in the dish [4-7]. However, the basic culture conditions for growing neurons *in vitro* have not been updated to reflect fundamental principles of brain physiology. Currently, most human neuronal cultures are grown in media based on DMEM/F12 [4, 5, 7-24], Neurobasal [25-30], or a mixture of DMEM and Neurobasal (DN) [31-34]. To promote neuronal differentiation and survival, a variety of supplements, such as serum, growth factors, hormones, proteins, and antioxidants, are typically added to these basal media. Although these media were designed and optimized to promote neuronal survival *in vitro*, they were not tested for their ability to support fundamental neuronal functions. Using several electrophysiological techniques such as patch clamping, calcium imaging, and multi-electrode arrays, we found that widely used tissue culture media (e.g., DMEM basal, Neurobasal, serum) actually impaired neurophysiological functions.

We identified several neuroactive components in these media that acutely interfered with neuronal function. To solve these issues, we designed a new chemically defined basal medium: BrainPhys basal. We used human neurons *in vitro* to demonstrate in a series of experiments that this new medium, combined with the appropriate supplements, better supports important neuronal functions while sustaining cell survival *in vitro*. Notably, BrainPhys-based medium better mimics the environment present in healthy living brains, unlike previous media based on DMEM, Neurobasal or serum. Although BrainPhys basal was specifically designed for the culturing of mature human neurons,



← **Figure 1. DMEM/F12 basal and Neurobasal media impair neuronal and synaptic activity.** (A) Human neurons derived from iPSCs or ESCs were plated directly on glass coverslips. The basic neuronal functions of single mature neurons were tested and compared in two classic extracellular basal media (DMEM/F12 or Neurobasal-A). The perfusion of extracellular media was switched from ACSF to DMEM/F12 or Neurobasal-A, and back to ACSF for recovery. (B) Calcium imaging analysis showing the number of spontaneously active cells in 10-min movies of the same fields of view recorded in ACSF, DMEM/F12 ($n = 9$), and back to ACSF ($n = 8$). Inorganic salt concentration and the osmolarity of ACSF were matched to DMEM. The gray lines represent individual fields of view. Statistics were performed with Wilcoxon test two-tailed, $**P = 0.004$; $*P = 0.016$). See also detailed calcium spikes analysis in CHAPTER 3A of this thesis. (C) *Top traces*: Current-clamp recordings reveal that spontaneous APs observed in ACSF quickly disappear in DMEM, together with a large increase in resting membrane potential (see also Supplementary Figure 1). *Middle traces*: Voltage-clamp recordings (at -70 mV, Cl^- reversal potential) reveal that DMEM-based medium induces a large Na^+ inward current and completely blocks spontaneous AMPA synaptic activity. *Bottom traces*: Voltage-clamp recordings (at 0 mV, Na^+ reversal potential) reveal that DMEM-based medium induces large Cl^- currents and completely blocks spontaneous GABA synaptic activity. (D) *Upper traces* (synaptic activity): Voltage-clamp recordings reveal that Neurobasal-A based medium strongly reduces spontaneous AMPA synaptic activity. In addition, it induces large tonic Cl^- currents and completely blocks spontaneous GABA synaptic activity. *Lower traces* (APs): Neurobasal-A does not affect the resting membrane potential but it strongly impairs evoked APs (500-ms step of current), voltage-dependent Nav/Kv currents ($I-V$ traces, clamp -70 mV, voltage steps 5 mV), and spontaneous APs.

our studies also showed that BrainPhys provided a functional environment for *ex vivo* brain slices and for culturing rodent primary neurons.

RESULTS

Electrophysiological recordings and calcium imaging are the prevailing techniques used to assess the functionality of neurons and may be acutely performed in perfusate of artificial cerebrospinal fluid (ACSF) *in vitro* or *ex vivo*. ACSF is a basic solution with neurophysiological concentrations of important inorganic salts, energy substrates and phosphate buffers. We used ACSF as a standard to define the level of optimal electrophysiological activity *in vitro*. Although ACSF can sustain healthy neurophysiological activity for a few hours, it does not sustain long-term survival or growth *in vitro*. Indeed, even with the addition of growth factors and other supplements, we found that neurons did not survive in ACSF for more than a few days. Therefore, neurons are usually cultured in an incubator with basal medium such as DMEM/F12 or neurobasal and supplements.

DMEM/F12 Basal Medium Impairs Basic Activity of Human Neurons

We first examined the calcium activity of human neurons in fresh DMEM/F12+sup medium, in which they were cultured for several weeks (**Figure 1A**). We found very few active human neurons in this culture medium, but many more cells became spontaneously active when we imaged the same fields of view in ACSF. The composition of ACSF

solutions may vary slightly between laboratories. For example, the calcium concentration is often two or three times higher than physiological levels (~1.2 mM) [35-37], which artificially increases synaptic release and therefore may also increase calcium neural network activity. To address this issue, we matched the inorganic salt concentration, the pH, and osmolarity of our ACSF to those in DMEM and repeated the calcium-imaging experiments. Even with these precautions, we confirmed that the number of active cells was significantly lower in DMEM basal than in ACSF (**Figure 1B**; see also logg.salk.edu/files/BrainPhys_movies.pptx).

We then used patch-clamping methods to determine how DMEM basal reduced the overall spontaneous activity of human neurons *in vitro* (**Figure 1C**). We found that DMEM basal consistently depolarized the resting potential of neurons ($n = 22/22$ cells, by 23 ± 3 mV; **Supplementary Figure 1A-D**). When we examined spontaneously active neurons in ACSF, we found that, on rare occasions, depolarization induced by DMEM increased the firing frequency without saturation ($n = 2/14$; **Supplementary Figure 1A**), but in most cases, it saturated and completely silenced the firing of the cells ($n = 12/14$; **Figure 1C** and **Supplementary Figure 1B,C**).

To investigate the influence of DMEM basal on synaptic function, we performed voltage-clamp experiments at the reversal potential of Cl^- (-70 mV) or Na^+ (0 mV) and found that both spontaneous AMPA and GABA synaptic events, which were recorded in ACSF, completely disappeared in DMEM (**Figure 1C**). This effect occurred in every tested neuron and was reversible when switched back to ACSF ($n = 6/6$ mature cells).

At the same time, we also observed that perfusion of DMEM consistently induced large depolarizing Na^+ and Cl^- influxes into the cells (**Supplementary Figure 1D,E**). Therefore, we suspected that components present in the DMEM but not in ACSF activate Na^+ and Cl^- channels on the neurons. To identify those components, we removed either all of the vitamins, all of the amino acids, or all of the extra components of DMEM. By removing the bulk of the amino acids, we avoided the DMEM-induced depolarization that caused the impairment of APs and synaptic activity. However, this solution failed to sustain the development and survival of neurons for more than a week (**Supplementary Figure 2**; see also logg.salk.edu/files/BrainPhys_movies.pptx). From these experiments, it was clear that neuroactive amino acids in DMEM impaired basic neuronal functions, including action potential propagation and synaptic communication.

Neurobasal Medium Reduces Synaptic Communication and Action Potential Firing

DMEM/F12 basal has been previously modified to optimize the survival of rat primary neurons in culture [38]. In this modified version, called Neurobasal (NB or NBA), the developers essentially removed some excitatory amino acids and ferrous sulfate and reduced the osmolarity; they found that NB+B27 improved rat neuron survival *in vitro*

in comparison with DMEM or DMEM/F12. Although cell survival is a critical parameter of cell culture, fundamental electrophysiological properties were not tested in these media. Therefore, we used human iPSC- and embryonic stem cell (ESC)-derived neurons to examine the acute effects of Neurobasal on neuronal functions (**Figure 1A**). Unlike DMEM/F12, NBA did not depolarize the resting membrane potential, and at least some excitatory synaptic events could be observed. Nevertheless, NBA strongly reduced or abolished the spontaneous excitatory synaptic activity observed in ACSF (**Figure 1D**; $n = 8/8$ mature neurons; **Figure 2E**). In Neurobasal, the concentrations of inorganic salts (e.g., NaCl) were almost half of what the neurons are naturally exposed to in the brain [39]. As a consequence, we found that Neurobasal significantly reduced the voltage-dependent sodium currents and rapidly inactivating potassium currents, and it impaired the amplitude and frequency of evoked and spontaneous APs ($n = 8/8$ cells; **Figure 1D**). We tried to increase the concentration of NaCl in Neurobasal [40]; however, despite some slight improvements in the amplitude of sodium currents through voltage-gated channels, this modification was not sufficient to optimize the electrical and synaptic activity of the human neurons. We suspected that, as in DMEM, several neuroactive components in Neurobasal were preventing optimal neurophysiological activity. Because neither DMEM nor Neurobasal was capable of optimally sustaining essential electrophysiological neuronal properties, we decided to design a new basal neuro-medium (BrainPhys basal) that would be better adapted to neuronal function in culture.

BrainPhys Basal Medium Supports Functional Action Potential Firing

Neuronal function is fundamentally based on the generation and propagation of APs. Voltage-gated sodium (Nav) and potassium (Kv) channels are crucial to achieving high firing rates of APs. APs reflect the sequential activation of Nav and Kv channels that trigger a large influx of sodium and an efflux of potassium, largely depending on ionic gradients. Therefore, the concentrations of major inorganic salts (Na^+ , Cl^- , K^+) in the media are critical for APs generation and propagation. The concentrations of NaCl are higher in BrainPhys basal, ACSF and DMEM compared to Neurobasal (~120 mM vs. ~70 mM) and closer to neurophysiological levels. As a result, the observed amplitudes of voltage-activated currents through important ionotropic channels (Nav, Kv) did not differ with perfusion of ACSF or BrainPhys medium (**Figure 2A**). In contrast, the amplitudes of Nav and rapidly inactivating Kv currents were systematically larger when recorded in BrainPhys or ACSF compared to Neurobasal (**Figure 2A** and **Supplementary Figure 3A**).

Consequently, the spontaneous and evoked APs in single neurons tested in BrainPhys reached levels similar to those measured in ACSF (**Figure 2B** and **Supplementary Figure 3B**). APs were also tested with optogenes such as channelrhodopsin (ChR2 or ChETA) that could be efficiently expressed in human neurons, allowing us to control the firing

← **Figure 2. BrainPhys basal medium supports optimal APs and synaptic activity.** (A) Voltage-activated (VA) sodium and potassium currents (measured from *I-V* curves, voltage clamp -70 mV, steps 5 mV) were similar in BrainPhys and ACSF. In the three panels on the left, each neuron ($n = 6$) was tested both in ACSF and BrainPhys (black vs. blue dots). The maximum amplitudes of voltage-activated sodium and potassium currents of single neurons ($n = 10$) were significantly reduced in Neurobasal media (red dots, Neurobasal-A; orange dots, Neurobasal) compared with BrainPhys (blue dots) or ACSF (black dots). Non-parametric paired Wilcoxon test *P* values are shown in italics. Histograms represent the mean \pm SE. (B) The resting membrane potential, spontaneous and evoked APs were the same in ACSF and BrainPhys. (C) Typical mature patched neuron expressing an optogene (synapsin:ChETA-YFP, green) and filled with rhodamine from the patch pipette (red shadow on the right). Optogenetic control of neuronal activity can be reliably achieved in BrainPhys. Recordings from single mature neurons tested with the same parameters in BrainPhys or Neurobasal-A illustrate that optogenetic control is dramatically improved in BrainPhys. Bottom part of the graphs in C are the corresponding raster plots. (D) Both excitatory (AMPA) and inhibitory (GABA) synaptic activities were clearly apparent in BrainPhys medium. The perfusion of different media while recording the spontaneous synaptic activity of single neurons revealed that BrainPhys better supported synaptic function compared with other media such as Neurobasal-A. Twelve sweeps are super-imposed in gray and one of them is highlighted in black for clarity. See also Supplementary Figure 3. (E) Patched neurons ($n = 22$) tested in different basal media are represented by the paired dots. Quantification of AMPA receptor-mediated spontaneous synaptic activity shows similar properties in ACSF and BrainPhys basal ($n = 8$). Neurobasal significantly reduces AMPA synaptic events ($n = 8$). Both Neurobasal and DMEM completely block GABA synaptic events ($n = 4 + 2$). All tests comparing spontaneous activity were performed without any synaptic antagonists or voltage-gated sodium channel blockers. Voltage clamp at respective reversal potential of chloride and sodium were used to distinguish glutamatergic from GABAergic events. Synaptic blockers (NBQX or Gabazine) were used in a subset of cells ($n = 15$) only to confirm the nature of the receptors mediating the observed synaptic activity. Wilcoxon *P* values are shown in italics. Two-tailed tests were used except for recovery from 0 Hz; then a one-tail test was used.

activity of the neurons in BrainPhys with brief flashes of light (**Figure 2C**). In sharp contrast, light stimulations of neurons that successfully evoked responses in BrainPhys failed to evoke clear APs in either NBA- or DMEM/F12-based medium (**Figure 2C** and **Supplementary Figure 3D**). Consistent with the overall improvement in firing activity, the Ca^{2+} activity measured with imaging in ACSF and BrainPhys was comparable (see logg.salk.edu/files/BrainPhys_movies.pptx).

BrainPhys Basal Medium Supports Functional Excitatory and Inhibitory Synaptic Activity

Integrated neural network activities—and ultimately, brain functions—are achieved by synaptic communication. Therefore, we optimized BrainPhys basal to support both excitatory and inhibitory synaptic function. The summation of excitatory and inhibitory synaptic neurotransmitters determines the probability of neurons to fire APs. Thus, simply optimizing AP firing may act indirectly to promote synaptic activity. In addition, specific ions such as calcium play critical roles in the synapse in triggering fast vesicular release. Aiming to reduce the differences between culture media and *in vivo* brain conditions, we

adjusted calcium levels in BrainPhys basal to be close to those in human cerebrospinal fluid *in vivo* ($\text{Ca}^{2+} \sim 1.1 \text{ mM}$) [37].

In the brain, glutamate mediates most of the synaptic excitation, whereas GABA mediates most of the synaptic inhibition. To examine whether the different media could support basic synaptic functions, we tested the levels of spontaneous synaptic activity mediated by AMPA or GABA_A receptors (identified in voltage clamp by reversible blockades with their respective antagonists, NBQX or SR95531). Voltage-clamp experiments showing strong Na^+ and/or Cl^- currents on perfusion of DMEM/F12 or NBA in comparison with ACSF strongly suggested that synaptic function was impaired by the presence of receptors agonists present in these media but not in ACSF. Therefore, in BrainPhys basal we excluded or reduced the neuroactive amino acids in classic media that could directly influence glutamatergic synaptic activity (glutamate, aspartate) and/or GABAergic synaptic activity (glycine, alanine, serine). Although spontaneous activity might be variable between cells, paired analysis of single mature neurons clearly showed that, unlike in DMEM or Neurobasal, the levels of spontaneous AMPA synaptic activity in BrainPhys were not significantly different from those in ACSF (**Figure 2E**) and were therefore largely improved compared with DMEM or Neurobasal (**Figure 1** and **Figure 2D,E**). We also observed that both Neurobasal and DMEM completely blocked synaptic inhibitory phasic activity ($n = 6/6$ cells, 4 in NBA, 2 in DMEM/F12; **Figure 2D,E**). The blockade of GABA synaptic activity was systematically accompanied by a large influx of chloride currents (**Supplementary Figure 1D,E**) and therefore suggested that the blockade was due to the excess of neuroactive components activating chloride channels (e.g. glycine, alanine, serine). Following up on this hypothesis, we indeed found that reducing the concentration of these elements improved GABAergic synaptic activity. The spontaneous GABA synaptic activity was proven to be functional in BrainPhys ($n = 7$ cells; **Figure 2D** and **Supplementary Figure 1F**) and comparable to levels of activity in ACSF ($n = 3$ cells tested in both solutions; **Figure 2D,E**).

BrainPhys Basal Medium Better Mimics the Human Brain's Energy Levels and Osmolarity

Our primary objective was to design a medium that better supported neuronal functions and activity *in vitro*. In addition, we wanted to design a medium that would better mimic the basic conditions in the human brain. These improvements are not only important for studies directly focusing on neuronal activity but are also potentially critical for disease modeling studies, as more realistic experimental models will increase translational success for patients.

To reach that goal, it seems essential to provide energy levels close to those observed in healthy human brains, especially since several neurological disorders have been shown to be related to mitochondria dysfunctions [41-43]. The mammalian brain uses glucose

as one of its main sources of energy [44]. Disruption of glucose homeostasis may affect brain physiology and lead to brain disorders [44, 45]. Surprisingly, the glucose levels in DMEM and Neurobasal are at least two to five times higher than those in the brain of hyperglycemic patients [46]. We therefore balanced the energetic components in BrainPhys to provide glycemic levels similar to those reported for the brains of healthy individuals (~2.5 mM).

Furthermore, we set the osmolality to approximate that of typical human cerebrospinal fluid (~300 mOsmol/L). In contrast, the osmolality of NBA is ~30% lower than neurophysiological levels (~220 for NB or 250 mOsmol/L for NBA). It is important to note that evaporation of the medium often occurs in tissue culture and can drastically affect the osmolality. To reduce cytotoxic osmotic stress, the cells were kept in an appropriately humidified incubator and half of the medium was replaced at least two or three times a week.

Supplements to Basal Medium That Do Not Acutely Impair Neuronal Activity

To sustain cell survival and/or neural differentiation *in vitro*, supplements such as antioxidants, growth factors, hormones and proteins should be added to basal media. Many of these molecules can be found naturally in serum; therefore, serum is often used. However, in addition to introducing variability from batch to batch, we found that serum (both chemically defined and fetal bovine serum (FBS)) dramatically impaired neuronal activity (**Supplementary Figure 4**). As an alternative to serum, we tested the effect of a combination of several chemically defined supplements on acute and long-term neural activity of human neurons *in vitro*. We patch-clamped single mature human neurons in various perfusates and found that adding a defined set of supplements (N2, B27, retinoic acid, brain-derived neurotrophic factor (BDNF), glia cell-derived neurotrophic factor (GDNF), ascorbic acid, cAMP, laminin, and cholesterol; **Supplementary Table 6**) to BrainPhys basal medium did not acutely affect the firing rate or excitatory/inhibitory synaptic activity ($n = 6$ cells; **Supplementary Figure 4**).

BrainPhys Basal+sup Supports Stable Long-Term Electrical Activity of Mature Human Neurons *In Vitro*

Our patch-clamping experiments comparing the effect of different media on neuronal activity demonstrated that BrainPhys basal medium was more efficient to support physiological action potential firing and synaptic activity. To study brain disorders *in vitro*, mature neurons should be cultured in a medium that supports their basic function for some time before analysis.

To find out whether BrainPhys basal + supplements was adequate to sustain long-term neuronal activity, we recorded the electrical activity of mature human neuronal cultures over several weeks. In those experiments, we used a multi-electrode array (MEA) system

integrated on a 48-well tissue culture plate (Axion; 16 electrodes per well; **Figure 3A,B**). We obtained frozen stocks of characterized and purified mature human iPSC neurons (commercially available iCell neurons from Cellular Dynamics), thawed the purified neurons directly onto the MEA plates, and started recording their activity two days later, every day, for about two weeks. With this system, we compared the electrical activity of neurons over time in BrainPhys+sup and other media. First, regardless of the basal media used, we confirmed our patch-clamping results that adding serum dramatically impaired neuronal activity. The impairment caused by serum lasted more than two weeks (**Figure 3C**). In NBA+sup or NBA/DMEM mixture+sup, we found that the overall neuronal activity was relatively low in the first few days and deteriorated rapidly within about a week (**Figure 3E**). The poor neuronal performance in the NBA+sup could be rescued, at least partially, within a few days when replaced with BrainPhys+sup (**Figure 3E**). In DMEM+sup, the neuronal activity was somewhat higher than in NBA+sup or NBA/DMEM+sup but also highly fluctuated over time, presumably in synchrony with the feeding cycles (twice a week in those experiments; **Figure 3D**). In contrast, BrainPhys+sup (with the same serum-free supplements) significantly improved the spontaneous activity of human neurons within a few days and, importantly, kept it stable for more than two weeks (**Figure 3A-E**, see statistics in **Supplementary Table 1**; for illustration of live MEA activity in different media, see also logg.salk.edu/files/BrainPhys_movies.pptx).

Taken together, these results demonstrate the effectiveness of BrainPhys basal with serum-free supplements to keep mature human neurons physiologically active and stable.

Characterization of Human Neurons Maturing Several Weeks in BrainPhys+sup Medium

The methodological variations to obtain different types of human neurons *in vitro* (e.g., cortical, dopaminergic) are probably countless. Generally, somatic cells (e.g., fibroblasts) or pluripotent stem cells (e.g., iPSCs, ESCs) are converted to neural progenitor cells (NPCs) in neural conversion media with various growth factors. Then, those NPCs are cultured in neuronal media and usually kept in the same media until analysis. We found that BrainPhys+sup was efficient for use with mature neurons obtained with different protocols, and it provided an active and neurophysiological environment for several weeks or more. Then we tested whether neural progenitors could be switched gradually from their neural progenitor medium to neuronal medium based on BrainPhys with supplements.

After expansion, human NPCs (derived from ESCs or iPSCs) were seeded on transferable, treated glass coverslips (no feeder layer). Within 24 h after plating the cells, the NPC medium was gradually replaced with subsequent feeds of BrainPhys+sup (**Figure 4A**). We found that human NPCs plated in BrainPhys+sup differentiated effectively into mature and synaptically active neurons within three to five weeks. After about four weeks in

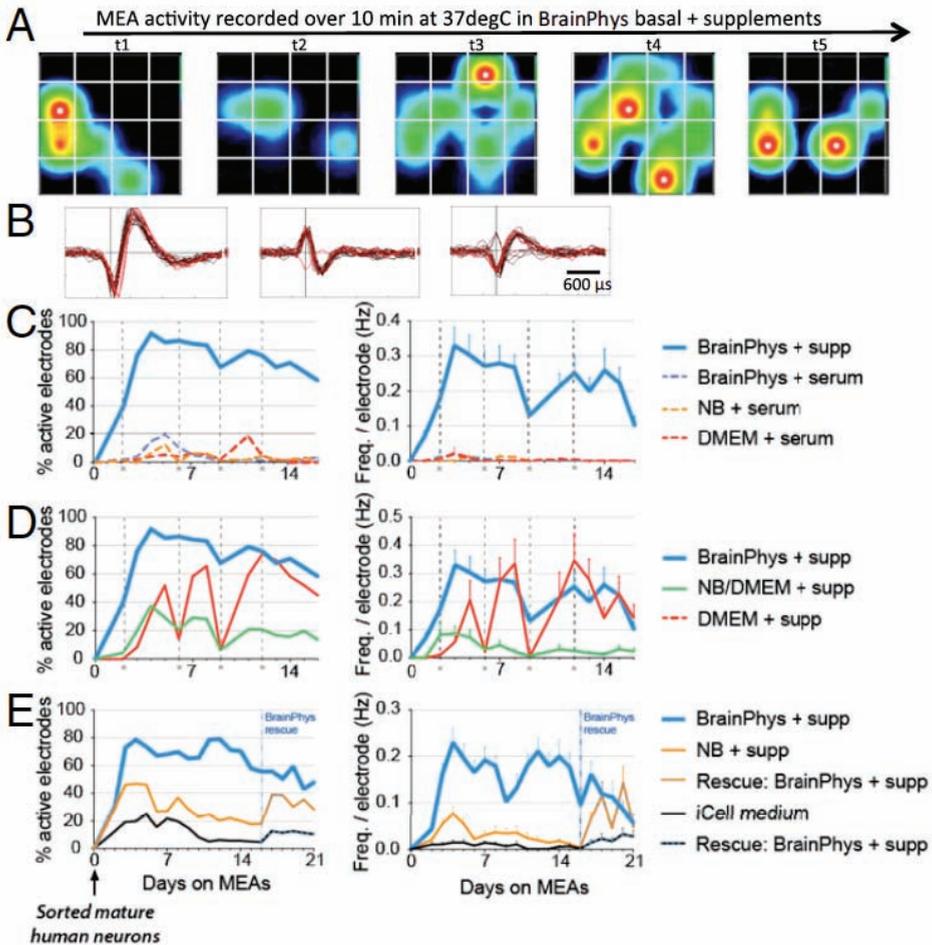


Figure 3. Long-term exposure of mature human neurons to BrainPhys basal (+supplements without serum) supports optimal and stable neuronal activity on MEAs. Electrical activity of purified mature human iPSC-derived neurons (iCell neurons from Cellular Dynamics) was recorded in 48-well multi-electrode array (MEA) plates (Axion) in different neuronal maturation media at 37 °C every day for 2-3 weeks. **(A)** Example illustrating the heatmap activity of seven snapshots of the same array in a well. Recordings (10 min) from neurons in BrainPhys with supplements (for 5 days) (see logg.salk.edu/files/BrainPhys_movies.pptx). **(B)** Examples of spikes detected by individual electrodes in BrainPhys+sup medium. **(C-E)** Percentage of active electrodes (>0.005 Hz) and frequency of detected spikes per electrode averaged from daily 10-min recordings. Each value represents the average of at least 96-192 electrodes spread in four to eight wells per condition. The same batches of cells were compared in the 48-well plates with different media. Dashed lines in **C** and **D** mark the feeding dates. Recordings were performed before feeding. Each experiment was repeated at least twice independently with different set of cells and combined in the figures. For clarity, the experiment in panel **C** and **D** was split into two graphs, but the same control (BrainPhys+sup) is represented twice. See analysis of spikes properties in Supplementary Table 1.

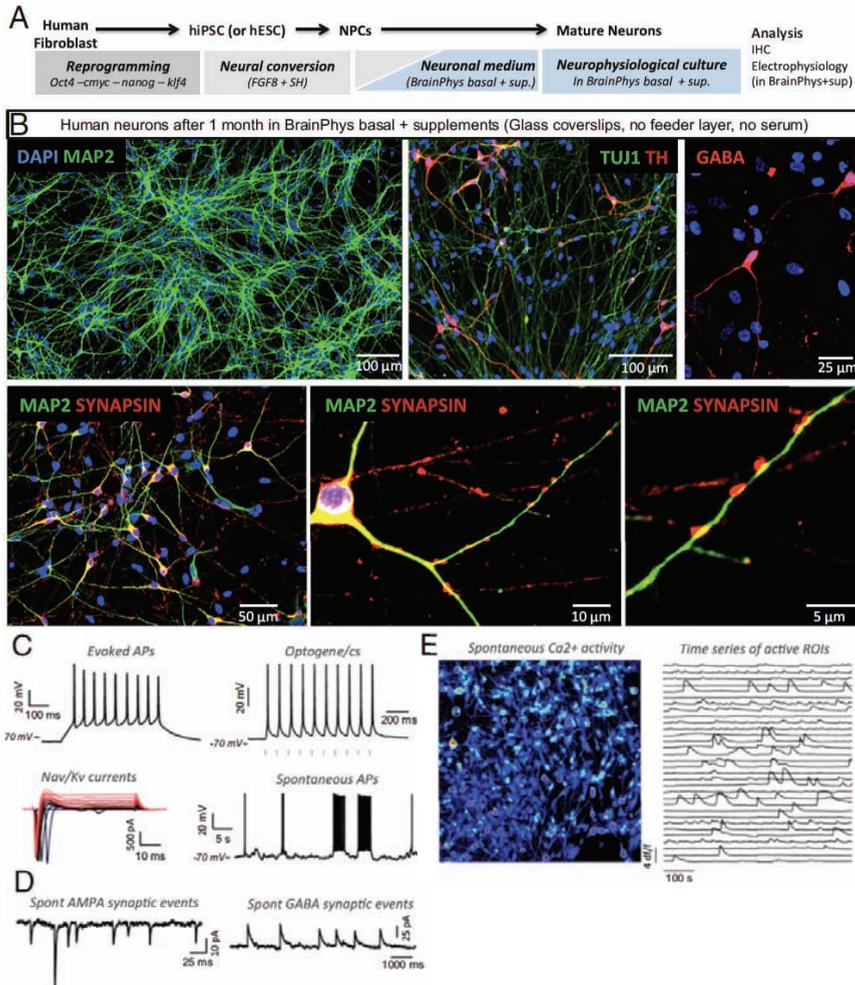


Figure 4. Characterization of human neurons cultured for several weeks in BrainPhys basal + supplements. (A) Human neurons derived from iPSCs or ESCs were plated directly on glass coverslips. Human NPCs were matured in neuronal medium (BrainPhys + sup) for 3–6 weeks. Analysis of electrophysiological properties was performed in the same neuronal medium (BrainPhys + sup). (B) Immunostaining of typical human iPSC-derived neuronal cultures grown in BrainPhys with supplements for 4 weeks. (C–D) Electrophysiological activity of typical functional neurons after 3–6 weeks in BrainPhys-based neuronal medium. Patch-clamp recordings were also performed in BrainPhys medium with supplements. (C) From left to right: Train of APs evoked by a small 500-ms depolarizing step of current or brief flashes of light (syn:ChETA-YFP). *I-V* traces (clamp -70 mV, steps 5 mV) showing typical voltage-activated Na^+ and K^+ currents. Spontaneous APs recorded in current clamp. (D) Spontaneous synaptic events mediated by AMPA receptors (sensitive to NBQX, voltage clamp at -70 mV) and GABA receptors (sensitive to Gabazine, voltage clamp at 0 mV). (E) Calcium imaging showing typical activity of neuronal cultures grown and recorded in BrainPhys + supplements. Time series analysis is plotted for the active regions of interest (white circles; see also logg.salk.edu/files/BrainPhys_movies.pptx). 91% of the active ROIs showed clear calcium spikes; the remaining showed slow calcium waves. Statistics for the calcium spikes of 41 active ROIs (mean \pm SE): 8 ± 1 calcium spikes per 10 min, $df/f = 1.1 \pm 0.1$; rise 10–90% = 3 ± 0.3 s; decay 37% = 11 ± 1 s; half-width = 12 ± 1 s.

neuronal medium, the cultures formed dense neuronal networks and stained abundantly for dendritic/axonal markers (MAP2 and TUJ1), synaptic markers (Synapsin) and neuronal markers (TH, GABA, VGlut) (**Figure 4B**). Consistent with our previous observations, human NPCs differentiated not only into neurons but also into a population of astrocytes staining positively for GFAP (**Supplementary Figure 5**).

The neurons survived on glass coverslips in BrainPhys+sup for several weeks or even several months for the best cell lines (tested up to five months). We tested the neurophysiology of these BrainPhys cultures by patch-clamping directly in BrainPhys+sup, in which they were cultured, and we found many mature neurons firing healthy trains of APs (**Figure 4C**) and receiving active AMPA and GABA synapses (**Figure 4D**; see summary of electrophysiological properties of large sample, $n = 106$, in **Supplementary Table 2**). For the best cell lines, those types of neurons were found around three weeks after plating in maturation medium and remained active for several months (tested up to five months). Similarly, calcium imaging experiments supported the finding that BrainPhys-based neural cultures generated and sustained functionally active neural networks (**Figure 4E**; see also logg.salk.edu/files/BrainPhys_movies.pptx).

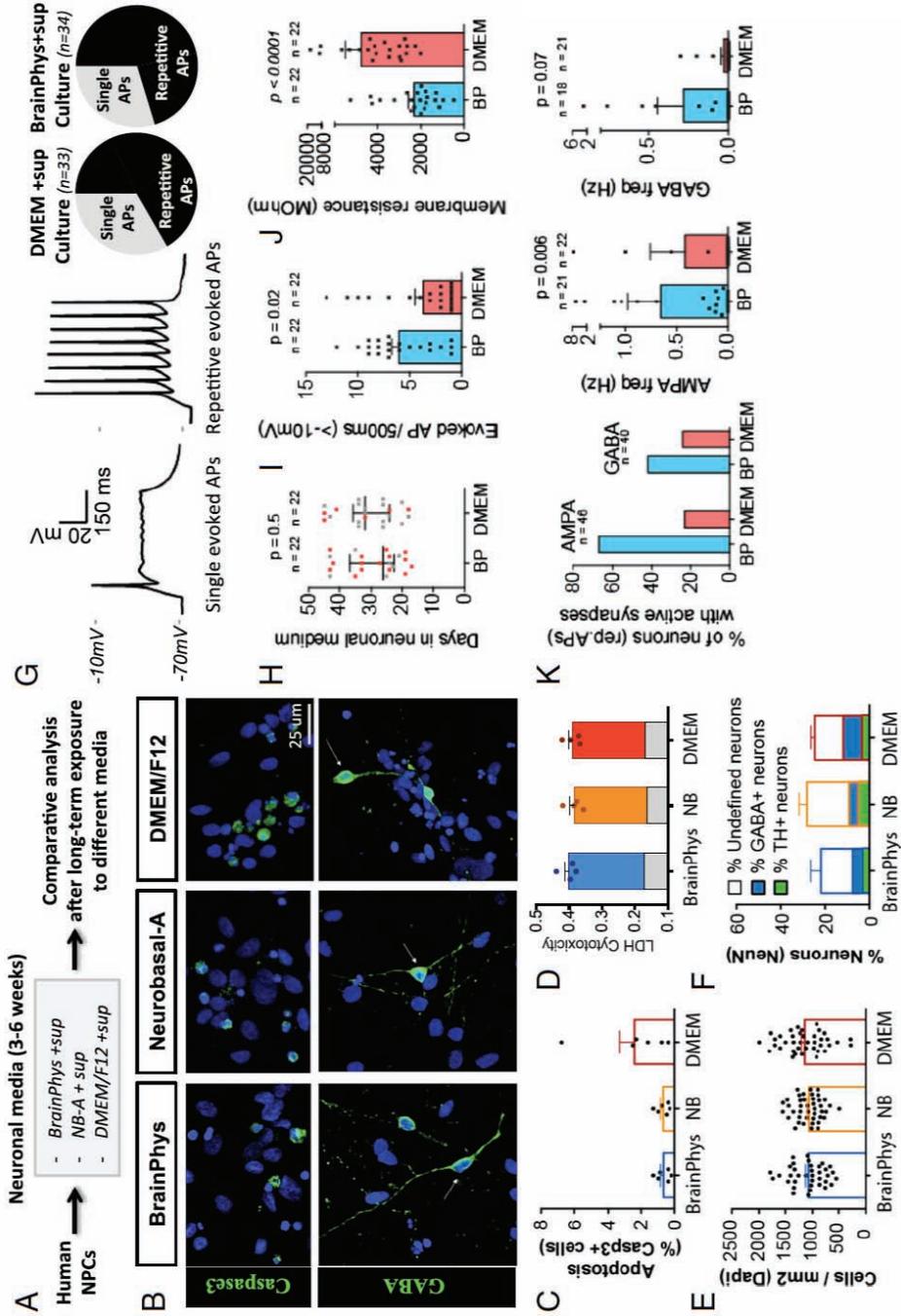
BrainPhys+sup Medium Supports Brain Cell Survival *In Vitro*

DMEM and Neurobasal media were specifically optimized to promote the survival of cells *in vitro*. To test the viability of the cells in different basal media (DMEM/F12, NBA and BrainPhys basal, all with the same set of serum-free supplements), we performed matched experiments using the same cell lines, plated at similar densities and at the same time, but fed for several weeks with different media (**Figure 5A**). We found that, after one month in either medium, the proportion of apoptotic cells (active caspase 3), the overall cell density (DAPI) and the concentration of lactate dehydrogenase released in the supernatant by dying cells did not change significantly between the three groups (**Figure 5B-E**).

BrainPhys+sup Supports the Basic Function of Many Classes of Neurons

Human cell reprogramming is used to differentiate a multitude of neuronal classes. We collected electrophysiological evidence that BrainPhys+sup is suitable to culture many classes of human neurons derived from iPSCs (**Supplementary Table 4**). Importantly, after ≥ 2 weeks in BrainPhys+sup, each of these classes held mature functional neurons (evoked APs $> 10\text{Hz}$ and spontaneously active synapses).

When we compared the differentiation of NPCs in BrainPhys+sup and other media (DMEM+sup and NBA+sup), we did not find significant differences in the proportion of NeuN-positive neurons (**Figure 5F**). Our iPSC differentiation protocol generated a variety of neuronal classes in the same culture (glutamatergic, GABAergic, dopaminergic). The proportion of NeuN+ neurons that positively stained for GABA or TH and the



← **Figure 5. Comparison of human neurons cultured in BrainPhys vs. classic media based on DMEM or Neurobasal.** (A) In these experiments, we seeded human NPCs on coated glass coverlips (no feeder layer) in neural progenitor medium (see Supplementary Methods). We then gradually switched to three different media: BrainPhys basal, DMEM/F12 or Neurobasal-A with the same serum-free supplements (see Supplementary Methods). Half-media changes were performed every 3 days for 3-6 weeks before analysis. (B-F) Representative immunostaining and analysis of cells cultured in the same plate with three different conditions. Nuclei were stained with DAPI (blue). Despite some variability, no clear differences in cell survival or proportion of neuronal subtypes (i.e., dopaminergic, GABAergic) were found between the three conditions. Quantification on all FOV with 20× objective (tiles 2 × 2) (for TH, $n = 23$ FOV; for NeuN, $n = 20$ FOV; for GABA, $n = 23$ FOV). (G) To characterize the influence of BrainPhys basal on functional properties, we compared the neurons matured in BrainPhys+sup with those in DMEM/F12+sup for 2-6 weeks. Alternative blind patch-clamping of the two groups was performed in ACSF for unbiased comparison. Electrophysiological types of neurons were defined based on their optimal firing patterns in response to 500-ms depolarizing steps of currents in ACSF. Similar proportion of cells with single or repetitive APs was found in the two media cultures. (H-K) For better sample homogeneity, the analysis was focused on neurons with repetitive APs. (H) The cells with repetitive APs spanned throughout the time windows of recording in both groups. Red dots represent neurons with evoked firing >10 Hz (threshold -10 mV). Median and interquartile ranges are shown. (I) The evoked firing frequency (threshold -10mV) of neurons with repetitive APs was significantly higher in the BrainPhys+sup cultures. (J) The membrane resistance of neurons with repetitive APs was significantly lower in the BrainPhys+sup cultures. (K) BrainPhys+sup culture also showed higher proportions of neurons receiving active AMPA synaptic inputs (NBQX-sensitive) and GABA synaptic inputs (Gabazine-sensitive). The presence of active synapses was determined by detecting at least four spontaneous synaptic events with distinctive kinetics over a 4-min recording in voltage clamp (-70 mV or 0 mV, respectively).

remaining NeuN+ neurons (presumably mostly glutamatergic) did not vary significantly in different neuronal media (**Figure 5F**). Therefore, when used for approximately two to four weeks at the latest stage of the culture process (**Figure 4A**), these three basal media did not appear to greatly influence the neuronal identity. Instead, it is more likely that early steps of the neuronal conversion from iPSC to NPC, and/or specific supplements, played a more critical role in the neuronal identity fate.

In addition, in BrainPhys+sup, NPCs differentiated not only into neurons but also into astrocytes that expressed GFAP protein (immunostaining). We also used GFAP promoter linked with fluorescent reporters to reveal astrocyte-like electrophysiological properties (low resting potentials, absence of sodium currents).

Culture in BrainPhys+sup Can Increase Action Potential Firing Frequencies and the Proportion of Synaptically Active Human Neurons

Despite the failure of classical basal media to support optimal neurophysiological activity, when combined with the right supplements, these media manage to support survival and neuronal differentiation *in vitro*.

Electrical activity is known to play an important role in neuronal development and synaptic function [47-49]. Typically in the brain, most newly generated glutamatergic synapses lack functional AMPA receptor mediated transmission. Over time, these synapses are eliminated if kept silent, whereas those exposed to correlated electrical activity

will mature and prevail [47]. We asked whether BrainPhys+sup could improve the synaptic function of mature neurons. To address that question, we randomly patched a homogeneous sample of 65 synapsin-GFP-positive neurons in cultures growing side by side in DMEM/F12 or BrainPhys basal (with the same supplements). To avoid possible bias due to tissue culture variability, we compared the effects of the media on neurons from the same NPCs, plated at the same density ($\sim 2 \times 10^5$ cells per well; 190 mm^2), in the same plate, at the same time, and we discarded the few coverslips in each groups where some cells detached from the coverslips and/or formed large clumps. We blindly alternated the patching of neurons that were cultured in BrainPhys or DMEM/F12 (with supplements); therefore, the time that the neurons spent in either neuronal medium was virtually identical in both groups (average 28 and 29 days, respectively, ranging from 17 to 45 days). The neurons that were patch-clamped had on average the same number of primary neurites (BrainPhys: 3.5 ± 0.2 , DMEM: 3.5 ± 1.2 ; **Supplementary Table 3**). The patched neurons were cultured in either DMEM/F12+sup or BrainPhys+sup media, but because patching the cells in DMEM impairs activity, we acutely assessed their functional properties in ACSF. In these conditions, we were reasonably confident that the only difference between the coverslips ($n = 22$) selected for analysis was the basal medium used for feeding. Despite our efforts to homogenize the sampling of recorded neurons, we obtained more neurons with high APs firing rates in the BrainPhys+sup cultures. To quantify this result, we defined different electrophysiological types of neurons based on their firing patterns in response to 500-ms depolarizing steps of currents (**Figure 5G**). We divided the neurons into two groups: those with single evoked APs and those with repetitive evoked APs. After two to six weeks in maturation media, electrophysiological testing in ACSF revealed a similar proportion of cells with single or repetitive APs. Overall neurons with particularly high evoked APs frequencies were found throughout the time window of our analysis in both culture groups (**Figure 5H**). However, the cells with repetitive firing showed significantly higher firing frequency when cultured in BrainPhys+sup instead of DMEM+sup (**Figure 5I**). Consistently, in neurons from BrainPhys-based cultures, the Nav and rapidly inactivating Kv currents were significantly higher, but the voltage activation threshold was similar to those of neurons in DMEM-based cultures (**Supplementary Table 3**). In addition, the membrane resistance of BrainPhys-cultured neurons was significantly lower (**Figure 5J**) and the capacitance was significantly higher than in DMEM-cultured neurons, characteristics that also typically correlate with more mature-like properties (**Supplementary Table 3**).

The numbers of neurons with repetitive evoked APs receiving active excitatory AMPA and inhibitory GABA synaptic inputs were about two to three times higher in the BrainPhys cultures (**Figure 5K**). Interestingly, these synaptic enhancements occurred without obvious differences in the number of proximal synaptic puncta measured by immunohistochemistry (IHC; **Supplementary Figure 6**). This finding suggests that spontaneous

electrical activity in BrainPhys+sup activates/strengthens silent synaptic contacts rather than increasing the formation of additional synapses.

BrainPhys+sup Medium Can Be used in Direct Neuronal Conversion Protocols (Induced Neurons)

Most of our tests were performed on human iPSC- or ESC-derived neurons. However, direct conversion of human fibroblasts to induced neurons (iNs), without passing through the pluripotent stage, represents an attractive alternative (**Figure 6A**). To that end, we tested whether BrainPhys basal medium could support neural conversion and/or maturation of human iNs (**Figure 6B**). Human iNs in BrainPhys+sup displayed healthy neuronal morphologies and positively stained for standard neuronal markers (**Figure 6B**). The efficiency of neural conversion using BrainPhys basal instead of classic basal media (DN: 50:50 mixture DMEM/F12 and NBA) was not significantly different (~60% of neurons with both media, **Figure 6C**). iNs cultured in BrainPhys +supplements for two weeks after conversion displayed healthy functional AP properties (**Figure 6D,E**). These results demonstrate the proof of concept that BrainPhys basal can be used with direct conversion methods.

BrainPhys+sup Medium Increases ARC Protein Expression in Human Neurons Directly Induced from Fibroblasts (iNs)

To further examine the influence of long-term exposure to BrainPhys+sup on human neurons using an alternative measure, we examined the cellular level of ARC protein. Transcription of the immediate early gene *ARC* is known to be increased by neuronal activity, or natural activating stimuli *in vivo* [50, 51]. An increase in neuronal ARC protein levels triggers critical molecular mechanisms that strengthen synaptic communication and play a role in memory consolidation [51, 52]. We compared iN cultures exposed to different basal media (BrainPhys basal vs. DN basal) with the same supplements (N2, B27, ascorbic acid, retinoic acid, laminin, BDNF and GDNF). Strikingly, within three weeks, BrainPhys basal significantly increased the levels of ARC protein in the human iNs (**Figure 6F**). These results demonstrate that the use of BrainPhys basal in direct neuronal reprogramming techniques does not directly influence the reconversion efficiency but increases the neuronal levels of functional activity.

BrainPhys Basal Is Also Applicable to Rodent Neurons

Although we developed BrainPhys basal to culture human neurons, we demonstrated that it could also be utilized for electrophysiological recordings in acute mouse brain slices or to culture rat primary neurons. Mouse hippocampal granule cells patched in BrainPhys basal were fully functional (**Supplementary Figure 7**), and their properties recorded in BrainPhys basal were virtually identical to those in ACSF ($n = 4$ cells). We also

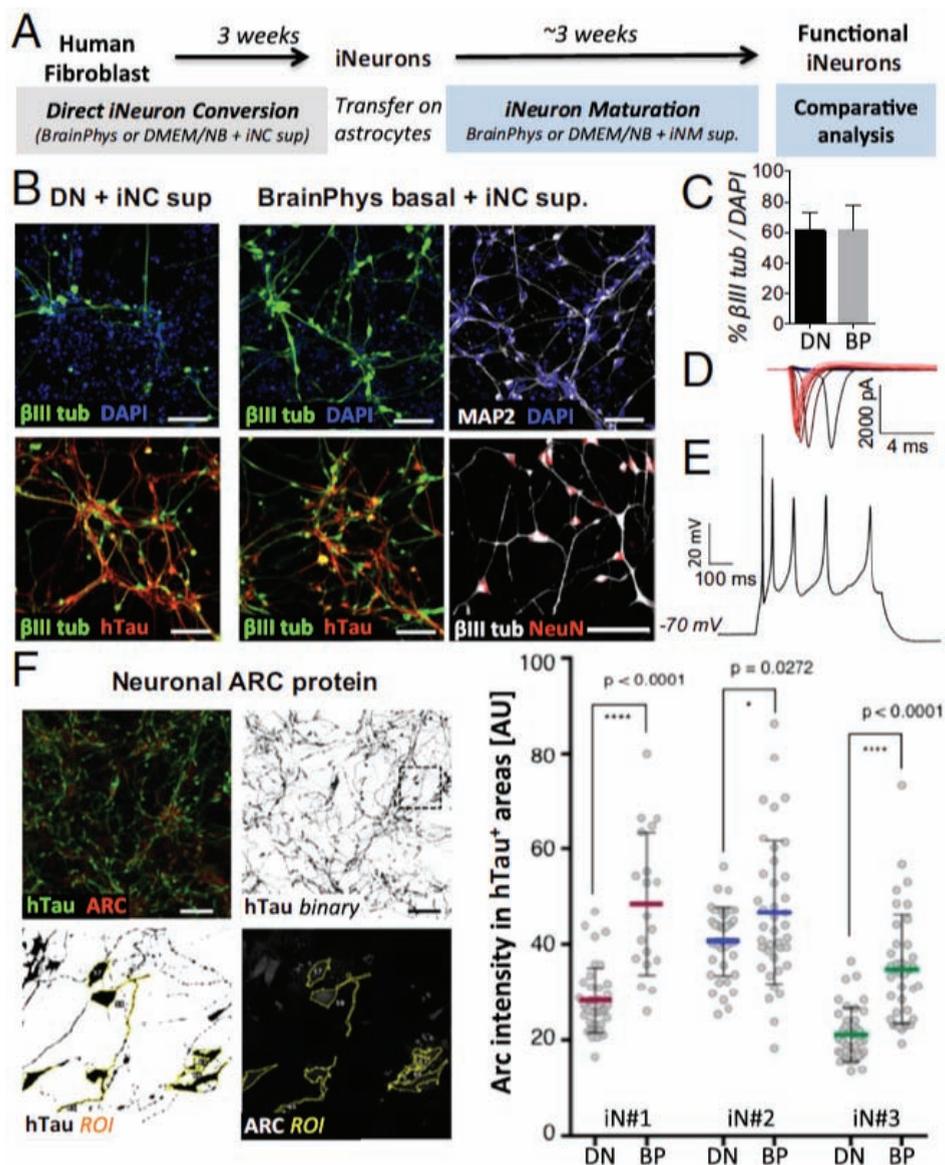


Figure 6. BrainPhys basal with the appropriate supplements permits efficient direct neuronal conversion as well as development of human iNs, and increases ARC protein expression. (A) Experimental design: iN-competent human dermal fibroblasts were converted in neuronal conversion (NC) medium for 3 weeks. For maturation iN cells were relocated onto astrocytes and further cultured in neuronal maturation (NM) medium (see Supplementary Methods for details). (B) Immunofluorescence staining for β -III-tubulin, human Tau (hTau), MAP2ab and NeuN following 3-week conversion in media based on DMEM/F12:Neurobasal (Left) or BrainPhys (Right) with the same supplements. (Scale bar, 50 μ m.) (C) The percentage of TUJ1+/DAPI was counted after 3 weeks of conversion in different basal media with the same supplements. No significant difference was observed in the efficiency of neuronal conversion when either basal medium was used. (D and E) Following 3 weeks of conversion, the iNs were replated on mouse astrocytes

for 2 weeks in BrainPhys basal+sup before patching and showed strong sodium/potassium currents and repetitive evoked APs. (F) Representative image of immunofluorescence staining for hTau and ARC in iN cells after maturation on astrocytes (3-weeks conversion + 3-weeks maturation). Quantification using ROI selection and measurement in ImageJ. (Scale bar, 100 μm .) Plot illustrating the quantification of neuronal ARC protein levels in iN cultures derived from three donors. Gray points represent individual cells (average $n = 38$ per group, minimum $n = 19$ per group). Bar show means \pm SD of each group. Unpaired t test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

checked that rat primary neurons could grow successfully in BrainPhys basal medium (supplemented with sm1) and found that they formed active neural networks within two weeks *in vitro* (Figure 7, $n = 7/7$ patched rat primary neurons had strong evoked AP firing ($>10\text{Hz}$) and were synaptically active).

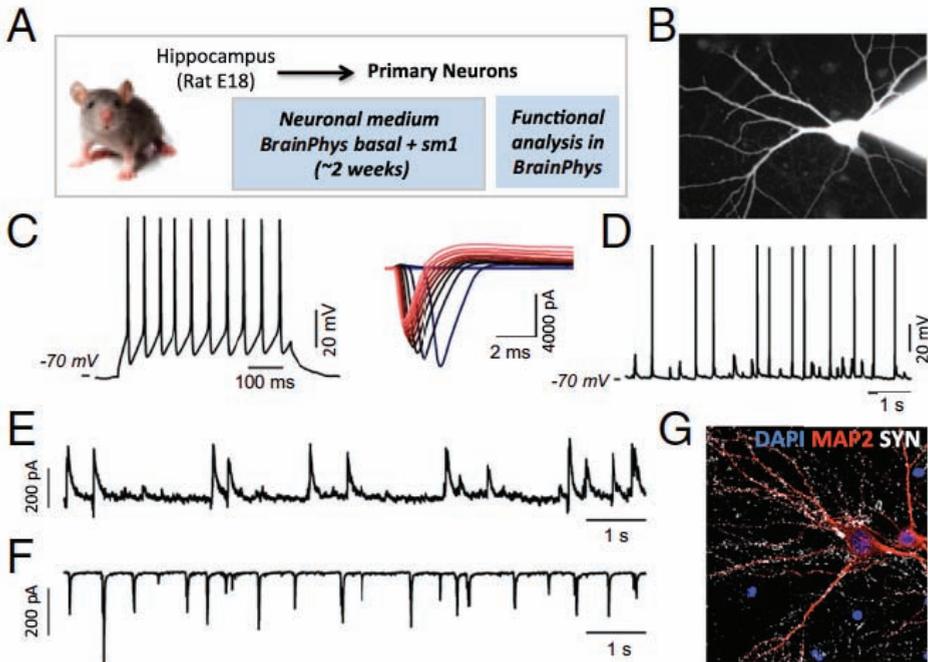


Figure 7. Differentiation of healthy and active rat primary hippocampal neurons in BrainPhys basal+sm1. (A) Rat hippocampal neurons were plated on coated glass coverslips (no feeder layer, no serum) in BrainPhys basal medium (with sm1) for ~ 2 weeks before characterization. (B) Images of a typical patch-clamped neuron filled with Rhodamine. On the right, the shadow is the Rhodamine-filled patch pipette. Electrophysiological characterization was performed in BrainPhys. (C) Train of APs evoked by a small 500-ms depolarizing step of current. I - V traces (clamp -70 mV, steps 5 mV) showing typical voltage-dependent Na^+ and K^+ currents. Colored traces correspond to increasing voltage steps. (D) Spontaneous APs recorded in current clamp (0 pA). (E) Spontaneous GABA synaptic events recorded in voltage clamp (at 0 mV). (F) Spontaneous AMPA synaptic events recorded in voltage clamp (-70 mV). (G) Coverslips were fixed after patching and processed for immunohistochemistry. Dendrites were stained with MAP2, presynaptic terminals with Synapsin, and nuclei with DAPI.

DISCUSSION

Improving Electrical and Synaptic Activity of Mature Human neurons in BrainPhys Basal with Supplements

In vivo, electrical activity is the essence of nervous system function. In this study, classic cell culture media were fully tested for their influence on fundamental neuronal activity. We discovered that many crucial neurophysiological properties were altered in virtually all classic media based on DMEM, Neurobasal, or serum. Less physiological conditions certainly reduce the clinical relevance of experimental models. To mimic electrical activity *in vivo* [53, 54], we designed a new neuronal culture medium (**Supplementary Table 5**). Over time, the enhancement of neuronal activity in BrainPhys medium was not detrimental to the survival of human neurons. To the contrary, we found that long-term exposure to BrainPhys medium enhanced neuronal synaptic function. Unlike other classical media, BrainPhys basal with a mixture of serum-free supplements optimally supported fundamental electrical and synaptic activity.

Maturation of Human Neurons in More Physiological Conditions *In Vitro*

We did not design BrainPhys basal to specifically accelerate maturation. However, the superiority of BrainPhys basal in supporting spontaneous electrical and synaptic activity may indirectly enhance neuronal development [40, 47-49].

In classic medium, without a feeder layer, the probability of patching mature and synaptically functional human neurons can be relatively low [4, 55]. Like other basal media, BrainPhys basal needs to be used with appropriate supplements. Further investigation might reveal new sets of growth factors that will better foster neuronal development. In our hands, BrainPhys basal with a specific set of supplements (**SUPPLEMENTARY METHODS** and **Supplementary Table 6**) was able to support the culturing of mature human neurons from several cell lines (greater than eight) obtained from a variety of patients. It is important to note, however, that we also encountered some cell lines, clones or batches that failed to generate mature and active neuronal cultures. Whenever these cell lines failed to meet our neurogenic standards, they also failed in other media. In addition, it is possible that enhancement of neuronal activity affects, to different degrees, various neuronal populations based on their maturity and synaptic profiles. Nevertheless, conditions supporting basic neuronal function *in vitro* will have overall important cellular implications. For example, *in vivo*, both ARC and synaptic activity play vital roles in the molecular/cellular mechanisms underlying memory formation and consolidation. As a result, ARC downregulation has been implicated in multiple neurological disorders [51]. We found that ARC protein expression was significantly increased in human neurons exposed to BrainPhys+sup medium. ARC expression is known to correlate with neuronal activity and regulate synaptic strength [51]; therefore, an increase in ARC in

healthy neurons exposed to BrainPhys+sup might play a role in the unsilencing of synapses of mature neurons. Although further investigation will be needed to clearly identify the neurodevelopment mechanisms involved, our results highlight the need to use more physiological media to differentiate brain cells *in vitro*.

Modeling Neurological Disorders in a Dish with More Realistic Conditions

We report that classic media, which we and others have used to differentiate and culture human neurons, are suboptimal to support electrical activity and, therefore, provide very different conditions than those observed in the living brain [53, 56]. Most neurological disorders are chronic and progressive and are very closely related to neuronal activity and synaptic function; thus, when modeling human neurological and psychiatric disorders *in vitro*, the lack of accurate physiological conditions may mask the real mechanisms of the pathologies. Although relevant phenotypes have been found between patient and control iPSC-derived neuronal cultures differentiated in currently used media [3], we predict that new phenotypes might be revealed from studying neurons in conditions that promote their electrophysiological activity. Neural models closely mimicking the living brain will be more likely to recapitulate the dysfunctions occurring in patients' brains and, in turn, lead to the discovery of more effective treatments for neurological and psychiatric disorders. The development of a new neuronal medium, such as BrainPhys basal, that sustains physiological neural activity *in vitro* takes us one step closer to this goal.

METHODS

Human iPSC-derived neurons were obtained with previously described protocols, using the four Yamanaka factors. The BrainPhys basal medium tested in this study was custom-made. Detailed methods can be found in the SUPPLEMENTARY METHODS.

CONFLICT OF INTEREST STATEMENT

The Salk Institute, C.B., and F.H.G. have filed a patent for the new BrainPhys medium described in this paper.

ACKNOWLEDGMENTS

Technical support, training and advice were obtained from several core facilities at the Salk Institute: Dr. Travis Berggren at the stem cell core, Dr. James Fitzpatrick at the bio-

photonic core, and the media core. We thank Bobbi Miller, Roberto Japelli and Lynne Moore for providing viral vectors. Dr. Jean-Michel Saffin and Romuald Arnaud provided technical advice for rat primary neuron culture. We thank Karen Carrasco and Alex Bryant for their help in some of the IHC analysis. We thank Dr. Tiago Goncalves, many members of the F.H.G. laboratory, and Dr. Rosa Cossart for fruitful discussions and feedback. This work was supported by the FP7 Marie Curie International Outgoing Fellowship for Career Development (to C.B.); the Limb Girdle Muscular Dystrophy 2I Research, Erna and Andrew Viterbi, and the Viterbi Family Foundation of the Jewish Community Foundation San Diego Neuroscience Research Initiative (to A.G.B.); Cancer Center Support Grant P30 CA014195 (to S.S.); and Ipsen Pharma, Annette C. Merle-Smith, The Leona M. and Harry B. Helmsley Charitable Trust grant, Bob and Mary Jane Engman, the JPB Foundation, G Harold and Leila Y. Mathers Foundation, and the National Institute of Mental Health Grant NIH MH095741 (to F.H.G.).

REFERENCES

1. 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).
2. Marchetto MC, Winner B, Gage FH. Pluripotent stem cells in neurodegenerative and neurodevelopmental diseases. *Hum. Mol. Genet.* 19(R1), R71-76 (2010).
3. Bellin M, Marchetto MC, Gage FH, Mummery CL. Induced pluripotent stem cells: the new patient? *Nat. Rev. Mol. Cell Biol.* 13(11), 713-726 (2012).
4. Brennand KJ, Simone A, Jou J *et al*. Modelling schizophrenia using human induced pluripotent stem cells. *Nature* 473(7346), 221-225 (2011).
5. 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).
6. Marchetto MC, Carroumeu C, Acab A *et al*. A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* 143(4), 527-539 (2010).
7. Nguyen HN, Byers B, Cord B *et al*. LRRK2 mutant iPSC-derived DA neurons demonstrate increased susceptibility to oxidative stress. *Cell Stem Cell* 8(3), 267-280 (2011).
8. Choi SH, Kim YH, Hebisch M *et al*. A three-dimensional human neural cell culture model of Alzheimer's disease. *Nature* 515(7526), 274-278 (2014).
9. Shao W, Zhang SZ, Tang M *et al*. Suppression of neuroinflammation by astrocytic dopamine D2 receptors via alphaB-crystallin. *Nature* 494(7435), 90-94 (2013).
10. Vilchez D, Boyer L, Morantte I *et al*. Increased proteasome activity in human embryonic stem cells is regulated by PSMD11. *Nature* 489(7415), 304-308 (2012).
11. Boyer LF, Campbell B, Larkin S, Mu Y, Gage FH. Dopaminergic differentiation of human pluripotent cells. *Curr. Protoc. Stem Cell Biol.* Chapter 1 Unit1H 6 (2012).
12. Giorgetti A, Marchetto MC, Li M *et al*. Cord blood-derived neuronal cells by ectopic expression of Sox2 and c-Myc. *Proc. Natl. Acad. Sci. U. S. A.* 109(31), 12556-12561 (2012).
13. Braz JM, Sharif-Naeini R, Vogt D *et al*. Forebrain GABAergic neuron precursors integrate into adult spinal cord and reduce injury-induced neuropathic pain. *Neuron* 74(4), 663-675 (2012).
14. Marchetto MC, Brennand KJ, Boyer LF, Gage FH. Induced pluripotent stem cells (iPSCs) and neurological disease modeling: progress and promises. *Hum. Mol. Genet.* 20(R2), R109-115 (2011).
15. Son EY, Ichida JK, Wainger BJ *et al*. Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell* 9(3), 205-218 (2011).
16. Caiazzo M, Dell'anno MT, Dvoretzkova E *et al*. Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature* 476(7359), 224-227 (2011).
17. Pang ZP, Yang N, Vierbuchen T *et al*. Induction of human neuronal cells by defined transcription factors. *Nature* 476(7359), 220-223 (2011).
18. Qiang L, Fujita R, Yamashita T *et al*. Directed conversion of Alzheimer's disease patient skin fibroblasts into functional neurons. *Cell* 146(3), 359-371 (2011).
19. Soldner F, Laganier J, Cheng AW *et al*. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* 146(2), 318-331 (2011).
20. Pfisterer U, Kirkeby A, Torper O *et al*. Direct conversion of human fibroblasts to dopaminergic neurons. *Proc. Natl. Acad. Sci. U. S. A.* 108(25), 10343-10348 (2011).
21. Eiraku M, Takata N, Ishibashi H *et al*. Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* 472(7341), 51-56 (2011).
22. Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Sudhof TC, Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463(7284), 1035-1041 (2010).

23. Soldner F, Hockemeyer D, Beard C *et al.* Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell* 136(5), 964-977 (2009).
24. Okabe S, Forsberg-Nilsson K, Spiro AC, Segal M, McKay RD. Development of neuronal precursor cells and functional postmitotic neurons from embryonic stem cells in vitro. *Mech. Dev.* 59(1), 89-102 (1996).
25. Liu Y, Liu H, Sauvey C, Yao L, Zarnowska ED, Zhang SC. Directed differentiation of forebrain GABA interneurons from human pluripotent stem cells. *Nat. Protoc.* 8(9), 1670-1679 (2013).
26. Zhang Y, Pak C, Han Y *et al.* Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* 78(5), 785-798 (2013).
27. Egawa N, Kitaoka S, Tsukita K *et al.* Drug screening for ALS using patient-specific induced pluripotent stem cells. *Sci. Transl. Med.* 4(145), 145ra104 (2012).
28. Johnson MA, Weick JP, Pearce RA, Zhang SC. Functional neural development from human embryonic stem cells: accelerated synaptic activity via astrocyte coculture. *J. Neurosci.* 27(12), 3069-3077 (2007).
29. Li XJ, Du ZW, Zarnowska ED *et al.* Specification of motoneurons from human embryonic stem cells. *Nat. Biotechnol.* 23(2), 215-221 (2005).
30. Hermann A, Gastl R, Liebau S *et al.* Efficient generation of neural stem cell-like cells from adult human bone marrow stromal cells. *J. Cell Sci.* 117(Pt 19), 4411-4422 (2004).
31. Koch P, Opitz T, Steinbeck JA, Ladewig J, Brustle O. A rosette-type, self-renewing human ES cell-derived neural stem cell with potential for in vitro instruction and synaptic integration. *Proc. Natl. Acad. Sci. U. S. A.* 106(9), 3225-3230 (2009).
32. Lancaster MA, Renner M, Martin CA *et al.* Cerebral organoids model human brain development and microcephaly. *Nature* 501(7467), 373-379 (2013).
33. Reinhardt P, Schmid B, Burbulla LF *et al.* Genetic correction of a LRRK2 mutation in human iPSCs links parkinsonian neurodegeneration to ERK-dependent changes in gene expression. *Cell Stem Cell* 12(3), 354-367 (2013).
34. Shi Y, Kirwan P, Smith J, Robinson HP, Livesey FJ. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nat. Neurosci.* 15(3), 477-486, S471 (2012).
35. Jones HC, Keep RF. Brain fluid calcium concentration and response to acute hypercalcaemia during development in the rat. *J. Physiol.* 402 579-593 (1988).
36. Brocard F, Shevtsova NA, Bouhadfane M *et al.* Activity-dependent changes in extracellular Ca²⁺ and K⁺ reveal pacemakers in the spinal locomotor-related network. *Neuron* 77(6), 1047-1054 (2013).
37. Decker CF, Aras A, Decker LE. Determination of Magnesium and Calcium in Cerebrospinal Fluid by Atomic Absorption Spectroscopy. *Anal. Biochem.* 8 344-348 (1964).
38. Brewer GJ, Torricelli JR, Evege EK, Price PJ. Optimized survival of hippocampal neurons in B27-supplemented Neurobasal, a new serum-free medium combination. *J. Neurosci. Res.* 35(5), 567-576 (1993).
39. H. D, Segal MB. *Physiology of the CSF and Blood-Brain Barriers.* CRC Press, Boca Raton, FL. (1996).
40. Schaarschmidt G, Schewtschik S, Kraft R *et al.* A new culturing strategy improves functional neuronal development of human neural progenitor cells. *J. Neurochem.* 109(1), 238-247 (2009).
41. Itoh K, Nakamura K, Iijima M, Sasaki H. Mitochondrial dynamics in neurodegeneration. *Trends Cell Biol.* 23(2), 64-71 (2013).
42. Cooper O, Seo H, Andrabi S *et al.* Pharmacological rescue of mitochondrial deficits in iPSC-derived neural cells from patients with familial Parkinson's disease. *Sci. Transl. Med.* 4(141), 141ra190 (2012).
43. Knott AB, Perkins G, Schwarzenbacher R, Bossy-Wetzl E. Mitochondrial fragmentation in neurodegeneration. *Nat. Rev. Neurosci.* 9(7), 505-518 (2008).
44. Mergenthaler P, Lindauer U, Dienel GA, Meisel A. Sugar for the brain: the role of glucose in physiological and pathological brain function. *Trends Neurosci.* 36(10), 587-597 (2013).

45. Zilberter Y, Zilberter T, Bregestovski P. Neuronal activity in vitro and the in vivo reality: the role of energy homeostasis. *Trends Pharmacol. Sci.* 31(9), 394-401 (2010).
46. Gruetter R, Novotny EJ, Boulware SD *et al.* Direct measurement of brain glucose concentrations in humans by ¹³C NMR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 89(3), 1109-1112 (1992).
47. Hanse E, Seth H, Riebe I. AMPA-silent synapses in brain development and pathology. *Nat. Rev. Neurosci.* 14(12), 839-850 (2013).
48. Spitzer NC. Electrical activity in early neuronal development. *Nature* 444(7120), 707-712 (2006).
49. Blankenship AG, Feller MB. Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nat. Rev. Neurosci.* 11(1), 18-29 (2010).
50. Guzowski JE, McNaughton BL, Barnes CA, Worley PF. Environment-specific expression of the immediate-early gene Arc in hippocampal neuronal ensembles. *Nat. Neurosci.* 2(12), 1120-1124 (1999).
51. Korb E, Finkbeiner S. Arc in synaptic plasticity: from gene to behavior. *Trends Neurosci.* 34(11), 591-598 (2011).
52. McIntyre CK, Miyashita T, Setlow B *et al.* Memory-influencing intra-basolateral amygdala drug infusions modulate expression of Arc protein in the hippocampus. *Proc. Natl. Acad. Sci. U. S. A.* 102(30), 10718-10723 (2005).
53. Contreras D. Electrophysiological classes of neocortical neurons. *Neural Netw.* 17(5-6), 633-646 (2004).
54. Cang J, Isaacson JS. In vivo whole-cell recording of odor-evoked synaptic transmission in the rat olfactory bulb. *J. Neurosci.* 23(10), 4108-4116 (2003).
55. Tang X, Zhou L, Wagner AM *et al.* Astroglial cells regulate the developmental timeline of human neurons differentiated from induced pluripotent stem cells. *Stem Cell Res.* 11(2), 743-757 (2013).
56. Moore AR, Filipovic R, Mo Z, Rasband MN, Zecevic N, Antic SD. Electrical excitability of early neurons in the human cerebral cortex during the second trimester of gestation. *Cereb. Cortex* 19(8), 1795-1805 (2009).
57. Ladewig J, Mertens J, Kesavan J *et al.* Small molecules enable highly efficient neuronal conversion of human fibroblasts. *Nat. Methods* 9(6), 575-578 (2012).
58. Mertens J, Stuber K, Poppe D *et al.* Embryonic stem cell-based modeling of tau pathology in human neurons. *Am. J. Pathol.* 182(5), 1769-1779 (2013).

SUPPLEMENTARY METHODS

Human Neurons

Human dermal fibroblasts were reprogrammed into pluripotent cells with the four Yamanaka factors (Oct3/4, Sox2, Klf4 and c-Myc) either in a retroviral vector or a non-integrating Sendai viral vector. Human iPSCs and ESCs were differentiated into NPCs as previously described [11]. For neural maturation, NPCs were plated on glass coverslips (Fisher Scientific, Cat. No. 12-545-80) coated with poly-ornithine (Sigma, Cat. No. P3655) and laminin (Invitrogen, Cat. No. 23017-015) and cultured in neuronal maturation medium (NM, see “**Media and Extracellular Solutions**” below) in 24-well plates. Half of the neuronal medium was gently replaced two to three times a week. The plates were kept in a humidified incubator at 37 °C with 5% CO₂ and 21% O₂. The pH (~7.3-7.4) and the osmolarity (~300-305 mOsmol/L) of medium were kept constant over time.

Direct Conversion of Human Dermal Fibroblasts into Induced Neurons (iNs)

Primary human dermal fibroblasts were established from skin biopsies from healthy donors and cultured in DMEM medium containing 15% FBS and 0.1% NEAA (all Gibco). Fibroblast cell line #1 was obtained from ATCC (BJ-CRL-2522™ foreskin fibroblasts); lines #2 and #3 were obtained from the Coriell Institute (catalog IDs GM22159 and AG08498). Direct Ngn2/Ascl1-based conversion was performed as previously described by Ladewig *et al.* with slight modifications [57]. Coding sequences for human Ascl1 and Ngn2 were linked by a 2A peptide sequence and cloned into pLVX-Tight-Puro construct (Clontech), resulting in the pLVXTP-N2A. Lentiviral particles for pLVX-EtO and pLVXTP-N2A were produced in HEK-293FT cells and enriched by centrifugation. Forty-eight hours following transduction, transgenic iN-competent fibroblasts were further passaged in the presence of G418 (200 µg/ml; Gibco) and puromycin (1 µg/ml; Sigma-Aldrich) in tetracycline-free FBS-containing media. To generate induced neurons, the medium was changed to induced neuron conversion (iNC) medium based on either DMEM:F12/Neurobasal (DN; 1:1 v/v) or BrainPhys (BP) for three weeks. iNC contains the following supplements: N2 supplement and B27 supplement (both 1x; Gibco), doxycycline (2 µg/ml; Sigma-Aldrich), Laminin (1 µg/ml, Life Technologies), dibutyryl cyclic-AMP (500 µg/ml; Sigma-Aldrich), human recombinant noggin (150 ng/ml; Peprotech), LDN-193189 (5 µM; Cayman Chemical Co) and A83-1 (5 µM; Stemgent), CHIR99021 (3 µM; LC Laboratories), forskolin (5 µM; LC Laboratories) and SB-431542 (10 µM; Cayman Chemical Co). The iNC medium (DN-iNC or BP-iNC) was changed every three days. To test whether direct iN conversion could be performed in BrainPhys-based media, iN cultures that converted for three weeks in either DN-iNC or BP-iNC were fixed with 4% PFA and stained for the neuronal markers β-III-tubulin (rabbit, 1:3000; Millipore), hTau (PHF1, mouse, 1:500; kind gift from Peter Davies), NeuN (mouse, 1:100; Millipore) and

Map2ab (chicken, 1:300; Abcam); see **Figure 6B**. For full functional maturation, iN cells were gently dislodged from their conversion plate using TrypLE and relocated on top of a monolayer culture of mouse astrocytes (see Mertens *et al.* AJP 2013 [58]) and further cultured in iN maturation (iNM) medium containing the supplements GDNF, BDNF (both 20 ng/ml; R&D), dibutyryl cyclic-AMP (500 µg/ml; SigmaAldrich), doxycycline (2 µg/ml; Sigma-Aldrich) and laminin (1 µg/ml; Life Technologies). iNM-containing media were again either based on DMEM:F12/Neurobasal (DN; 1:1 v/v) or BrainPhys (BP), resulting in DN-iNM or BP-iNM, which were changed every three days. Following three weeks of maturation (six weeks of conversion in total), cultures were fixed with 4% PFA and stained for hTau and ARC (rabbit, 1:200; Acris); see **Figure 6F**.

Rodent Neurons

Rat hippocampal dissociated cultures were prepared at embryonic day 18. Hippocampal dissections were briefly treated with Papain, then gently dissociated and plated in BrainPhys basal + Neurocult sm1 supplement (Stem Cell Technologies, Cat. No. 07511) on glass coverslips coated with poly-L-Lysine. To test BrainPhys directly on the brain *ex vivo* rather than in artificial cultures, we prepared 300-µm thick slices of the mouse hippocampus (C57Bl6).

Media and Extracellular Solutions

Neural media (NM) consisted of basal media and supplements. We used the following basal media: BrainPhys basal (BP, custom-made), Dulbecco's Modified Eagle's Medium basal (DMEM/F12; Gibco, Cat. No. 10565-018), Neurobasal-A (NB; Gibco, Cat. No. 10888-022) or a mixture of DMEM/F12 and Neurobasal-A (DN, 50:50). Most of the DMEM used in these experiments was obtained from Life Technologies; a few batches were custom-made in house. Neurobasal was always obtained from Life Technologies. ACSF and BrainPhys basal were custom-made in house. To evaluate the influence of DMEM in this study, we always used DMEM/F12 and referred to it as DMEM or DMEM/F12. To evaluate the influence of Neurobasal in this study, we almost always used Neurobasal-A and referred to it as Neurobasal, NB or NBA. We also tested other commercial versions of Neurobasal with unknown formulations (e.g., ref# A1413701 or ref# 21103-049 from Life Technologies) and found no improvement in acute electrophysiology. For human iPSCs/ESCs, basal media were complemented with the same cocktail of supplements: 1x N2 (Gibco, Cat No. 17502-048), 1x B27 (Gibco, Cat No. 17504-044), Brain-derived Neurotrophic Factor (BDNF, 20 ng/ml; Peprotech, Cat No. 450-02), Glia-derived Neurotrophic Factor (GDNF, 20 ng/ml; Peprotech, Cat No. 450-10), ascorbic acid (AA, 200 nM; Sigma, Cat No. A0278), dibutyryl cyclic AMP (cAMP, 1mM Sigma, Cat No. D0627) and laminin (1 µg/ml; Invitrogen, Cat No 23017-015). For Neurobasal, Glutamax was also added (0.5 mM, Gibco, Cat. No. 35050061). In addition to this cocktail of supplements

used for neuronal maturation, we occasionally tested the influence of cholesterol (10 $\mu\text{g}/\text{ml}$; Sigma) and fetal bovine serum (FBS 10%) on acute neuronal activity. The composition of ACSF was adjusted to match the inorganic salt concentration and osmolarity of the DMEM and BrainPhys basal when acutely compared. ACSF contained (in mM) 121 NaCl, 4.2 KCl, 1.1 CaCl₂, 1 MgSO₄ (or 0.4 MgSO₄ and 0.3 MgCl), 29 NaHCO₃, 0.45 NaH₂PO₄-H₂O, 0.5 Na₂HPO₄ and 20 glucose (all chemicals from Sigma). Experiments to measure electrophysiological activity were performed in ACSF or in basal media with or without supplements. Long-term experiments were always performed with basal media and supplements.

Patch Clamping

Across all experiments, 292 neurons were patched and analyzed. For whole-cell patch-clamp recordings, individual coverslips were transferred into a heated recording chamber and continuously perfused (1 ml·min⁻¹) with either basal media or artificial cerebrospinal fluid (ACSF) bubbled with a mixture of CO₂ (5%) and O₂ (95%) and maintained at 25 °C. For targeted whole-cell recordings, we used a 40× water-immersion objective, differential interference contrast filters (all Olympus), an infrared digital camera (Rolera XR - Qimaging), a digidata 1440A/Multiclamp 700B and Clampex 10.3 (Molecular Devices). Patch electrodes were filled with internal solutions containing 130 mM K-gluconate, 6 mM KCl, 4 mM NaCl, 10 mM Na-HEPES, 0.2 mM K-EGTA; 0.3 mM GTP, 2 mM Mg-ATP, 0.2 mM cAMP, 10 mM D-glucose, 0.15% biocytin and 0.06% rhodamine. The pH and osmolarity of the internal solution were close to physiological conditions (pH 7.3, 290–300 mOsmol). Data were all corrected for liquid junction potentials (10 mV). Electrode capacitances were compensated on-line in cell-attached mode (~7 pF). Recordings were low-pass filtered at 2 kHz, digitized, and sampled at intervals of 50 ms (20 kHz). To control the quality and the stability of the recordings throughout the experiments, access resistance, capacitance and membrane resistance were continuously monitored on-line and recorded. The resistance of the patch pipettes was between 3 and 5 MOhm. The access resistance of the cells in our sample was ~40 MOhm on average. Patch-clamping results of every tested solution were confirmed by recovery to level comparable to the control. Synaptic antagonists were only used on a subset of neurons to confirm the nature of the spontaneous synaptic events detected. All events showing typical synaptic AMPA-receptor-mediated kinetics were blocked by NBQX (10 μM NBQX, Sigma Ref#N183) and were observed exclusively in voltage clamp at -70 mV (close to Cl⁻ reversal potential). All events showing typical synaptic GABA_A receptor-mediated kinetics were blocked by gabazine (10 μM SR95531, Sigma Ref#S106) and were observed exclusively in voltage clamp at 0 mV (close to Na⁺ reversal potential).

Multi-Electrode Arrays (MEAs)

Multi-Well MEA Plates. MEA plates were composed of 48 wells with each well containing an array of 16 individual embedded nano-textured gold microelectrodes (~40-50 μm diameter; 350 μm center-to-center spacing) with 4 integrated ground electrodes, for a total of 768 channels (Axion Biosystems). Prior to culture, each well was coated with 200 μl of a filter-sterilized 0.1% solution of polyethylenimine (Sigma Aldrich) in borate buffer (3.10 g of boric acid (Fisher Scientific) and 4.75 g of sodium tetraborate (Sigma Aldrich) in 1 L of distilled water) at room temperature. The solution was removed after one hour, and the wells were washed four times with sterile water and air-dried overnight in a sterile biological safety cabinet.

Cell Culture on MEAs. One vial of cryopreserved human iPSC-derived iCell neurons (Cellular Dynamics International) was thawed and pelleted by centrifugation according to the manufacturer's protocol, and resuspended at a concentration of 28,000 viable neurons/ μl in iCell Neuron Maintenance Medium with 10 $\mu\text{g}/\text{ml}$ laminin (Sigma Aldrich). A 5- μl droplet of the cell suspension (140,000 neurons) was added to the center of each well in the MEA, directly over the electrode area. Sterile water was added to the area around the wells to reduce evaporation, and the MEA was incubated with the seeded neurons in a cell culture incubator at 37°C and 5% CO₂ for one hour. 300 μl of iCell Neuron Maintenance Medium was then gently added to each well, and the water surrounding the wells was removed. Two days after the cell plating (day *in vitro* (DIV) 2), the medium was replaced with one of eight different media: Neurobasal-A (Life Technologies) with 10% defined fetal bovine serum (FBS; HyClone); BrainPhys with 10% FBS; DMEM/F12 (Life Technologies) with 10% FBS; Neurobasal-A supplemented with N2 (Gibco), B27 (Gibco), BDNF (20 ng/ml; Peprotech), GDNF (20 ng/ml; Peprotech), ascorbic acid (200 nM; Sigma), cAMP (1mM; Sigma), and laminin (1 $\mu\text{g}/\text{ml}$; Life Technologies); BrainPhys with the aforementioned supplements; DMEM/F12 (Life Technologies) with the aforementioned supplements; a 1:1 mixture of Neurobasal-A/DMEM with the aforementioned supplements; or fresh iCell Neuron Maintenance Medium (six wells each medium). At DIV 6, 9, and 13, 50% of the medium in each well was exchanged with fresh medium. On DIV 16, all medium was replaced with BrainPhys basal+sup.

MEA Recording and Data Analysis. Between DIV 2 and 21, spontaneous activity was recorded for 10 min each day at 37 °C using the Maestro MEA system (Axion Biosystems) and the associated Axion Integrated Studio (AxIS 1.8.1.5). The 768 channels of the MEA were sampled simultaneously with a gain of 1200X and a sampling rate of 12.5 kHz/channel. For all recordings, a Butterworth band-pass filter (200 Hz – 3 kHz) was applied along with an adaptive threshold spike detector set at 6 \times standard deviation. Data from the recordings were saved to three different file types simultaneously; a raw data file (*.raw file) that included all data, a spike counts file (*.csv file) that included the spikes per electrode with a 1 s bin time, and an alpha map (*.map file) that included spike timing

and profile information. The frequency reported in the analysis represent the average frequency of 96 to 192 electrodes (6 to 12 wells) for each tested condition. The percentage of active electrode represents the number of electrode with a spike frequency >0.005 Hz.

Analysis and Statistics

Statistical analysis of electrophysiology data and calcium imaging was assisted with Clampfit 10.3, Matlab 2011b, Igor Pro 6, Prism 5, MiniAnalysis and Microsoft Excel. Standard errors of the mean were reported. Statistical significance was assessed with two-tailed non-parametric paired (Wilcoxon) or unpaired (Mann Whitney) tests.

Calcium Imaging

Neurons attached on glass coverslips were incubated for ~20 min (37 °C, 5% CO₂, and 95% humidity) with 4 μM calcium-sensitive dye Fluo-4 AM (Invitrogen, Cat No. F14201) in maturation medium. Fluo-4 was washed and coverslips were transferred into a heated recording chamber and continuously perfused (1 ml·min⁻¹) with either basal media or ACSF bubbled with a mixture of CO₂ (5%) and O₂ (95%) and maintained in the chamber at 35 °C. We waited at least 15 min after the coverslip transfer before starting any recording. Calcium imaging movies were acquired with a laser-scanning microscope at 471 nm (Olympus, Fluoview FV1000MPE) and a 25× objective (Olympus, XLPLN NA 1.05). At least two consecutive five-minute movies were recorded in each condition (control, tested medium, recovery). In many cases the extracellular medium defined as the control and recovery medium was ACSF, but similar results were obtained when the order of media perfusion was changed (e.g., control = BrainPhys, tested = ACSF, recovery = BrainPhys). When necessary, gain sensitivity and focus were adjusted between movies. Laser power was not changed (~2%). After changing each perfusate, we also waited at least five minutes to make sure the entire bath solution was replaced before initiating the next recording. We analyzed no more than three fields of view per coverslip. Regions of interest (ROIs) were categorized as spontaneously active when at least one clear neuronal calcium event was detected on a soma. Neuronal calcium events were defined as a sharp transient increase in fluorescence intensity (Fluo-4 AM, $dF/F > 5\%$, fast rise, slower decay). The time series for each ROI were calculated with the Olympus Fluoview FV1000 software.

Immunohistochemistry

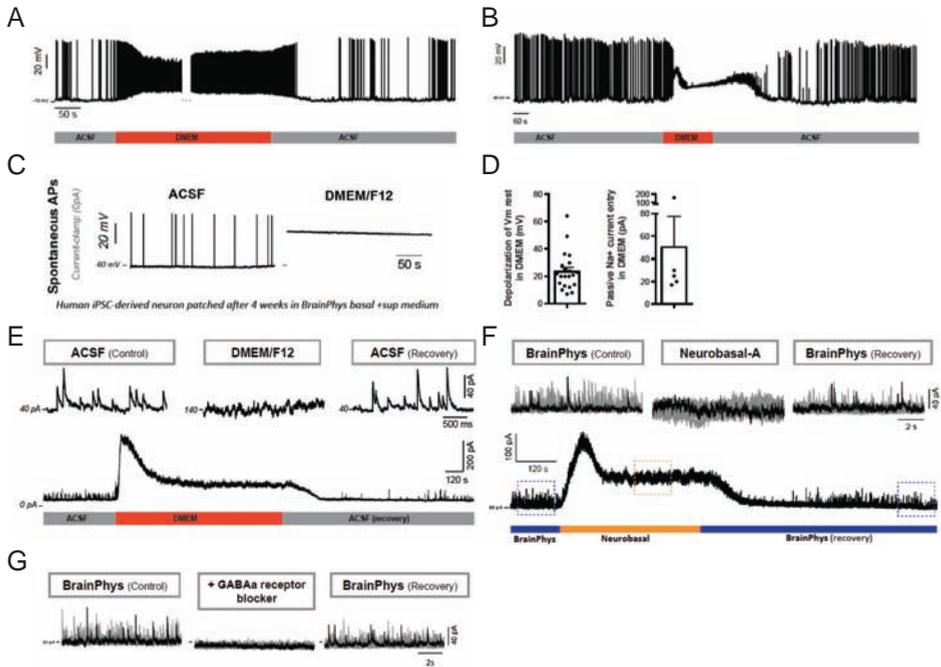
Immunohistochemistry experiments were performed on neurons plated on glass coverslips in 24-well plates. Comparison between media conditions were made on the same cell line, growing side by side in the same plate with the same original cell density. Standard immunohistochemistry protocols were used. Coverslips were stained with DAPI and a combination of the following antibodies: mouse-Map2(2a+2b) (1:500, Sigma), mouse-TUJ1 (1:1000, Covance), rabbit-Synapsin1 (1:500, Calbiochem), rabbit-TH (1:500, Pel-

Freez), rabbit-GFAP (1:200, Dako), rabbit-GABA (1:500, Sigma), goat-DCX (1:500, Santa Cruz).

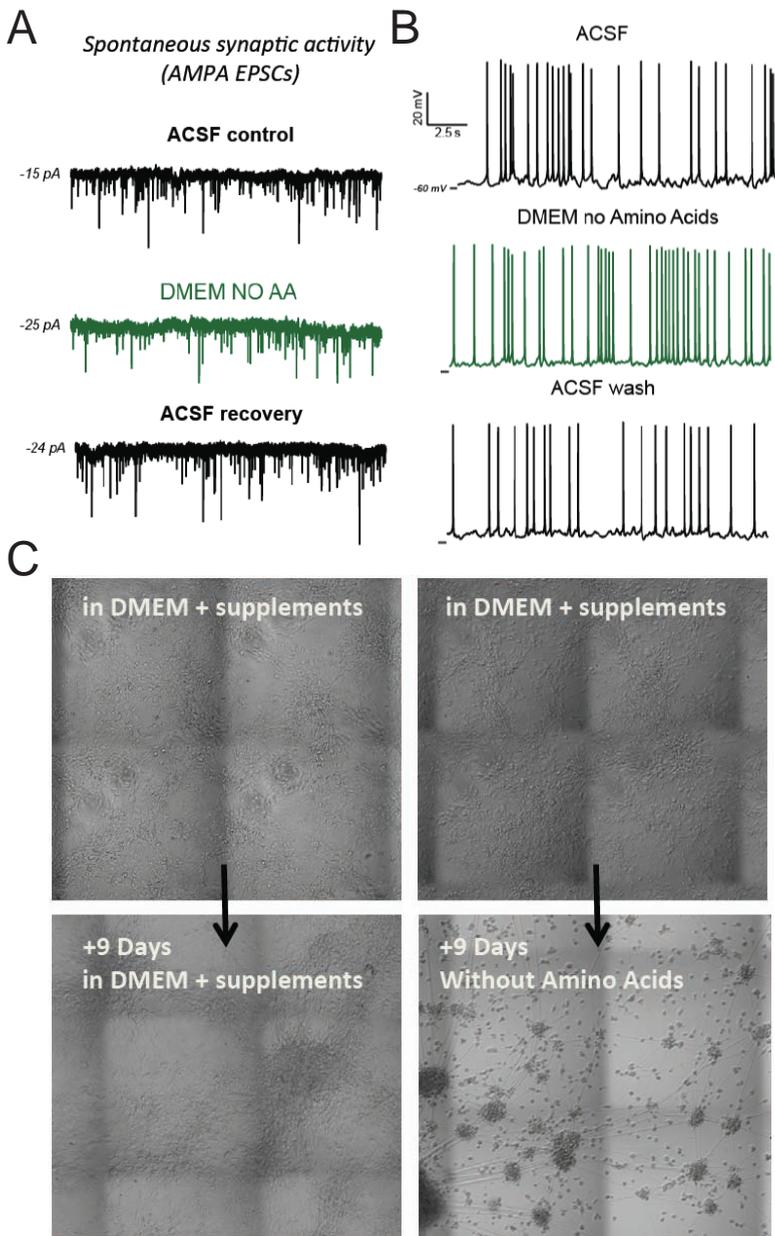
Optimal Use of BrainPhys and Cautionary Notes

BrainPhys basal was specifically designed to support long-term basic synaptic and electrical activity of mature neurons *in vitro*. BrainPhys basal + supplements should be used on mature neurons for at least one or two weeks before phenotyping the cells. Enhancing electrical activity may also have indirect beneficial influence on neuronal development, and we have shown that BrainPhys basal can also be used successfully in some neural maturation protocols. However, optimal maturation and survival can only be achieved with appropriate supplements, physical supports and healthy neuronal progenitors. These latest technical aspects should be carefully controlled and adapted based on specific experimental objectives. It is also important to note that, despite the obvious benefits of enhancing neurophysiological activity, there is a possibility that, at least in some cases, increasing cellular metabolism makes the cells more sensitive to stress (e.g., if metabolically active during *in vitro* manipulation). In addition, hyper-osmotic solution may trigger neurodegeneration. The osmolarity of BrainPhys basal is close to physiological values (~300 mOsm/L); as evaporation of the medium may induce detrimental osmotic stress for the cells, special care should be applied to limit evaporation in the incubator. If BrainPhys basal is used in maturation medium, the switch from neural progenitor medium to BrainPhys+sup should be gentle and gradual.

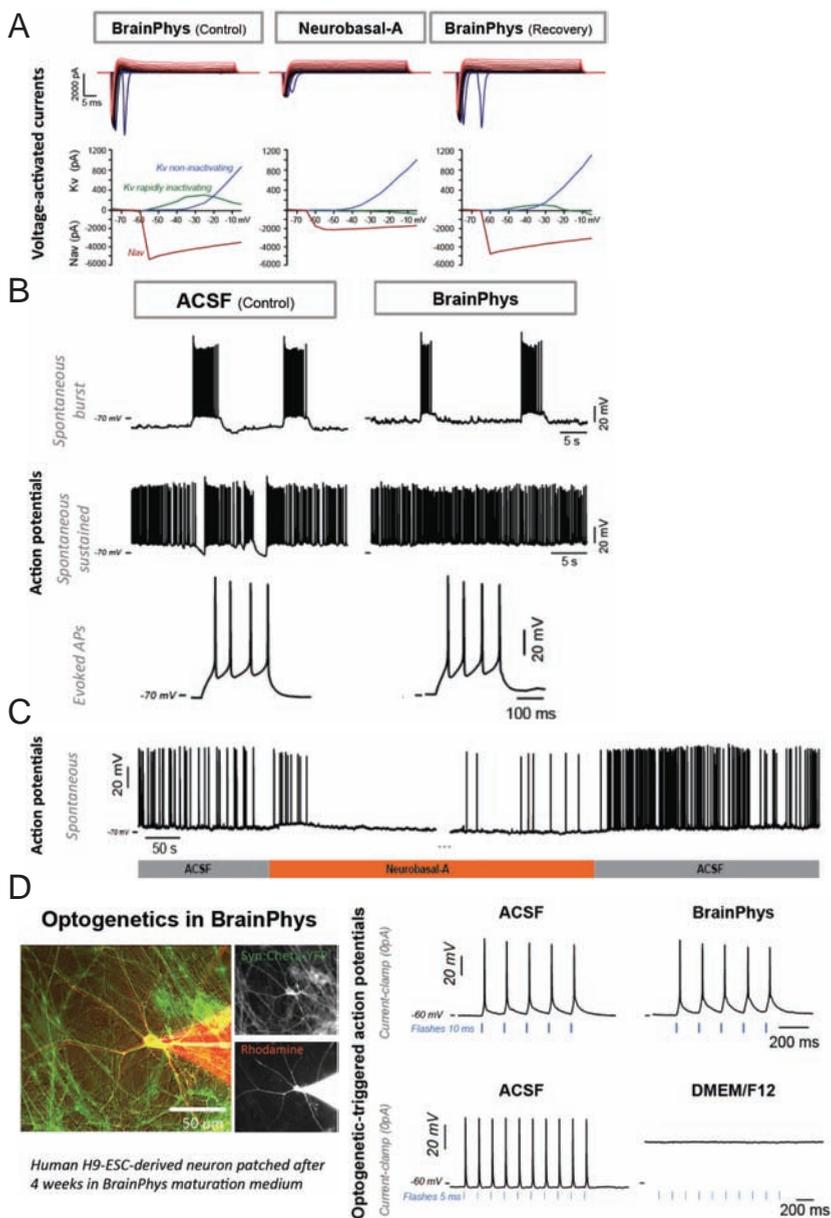
SUPPLEMENTARY INFORMATION



Supplementary Figure 1. (related to Figure 1 and 2). **Influence of different basal media on the activity of mature human neurons.** DMEM depolarized the membrane potential of every examined neuron and, in most cases, it also saturated and silenced spontaneous action potential firing. The traces were obtained from whole-cell current-clamp spontaneous recordings of stem cell-derived human neurons in ACSF, DMEM or Neurobasal. **(A)** Rare example in which DMEM did not saturate or silence AP firing. **(B,C)** Two more typical examples, where DMEM-induced depolarization silenced spontaneous AP firing. **(D)** Neurons recorded in current clamp in ACSF were consistently depolarized by perfusion of DMEM/F12 ($n = 22$). This membrane potential depolarization was associated with passive sodium influx measured in voltage clamp at -70 mV ($n = 5$). **(E,F)** Voltage-clamp (0 mV) recording showed that DMEM/F12 and Neurobasal induced large Cl^- currents and blocked spontaneous inhibitory synaptic activity. **(G)** Typical GABAergic synaptic activity in BrainPhys that could be blocked by SR95531 (Gabazine) and recovered.



Supplementary Figure 2. DMEM without amino acids enhanced electrical activity but did not sustain long-term survival and maturation *in vitro*. (A) Spontaneous EPSCs (AMPA) were similar in ACSF and custom-made DMEM without amino acids. (B) Resting membrane potential and action potentials were similar in ACSF and custom-made DMEM without amino acids. (C) Custom-made DMEM without amino acids (+ standard supplements) failed to support neuronal survival *in vitro* within about a week.

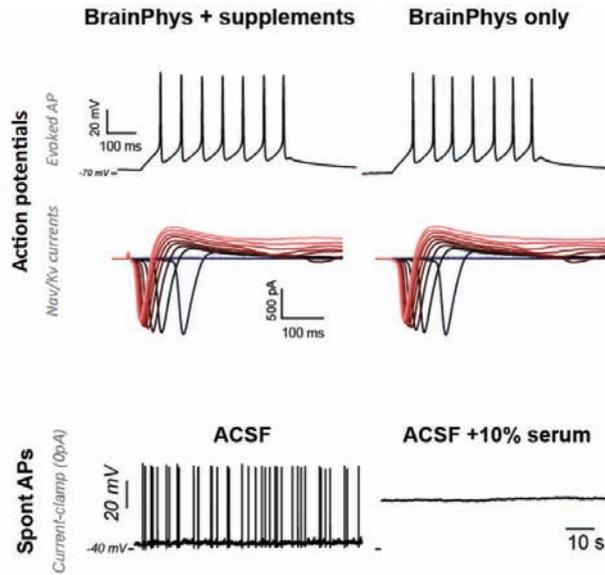


Supplementary Figure 3. (related to Figure 1 and 2). **Influence of different basal media on the activity of human neurons.** (A) *Top:* Voltage-clamp I - V traces (clamp -70 mV, steps 5 mV) show the voltage-dependent Na^+ and K^+ currents of a typical neuron recorded in BrainPhys, Neurobasal and back to BrainPhys. *Bottom:* I - V curves of analyzed Nav and Kv peak currents. (B) Spontaneous and evoked APs were similar in BrainPhys and ACSF. (C) Resting membrane potential was unaffected in Neurobasal but the spontaneous firing rate was strongly reduced. (D) The activity of human ESC- or iPSC-derived neurons maturing in BrainPhys and expressing optogenes was effectively controlled by light either in ACSF or BrainPhys but not in DMEM/F12 or Neurobasal.

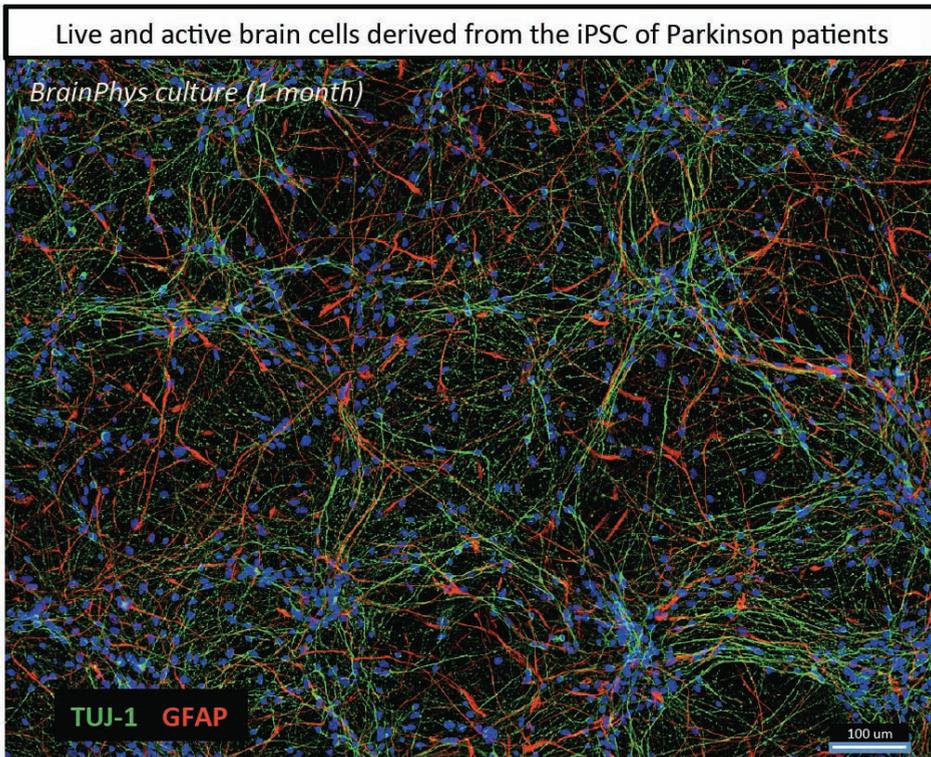
Supplements tested with
BrainPhys basal medium

- N2
- B27
- BDNF
- GDNF
- cAMP
- Acid ascorbic
- Laminin
- Cholesterol

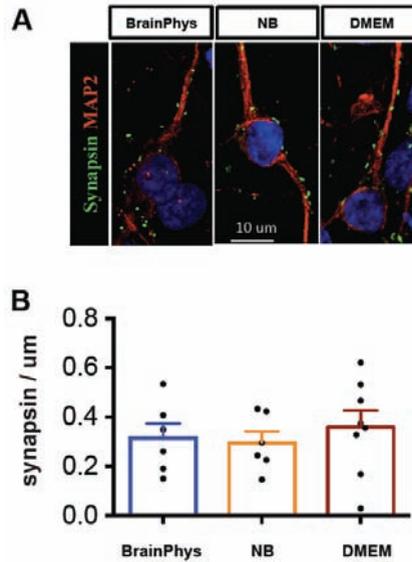
Those supplements did not acutely affect synaptic activity either



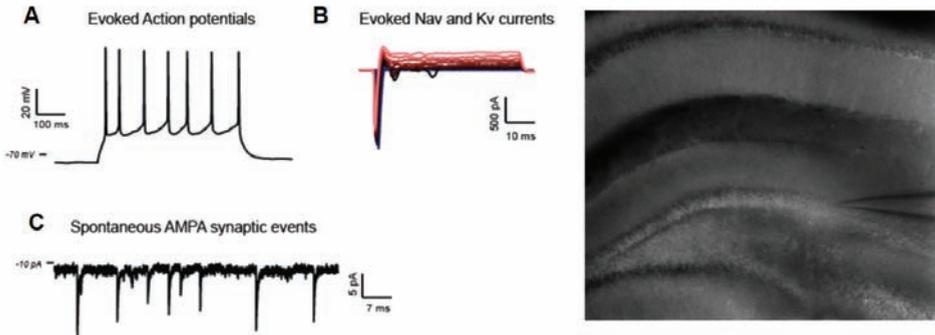
Supplementary Figure 4. (related to Figure 3). **Supplements tested with BrainPhys basal did not acutely affect the electrical activity. In contrast, serum completely saturated and blocked basic neuronal activity.**



Supplementary Figure 5. (related to Figure 4). **BrainPhys basal + serum-free supplements supports survival and growth of both human neurons and astrocytes.**



Supplementary Figure 6. (related to Figure 5). **Proximal synaptic puncta.** (A,B) Immunostaining to quantify the number of synapsin puncta on proximal MAP2+ dendrites (within $\sim 100 \mu\text{m}$ from soma) did not reveal any significant difference between the tested human neurons maturing in BrainPhys or other media.



Supplementary Figure 7. (related to Figure 7). **Patch-clamp recordings of mouse brain slices *ex vivo* in BrainPhys.** (A-C) Typical functional granule neuron patched from the adult mouse hippocampus in BrainPhys. (A) APs evoked with a 500-ms step of current from resting membrane potential (current clamp 0 pA). (B) *I-V* traces of incremental 5-mV steps (voltage clamp -70 mV) showing Nav and Kv currents. (C) Spontaneous AMPA synaptic events (voltage clamp -70 mV).

Supplementary Table 1. (related to Figure 3). **Properties of spontaneous firing activity of mature human iPSC neurons exposed over time to BrainPhys basal+sup medium on multi-well MEAs.** Data recorded from commercial iCell neurons (Cellular Dynamics) after 4 days and 12 days in BrainPhys medium.

MEAs 48-well plates, 4 x 4 electrodes per well		4 days		12 days	
		Mean	S.E.M.	Mean	S.E.M.
Total number of recording electrodes		96	N/A	96	N/A
Number of wells		6	N/A	6	N/A
% active electrodes (>0.005 Hz)		88	N/A	73	N/A
Spontaneous spike frequency (Hz)		0.330	0.053	0.252	0.049
Spike amplitude (mV)		0.0263	0.0006	0.0244	0.0006
Burst Analysis	Number of bursts per minute	0.71	0.10	0.59	0.11
	% spikes in burst	32	3	27	4
	Burst duration (s)	6.50	1.33	6.78	0.86
	Number of spikes per burst	10	1	13	3
	Inter-spike interval during burst (s)	0.96	0.19	0.99	0.20
	Spike frequency during burst (Hz)	2.90	0.31	3.11	0.55
	Peak frequency during burst (Hz)	20.47	3.90	29.24	7.90
	Interburst interval (s)	107.73	13.58	101.49	14.98

Supplementary Table 2. (related to Figure 4). **Electrophysiological whole-cell properties of single mature human neurons cultured several weeks in BrainPhys basal + supplements.** Descriptive statistical analysis was performed on $n = 106$ neurons with evoked repetitive firing (>10 Hz) obtained from four different healthy human iPSC-derived lines and one human ESC-derived line (H9). NPCs were plated on coated glass coverslips (no feeder layer) and matured in BrainPhys basal+sup for several weeks before patch clamping.

	Median	25 th percentile	75 th percentile	Mean	Minimum	Maximum
Days in BrainPhys basal+sup	31	25	40	41	17	152
Membrane "passive" properties						
Capacitance (pF)	47	32	72	61	14	261
Membrane resistance (MOhm)	1399	949	2048	1610	158	4357
Estimated V_m rest at 0 pA (membrane potential mV)	-50	-58	-43	-50	-75	-19
Sodium and potassium voltage-dependent currents						
Threshold to activate Nav max peak (membrane potential mV)	-40	-45	-35	-39	-55	-5
Nav currents (max peak in pA)	-2828	-4674	-2071	-3575	-9848	-1074
Kv rapidly inactivating currents (pA, max peak)	126	53	253	164	1	1043
Kv slowly or non-inactivating currents (pA, at -10 mV)	386	258	546	416	59	922
Evoked action potential properties						
AP max amplitude (membrane potential mV)	22	16	30	23	5	45
Firing frequency above -10 mV (Hz)	14	12	18	15.6	10	44
Spike duration (ms)	5.9	5	7.1	6.4	3.4	24
Max rise slope (mV/ms)	86	64	123	97	27	240
Max decay slope (mV/ms)	-29	-35	-24	-29	-52	-11
Afterhyperpolarization (mV)	-9.9	-13	-6	-9.6	-23	-0.45
Threshold evoked AP (membrane potential mV)	-38	-40	-33	-37	-49	-25
AMPA spontaneous events – functional excitatory synapses						
Frequency spontaneous EPSCs (Hz)	1.3	0.2	4.1	3	0.02	30
Amplitude median (pA)	12	9.5	14	13	6.1	30
Amplitude mean (pA)	14	11	18	15	7.4	40
Rise time 10-90% (ms)	0.88	0.76	1.1	0.96	0.41	2.4
Decay time (ms)	1.7	1.4	2.1	1.9	0.73	5.5
Halfwidth (ms)	1.8	1.5	2.2	1.9	0.82	5.2
Patch						
Pipette resistance (MOhm)	4.5	3.8	5	4.4	3	6
Access resistance (MOhm)	24	18	33	26	11	58

Supplementary Table 3 (related to Figure 5). **Electrophysiology analysis of neurons cultured side by side for 2-6 weeks in BrainPhys+sup or DMEM/F12+sup medium (patched in ACSF for fair comparison).** sAP, cells with single evoked action potential; rAP, cells with repetitive evoked action potentials. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

	DMEM Culture Neurons	BrainPhys Culture Neurons	% Change BrainPhys vs. DMEM	P Value (Mann Whitney)
Blind whole-cell recordings of neurons cultured in the same plate in either medium				
Number of coverslips used	10	12		■
Number of patched neurons	33	34		■
Cell lines	H9 ($n = 14$), iPSC ($n = 19$)	H9 ($n = 20$), iPSC ($n = 14$)		■
Time spent in neural differentiating medium	Mean = 29 d (range 18-45)	Mean = 28 d (range 17-43)		■ 0.5037
Time spent in neural differentiating medium (average) (rAP)	Mean = 31 d (range 18-45)	Mean = 29 d (range 17-43)		■ 0.5588
Action Potential types				
Cells with single evoked action potential (sAP)	33% ($n = 11$)	29% ($n = 10$)	88%	■
Cells with repetitive evoked action potentials (rAP)	67% ($n = 22$)	71% ($n = 24$)	106%	■
Neuronal morphology of the patched neurons (rAP)				
	$n = 22$	$n = 24$		
Number of primary neurites	3.8 ± 0.3	3.9 ± 0.2	103%	■ 0.7155
Soma diameter (μm) *	13 ± 0.5	15.5 ± 0.6	120%	↑ 0.01
Membrane "passive" properties (rAP)				
	$n = 22$	$n = 24$		
Capacitance (pF) **	23 ± 2.1	33 ± 2.5	143%	↑ 0.0032
Membrane resistance (MOhm) ****	4787 ± 734	2396 ± 241	50%	↓ <0.0001
Estimated V_m at 0 pA (mV)	-46 ± 2	-48 ± 3	104%	■ 0.5889
Sodium and potassium voltage-dependent currents (rAP)				
	$n = 22$	$n = 24$		
Threshold to activate Nav max peak (mV)	-35 ± 2.3	-36 ± 2.4	103%	■ 0.729
Nav currents (max peak in pA) *	-1595 ± 362	-2133 ± 270	134%	↑ 0.017
Kv rapidly inactivating currents (pA, max peak) *	87 ± 18	130 ± 16	149%	↑ 0.025
Kv slowly or non-inactivating (pA, at -10 mV)	223 ± 20	297 ± 32	133%	↗ 0.0874
Evoked action potential properties (rAP)				
	$n = 22$	$n = 24$		
AP max amplitude (mV)	11 ± 2.8	17 ± 2.6	155%	↗ 0.0718
Firing frequency above -10 mV (Hz) *	7.4 ± 0.8	11.8 ± 0.6	159%	↑ 0.022
Spike duration (ms)	9.3 ± 0.9	8.8 ± 1.2	95%	■ 0.3021
Max rise slope (mV/ms)	51 ± 9	68 ± 9	133%	↗ 0.0965
Max decay slope (mV/ms)	-19 ± 3	-24 ± 2	126%	↗ 0.0544
Afterhyperpolarization (mV)	-8.7 ± 0.9	-8.5 ± 1.2	98%	■ 0.9434

Supplementary Table 3 (related to Figure 5). **Electrophysiology analysis of neurons cultured side by side for 2-6 weeks in BrainPhys+sup or DMEM/F12+sup medium (patched in ACSF for fair comparison).** sAP, cells with single evoked action potential; rAP, cells with repetitive evoked action potentials. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. (continued)

	DMEM Culture Neurons	BrainPhys Culture Neurons	% Change BrainPhys vs. DMEM		P Value (Mann Whitney)
Spontaneous action potentials firing					
Silent (sAP)	100% ($n = 10/10$)	100% ($n = 9/9$)	100%	■	
Silent (rAP)	45% ($n = 10/22$)	33% ($n = 8/24$)	73%	↘	
Single spikes (rAP)	9% ($n = 2/22$)	17% ($n = 4/24$)	188%	↗	
Burst and/or regular firing (rAP)	45% ($n = 10/22$)	50% ($n = 12/24$)	111%	↗	
AMPA spontaneous events – functional excitatory synapses					
Frequency spontaneous EPSCs (average all sAP) (Hz)	0	0	100%	■	1
Frequency spontaneous EPSCs (average all rAP) (Hz) **	0.4 ± 0.3	0.7 ± 0.3	175%	↑	0.0017
Proportion of rAP cells with clear spontaneous EPSCs	23% ($n = 5/22$)	67% ($n = 16/24$)	291%	↑	
Frequency (average connected cells only) (Hz)	1.8 ± 1.4	1.0 ± 0.4	56%	■	0.8794
Amplitude (pA)	16 ± 1.7	15 ± 1.4	94%	■	0.4455
Rise time 10-90% (ms)	0.73 ± 0.16	0.78 ± 0.05	107%	■	0.2977
Decay time (ms)	1.7 ± 0.3	1.8 ± 0.08	106%	■	0.3528
GABA spontaneous events – functional inhibitory synapses					
Frequency spontaneous IPSCs (average all sAP) (Hz)	0	0	100%	■	1
Frequency spontaneous IPSCs (average all rAP) (Hz)	0.03 ± 0.02	0.26 ± 0.15	867%	↗	0.1366
Proportion of rAP cells with clear spontaneous IPSCs	24% ($n = 5/21$)	42% ($n = 8/19$)	175%	↑	
Frequency (average connected cells only) (Hz)	0.11 ± 0.05	0.62 ± 0.33	564%	↗	0.1663
Amplitude (pA)	23 ± 5	25 ± 2	109%	■	0.6775
Rise time 10-90% (ms)	2.2 ± 0.3	2.5 ± 0.2	114%	■	0.5828
Decay time (ms)	10 ± 0.3	9.2 ± 0.7	92%	■	0.3528

Supplementary Table 4. BrainPhys basal + supplements medium supports the culture of a large variety of neurons and astrocytes. This table summarizes the different cell types we have tested to date in culture with BrainPhys basal+sup. Evoked APs with peak amplitudes overshooting -10 mV and frequencies above 10 Hz for 500 ms were defined as “mature APs”. Cells were described with active synapses if more than three clean spontaneous events (typical kinetics reversibly blocked by NBQX) were recorded over 5 min in voltage clamp at -70 mV.

	Tested with BrainPhys basal + sup	Experimental evidences of cell identity	Properties after >2 weeks in BrainPhys basal + sup
Cell types			
Neurons	✓	Map2, Tuj1, NeuN; Electrophysiology	Mature APs + active synapses
Astrocytes	✓	GFAP immunostaining, GFAP viral vector	ePhys typical of astrocytes
Neuronal classes			
Glutamatergic	✓	Active AMPA synapses	Mature APs + active synapses
GABAergic	✓	Active GABA synapses; GABA immunostaining	Mature APs + active synapses
Dopaminergic	✓	TH immunostaining	Mature APs + active synapses
Serotonergic	✓	TPH2 immunostaining, TPH2 viral vector	Mature APs + active synapses
Motor neurons	✓	Hb9 viral vector	Mature APs + active synapses
Dentate gyrus granule neurons	✓	Prox1 viral vector	Mature APs + active synapses
Cell lines tested from different origins			
Human iPSCs (fibroblast-derived)	✓		
Human iNs (fibroblast-derived)	✓		
Human ESCs (H9)	✓		
Rat hippocampal primary cells	✓		
Mouse hippocampal slices	✓		

Supplementary Table 5. Composition of BrainPhys basal designed and tested in the current study.

Compound	Property	BrainPhys basal used in study (mM)	BrainPhys basal suggested range (mM)	Neurobasal-A (mM)	Neurobasal (mM)	DMEM/F12 Ref# 10565 (mM)
Inorganic Salts						
Sodium Chloride (NaCl)	Neuroactive	121	>70 and <150	68.97	51.72	120.68
Potassium Chloride (KCl)	Neuroactive	4.2	<5	5.33	5.33	4.16
Calcium Chloride (CaCl ₂) (anhyd.)	Neuroactive	1.1	>0.8 and <1.5	1.8	1.8	1.05
Magnesium Sulfate (MgSO ₄) (anhyd.)	Neuroactive	1.0	<2	0.0	0.0	0.407
Magnesium Chloride (MgCl ₂) (anhyd.)	Neuroactive	0.0	<2	0.814	0.814	0.301
Ferric Nitrate (Fe(NO ₃) ₃ · 9H ₂ O)	Neuroactive	0.000124	<0.0004	0.000248	0.000248	0.000124
Zinc Sulfate (ZnSO ₄ · 7H ₂ O)	Neuroactive	0.0015	<0.002	0.000674	0.000674	0.0015
Cupric Sulfate (CuSO ₄ · 5H ₂ O)	Neuroactive	0.0	<0.00001	0.0	0.0	0.0000052
Ferric Sulfate (FeSO ₄ · 7H ₂ O)	Neuroactive	0.0	<0.0015	0.0	0.0	0.0015
Sodium Bicarbonate (NaHCO ₃)	pH	29	<35	26.19	26.19	29.02
Sodium Phosphate Dibasic (Na ₂ HPO ₄) (anhyd.)	pH	0.5	<1	0.0	0.0	0.5
Sodium Phosphate Monobasic (NaH ₂ PO ₄ · H ₂ O)	pH	0.45	<1	0.906	0.906	0.453
Amino Acids						
Glycine	Neuroactive	0.002	<0.05	0.4	0.4	0.25
L-Alanine	Neuroactive	0.002	<0.05	0.0225	0.0225	0.05
L-Aspartic Acid	Neuroactive	0.0	<0.003	0.0	0.0	0.05
L-Glutamic Acid	Neuroactive	0.0	<0.02	0.0	0.0	0.05
L-Serine	Neuroactive	0.002	<0.03	0.4	0.4	0.25
L-Alanyl-L-Glutamine		0.5	<0.5	0.0	0.0	2.5
L-Arginine Hydrochloride		0.3	<0.5	0.398	0.398	0.699
L-Asparagine-H ₂ O		0.05	<0.05	0.00553	0.00553	0.05
L-Cysteine Hydrochloride-H ₂ O		0.1	<0.15	0.26	0.26	0.0998
L-Cystine 2HCl		0.0	0.0	0.0	0.0	0.1
L-Histidine Hydrochloride-H ₂ O		0.15	<0.2	0.2	0.2	0.15
L-Isoleucine		0.416	<0.8	0.802	0.802	0.416
L-Leucine		0.451	<0.8	0.802	0.802	0.451
L-Lysine Hydrochloride		0.499	<0.8	0.798	0.798	0.499
L-Methionine		0.116	<0.2	0.201	0.201	0.116
L-Phenylalanine		0.215	<0.4	0.4	0.4	0.215
L-Proline		0.06	<0.1	0.0675	0.0675	0.15
L-Threonine		0.449	<0.7	0.798	0.798	0.449
L-Tryptophan		0.0441	<0.07	0.0784	0.0784	0.0442
L-Tyrosine Disodium Salt Dihydrate		0.214	<0.4	0.398	0.398	0.214
L-Valine		0.452	<0.7	0.803	0.803	0.452

Supplementary Table 5. Composition of BrainPhys basal designed and tested in the current study. (continued)

Compound	Property	BrainPhys basal used in study (mM)	BrainPhys basal suggested range (mM)	Neurobasal-A (mM)	Neurobasal (mM)	DMEM/F12 Ref# 10565 (mM)
Energetic Substrates						
D-Glucose (Dextrose)	Energy	2.5	<5	25.0	25.0	17.51
Sodium Pyruvate	Energy	0.5	<1	0.227	0.227	0.5
Vitamins						
Choline Chloride		0.0641	<0.07	0.0286	0.0286	0.0641
D-Calcium Pantothenate (B5)		0.0047	<0.006	0.00839	0.00839	0.0047
Folic Acid (B9)		0.00601	<0.006	0.00907	0.00907	0.00601
i-Inositol		0.07	<0.07	0.04	0.04	0.07
Niacinamide (B3)		0.0166	<0.02	0.0328	0.0328	0.0166
Pyridoxine Hydrochloride		0.00986	<0.010	0.0196	0.0196	0.00986
Thiamine Hydrochloride		0.00644	<0.007	0.0119	0.0119	0.00644
Vitamin B12 (Cyanocobalamin)		0.000502	<0.0006	0.000005	0.000005	0.000502
Riboflavin (B2)		0.000582	<0.0006	0.00106	0.00106	0.000582
Other						
HEPES (pH buffer)	pH	5.0	<10	10.92	10.92	0.0
Phenol Red (pH indicator)	pH	0.0215	<0.07	0.0736	0.0215	0.0215
Excluded from BrainPhys Basal (* = removed because in B27 supplement)						
Hypoxanthine Na (other component)		0	<0.015	0.0	0.0	0.015
Lipoic Acid (organosulfur compound)		0	<0.0006	0.0	0.0	0.00051
Thymidine (DNA nucleoside)		0	<0.002	0.0	0.0	0.00151
* Linoleic Acid (essential fatty acid)		0	<0.00015	0.0	0.0	0.00015
* Putrescine 2HCl (organic chemical compound)		0	<0.15	0.0	0.0	0.000503
* Biotin (B7) (vitamin)		0	<0.0001	0.0	0.0	0.0000143
Properties						
Osmolarity	QC	305 mOsmol/L	280-330 mOsmol/L	257 mOsmol/L	220 mOsmol/L	315 mOsmol/L
pH	QC	7.4	7.3-7.5			

Supplementary Table 6. Composition of BrainPhys supplements tested in the current study.

Compound	Compound class	Notes, Properties	Concentration used in study	Suggested range for use with BrainPhys
Brain Derived Neurotrophic Factor (BDNF)	Protein (Neurotrophic Factors)		20 µg/L	<40 µg/L
Glial Cell Derived Neurotrophic Factor (GDNF)	Protein (Neurotrophic Factors)		20 µg/L	<40 µg/L
Laminin	Protein		~1 mg/L	<20 mg/L
Bovine Serum Albumin (BSA), Fatty Acid-Free Fraction V	Protein	part of B27	B27	
Catalase	Protein	part of B27	B27	
Human Recombinant Insulin	Protein	part of B27	B27	
Human Transferrin	Protein	part of B27	B27	
Superoxide Dismutase	Protein	part of B27	B27	
Human Transferrin (Holo)	Protein	part of N2	0.0010000 mM	
Insulin Recombinant Full Chain	Protein	part of N2	0.0008610 mM	
Corticosterone	Steroid Hormone	part of B27	B27	
Progesterone	Steroid Hormone	part of N2/B27	0.0000200 mM + B27	
T3 (triiodo-L-thyronine)	Thyroid Hormone	part of B27	B27	
Linoleic Acid	Essential Fatty Acid	part of B27	B27	
Linolenic Acid	Essential Fatty Acid	part of B27	B27	
Cholesterol	Lipid		0 or 10 µg/mL	<6 µg/mL
Ascorbid Acid (Vitamin C)	Vitamin		0.2 µM	<0.3
Biotin (B7)	Vitamin	part of B27	B27	
DL Alpha-Tocopherol Acetate	Vitamin	part of B27	B27	
DL Alpha-Tocopherol	Vitamin	part of B27	B27	
Vitamin A (Retinoic acid)	Vitamin	part of B27	B27	
Selenite	Sulfate Mineral	part of N2/B27	0.00003 mM + B27	
Putrescine 2HCl	Organic Chem. Comp.	part of N2/B27	0.1001 mM + B27	
D-Galactose	Monosaccharide	part of B27	B27	
Dibutyryl cAMP Sodium Salt	Nucleotide		0.0010000 mM	<0.002 mM
Serum	Supplement	Neuroactive	0 mM	0 mM CRITICAL TO EXCLUDE



CHAPTER 4A

Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology

Cedric Bardy¹, Mark van den Hurk^{1,2,*}, Boyko Kakaradov^{3,*}, Jennifer A. Erwin¹, Baptiste N. Jaeger¹, Ruben V. Hernandez¹, Tameji Eames¹, Andres A. Paucar¹, Mark Gorris¹, Cynthia Marchand¹, Roberto Jappelli¹, Jerika Barron¹, Alex K. Bryant¹, Mariko Kellogg¹, Roger S. Lasken⁴, Bart P.F. Rutten², Harry W.M. Steinbusch², Gene W. Yeo^{3,5}, Fred H. Gage¹

¹ Salk Institute for Biological Studies, Sanford Consortium for Regenerative Medicine, La Jolla, CA, USA

² Division of Translational Neuroscience, Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, The Netherlands

³ Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA, USA

⁴ J. Craig Venter Institute, La Jolla, CA, USA

⁵ Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore and Molecular Engineering Laboratory, A*STAR, Singapore, Singapore.

* These authors contributed equally to this work.

Molecular Psychiatry 2016; 21:1573-88

doi: 10.1038/mp.2016.158

ABSTRACT

Human neural progenitors derived from pluripotent stem cells develop into electrophysiologically active neurons at heterogeneous rates, which can confound disease-relevant discoveries in neurology and psychiatry. By combining patch clamping, morphological and transcriptome analysis on single-human neurons *in vitro*, we defined a continuum of poor to highly functional electrophysiological states of differentiated neurons. The strong correlations between action potentials, synaptic activity, dendritic complexity and gene expression highlight the importance of methods for isolating functionally comparable neurons for *in vitro* investigations of brain disorders. Although whole-cell electrophysiology is the gold standard for functional evaluation, it often lacks the scalability required for disease modeling studies. Here, we demonstrate a multimodal machine-learning strategy to identify new molecular features that predict the physiological states of single neurons, independently of the time spent *in vitro*. As further proof of concept, we selected one of the potential neurophysiological biomarkers identified in this study—GDAP1L1—to isolate highly functional live human neurons *in vitro*.

INTRODUCTION

The unprecedented access to live neurons from patients via human induced pluripotent stem cell (iPSC) models is revolutionizing medical research opportunities in neurology and psychiatry. The translational success of these models depends on the ability to recapitulate *in vitro* the complexity of functional human brain circuits. Rapidly growing technological advances, which build on decades of elegant investigations on neurodevelopmental processes in animal models [1, 2], offer a panoply of protocols to drive cellular fates towards neurons of particular neurotransmitter classes (e.g., dopaminergic, serotonergic) and brain region identities (e.g., cortical neurons, midbrain neurons, motor neurons) [3-7]. Many important studies have also shown that electrophysiologically active neurons can be generated from iPSCs or fibroblast-direct conversion [5, 8-13]. Despite the clear success of generating highly functional neurons, it is also evident that such human cultures often comprise neuronal populations of heterogeneous electrophysiological states [14]. Indeed, patch-clamping experiments have reported an important variability of functional maturity among cell lines, cell batches and even within the same culture dish [15-17]. Co-culture with astrocytes or lengthy periods of time spent in neuronal medium have been reported to increase neuronal maturity on average but may also increase tissue culture variability [18]. In addition, the length of time required to reach functional maturation significantly varies among numerous published reports from 3 weeks to more than 5 months [18, 19]. Such wide ranges may depend on many technical aspects such as loose criteria defining maturity, discrepancies in tissue culture protocols, or inherent differences among batches of cells [20]. Patch clamping is the current gold standard to demonstrate the functionality of a neuronal culture. However, patch clamping is low throughput and provides information for only a handful of neurons selected from several hundreds of thousands of cells. This technical limitation precludes a thorough characterization of the functional maturity of the actual neurons used with a variety of read outs for identifying the particular traits of patients' cell lines (e.g., biochemistry, morphology, cell survival). In this study, we demonstrate a strategy to define functional states of human neurons *in vitro*, independently of time spent in culture. We bridge the gap between electrophysiology and molecular profile by successfully combining patch clamping, morphology and RNA-seq on single human neurons (Patch-seq). Using a machine-learning classifier trained on our multimodal data set, we reveal new biomarkers that efficiently predict which neurons are highly functional. These biomarkers allow for the functional classification of a large number of neurons without patch clamping and can be used to stratify functional heterogeneity.

METHODS

Human Neuronal Culture

Human dermal fibroblasts were reprogrammed into pluripotent cells with the four Yamanaka factors (Oct3/4, Sox2, Klf4 and c-Myc), either in a retroviral vector or a non-integrating Sendai viral vector. Human iPSCs and embryonic stem cells (ESCs) were differentiated into neural progenitor cells (NPCs) as previously described [6]. NPCs were expanded for 3 to 5 passages (split 1:2 or 1:3 per passage) and stored at -80°C . Then they were thawed and expanded for at least one more passage in neural progenitor medium (DMEM/F12 supplemented with FGF8, SHH, B27 and N2, without retinoic acid). Dissociated NPCs were directly plated on glass coverslips (Fisher Scientific, Cat. No. 12-545-80) coated with poly-ornithine (Sigma, Cat. No. P3655) and laminin (Invitrogen, Cat. No. 23017-015) in 24-well plates. Twenty-four hours later, the cells were switched gradually (half medium change) to neuronal medium: BrainPhys basal [13] supplemented with 1x N2 (Gibco, Cat No. 17502-048), 1x B27 (Gibco, Cat No. 17504-044), brain-derived neurotrophic factor (BDNF, 20 ng/ml; Peprotech, Cat No. 450-02), glia-derived neurotrophic factor (GDNF, 20 ng/ml; Peprotech, Cat No. 450-10), ascorbic acid (AA, 200 nM; Sigma, Cat No. A0278), dibutyryl cyclic AMP (cAMP, 1mM Sigma, Cat No. D0627), and laminin (1 $\mu\text{g}/\text{ml}$; Invitrogen, Cat No 23017-015). Half of the neuronal medium was gently replaced two to three times a week. The plates were kept in a humidified incubator at 37°C with 5% CO_2 and 21% O_2 . The pH (~ 7.3 - 7.4) and the osmolarity (~ 300 - 305 mOsmol/L) of medium were maintained constant over time.

Patch Clamping

For whole-cell patch-clamp recordings, individual coverslips were transferred into a heated recording chamber and continuously perfused ($1\text{ ml}\cdot\text{min}^{-1}$) with either BrainPhys basal medium or artificial cerebrospinal fluid (ACSF) bubbled with a mixture of CO_2 (5%) and O_2 (95%) and maintained at 25°C . The composition of ACSF was adjusted to match the inorganic salt concentration and osmolarity of BrainPhys basal. ACSF contained (in mM) 121 NaCl, 4.2 KCl, 1.1 CaCl_2 , 1 MgSO_4 (or 0.4 MgSO_4 and 0.3 MgCl), 29 NaHCO_3 , 0.45 $\text{NaH}_2\text{PO}_4\cdot\text{H}_2\text{O}$, 0.5 Na_2HPO_4 and 20 glucose (all chemicals from Sigma).

For single-cell patch-clamp electrophysiological recordings, we used a digidata 1440A/Multiclamp 700B and Clampex 10.3 (Molecular Devices). Patch electrodes were filled with internal solutions containing 130 mM K-gluconate, 6 mM KCl, 4 mM NaCl, 10mM Na-HEPES, 0.2 mM K-EGTA; 0.3mM GTP, 2 mM Mg-ATP, 0.2 mM cAMP, 10 mM D-glucose, 0.15% biocytin and 0.06% rhodamine. The pH and osmolarity of the internal solution were close to physiological conditions (pH 7.3, 290–300 mOsmol). Data were all corrected for liquid junction potentials (10 mV). Electrode capacitances were compensated on-line in cell-attached mode (~ 7 pF). Recordings were low-pass filtered at 2

kHz, digitized, and sampled at intervals of 50 ms (20 kHz). To control the quality and the stability of the recordings throughout the experiments, access resistance, capacitance and membrane resistance were continuously monitored on-line and recorded. The resistance of the patch pipettes was between 3 and 5 MOhm. The access resistance of the cells in our sample was ~40 MOhm on average. Spontaneous synaptic AMPA events were recorded at the reversal potential of Cl⁻ and could be reversibly blocked by AMPA receptor antagonist (10 μM NBQX, Sigma Ref#N183). Spontaneous synaptic GABA events were recorded at the reversal potential of cations and could be reversibly blocked with GABA_A receptor antagonist (10 μM SR95531, Sigma Ref#S106). Statistical analysis of electrophysiology data was assisted with Clampfit 10.3, Matlab 2014b, Igor Pro 6, Prism 6, MiniAnalysis, Excel and custom-made Python programs.

Electrophysiological Recording Protocol

The same electrophysiological protocol was applied to all neurons ($n = 290$). Whole cell patch-clamp recordings were obtained in the absence of any receptor antagonists in ACSF or BrainPhys basal. After breaking the membrane seal, we first maintained the cell in voltage clamp (VC) at -70 mV by injecting small amounts of current when necessary (“Baseline at -70 mV”). We applied a test pulse of -5 mV to measure passive cell properties (membrane resistance, access resistance, capacitance). Voltage-dependent sodium and potassium channels properties were examined with an IV curve. Spontaneous synaptic events mediated by AMPA receptors were recorded for at least 5 min at -70 mV (close to Cl⁻ reversal potential). The nature of AMPA synaptic events was confirmed on a small subset of cells ($n = 15$, excluded from RNA-seq) by reversible blockade with AMPA receptor antagonist (10 μM NBQX). The patch was then switched to current clamp. Small currents were injected if necessary to maintain the cell resting potential around -70 mV. Current steps (increments 2 pA, $n = 15$, 500 ms) were applied to measure the properties of evoked action potentials (APs). Next, spontaneous APs were recorded for 1 to 5 min at -70 mV and then at resting potential without injecting any current (CC 0 pA). The patch was then reversed to voltage clamp and slowly increased to the reversal potential of cations (~0 mV). Once the baseline stabilized, we recorded spontaneous GABAergic synaptic activity for 5 to 10 min. The nature of GABA synaptic events was confirmed on a small subset of cells ($n = 5$, excluded from RNA-seq) by reversible blockade with GABA_A receptor antagonist (10 μM Gabazine).

AP Types Classification

“Type 0 cells” did not express voltage-dependent sodium currents and were excluded from analysis. “Type 1 neurons” expressed small Nav currents but were not able to fire APs above -10 mV. The arbitrary limit of -10 mV was chosen as it is close to the reversal potential of cations (0 mV). Healthy APs usually reach or overshoot the reversal potential

of cations. “Type 2 neurons” fired one AP above -10 mV, which was followed by a plateau. “Type 3 neurons” also fired one AP above -10 mV and one or a few aborted spikes below -10 mV. “Type 4 neurons” fired more than one AP above -10 mV but at a frequency below 10 Hz. “Type 5 neurons” fired APs above -10 mV at 10 Hz or more. Our categorization of functional types of neurons followed a continuum that relates to the stage of maturity of the neurons. While Type 1 neurons are considered immature, Type 5 neurons are considered more mature and functionally active. Although after ~3-6 weeks of maturation we found on average ~75% of Type 4-5 neurons, we found every AP type of cell at most differentiation time points we looked at (range of 2 weeks to 5 months after NPC maturation). This finding suggests some degree of variability in the electrophysiological maturity of neurons even within cultures of the same age. Remarkably, in our samples, the large majority of cells receiving active excitatory synapses were Type 5 neurons and we almost never found clear, spontaneously active AMPA and GABA synaptic inputs in Type 1 neurons and rarely in Type 2.

Neuromorphometry

Images of the patch-clamped cells were taken with a 40× water-immersion objective, differential interference contrast filters (all Olympus), an infrared digital camera (Rolera XR -Qimaging), and an epifluorescence mercury lamp. Mosaics of the several fields of views to cover the span of the neurites were taken and stitched later on Neurolucida (MBF Biosciences). Before patch clamping, the targeted cells were imaged with DIC and epifluorescence (most cells patched expressed Synapsin-GFP). The entire cells were filled with a red fluorescent dye (rhodamine) and another set of images was taken rapidly after completion of the electrophysiological recordings. All the images were stitched and analyzed on Neurolucida. The soma size and shape were measured on the DIC images before patching to avoid deformation by the patch pipette. The images from the intracellular rhodamine dye, occasionally supplemented by GFP vectors, were used to guide the neuromorphological reconstruction. For all cells, the morphology was reconstructed within a radius of 150 μm . Many cell processes reached beyond that distance but, for homogeneity of the sample, we restricted the analysis to that arbitrary radius. Occasionally, patched cells with pictures of insufficient quality were excluded from the analysis. When necessary, some processes partially hidden by the patch electrode were extrapolated.

Single-Cell Collection and cDNA Preparation

Following electrophysiological recording, we applied slight additional negative pressure to establish a stronger connection between cell and patch electrode, but without completely aspirating the cell in the pipette. The neuron, including its processes (axon and dendrites), was then transferred in a volume of ~2 μl of internal patch solution into a PCR tube containing 8.0 μl sample buffer by slowly retracting the patch pipette

from the chamber bath and breaking the tip of the electrode along the inside wall of the tube. Sample buffer was made according to the protocol supplied by Clontech with the SMARTer™ Ultra Low RNA Kit and included both ERCC spike-ins (Life Technologies) and ArrayControl spikes 1, 4 and 7 (RNA standard for the Fluidigm C1) for the purpose of quality control (QC). Successful removal of the cell from the coverslip was always confirmed by DIC optics. Collected single cells were immediately spun down and put on ice for SMARTer™ cDNA synthesis following manufacturer's instructions (Clontech). Briefly, first-strand cDNA was synthesized from poly(A)⁺ RNA by incubation with 1 µl of 3' SMART CDS Primer II A (24 µM) for 3 min at 72°C, followed by reverse transcription in a 20-µl final reaction volume using 200 units of SMARTScribe™ Reverse Transcriptase for 90 min at 42 °C and inactivation for 10 min at 70°C. First-strand cDNA was then purified using Agencourt AMPure XP SPRI Beads (Beckman Coulter Genomics) and amplified by long-distance PCR using the Advantage® 2 PCR Kit (Clontech) with the following PCR thermocycler program: 95°C for 1 min, 18 cycles of 95°C for 15 s, 65°C for 30 s, 68°C for 6 min, and 72°C for 10 min. PCR-amplified double-stranded (ds) cDNA was immobilized onto SPRI beads, purified by two washes in 80% ethanol, and eluted in 12 µl of purification buffer (Clontech). The quality (Agilent 2100 Bioanalyzer High Sensitivity DNA Kit; Agilent) and quantity (Qubit dsDNA High Sensitivity Assay Kit; Invitrogen) of each ds cDNA sample were assessed prior to library preparation.

qPCR Assays

To select the single-cell cDNA samples suitable for mRNA-seq, we determined the level of expression of *GAPDH* and *ACTB* in each sample by TaqMan real-time PCR. Samples with Ct values ≤ 30 for both housekeeping genes were typically considered positive for library preparation. For each gene, duplicate 10-µl PCR reactions were performed on an ABI Prism 7900 Sequence Detector (Applied Biosystems) using 0.50 µl of 1:5-diluted ds cDNA template in standard TaqMan Gene Expression Assay with FAM reporter. Real-time PCR assays for detection of the ERCCs and ArrayControl RNA spikes were performed using, respectively, standard TaqMan Gene Expression Assays (Life Technologies) and SYBR Green PCR Master Mix (Applied Biosystems) with custom primers (Fluidigm).

Illumina Transcriptome Library Preparation and Sequencing

Construction of single-cell mRNA-seq libraries was typically performed with 0.25 ng of input cDNA using the Nextera XT DNA sample prep kit (Illumina) with modified protocol. Briefly, cDNA was tagmented for 5 min at 55°C in a 5-µl reaction containing 2.5 µl of Tagment DNA Buffer and 1.25 µl of Amplicon Tagment Mix; tagmentation was neutralized with 1.25 µl of Neutralize Tagment Buffer for 5 min. Tagmented DNA was then subjected to 12-cycle PCR amplification using 3.75 µl of Nextera PCR Master Mix and 1.25 µl each of index 1 (i7) and index 2 (i5) library-identifying (barcoded) sequenc-

ing primers. The constructed libraries were run on a 1.5% agarose gel in Tris-borate/EDTA buffer, stained with SYBR Gold (Invitrogen), and size selected for ~300-400 or ~300-650 bp (insert size of ~165-265 or ~165-515 bp, respectively). Gel-excised library fragments were purified with the Wizard SV Gel and PCR Clean-Up System (Promega), eluted in 40 μ l of nuclease-free water, and concentrated by speedvacuum centrifugation. Each library was then quantified (Qubit dsDNA High Sensitivity Assay Kit; Invitrogen) and examined for correct size (Agilent 2200 TapeStation High Sensitivity D1K Screen-Tape Assay; Agilent), after which equimolar amounts of uniquely barcoded libraries were pooled together and used for cluster generation and 100-bp paired-end sequencing on a HiSeq 2000 or 2500 sequencer (Illumina).

Bioinformatic Analysis of Single-Cell Transcriptomes

Single cell mRNA sequencing data from $n = 56$ patched human neurons, which passed a series of QC, allowed us to correlate electrophysiological profiles with gene expression profiles. For each of $n = 56$ neurons, raw sequencing reads were mapped to the human reference transcriptome (Gencode v19) using gapped-alignment strategies. Alignment was performed by STAR (version 2.3.0) followed by gene-level quantification with HTseq (version 0.6.1). Per-gene expression outputs were scaled to transcripts per million (tpm) units.

Data Transformation and Dimensionality Reduction for Transcriptome Principal Component Analysis

Whole-gene expression tpm counts were log-transformed: $\log(\text{tpm}+1)$ to normalize their distribution. PCA was performed on the log-transformed expression matrix E (cells = rows, genes = columns). Prior to PCA dimensionality reduction, the expression of each gene (column) was standardized by subtracting the mean expression of that gene across all 56 cells and dividing by its standard deviation. All 56 cells were scatter-plotted against the first two principal components of the expression matrix E . While the PCA of the transcriptomes was unsupervised, each cell was later colored on the plot by its respective AP Type, allowing us to visually assess any functionally significant clustering. To formalize this, we also performed hierarchical agglomerative clustering (Euclidean distance, average linkage) of the cell-cell covariance matrix (E^*E').

Differential Expression Between Intermediate and Functional Neuron States

The five AP Types of differentiated neurons were stratified into immature (Types 1-3), transitional (Type 4) and highly functional (Type 5) neurons. To identify differentially expressed genes in these neuron groups, two non-parametric statistical tests were used: Mann-Whitney U (MWU) and Kolmogorov-Smirnoff 2-sided (KS2). The reported P -values are presented, but the threshold for significance is Bonferroni-corrected from $P < 0.05$ to $P < 3.1e-6$.

Extremely Randomized Trees Classifier

The algorithm learned to classify the functional states of neurons solely based on their molecular phenotypes (gene expression). We stratified the cells into two classes based on electrophysiological states. The classifier randomly selected 90% of the cells in our sample for training. The testing was performed on the remaining 10%. To obtain a predictive value for each cell, we reiterated the learning phase 10 times. Using the extremely randomized trees classifier [21] with 10-fold cross-validation, we achieved good validation performance with >85% generalization accuracy, as shown by a high area under the Receiver Operating Characteristic (ROC) curve (area under the curve = 0.93). The maximum number of features was 200, and we used 900 estimators.

Construction of GDAP1L1-EGFP Lentiviral Reporter Vector

A transcriptional reporter lentivector encoding EGFP was designed to track human cells expressing gene GDAP1L1 (ganglioside-induced differentiation associated protein 1-like 1). A 1350 bp 5' gene region—chr20:44,245,985-44,247,334—UCSC Human Genome Browser Dec. 2013 (GRCh38/hg38) Assembly, located precisely upstream to the gene's translation start codon and including a 77 bp 5' UTR, was used as a promoter. The sequence was amplified from male genomic DNA (Promega) with Phusion High-Fidelity DNA Polymerase (NEB) using the forward primer G220 (5'–CGTTATCGATGTATGCT-GAACCAGGGAGGCT), which adds a *Cla*I site, and the reverse primer G221 (5'–attactc-gaggacagcccgaatcagagga), which adds a *Xho*I site. The PCR product was introduced as a promoter into the LV pCSC-Syn-mcs-EGFP, a derivative of pCSC-SP-PW-EGFP [22]. The construct (pCSC-GDAP1L1-EGFP) was verified by DNA sequencing.

Immunohistochemistry

Immunohistochemistry experiments were performed on neurons plated on glass coverslips. Standard immunohistochemistry protocols were used. Coverslips were stained with DAPI and a combination of the following antibodies: mouse-Map2(2a+2b) (1:500, Sigma), mouse-TUJ1 (1:1000, Covance), Rabbit-GFAP (1:200, Dako) and GDAP1L1 (1:250, Origene).

FACS Analysis

Adherent neuronal cultures were thoroughly washed with DPBS without Mg⁺ and Ca⁺ before dissociating with Accutase containing DNase (100 U mL⁻¹) for 5 min at room temperature. Cell suspensions were collected with DPBS containing 1% Knockout Serum Replacement (KOSR) and span down at 200 rfc for 5 minutes. The cell pellet was gently resuspended in DPBS containing 1% KOSR, 2mM EDTA, and DNase (100 U mL⁻¹). Cell suspension was passed through a 70- μ m nylon cell strainer and sorted using a Becton-Dickinson Influx™ cytometer. Gated samples were collected in 15-mL tubes containing

BrainPhys-based neuronal differentiation medium containing Rho kinase (ROCK) inhibitor (StemCell Tech.) and penicillin-streptomycin (Life Tech.). After sorting, we re-plated 50,000 neurons (GDAP1L1:GFP-positive and GFAP:dsRED-negative) per well (48-well plate) onto poly-ornithine/laminin-coated 8-mm glass coverslips containing established human astrocytes-cerebellar (ScienCell) at 70% confluence. The medium was changed 24 h post-sort to BrainPhys differentiation medium without ROCK inhibitor, and the re-plated cells were cultured as previously described [13].

Statistics

Unless specified clearly otherwise, throughout the manuscript we present the mean and standard error of the mean. Concerned to avoid the assumption of Gaussian distribution in our samples, we consistently used non-parametric statistics such as Spearman, Mann-Whitney and Wilcoxon. Statistics were two-tailed and Bonferroni corrected for multiple hypotheses tests. In several plots, to estimate the goodness of fit of the linear regression represented graphically, we reported the R^2 and the corresponding P -values, which estimates how significantly the slope deviates from zero.

RESULTS

Human neuronal circuits were established *in vitro* from iPSCs and ESCs (H9 line) (Figure 1A). IPSCs were reprogrammed from the fibroblast cells of four healthy individuals. IPSCs and ESCs were differentiated towards midbrain NPCs using an established protocol [6]. NPCs were expanded for 5-11 passages and re-plated in neuronal medium (BrainPhys basal + serum free supplements) that enabled differentiation and maturation under conditions supporting optimal electrophysiological activity [13]. Following this protocol, the human NPCs differentiated into cultures comprising about 50% astrocytes (GFAP+) and 50% neurons (TUJ1+/MAP2+). We did not use additional feeder layers of astrocytes for initial neuronal maturation. Based on immunostainings, we estimated that the neural circuits were formed with a majority of glutamatergic neurons (~62%) and smaller percentages of GABAergic neurons (~27%, GABA+) and dopaminergic neurons (~11%, TH+) [13]. Based on patch-clamping data, we also found that neurons in highly functional states received 2.3 times more spontaneously active AMPA-mediated synaptic inputs than GABAergic synaptic inputs. The neurons were analyzed after ~30 days on average in neuronal medium (the full range tested throughout the entire study was 0-30 weeks). Each cell was tested with a consistent electrophysiological patch-clamping protocol measuring voltage-gated Na^+/K^+ currents, evoked and spontaneous APs, membrane resistance, resting potential, capacitance and synaptic activity (AMPA/GABA)

(**Supplementary Figure 1**) and was filled with a dye (rhodamine) to allow morphological reconstruction. The same live cells were also collected for RNA-seq analysis.

Electrophysiological Traits of Human Neurons *In Vitro* Reveal a Continuum of Functional States

We used unbiased statistical analyses to identify the broad range of functional states of differentiated neurons *in vitro*. To define the functional properties of each neuron, we performed a series of patch-clamping tests (**Supplementary Figure 1**). Then, we applied a principal component analysis (PCA) on 25 electrophysiological measurements mostly reflecting the functional intrinsic properties of 246 patch-clamped neurons (**Figure 1B**; ePhys PCA features and corresponding loadings listed in **Figure 1C**). Such analysis aimed to describe objectively the inherent variation in the basic electrophysiological profiles of human iPSC-derived neurons that is routinely observed *in vitro* by many independent laboratories. The first principal component, ePhys PC1, explained 42% of the variation in the dataset (**Supplementary Figure 2A**), and revealed a continuum of functional neuronal states that we separated into three broad groups: “immature states,” “transition states” and “highly functional states.” These neuronal states appeared to reflect increasing stages of electrophysiological maturation. In support, we found significant correlations with several properties that are associated with neurophysiological development. The strongest correlation was obtained with AP firing properties. Indeed, the amplitude (“AP peak”) and the firing rate (“AP frequency above -10 mV”) of evoked APs stand out as two of the most influential loadings in PC1 (**Figure 1C**). We combined these two AP metrics to further categorize the neurons into five “AP Types” (**Figure 1D**, **Supplementary Figure 2B**), which significantly aligned with ePhys PC1 (**Figure 1E**, $P < 0.01$). Our analysis suggests that either metrics – “ePhys PC1” or “AP Types” – can be used objectively to determine the functional states of human neurons *in vitro*. Importantly, we demonstrated that these metrics largely reflect stages of neurophysiological maturation. For instance, we found significant correlations of AP Type classification with several other electrophysiological properties that typically associate with neuronal maturation, such as the cell membrane resistance, which is typically lower in more mature neurons (**Figure 1F**, **Supplementary Figure 2D**). Naturally, neurodevelopment also correlates with time spent in culture, and indeed we found a smaller fraction of AP Type 4 or 5 neurons in the first three weeks in neuronal medium than at later time points (**Figure 1G**). We also found a significant correlation of the PC1 variance with the time the cells spent in neuronal medium. However, the correlation between time in culture and the functional state of the neurons was rather weak (**Figure 1H**), which may be a consequence of the fact that single human neurons differentiate at variable rate *in vitro*. In addition, different tissue culture practice may influence the efficiency and speed of maturation. For example, we found that simply increasing the passage number of NPCs before neuronal differentia-

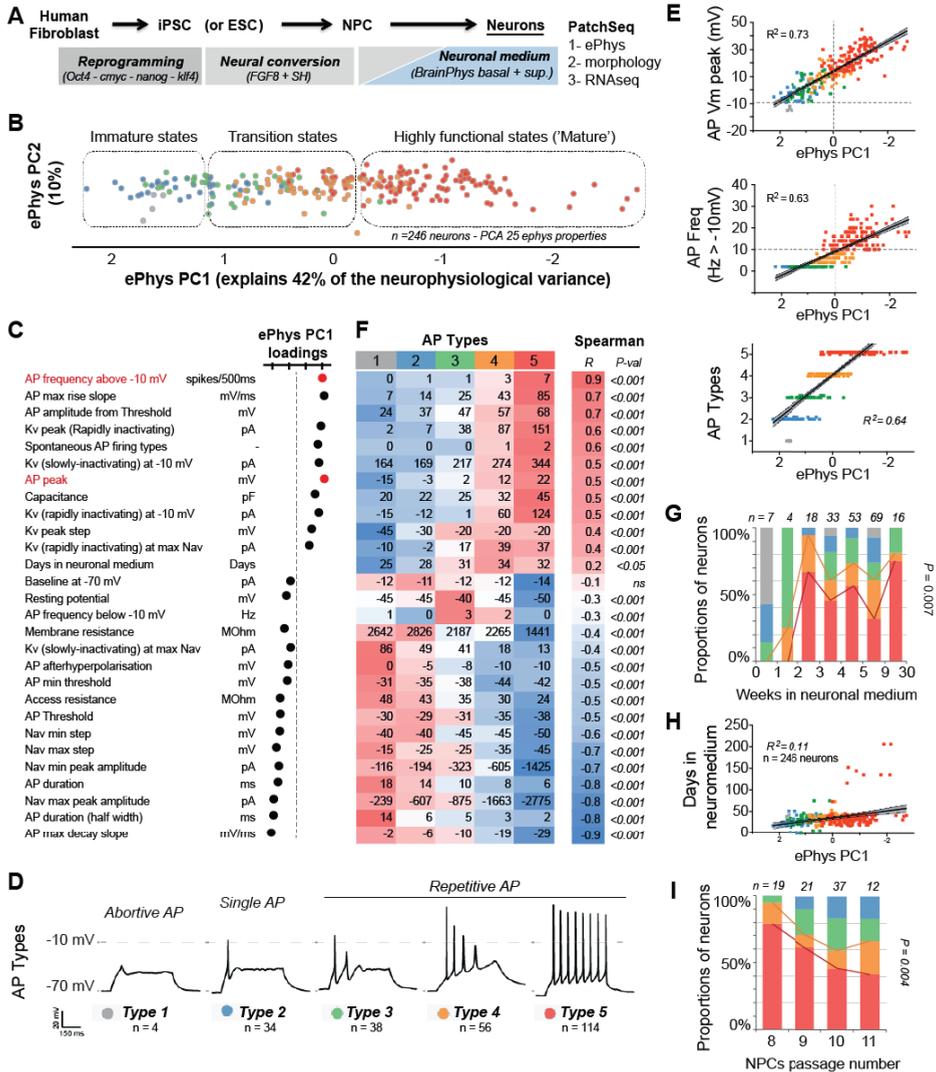


Figure 1. Human neurons in culture can be categorized into different electrophysiological states.

(A) Human ESCs (H9) or iPSCs from healthy subjects were converted to neuronal progenitor cells (NPCs). Neurons and astrocytes derived from the same progenitors were then cultured in neuronal medium (BrainPhys basal medium with supplements). See methods for more details. (B) Principal component analysis (PCA) of 25 electrophysiological properties (listed in C) measured with patch clamping of human neurons (n = 246). This PCA integrates basic cell-intrinsic electrophysiological properties such as passive membrane properties, voltage-dependent sodium and potassium currents and APs firing. Each dot represents a neuron. Astrocytes or any cells that did not express at least small voltage-dependent sodium currents were not included in the PCA. Cells that did not have unambiguous analysis of all the chosen 25 properties were also excluded from this PCA. (C) The dot graph represents the relative values of the loadings onto the first principal component (PC1) for each of the 25 properties used in the above PCA. The loadings highlighted in red correspond to the two measures that were used subsequently to define a continuum of five functional

states. **(D)** Representation of the typical heterogeneous neuronal responses to optimal depolarizing current steps for 500 ms (V_m rest clamped around -70mV). We classified those heterogeneous states of differentiated neurons into a continuum of five AP Types based on the combination of key electrophysiological properties identified with unbiased PCA: (1) the maximal peak of AP (V_m measured at the top of the best evoked AP), and (2) the frequencies of APs overshooting -10mV . The same neuronal color key throughout the figure corresponds to this AP Types classification (see also Supplementary Figure 2B). **(E)** The graphs show the PC1 value for each neuron ($n = 246$) against the key AP properties and the corresponding AP Type classification. EPhys PC1 values highly correlate with the maximum peak and frequencies of APs (evoked APs freq were counted only if overshooting -10mV in response to a square pulse of current of 500 ms from resting -70mV). Linear regression fits $\pm 95\%$ confidence intervals are shown. **(F)** Most measurements listed on each row significantly correlate with the AP Types classification (columns 1-5 of heat map). The values in the central heat map represent the median for the neurons in respective AP Type categories. To illustrate the direction of the measurement variations between AP Types, we color coded the heat map (values normalized by row with mean = 0 and variance = 1) from low to high values (blue, to white, to red). The last two columns show the Spearman's correlation coefficient (R) and its significance (P -val) between the measurements (rows) vs. the numerical class (1-5) of AP Types. The properties were sorted by decreasing correlation coefficient obtained. The P -values were corrected for multiple hypothesis testing (Bonferroni correction: $P < 0.05$ ($2\text{E-}03$), $P < 0.01$ ($3\text{E-}04$), $P < 0.001$ ($3\text{E-}05$)). **(G)** The differentiated neurons were patch clamped after different periods in neuronal medium. Days were counted from the switch of neuronal progenitor medium to neuronal medium (BrainPhys basal + supplements from 0 to up to 7 months). The proportion of more functional neurons (Type 4 and 5) significantly increased the first 3 weeks in neuronal medium, implying that an early maturation phase corresponds with the development of electrophysiological types; however, the proportions of Type 4 and 5 appeared to plateau after that period. The significance of this relationship in single neurons was measured with Spearman correlation between days in neuronal medium (x axis) and the numerical class (1-5) of AP Types. **(H)** We found a significant but poor correlation of the time spent in neuronal medium and functional states (measured by ePhys PC1 or AP Types). **(I)** After expansion (six passages) and storage at -80°C , NPCs derived from the same cell line (H9) were thawed and cultured for another two to five passages before re-plating in neuronal medium. The neurons ($n = 89$) were patch clamped after 4-6 weeks in neuronal medium and categorized into AP Types. The proportion of Type 5 neurons significantly decreased with high NPC passage numbers. The significance of this relationship in single neurons was measured with the Spearman correlation between passage numbers (x axis) and the numerical class (1-5) of AP Types.

tion significantly reduced the proportion of Type 5 neurons within subsequent cultures (**Figure 11**). Similarly, various cell lines may generate variable proportions of AP Type 5 neurons. We propose that, to identify accurate neuronal phenotypes between cell lines (e.g., patients vs. healthy subjects), it is important to compare neurons of equivalent functional states, rather than relying exclusively on the time the cells have spent maturing *in vitro*.

Synaptic Activity Correlates with AP Types

Synaptic activity is the essence of neuronal communication and its disruption is often the cause of behavioral defects in neurological disorders. Therefore, many recent studies have focused on measuring possible synaptic abnormalities in iPSC-derived neurons in the modeling of neural diseases *in vitro* [23-26]. As synaptic activity is associated with the functional maturity of neuronal circuits [27], we examined the relationship between

spontaneous synaptic activity and our AP Type categories (**Figure 2A**). Our culture protocol generated roughly two times more glutamatergic neurons than GABAergic neurons. This bias was strongly reflected in a higher likelihood to record spontaneous active excitatory inputs than GABAergic synaptic inputs. Despite those differences, we found that both AMPA and GABA activity significantly increased along the continuum of AP Types (**Figure 2B,C**). In our conditions, 95% of AP Type 5 neurons received active pre-synaptic inputs, whereas only ~10% of Type 2 neurons did. This result is particularly important to consider in iPSC disease models, which compare synaptic functions between patients. Overall, our analysis highlights the point that, if neural cell lines do not have similar proportions of AP Types, such variability will influence differences in the measured synaptic activity.

Neuromorphological Features of AP Types

To further investigate how our classification of AP Types reflects neurodevelopmental stages, we filled 256 patched neurons with an inert red dye (rhodamine) and reconstructed their morphology (**Figure 3A**). We analyzed 20 morphological features (listed in **Figure 3B**; **Supplementary Figure 3C**) and found that the size of the soma and the complexity of neurite arborization increased in higher AP Types (**Figure 3C,D**). The arborizations of Type 4 and 5 neurons were significantly more complex than those of Types 1, 2 or 3. Furthermore, the increase in soma size and distant arborization complexity were consistent with larger intracellular volume, which was estimated with the cell capacitance. Despite high cell-to-cell variabilities even within functional states, the most highly functional neurons (i.e., AP Type 5) tended to have overall significantly larger intracellular volume (**Figure 3E**). This multimodal analysis demonstrates that the electrophysiological states correlate with morphological neurodevelopmental phenotypes.

Single-Cell Transcriptome Analysis of Patch-Clamped Neurons

To examine the relationship between gene expression and electrophysiology, we generated single-cell transcriptome data from the same neurons that were analyzed for electrophysiology and morphology. At the end of the patch-clamping and imaging protocol, which lasted about 30 min, each cell was isolated with negative pressure applied to the same pipette that was used for the whole-cell recordings. Images acquired before and after single-cell collection confirmed that only the targeted single cells were taken out, leaving surrounding cells in place (**Figure 4A**). Importantly, we collected the entire neuron, including the soma, nucleus and neurites (**Figure 4A**, right panel). Poly-A selected mRNA molecules were reverse transcribed to cDNA and amplified immediately after patching and isolation to avoid RNA degradation. Before RNA-seq library preparation, we subjected the cDNA from each cell to a set of QC criteria to exclude potentially low-quality captures and amplifications (**Figure 4B**). We then generated RNA-seq libraries from

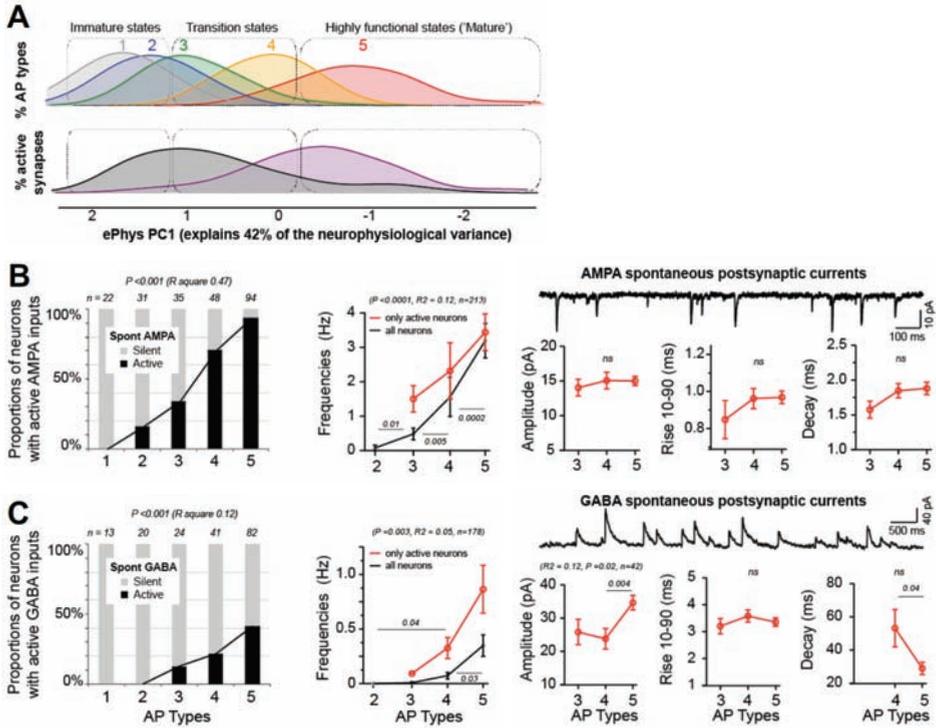


Figure 2. Synaptic input activity highly correlates with AP states. (A) Frequency distribution of AP Types and differentiated neurons with active synaptic inputs against the ePhys PC1 measure in Figure 1B. (B) The trace represents typical spontaneous (Spont) AMPA events (NBQX-sensitive). Patched neurons were classified as synaptically active (active synaptic inputs) in the left histogram if more than three clear glutamatergic spontaneous events were detected (with typical AMPA kinetics and amplitude above noise levels) within 5-min recordings in voltage clamp close to the reversal potential of Cl^- channels (-70 mV). (C) The trace represents typical spontaneous GABA events (Gabazine-sensitive). Patched neurons were classified as pre-synaptically active in the left histogram if more than three clear GABAergic spontaneous events were detected (with typical GABA kinetics and amplitude above noise levels) within 5-min recordings in voltage clamp close to the reversal potential of Na^+ channels (0 mV). (B,C) Mean \pm S.E.M. shown. For the spontaneous events amplitudes, the medians were 13% lower than mean but strongly correlated ($R^2 = 0.90$, $P < 0.0001$). For each graph a linear regression was fitted and the significant P -values were noted in brackets with R^2 and n . Non-significant regression fit P -values > 0.05 were noted as '(ns)'. Additional statistics between individual types were performed with Mann Whitney two-tailed test, and only the significant P -values from these tests were noted in the graphs above each compared group.

amplified cDNA. Highly expressed housekeeping genes such as *ACTB* and *GAPDH* were readily measured, and quantitative real-time PCR measurements significantly correlated with sequencing read counts (Figure 4C,D; Supplementary Figure 4A,B). The transcript abundances of artificial RNA spike-ins mixed in with the single-cell RNA (ERCCs and ArrayControl spike-in RNA standards) were also highly correlated with the normalized sequencing counts (FPKM) (Supplementary Figure 4C). These observations support the

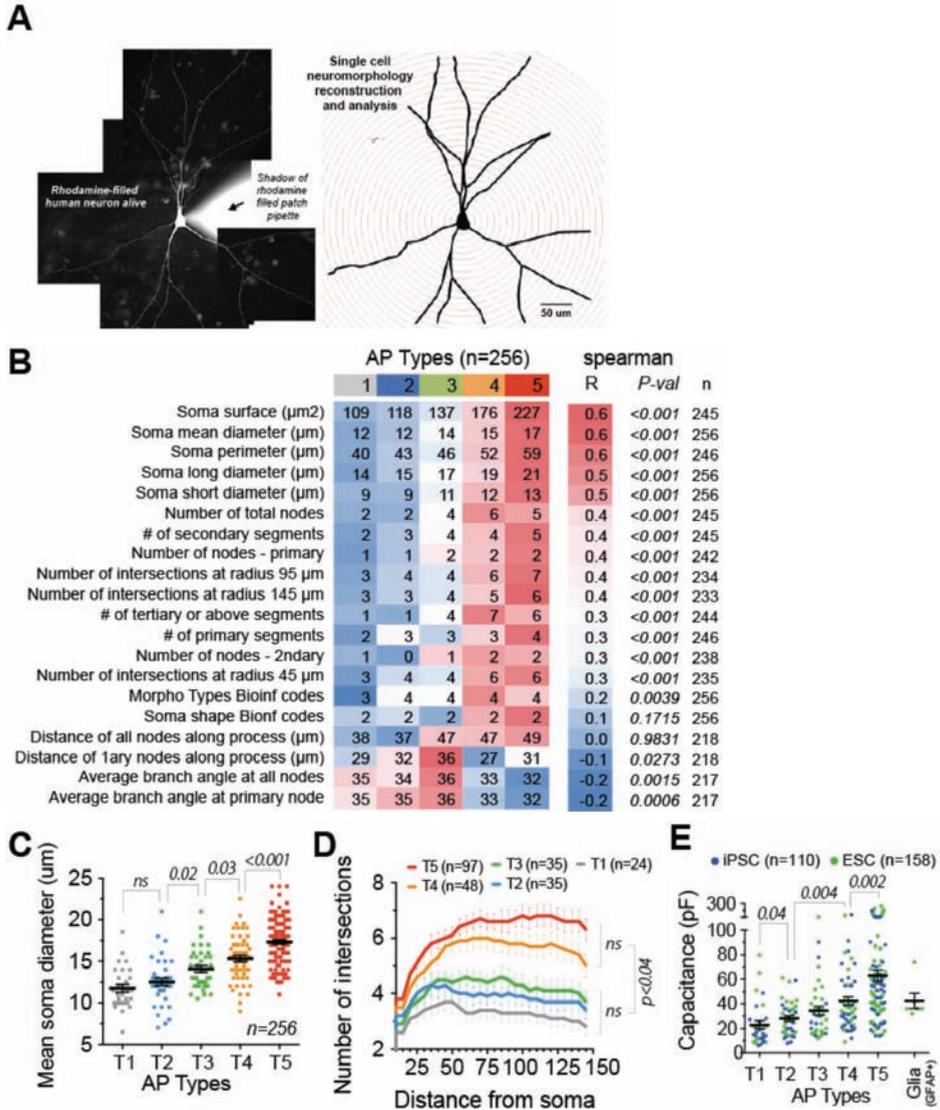


Figure 3. Morphological features that correlate with AP states. (A) Stacked photos of a live patched neuron, which was filled with rhodamine (left) and morphologically reconstructed with NeuroLucida (right). (B) Correlation of morphological measurements (rows) with AP Types (columns 1-5). The values in the central heat map represent the median for the neurons in respective AP Type categories. To illustrate the direction of the measurement variations between AP Types, we color coded the heat map (values normalized by row with mean = 0 and variance = 1) from low to high values (blue, to white, to red). The last two columns show the Spearman correlation coefficient (*R*) and significance (*P*-val) of Spearman's rank correlation between the electrophysiological measures (rows) and the numerical class (1-5) of ePhys Type. The properties were sorted by decreasing correlation coefficient. (C) Mean soma diameter significantly increased in more functional AP Types. (D) Sholl analysis revealed that more functional AP Types had significantly more complex dendritic/axonal arborization. Mann-Whitney *U* two-tailed tests were performed at 20, 50 and 100

μm between different types. The dendritic complexities of Type 4 and 5 neurons were not significantly different, and the complexity between Types 1, 2, 3 was not either. However, Types 4 and 5 were significantly more complex than Types 1, 2, 3. Mean \pm S.E.M. shown. **(E)** The capacitance significantly increased in more functional AP Types of neurons. Significance threshold was $P < 0.05$. Mann-Whitney U two-tailed P -values are shown in **C**, **D** and **E**.

quality of our single-cell transcriptomic data. To minimize erroneous conclusions that arise from lowly expressed, highly varying genes, we excluded reads that were detected below 1 tpm (transcripts per million). Additionally, a small number of cells ($n = 3$) with fewer than 2,000 genes detected above 5 tpm were excluded. Without these outliers, an average of $\sim 3,000$ genes per cell was detected above 5 tpm (**Figure 4E**). Interestingly, the number of genes detected was significantly higher in Type 5 neurons and astrocytes than in less active AP Types (**Figure 4E**). However, we found no significant correlation between the number of detected genes (>5 tpm) and the size of the cells (**Figure 4F**).

Molecular Signatures Distinguish Neuronal Functional States *In Vitro*

To examine variation in the single-cell gene expression, we excluded the outliers with trace of DNA contamination, abnormal expression level of housekeeping genes, damaged axons and dendrites from the collection, unstable physiological recordings before collection, abnormal cDNA quantity, low number of uniquely mapped reads and low number of detected genes. We then performed a hierarchical clustering analysis and PCA with the transcriptomes of 56 single cells (50 differentiated neurons, and 6 astrocytes) that passed all QC (**Figure 5A,B**). This unsupervised and unbiased single-cell RNA-seq analysis revealed a distinct transcriptomic cluster of neurons corresponding to highly functional states (83% AP Type 5 neurons, 17% AP Type 4 neurons). This highly functional cluster separated from astrocytes and a mixed group of neurons that represents “transitory” neuronal states (**Figure 5B**, **Supplementary Figure 5A**). Furthermore, the molecular profiles of cells in distinct functional states grouped together, independently of individual genetic background (cell lines from different subjects) and the time spent in culture (**Figure 5A,B**).

To better recapitulate the transcriptional dynamics of functional states, we also demonstrated the molecular trajectory of the electrophysiological maturation by applying monocle analysis [28]. Interestingly, AP Types 1 and 2 clustered at the beginning of the pseudotime and branched out towards astrocytes (GFAP+) or towards AP Type 3 and 4 neurons, and highly functional AP Type 5 neurons clustered at the end of the pseudotime (**Figure 5C**). Together, these results indicate that mRNA expression correlates strongly with neurophysiological states.

To obtain physiologically relevant neuronal circuits, we generated mixed cultures of neurotransmitter classes and astrocytes. The patch clamping was not targeting a specific subpopulation of neurons, and most patched cells were tagged with Synapsin:GFP lenti-

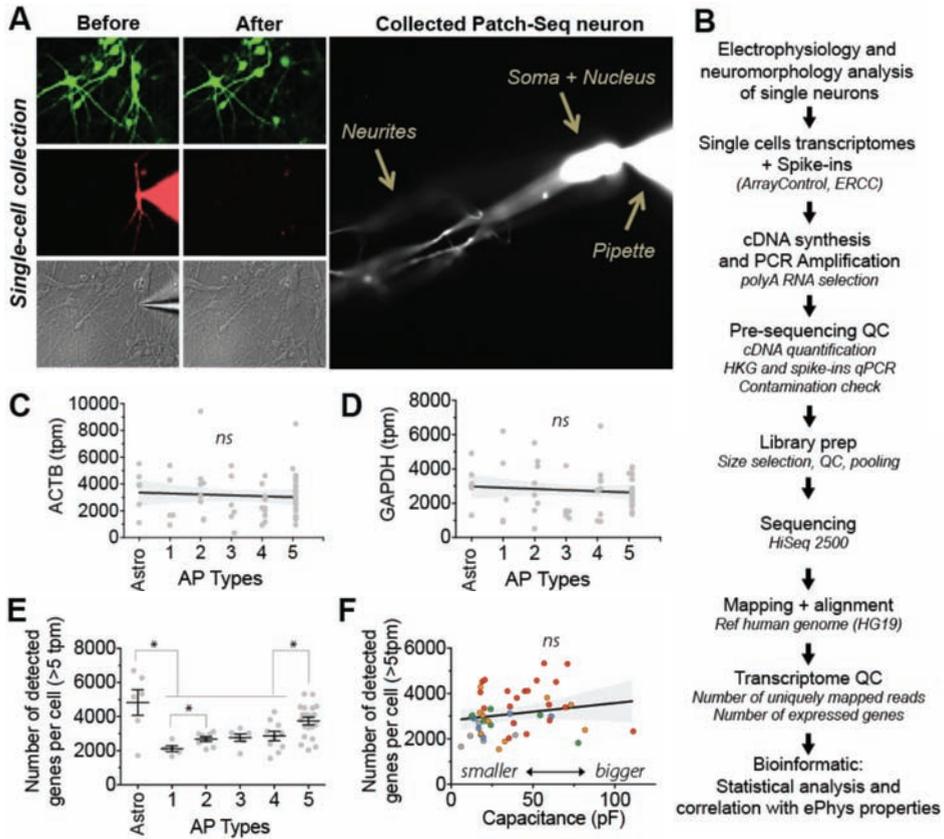


Figure 4. Whole single-cell RNA-seq of patch-clamped neurons (Patch-seq). (A) The photos show an example of neuronal culture stained with GFP before and after a single neuron was patched and collected for transcriptome analysis. For every cell included in the analysis, the entire neuron was collected, including the soma/nucleus and neurites. The photos permitted us to confirm that only the patched neuron filled with rhodamine was collected, leaving the surrounding tissue intact. (B) Following electrophysiological and morphological analyses of live neurons, the single cells were collected and their transcriptome processed for deep sequencing, bioinformatics processing and statistical analysis. (C,D) Housekeeping genes such as *ACTB* and *GAPDH* were detected in every cell. Their expression levels were not significantly different between types of neurons and astrocytes. (E) Significantly more genes were expressed in astrocytes and Type 5 neurons compared to the other neuronal types. Asterisks represent Mann Whitney P -values <0.05 . (F) The number of detected genes above 5 tpm did not significantly correlate with the size of the neurons, estimated here by capacitance. The AP Types are color coded in the graph (red: T5, orange: T4, green: T3, blue: T2, gray: T1).

vector. Therefore, to better characterize our sample, we isolated the AP Type 5 neurons that were sequenced ($n = 22$) and determined the nature of their neurotransmitter classes. Based on the normalized expression of five known genes, most Type 5 neurons segregated into different neurotransmitter classes (36% glutamatergic *VGLUT3*, 27% GABAergic *GAD1*, 14% dopaminergic *VMAT2* or *COMT*, 9% serotonergic *TPH2*, and

14% undefined neurons) (**Figure 5D**). These results demonstrate that our functional classification is sufficiently broad to incorporate more detailed subcategories of various neurotransmitter classes.

To further characterize the molecular profiles of differentiated neurons in various functional states, we compared the single-cell mRNA expression of 45 genes typically known for their importance to neuronal functions (**Figure 5E**, **Supplementary Figure 5B**). Not surprisingly, most well-known neuronal genes (e.g., pan neuronal, synaptic, APs, ion channels genes) were expressed in all the differentiated neurons in our sample. Typical astrocytic/glia genes (e.g., *AQP4*, *GFAP*, *S100B*, *GJA1*, *SLC1A3*, *EFNB1*) were expressed in astrocytes, which were identified based on LV marker GFAP:tdTomato, distinctive morphology and electrophysiology. Genes typical of early neuronal development (e.g., *SOX2*, *NES*) were expressed mostly in immature AP Types and some astrocytes. A few neuronal genes appeared expressed in all AP Types and even in astrocytes/glia (e.g., *MAP2*, *TUBB3*, *TUBB*). Interestingly though, many neuronal genes that were exclusive to neurons expressed at similar levels in all AP Types (e.g., *MAPT*, *NCAM1*, *ANK3*, *SYN1*, *SYP*). Therefore, we asked whether some genes may specifically express in highly functional AP Types of neurons. We first assessed the Spearman's correlation between gene expression and AP Type classification, which revealed 45 genes that most significantly correlated (**Figure 5F**). Some of these genes have known neuronal functions such as synaptic plasticity (e.g., *RAB3B* [29, 30], *PCLO/Piccolo*) or voltage-gated sodium channels (e.g., *SCN9A/Nav1.7*). However, interestingly, a majority of these genes have not been investigated before for specific neuronal functions and may be used as new potential biomarkers to identify functional human neurons *in vitro*. This finding then led us to propose that single-cell gene expression data has utility in the prediction of the electrophysiology of human neurons.

Machine-Learning Classifiers Integrate Transcriptome Features to Predict Functional States

To identify the single-cell transcriptomic signature important to predict functional states, we trained an Extremely Randomized Trees classifier using 56 single cells (**Figure 6A**). We used a randomized 10-fold cross-validation to evaluate the performance of the algorithm in classifying cells in different functional state categories (**Figure 6**). In **Figure 6**, we illustrated the results of three classifiers: the first one isolates AP Types 4-5 neurons with 92% accuracy in a mixed population including astrocytes (**Figure 6B**); the second isolates AP Type 5 neurons with 83% accuracy in a population of differentiated neurons only (**Figure 6C**); and the third isolates highly functional neuronal states (based on transcriptome PCA clustering in **Figure 5B**) with 86% accuracy in a population of differentiated neurons only (**Figure 6D**, **Supplementary Figure 6A**). For higher accuracy (more true positive), the classifiers can be trained differently by adjusting the “classifier score split”,

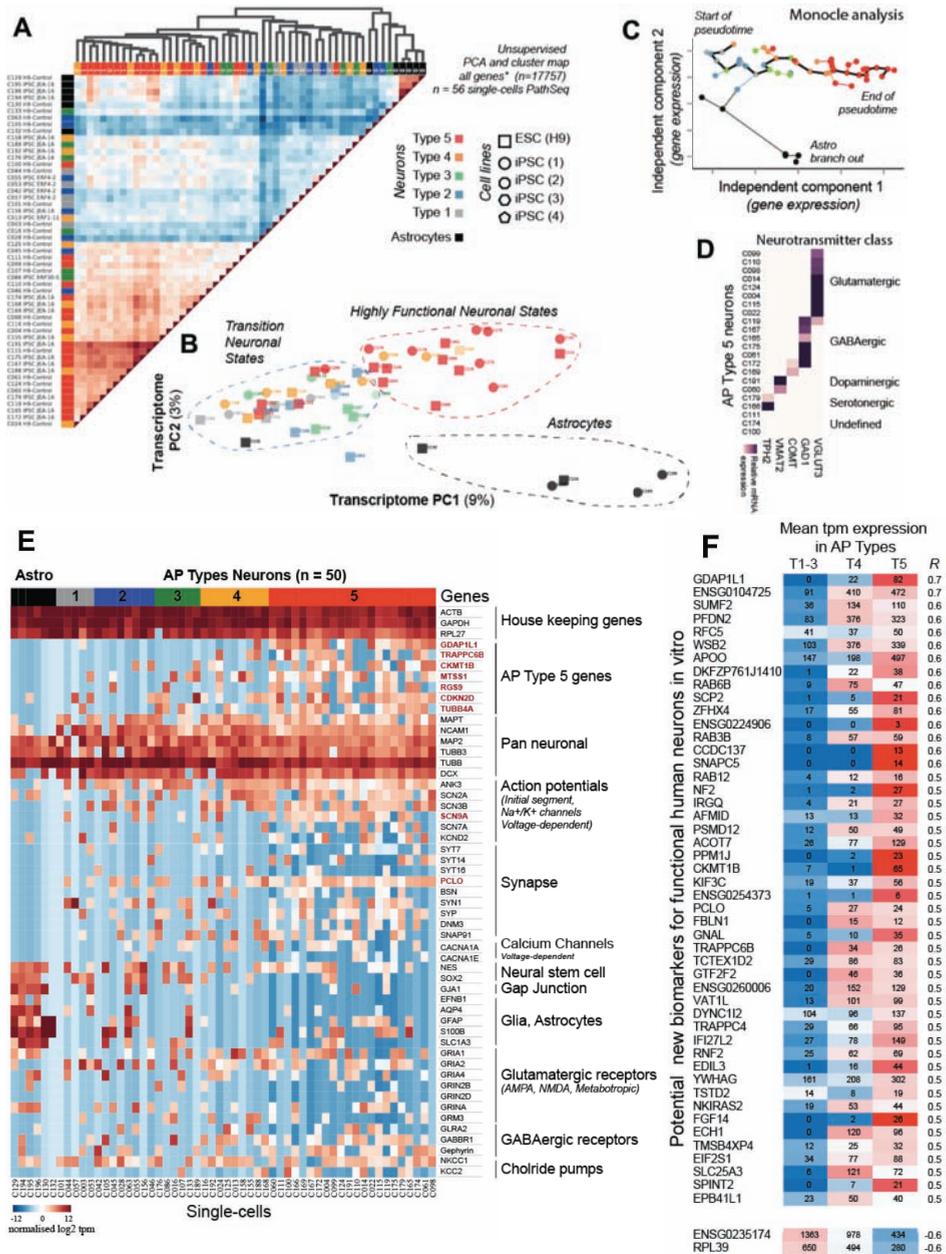


Figure 5. Single-cell transcriptomes segregate functional states of differentiated neurons. Unbiased and unsupervised analysis was performed on the transcriptome (17,757 genes detected >5 tpm in at least one cell) on a sample of 56 whole single cells (including nucleus, soma and distant neurites), which passed all QC. This sample comprised 50 differentiated neurons displaying Nav currents and 6 astrocytes expressing GFAP:GFP. **(A)** Unsupervised clustered heat map of cell-to-cell transcriptome correlations (Euclidean distances). The linkage distances of the hierarchical clustering represent an estimate of the quality of the unbiased/unsupervised clusters. **(B)** Unsupervised principal component analysis on the entire single-cell

transcriptomes. Cell transcriptome profiles (symbols) are represented in a two-dimensional principal component space. These unsupervised analyses reveal molecular segregation between groups of neurons in highly functional states (almost all AP Type 5 – 85%), less functional neurons in ‘transition states’ (mix of different AP Types with significantly less Type 5 – 15%) and astrocytes. Furthermore, the functional molecular clusters surpassed transcriptional differences between iPSC lines from different subjects and even an embryonic stem cell line. **(C)** Monocle analysis illustrates the progress through functional states by pseudotemporal ordering of single-cell mRNA expression profiles. Cell expression profiles (points) are represented in a two-dimensional independent component space. Lines connecting points represent edges of the minimum spanning tree constructed by monocle. Solid black line indicates the main diameter path of the minimum spanning tree and provides the backbone of monocle’s pseudotime ordering of the cells based on molecular profiles. **(D)** The expression of five key neurotransmitter genes in 20 sequenced Type 5 neurons was normalized and compared to determine their neurotransmitter identity. Most cells could be classified as either glutamatergic, GABAergic, dopaminergic or serotonergic, and a few cells remained undefined. **(E)** Heatmap of single-cell gene expression levels. The genes were selected and grouped based on known neuronal functions. For comparison, AP Type 5 genes identified in the present study are in red/bold. See also Supplementary Figure 5B. **(F)** Genes significantly correlating with numerical AP Types classification by Spearman’s rank correlation coefficient ($P < 0.001$).

but the algorithm may also generate more false negatives, resulting in lower sensitivity. The area under the curve (AUC) of the receiver-operating characteristic, which is a measure of the trade-off between sensitivity and specificity, was 0.93, clearly above random prediction (0.50) (**Supplementary Figure 6B**). The unbiased training algorithms selected and weighted the optimal genes to predict functional types (**Figure 6**).

Identification and Isolation of Functional Neurons with a *GDAP1L1* Promoter-Driven GFP Viral Vector

In addition to the machine-learning classifier, we performed an in-depth differential expression analysis to reveal a complementary set of genes specifically expressed in Type 5 neurons (**Figure 7A**). Altogether, we performed three independent unbiased analyses to identify a total of 165 genes (**Supplementary Table 1**) related to specific electrophysiological functional states (Spearman correlation in Figure 5F; machine learning in Figure 6B-D, differential expression in Figure 7A). Among the most highly ranked genes, 29 genes appeared in more than one analysis (**Supplementary Table 1**). Only one gene appeared in all of those independent analyses: *GDAP1L1*. *GDAP1L1*, which is the human paralog of the ganglioside *GDAP1*, had been previously suspected to have an important function in neuronal cells because of its high abundance in brain tissues [31]. Interestingly, *GDAP1L1* mRNA is expressed exclusively in brain tissues in human (**Supplementary Figure 7A**). We confirmed with immunostaining that *GDAP1L1* protein was present in human iPSC-derived neurons but not in astrocytes (**Figure 7B**). Our immunohistochemistry staining also suggested that *GDAP1L1* was only present in a subset of TUJ1+ neurons (**Figure 7C**), presumably highly functional neurons (Type 5). To confirm this hypothesis, we constructed a lentiviral vector expressing GFP under the control of the *GDAP1L1* promoter. We infected three different

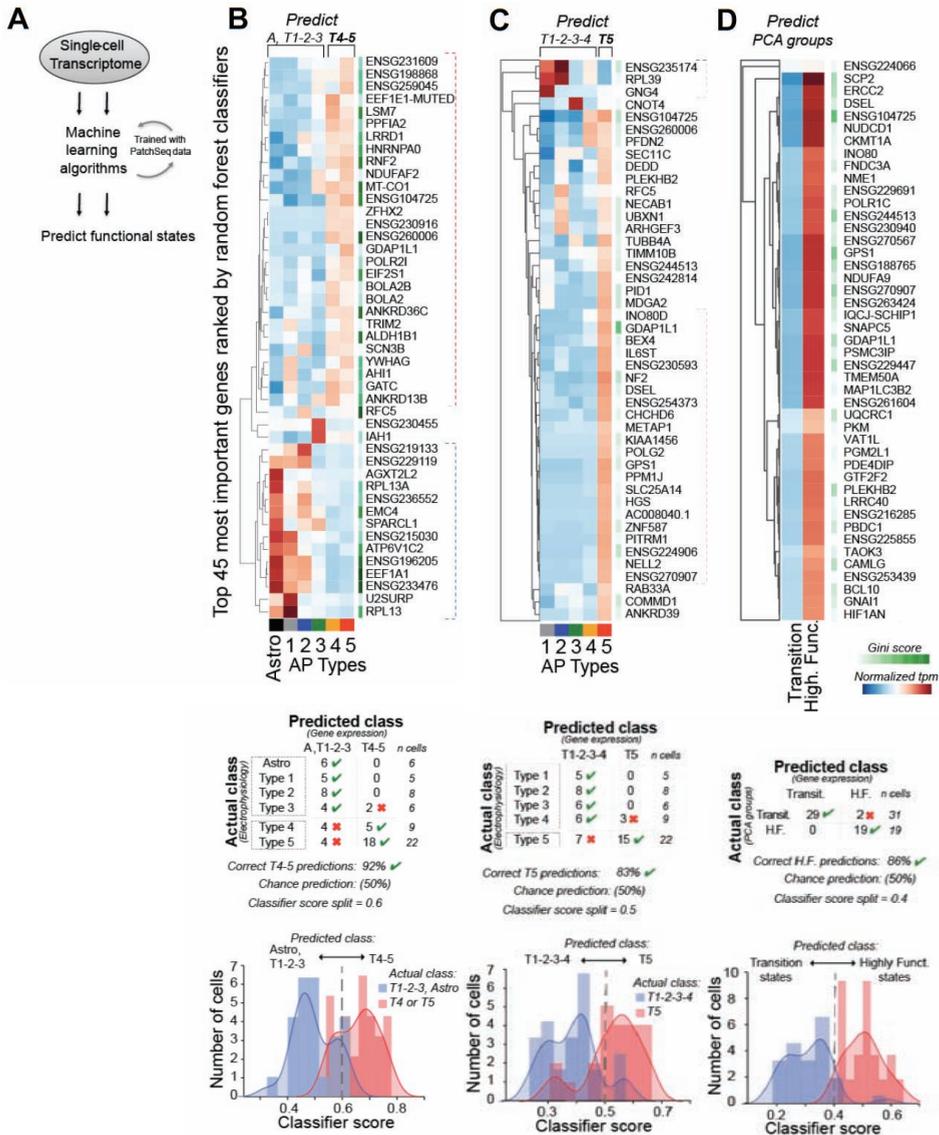


Figure 6. Machine-learning tree classifiers predict functional states of neurons based on single-cell transcriptome and reveal potential biomarkers. (A) Extremely Randomized Trees (ERT) classifier built with the transcriptome of 56 single cells and trained with electrophysiological data. Actual classes were attributed by electrophysiological measurements. Predicted classes were attributed by the machine-learning algorithm based on single-cell transcriptomes. (B-D) Each ERT classifier was trained to categorize the cells in two functional classes. Classifier B was trained to predict AP Types 4-5 in a mixed group of cells including differentiated neurons and astrocytes. Classifier C was trained to predict Type 5 neurons from other differentiated neurons. Classifier D was trained to predict highly functional (HF) neurons from other differentiated neurons in transitional states determined by the PCA clusters. The lists show the top 45 genes selected by the ERT classifier with the highest Gini scores. The green heat map columns represent the normalized importance of each gene attributed by the classifier (Gini score). The blue-red heat map matrices represent

the mean expression normalized (mean 0 and variance 1) of each gene from high (red) to low (blue) in actual cell types. The genes were ordered by Euclidean clustering of the gene expression by actual functional states. Confusion matrices are displayed below the heat maps. The confusion matrices values represent the numbers of cells in each category. Predictions were annotated with a green checkmark if correct and a red cross if false. The test-fold score for each cell was recorded and a score histogram was computed for each cell group (blue and red bars). Classes were predicted with a determined classifier score split, indicated by a gray dashed line in the histograms. See also Supplementary Figure 6.

cell lines of human neurons with the vector (**Figure 7D**, see also **Supplementary Figure 6** for additional IHC of vector). While GDAP1L1 mRNA and protein were not found in astrocytes, the GFP under the control of GDAP1L1 gene promoter was occasionally found in astrocytes (GFAP+). Nevertheless, we patched the neurons with the brightest GFP and clear neuronal morphology and found that all of them were highly functional neurons ($n = 10/12$ were Type 5, $n = 2/12$ were Type 4, and $n = 11/12$ neurons had clear spontaneous AMPA synaptic activity) (**Figure 7E,F**). Next, we asked whether this strategy could be used to isolate mature functional human neurons by a higher-throughput technique. With that aim, we derived NPCs from three different cell lines and differentiated them into neurons for ~4 weeks in neuronal maturation medium. Then, we infected the neuronal cultures with two lentiviral vectors (LV), GDAP1L1-GFP (to identify functional neurons) and GFAP-tdTomato (to identify astrocytes). A few days later we processed the cells with FACS. After exclusion of dead cells and GFAP+ cells, we isolated a population of neurons with high GDAP1L1-GFP expression (**Figure 7A**). This subset of neurons with high GDAP1L1-GFP also showed particularly high abundance of TRAPPC6B protein (measured with primary antibody) (**Figure 7A**). TRAPPC6B is another top gene candidate that appeared from our transcriptome analysis to be predictive of highly functional neurons. This finding further supported our premise that this subpopulation of high GDAP1L1-GFP neurons represents the electrophysiologically active state. Therefore, we repeated the FACS experiment with live cells differentiated for ~4 weeks and sorted out three subsets of live cells: (1) presumably highly functional neurons (GDAP1L1:GFP-highly positive and GFAP:tdTomato-negative), (2) astrocytes (GFAP:tdTomato-positive and GDAP1L1:GFP-negative) and (3) the rest of the cells, including less functional neurons (**Supplementary Figure 8**). We then re-plated 50,000 neurons with high GDAP1L1-GFP per well (48-well plate) onto a glass coverslip coated with a layer of human cerebellum astrocytes. Four days after re-plating, we patch-clamped the GDAP1L1:GFP-positive sorted cells and confirmed that they were highly functional Type 5 neurons (**Figure 7B**, $n = 5/5$). Altogether, these findings demonstrate a first proof of concept that the genes identified in our multimodal analysis may be used as biomarkers to help identify, enrich or isolate highly functional states of iPSC-derived human neurons *in vitro*.

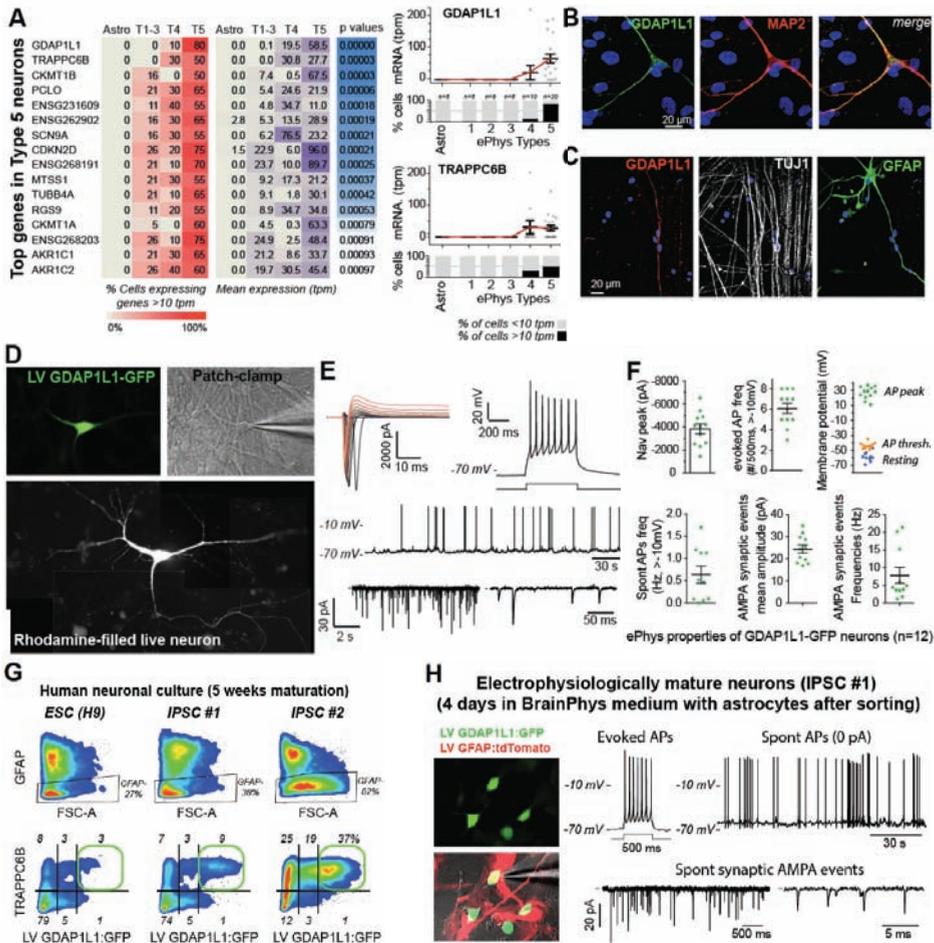


Figure 7. Biomarkers to isolate highly functional Type 5 neurons. (A) The top 16 genes expressed mostly in Type 5 neurons. The genes were selected based on the combination of several criteria: OFF (<10 tpm) in all astrocytes, OFF in >70% of Types 1-2-3 neurons, OFF in >50% of Type 4 neurons, ON (>10 tpm) in >50% Type 5 neurons and then by the genes significantly more expressed in Type 5 versus all other cell types (P -values from a Mann Whitney test). The selected genes were then ordered by P -values and the most significant 16 genes are shown ($P < 0.001$). Significance was tested with two-tailed Mann-Whitney U test between Type 5 neurons and all the other cells. The expression of the top two genes was plotted. Each gray point represents a single neuron. The red curve is the mean \pm S.E.M. In the bottom charts black bars represent the proportion of cells in each cell type group with ON expression (>10 tpm). (B,C) Immunostainings of fixed human neuronal cultures confirm the translation of GDAP1L1 at the protein level in some neurons (MAP2+, TUJ1+) but not in astrocytes (GFAP+). (D) Example of a live human neuron expressing GFP under GDAP1L1 promoter and filled with rhodamine with the patch-clamping pipette. (E) Whole-cell patch-clamp recordings from GDAP1L1-GFP neurons. The brightest GFP cells with neuronal morphology were selected for patch clamping after ~4 weeks in BrainPhys neuronal medium. The neurons expressed strong Nav/Kv currents (top left). The evoked APs were measured by slightly hyperpolarizing the cells to reduce spontaneous activity (top right). Spontaneous APs were recorded at resting membrane potential with zero current injected (middle). Spontaneous AMPA-mediated excitatory synaptic activity was recorded in volt-

age clamp at -70 mV (bottom). **(F)** The electrophysiological properties of the patched GDAP1L1-GFP neurons ($n = 12$) were mature and functional. Means \pm S.E.M. shown. The properties of single neurons are represented by each dot in the graphs. APs were counted only if amplitude was above -10 mV. Spontaneous AP frequencies were measured at resting membrane potential. **(G)** Three human neuronal cell lines, which matured for 5 weeks in BrainPhys basal + supplements, were infected with LV GDAP1L1:EGFP for 5 days before being dissociated, fixed and stained. Healthy cells not expressing GFAP (GFAP-neg; top graphs) were analyzed for their expression of TRAPPC6B protein and LV GDAP1L1:GFP expression (bottom graphs). The green rectangles highlight the cells expressing high levels of GFP under GDAP1L1 promoter and high levels of TRAPPC6B protein. The proportion of presumably mature ePhys types of neurons varied highly between the three cell lines. **(H)** Using two lentiviral vectors we sorted live astrocytes and live neurons from iPSC#1 after 5 weeks in BrainPhys and re-plated them on glass coverslips for 4 days before electrophysiological evaluation. The population of neurons expressing high levels of GFP under GDAP1L1 promoter (and no GFAP:tdTomato) were patched and all the cells were classified as highly functional neurons ($n = 5$) with on average high evoked firing frequencies (16 ± 2 Hz), low resting potentials (-58 ± 2 mV), and large AP amplitude (91 ± 4 mV).

DISCUSSION

Studies of post-mortem human brain tissues have revealed some molecular and anatomical features in neurodevelopment and in the adult brain [32]. However, studying the neurophysiological properties of live single neurons in the human brain is technically challenging and so far has been mostly restricted to animal models, occasionally to brain slices of aborted human fetus [33, 34], and more recently to human neuron cultures reprogrammed from fibroblasts [13, 16, 25, 26, 35, 36]. In this study, we have bridged human neurophysiology and gene expression by integrating single-cell electrophysiology, morphology and transcriptomics analysis of live human neural cells.

Single-Cell Patch-Seq: Multimodal Analysis of Electrophysiology, Morphology and Transcriptomic Profiles

Recent breakthroughs in single-cell transcriptomic methods have emphasized the importance of identifying different biological types of cells [37-47]. To resolve the diversity of cellular states, the use of single-cell analysis yields important insights that are masked or misrepresented in bulk RNA preparation from mixed cell populations [39, 48, 49]. For many years, only a few studies succeeded in applying gene expression analysis to patch-clamped neurons [50-56]. Here, we demonstrate the possibility of performing whole-transcriptome analysis on patch-clamped single human neurons derived from iPSCs or ESCs. Interestingly, many neuronal mRNAs are transported to the synapses and spines for rapid on-demand local translation [57, 58]. Unlike most single-cell methods relying on tissue dissociation and flow cytometry, the method developed in this study processes the entire neuron and therefore includes mRNA from the nucleus, the soma and distant axons/dendrites, which may provide a more accurate representation of the transcriptome [58]. This approach may also be relevant for specific neurological disorders that may af-

fect mRNA dendritic/axonal trafficking or the function of mRNA at synaptic sites [59, 60]. Most importantly, in addition to the transcriptomic profile, our method provides unique physiological and morphological information about the cells. Our results reveal the strong correlation between functional and molecular profiles, validate the biological relevance of our methods merging patch-clamping analysis and single cell RNA-seq, and demonstrate that transcriptome analysis can be used to predict the physiology of a cell. This approach can be applied to any kind of electrophysiological type of cell and may help to better define the cellular diversity in the brain [61-64]. Beyond the usefulness of such multimodal analysis in cellular neuroscience, this framework may also branch out to other fields of cell biology to correlate human cell physiology and molecular signatures.

Functional Neuronal States and Neurodevelopmental Timeline *In Vitro*

Typically, neurodevelopmental processes are described over a specific timeline. For example, newly born mouse neurons mature in pre-existing adult brain circuits by following a timed sequence of physiological events [65, 66]. However, at the single-cell level, intrinsic and extrinsic programs can dynamically influence the rate and speed at which different human neurons reach functional states *in vitro*. The prediction of this maturation process is sometimes challenging as it may vary between cell types, individuals and species [15-17]. For example, the human brain matures over a longer period of time than most other species studied in laboratories. Similarly, human neurons *in vitro* usually take a longer time than rodent neurons to reach the equivalent maturity [35]. In addition, the technical variability of tissue culture protocols between independent labs makes “time *in vitro*” an unreliable metric to objectively assess the functional maturity of neurons. As an alternative, we propose a method to efficiently quantify and isolate defined neurophysiological profiles *in vitro* based on data-driven features that do not require the measure of time *in vitro*. In this classification, we define the AP Type 5 neurons as a highly functional state of maturity; we define this stage as a broad minimal requirement to study functional human neurons *in vitro*. However, it is important to note that this stage is not necessarily terminal. For example, we have recorded the electrophysiological properties of neurons that were matured and kept in BrainPhys with supplements for more than 20 weeks and, while the broad properties used to define Type 5 neurons were indistinguishable from those of neurons maturing in the same medium for 6 weeks, we found that the membrane resistance significantly decreased over the longer time periods (e.g., Type 5 neurons R_m was 1612 ± 208 MOhm after 45 days in BrainPhys ($n = 17$) and 224 ± 26 MOhm after 150 days in BrainPhys ($n = 11$), see also ref. [35]). Furthermore, we based our functional state classification on a PCA that only integrated cell-intrinsic electrophysiological properties (25 features). Following this reasoning, we excluded synaptic input properties from that PCA because they most strongly depend on extrinsic neuronal network properties. This approach also allows us to use our classification model to compare synaptic function in

cells of equivalent AP Types with less bias. Nevertheless, adding healthy synaptic activity to the PCA expands mostly PC2 and does not affect PC1 or our functional state classification.

Neuronal Diversity

In the present study, we present an unbiased profiling of the electrophysiological states of human neurons *in vitro*. The framework that we propose here is sufficiently broad to include a wide heterogeneity of neuronal neurotransmitter classes (e.g., glutamatergic, GABAergic, dopaminergic, serotonergic and motorneurons). Similarly, we found that each of our defined AP Types could be obtained from neurons categorized in various morphological classes (e.g., bipolar, fusiform, pyramidal, multipolar; see **Supplementary Figure 3A,B**). Finally, the framework is also sufficiently broad to avoid masking relevant differences in disease modeling studies. While the present study points out the importance of resolving broad neurophysiological states *in vitro*, several other important studies have also highlighted the unique nature of cellular and brain region identities of neurons in rodents [55, 56, 62, 63, 67-70] and human brain [71, 72]. The complex, precise, heterogeneous and dynamic assembly of 100 billion cells in the human brain is the best living illustration of cellular diversity in functional tissues. Interestingly, each neuron has the potential to be unique even at the genomic level [73]. The depth of the classification one should reach remains arbitrary [62]. Further work will be needed to deepen our broad functional states classification to more precise sub-classification. For example, it is possible that within AP Type 5 neuron categories, different neurotransmitter classes have subtle different physiological properties that they may acquire over various periods *in vitro*. Regardless, our results point out the importance of stratifying neurons into equivalent functional states before investigating the differences among patients in disease models *in vitro*.

Models of Neurological Disorders *In Vitro* Will Benefit from Better Characterization and Quantification of Electrophysiologically Active Human Neurons

Advances in human cellular reprogramming have propelled a new wave of *in vitro* experiments to identify differences between healthy subjects and patients with neurological and psychiatric diseases [24, 74-77].

Studies of iPSC models *in vitro* have found that neurons obtained from patients affected by autism have significantly smaller soma, reduced dendritic complexity and less synaptic activity [78]. We found the same set of phenotypes when comparing less functional AP Types (Types 1-2-3) with more functional AP Types (Types 4-5). Together, these findings support the hypothesis that autistic neurons suffer from a general lack of functional maturity, which may cause an indirect defect in synapses. Being able to examine many

cellular phenotypes simultaneously in the same single neurons may help to discern the causes of a disease from subsequent pathophysiological cascades and may be critical to the discovery of effective treatments.

Despite the huge promises for translational research, the inherent technical and biological variability of iPSC technologies remain a significant hurdle [16, 24]. Disorders that reveal symptoms relatively late in adulthood, such as Parkinson's or Alzheimer's, are less likely to reveal early neurodevelopmental features. Therefore, it is important, at least in these latter models, to compare neurons of equivalent functional maturity, and efficient ways to identify electrophysiologically active neurons are needed. By revealing the molecular signature of highly functional states of differentiated neurons (AP Type 5), we have shown the possibility to predict functional profiles without the need of electrophysiological measurements. The ability to efficiently identify neurons in a specific functional state compensates for the heterogeneous proportion of functional neurons among cell lines and will strengthen comparative analysis in stem cell models of brain disorders. Ultimately, strategies to provide electrophysiologically homogeneous human neurons will assist large-scale drug screening, which can lead to the discovery of efficient treatments for neurological and psychiatric disorders.

ACKNOWLEDGMENTS

We are grateful to Elisabeth Santo and Sarah Marshall for help with the morphological reconstruction. We thank Gage lab members Bobbie Miller and Lynne Moore for preparation of viral vectors and Eunice Meija for immunohistochemistry. Thanks to Gage lab (Prattap Venepalli, Apua Paquola, Sara Linker, Son Pham) and Yeo lab (Olga Botvinnik) members for fruitful bioinformatics discussions on single-cell transcriptomics. We thank Mary Lynn Gage for edits on the manuscript. This study was supported by grants from Ipsen Pharma, Annette C. Merle-Smith, The Leona M. and Harry B. Helmsley Charitable Trust Grant #2012-PG-MED002, Bob and Mary Jane Engman, the JPB Foundation, G Harold and Leila Y. Mathers Foundation, and NIH Grants MH095741 (to F.H.G.); also by a Fay/Frank Seed Grant from the Brain Research Foundation and NIH Grants NS075449, HG004659, HG007005 (to G.W.Y.). G.W.Y. is an Alfred P Sloan Research Fellow. This work was also supported by NSF Graduate Research Fellowship (to B.K.), the George E. Hewitt Foundation for Medical Research (to J.E.) and the FP7 Marie Curie International Outgoing Fellowship for Career Development (to C.B.).

REFERENCES

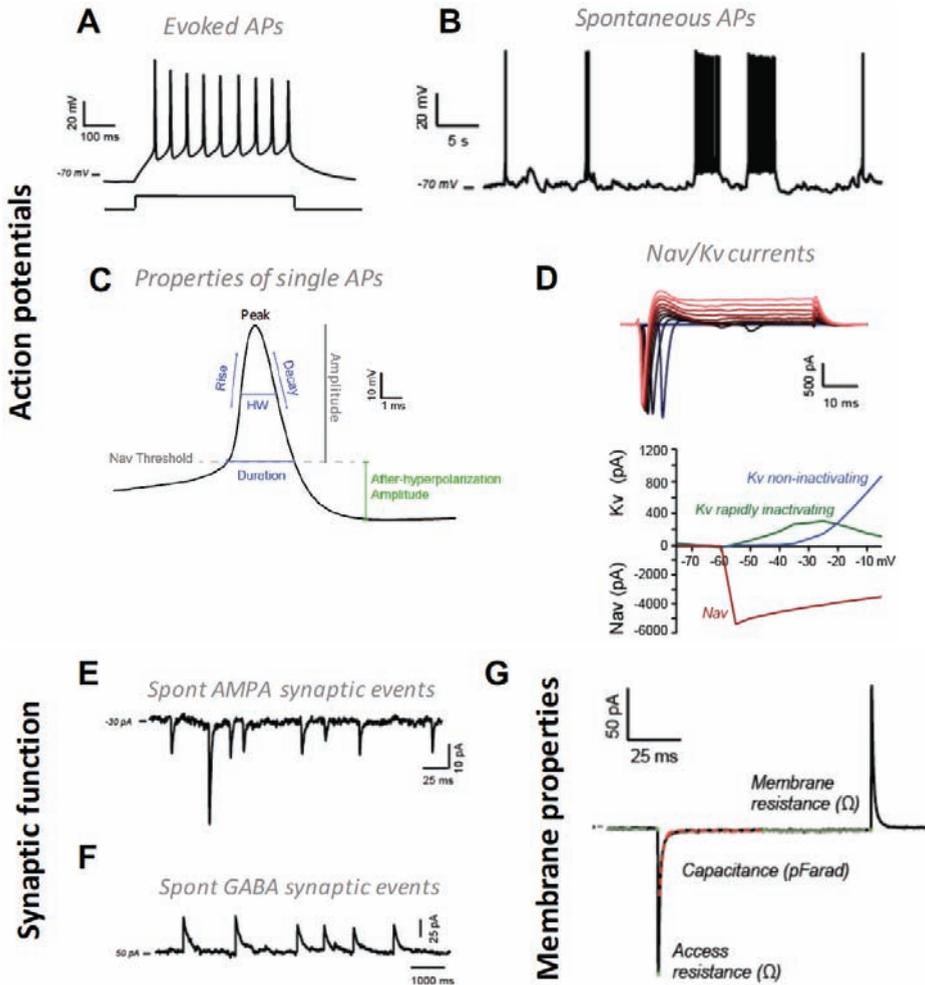
1. Greig LC, Woodworth MB, Galazo MJ, Padmanabhan H, Macklis JD. Molecular logic of neocortical projection neuron specification, development and diversity. *Nat. Rev. Neurosci.* 14(11), 755-769 (2013).
2. Molyneaux BJ, Arlotta P, Menezes JR, Macklis JD. Neuronal subtype specification in the cerebral cortex. *Nat. Rev. Neurosci.* 8(6), 427-437 (2007).
3. Cho MS, Hwang DY, Kim DW. Efficient derivation of functional dopaminergic neurons from human embryonic stem cells on a large scale. *Nat. Protoc.* 3(12), 1888-1894 (2008).
4. Vadodaria KC, Mertens J, Paquola A *et al.* Generation of functional human serotonergic neurons from fibroblasts. *Mol. Psychiatry* 21(1), 49-61 (2016).
5. Shi Y, Kirwan P, Smith J, Robinson HP, Livesey FJ. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nat. Neurosci.* 15(3), 477-486, S471 (2012).
6. Boyer LF, Campbell B, Larkin S, Mu Y, Gage FH. Dopaminergic differentiation of human pluripotent cells. *Curr. Protoc. Stem Cell Biol.* Chapter 1 Unit1H 6 (2012).
7. Mertens J, Marchetto MC, Bardy C, Gage FH. Evaluating cell reprogramming, differentiation and conversion technologies in neuroscience. *Nat. Rev. Neurosci.* 17(7), 424-437 (2016).
8. Kirwan P, Turner-Bridger B, Peter M *et al.* Development and function of human cerebral cortex neural networks from pluripotent stem cells *in vitro*. *Development* 142(18), 3178-3187 (2015).
9. Wernig M, Zhao JP, Pruszak J *et al.* Neurons derived from reprogrammed fibroblasts functionally integrate into the fetal brain and improve symptoms of rats with Parkinson's disease. *Proc. Natl. Acad. Sci. U. S. A.* 105(15), 5856-5861 (2008).
10. Pang ZP, Yang N, Vierbuchen T *et al.* Induction of human neuronal cells by defined transcription factors. *Nature* 476(7359), 220-223 (2011).
11. Chanda S, Ang CE, Davila J *et al.* Generation of induced neuronal cells by the single reprogramming factor ASCL1. *Stem Cell Reports* 3(2), 282-296 (2014).
12. Zhang Y, Pak C, Han Y *et al.* Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* 78(5), 785-798 (2013).
13. Bardy C, Van Den Hurk M, Eames T *et al.* Neuronal medium that supports basic synaptic functions and activity of human neurons *in vitro*. *Proc. Natl. Acad. Sci. U. S. A.* 112(20), E2725-2734 (2015).
14. Prilutsky D, Palmer NP, Smedemark-Margulies N, Schlaeger TM, Margulies DM, Kohane IS. iPSC-derived neurons as a higher-throughput readout for autism: promises and pitfalls. *Trends Mol. Med.* 20(2), 91-104 (2014).
15. Belinsky GS, Moore AR, Short SM, Rich MT, Antic SD. Physiological properties of neurons derived from human embryonic stem cells using a dibutyryl cyclic AMP-based protocol. *Stem Cells Dev.* 20(10), 1733-1746 (2011).
16. Wu H, Xu J, Pang ZP *et al.* Integrative genomic and functional analyses reveal neuronal subtype differentiation bias in human embryonic stem cell lines. *Proc. Natl. Acad. Sci. U. S. A.* 104(34), 13821-13826 (2007).
17. Hu BY, Weick JB, Yu J *et al.* Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc. Natl. Acad. Sci. U. S. A.* 107(9), 4335-4340 (2010).
18. Tang X, Zhou L, Wagner AM *et al.* Astroglial cells regulate the developmental timeline of human neurons differentiated from induced pluripotent stem cells. *Stem Cell Res.* 11(2), 743-757 (2013).
19. Weick JB, Johnson MA, Skroch SP, Williams JC, Deisseroth K, Zhang SC. Functional control of transplantable human ESC-derived neurons via optogenetic targeting. *Stem Cells* 28(11), 2008-2016 (2010).
20. Yang N, Ng YH, Pang ZP, Sudhof TC, Wernig M. Induced neuronal cells: how to make and define a neuron. *Cell Stem Cell* 9(6), 517-525 (2011).

21. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Machine Learning* 63(1), 3-42 (2006).
22. Marr RA, Guan H, Rockenstein E *et al.* Neprilysin regulates amyloid Beta peptide levels. *J. Mol. Neurosci.* 22(1-2), 5-11 (2004).
23. Brennand KJ, Simone A, Jou J *et al.* Modelling schizophrenia using human induced pluripotent stem cells. *Nature* 473(7346), 221-225 (2011).
24. Dolmetsch R, Geschwind DH. The human brain in a dish: the promise of iPSC-derived neurons. *Cell* 145(6), 831-834 (2011).
25. Shcheglovitov A, Shcheglovitova O, Yazawa M *et al.* SHANK3 and IGF1 restore synaptic deficits in neurons from 22q13 deletion syndrome patients. *Nature* 503(7475), 267-271 (2013).
26. Wen Z, Nguyen HN, Guo Z *et al.* Synaptic dysregulation in a human iPSC cell model of mental disorders. *Nature* 515(7527), 414-418 (2014).
27. Lledo PM, Alonso M, Grubb MS. Adult neurogenesis and functional plasticity in neuronal circuits. *Nat. Rev. Neurosci.* 7(3), 179-193 (2006).
28. Trapnell C, Cacchiarelli D, Grimsby J *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 32(4), 381-386 (2014).
29. Tsetsenis T, Younts TJ, Chiu CQ, Kaeser PS, Castillo PE, Sudhof TC. Rab3B protein is required for long-term depression of hippocampal inhibitory synapses and for normal reversal learning. *Proc. Natl. Acad. Sci. U. S. A.* 108(34), 14300-14305 (2011).
30. Lewis S. Synaptic plasticity: a key player in presynaptic plasticity. *Nat. Rev. Neurosci.* 12(10), 548 (2011).
31. Pedrola L, Espert A, Wu X, Claramunt R, Shy ME, Palau F. GDAP1, the protein causing Charcot-Marie-Tooth disease type 4A, is expressed in neurons and is associated with mitochondria. *Hum. Mol. Genet.* 14(8), 1087-1094 (2005).
32. Lewis DA. The human brain revisited: opportunities and challenges in postmortem studies of psychiatric disorders. *Neuropsychopharmacology* 26(2), 143-154 (2002).
33. Moore AR, Filipovic R, Mo Z, Rasband MN, Zecevic N, Antic SD. Electrical excitability of early neurons in the human cerebral cortex during the second trimester of gestation. *Cereb. Cortex* 19(8), 1795-1805 (2009).
34. Moore AR, Zhou WL, Jakovcevski I, Zecevic N, Antic SD. Spontaneous electrical activity in the human fetal cortex in vitro. *J. Neurosci.* 31(7), 2391-2398 (2011).
35. Nicholas CR, Chen J, Tang Y *et al.* Functional maturation of hPSC-derived forebrain interneurons requires an extended timeline and mimics human neural development. *Cell Stem Cell* 12(5), 573-586 (2013).
36. Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Sudhof TC, Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463(7284), 1035-1041 (2010).
37. Eberwine J, Sul JY, Bartfai T, Kim J. The promise of single-cell sequencing. *Nat. methods* 11(1), 25-27 (2014).
38. Saliba AE, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 42(14), 8845-8860 (2014).
39. Sandberg R. Entering the era of single-cell transcriptomics in biology and medicine. *Nat. Methods* 11(1), 22-24 (2014).
40. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat. Rev. Genet.* 14(9), 618-630 (2013).
41. Tang F, Barbacioru C, Wang Y *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6(5), 377-382 (2009).
42. Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* 10(11), 1096-1098 (2013).

43. Wu AR, Neff NF, Kalisky T *et al.* Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* 11(1), 41-46 (2014).
44. Islam S, Kjallquist U, Moliner A *et al.* Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* 21(7), 1160-1167 (2011).
45. Islam S, Kjallquist U, Moliner A *et al.* Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat. Protoc.* 7(5), 813-828 (2012).
46. Tang F, Barbacioru C, Nordman E *et al.* RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat. Protoc.* 5(3), 516-535 (2010).
47. Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* 9(1), 171-181 (2014).
48. Patel AP, Tirosh I, Trombetta JJ *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344(6190), 1396-1401 (2014).
49. Toriello NM, Douglas ES, Thaitrong N *et al.* Integrated microfluidic bioprocessor for single-cell gene expression analysis. *Proc. Natl. Acad. Sci. U. S. A.* 105(51), 20173-20178 (2008).
50. Belinsky GS, Rich MT, Sirois CL *et al.* Patch-clamp recordings and calcium imaging followed by single-cell PCR reveal the developmental profile of 13 genes in iPSC-derived human neurons. *Stem Cell Res.* 12(1), 101-118 (2014).
51. Farago N, Kocsis AK, Lovas S *et al.* Digital PCR to determine the number of transcripts from single neurons after patch-clamp recording. *Biotechniques* 54(6), 327-336 (2013).
52. Toledo-Rodriguez M, Markram H. Single-cell RT-PCR, a technique to decipher the electrical, anatomical, and genetic determinants of neuronal diversity. *Methods Mol. Biol.* 403 123-139 (2007).
53. Toledo-Rodriguez M, Blumenfeld B, Wu C *et al.* Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. *Cereb. Cortex* 14(12), 1310-1327 (2004).
54. Subkhankulova T, Yano K, Robinson HP, Livesey FJ. Grouping and classifying electrophysiologically-defined classes of neocortical neurons by single cell, whole-genome expression profiling. *Front. Mol. Neurosci.* 3 10 (2010).
55. Fuzik J, Zeisel A, Mate Z *et al.* Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nat. Biotechnol.* 34(2), 175-183 (2016).
56. Cadwell CR, Palasantza A, Jiang X *et al.* Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. *Nat. Biotechnol.* 34(2), 199-203 (2016).
57. Steward O, Wallace CS, Lyford GL, Worley PF. Synaptic activation causes the mRNA for the IEG Arc to localize selectively near activated postsynaptic sites on dendrites. *Neuron* 21(4), 741-751 (1998).
58. Cajigas JJ, Tushev G, Will TJ, Tom Dieck S, Fuerst N, Schuman EM. The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* 74(3), 453-466 (2012).
59. Bassell GJ, Warren ST. Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* 60(2), 201-214 (2008).
60. Bagni C, Greenough WT. From mRNP trafficking to spine dysmorphogenesis: the roots of fragile X syndrome. *Nat. Rev. Neurosci.* 6(5), 376-387 (2005).
61. Wichterle H, Gifford D, Mazzoni E. Neuroscience. Mapping neuronal diversity one cell at a time. *Science* 341(6147), 726-727 (2013).
62. Markram H, Muller E, Ramaswamy S *et al.* Reconstruction and Simulation of Neocortical Microcircuitry. *Cell* 163(2), 456-492 (2015).
63. Defelipe J, Lopez-Cruz PL, Benavides-Piccione R *et al.* New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nat. Rev. Neurosci.* 14(3), 202-216 (2013).

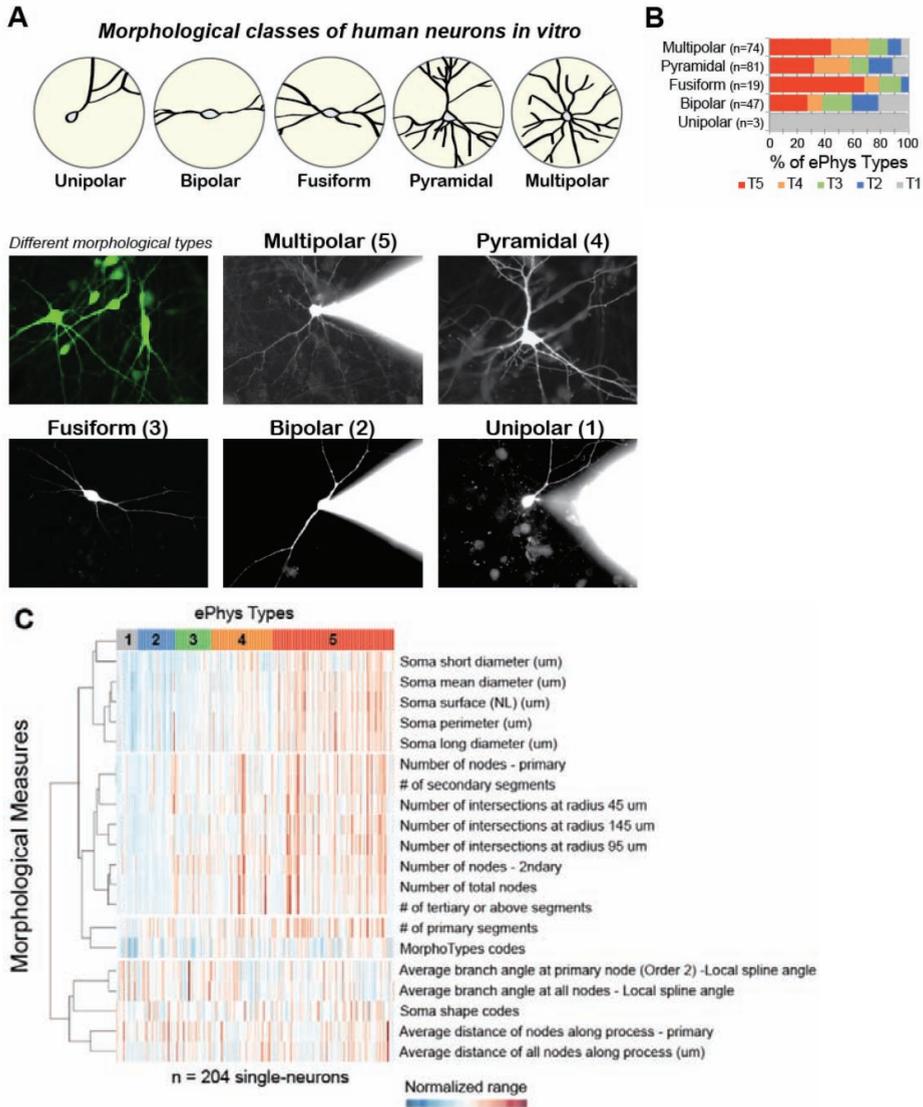
64. Gupta A, Wang Y, Markram H. Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex. *Science* 287(5451), 273-278 (2000).
65. Bardy C, Alonso M, Bouthour W, Lledo PM. How, when, and where new inhibitory neurons release neurotransmitters in the adult olfactory bulb. *J. Neurosci.* 30(50), 17023-17034 (2010).
66. Carleton A, Petreanu LT, Lansford R, Alvarez-Buylla A, Lledo PM. Becoming a new neuron in the adult olfactory bulb. *Nat. Neurosci.* 6(5), 507-518 (2003).
67. Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C. Interneurons of the neocortical inhibitory system. *Nat. Rev. Neurosci.* 5(10), 793-807 (2004).
68. Ascoli GA, Alonso-Nanclares L, Anderson SA *et al.* Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat. Rev. Neurosci.* 9(7), 557-568 (2008).
69. Druckmann S, Hill S, Schurmann F, Markram H, Segev I. A hierarchical structure of cortical interneuron electrical diversity revealed by automated statistical analysis. *Cereb. Cortex* 23(12), 2994-3006 (2013).
70. Tasic B, Menon V, Nguyen TN *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* 19(2), 335-346 (2016).
71. Hawrylycz M, Miller JA, Menon V *et al.* Canonical genetic signatures of the adult human brain. *Nat. Neurosci.* 18(12), 1832-1844 (2015).
72. Stein JL, De La Torre-Ubieta L, Tian Y *et al.* A quantitative framework to evaluate modeling of cortical development by neural stem cells. *Neuron* 83(1), 69-86 (2014).
73. McConnell MJ, Lindberg MR, Brennand KJ *et al.* Mosaic copy number variation in human neurons. *Science* 342(6158), 632-637 (2013).
74. Bellin M, Marchetto MC, Gage FH, Mummery CL. Induced pluripotent stem cells: the new patient? *Nat. Rev. Mol. Cell Biol.* 13(11), 713-726 (2012).
75. Marchetto MC, Brennand KJ, Boyer LF, Gage FH. Induced pluripotent stem cells (iPSCs) and neurological disease modeling: progress and promises. *Hum. Mol. Genet.* 20(R2), R109-115 (2011).
76. Pasca SP, Portmann T, Voineagu I *et al.* Using iPSC-derived neurons to uncover cellular phenotypes associated with Timothy syndrome. *Nat. Med.* 17(12), 1657-1662 (2011).
77. Yu DX, Marchetto MC, Gage FH. Therapeutic translation of iPSCs for treating neurological disease. *Cell Stem Cell* 12(6), 678-688 (2013).
78. Marchetto MC, Carromeu C, Acab A *et al.* A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* 143(4), 527-539 (2010).

SUPPLEMENTARY INFORMATION

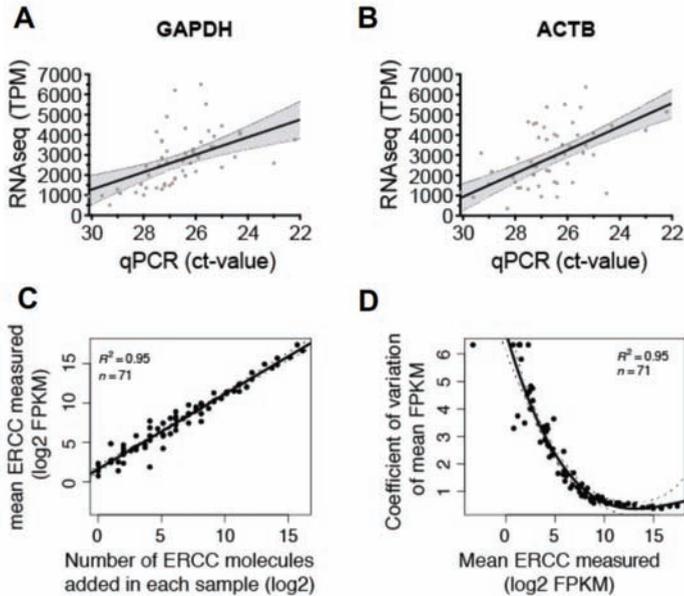


Supplementary Figure 1. (related to Figure 1). **Electrophysiological patch-clamping measurements of human neurons.** (A) Properties of evoked action potentials (APs) were measured in current clamp at an imposed resting potential around -70 mV. The best firing responses were chosen from a series of incremental depolarizing steps of currents (+2 pA, 500 ms). (B) Spontaneous APs were measured for 3 to 5 min in current clamp with 0 pA injected current. (C) The properties of single action potentials were analyzed from the first spike of the train of evoked APs as shown in A. (D) Voltage-dependent sodium (Nav) and potassium (Kv) currents were measured in voltage clamp at an imposed resting potential around -70 mV with a series of incremental depolarizing steps of voltage (+5 mV, 50 ms). (E,F) Spontaneous synaptic activity was measured in voltage clamp for 5 min or more. AMPA events were measured at -70 mV, close to the anions reversal potential. GABA events were measured at 0 mV, close to cations reversal potential. (G) Membrane properties were obtained from averaging the voltage clamp responses ($n = 15$) to a test pulse (-5 mV) at -70 mV.

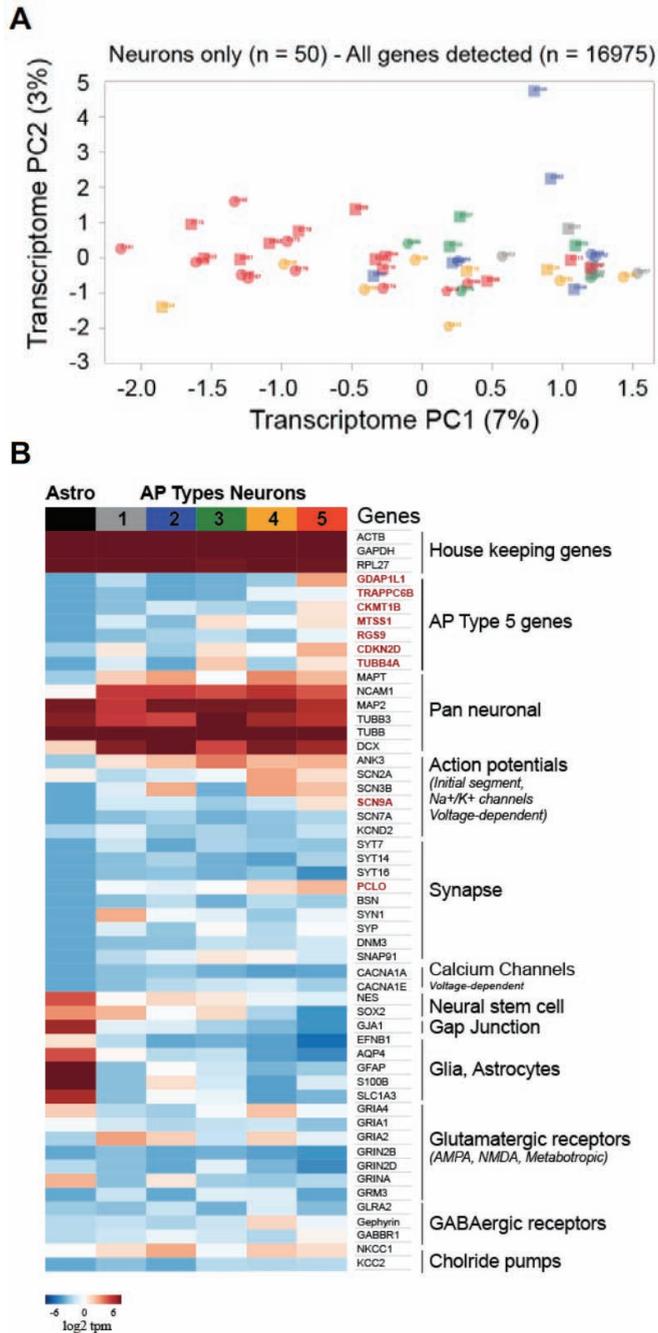
(C) We defined 5 electrophysiological types of neurons (AP Types) based on the combination of two key ePhys properties – the maximal peak and frequencies of evoked action potentials (APs). (D) The heat map represents the normalized values of electrophysiological properties (*y* axis) for each single human neuron patched (*x* axis). Euclidean hierarchical clustering was applied to the electrophysiological properties. The single neurons were ordered by ePhys type. Mann Whitney *P*-values were measured between Type 5 neurons and all the other types and corrected for multiple hypothesis testing (Bonferroni).



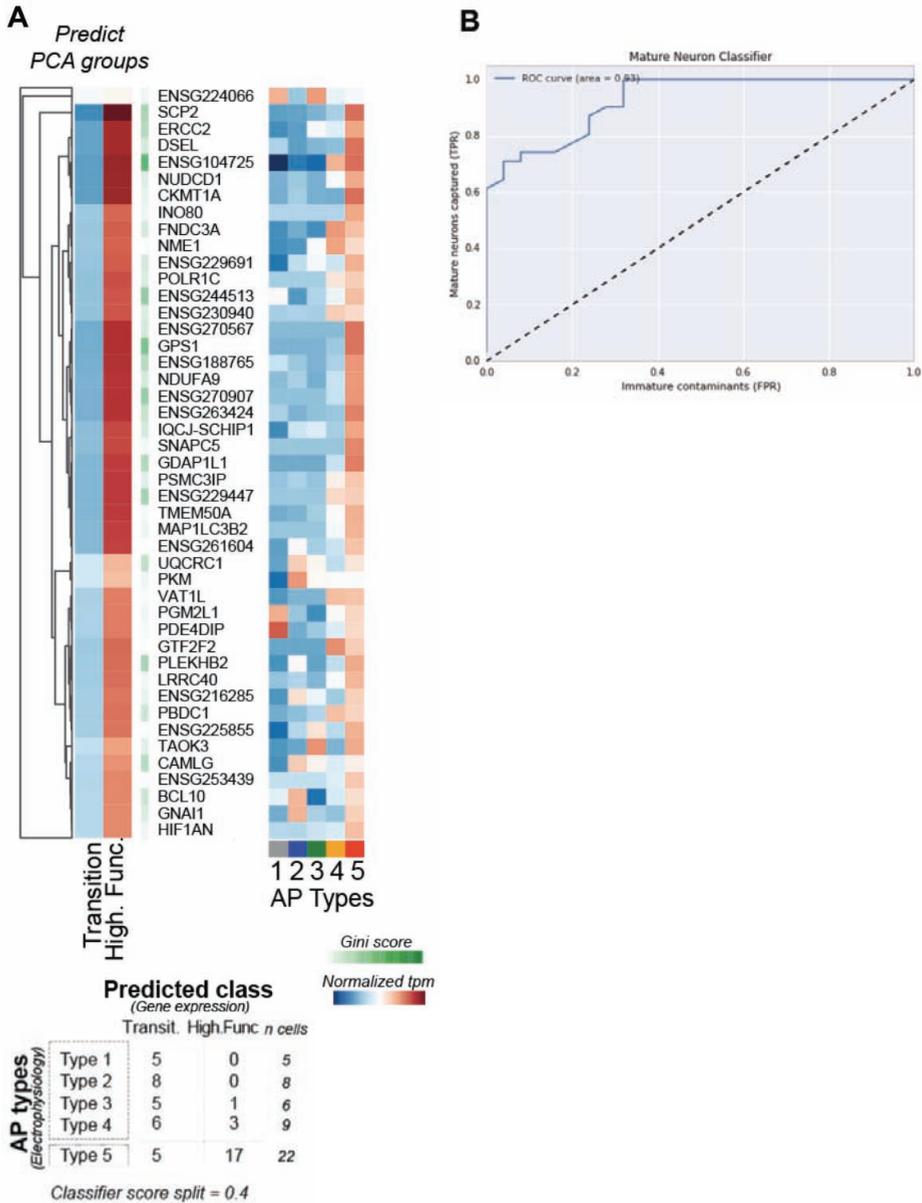
Supplementary Figure 3. (related to Figure 3). **Morphological measurements.** (A) Typical morphological classes of human neurons *in vitro*. Original pictures of morphological types used for the drawings. (B) Proportion of ePhys types found in each morphological category. Only neurons unambiguously classified were used in this sample ($n = 224$). We did not find any correspondence among specific morphological classes and the ePhys types. (C) The heat map represents the normalized values of morphological properties (y axis) for each patched single neuron (x axis). Euclidean hierarchical clustering was applied to the morphological properties. The single neurons were ordered by ePhys type.



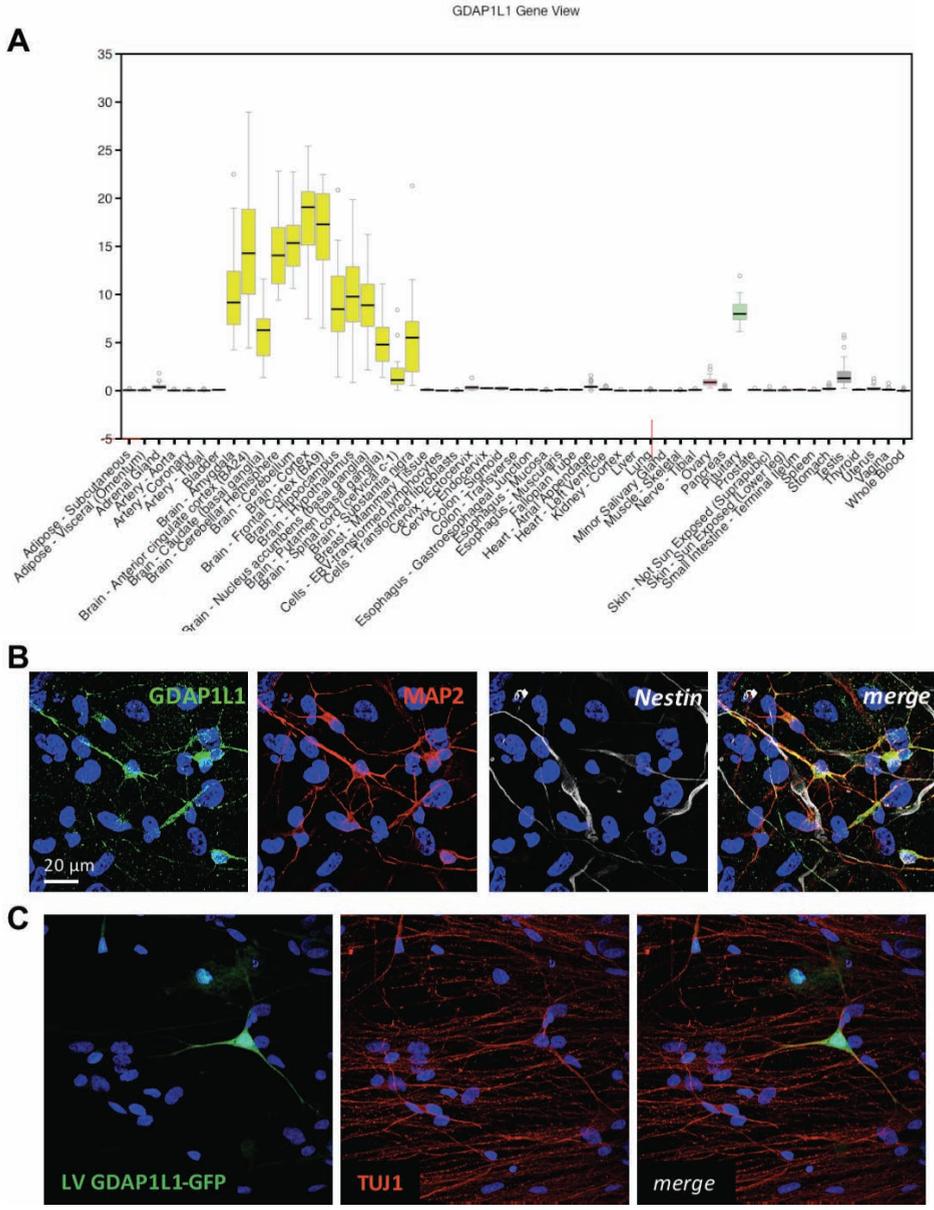
Supplementary Figure 4. (related to Figure 4). **Single-cell RNA-seq data of Patch-seq neurons.** (A,B) We found a significant correlation between single-cell gene expression measured by RNA-seq and qPCR. The solid lines represent the regression model fit to data with 95% confidence intervals. (C) Each cell sample was mixed with artificial transcripts of known abundance, which linearly correlated with the transcript levels measured with sequencing. 71/73 ERCC transcripts were detected across all cells ($n = 40$ cells). Each point represents the measure of each ERCC transcript averaged for all cells. (D) The variance of the measured transcript levels was the highest for transcripts of particularly low abundance.



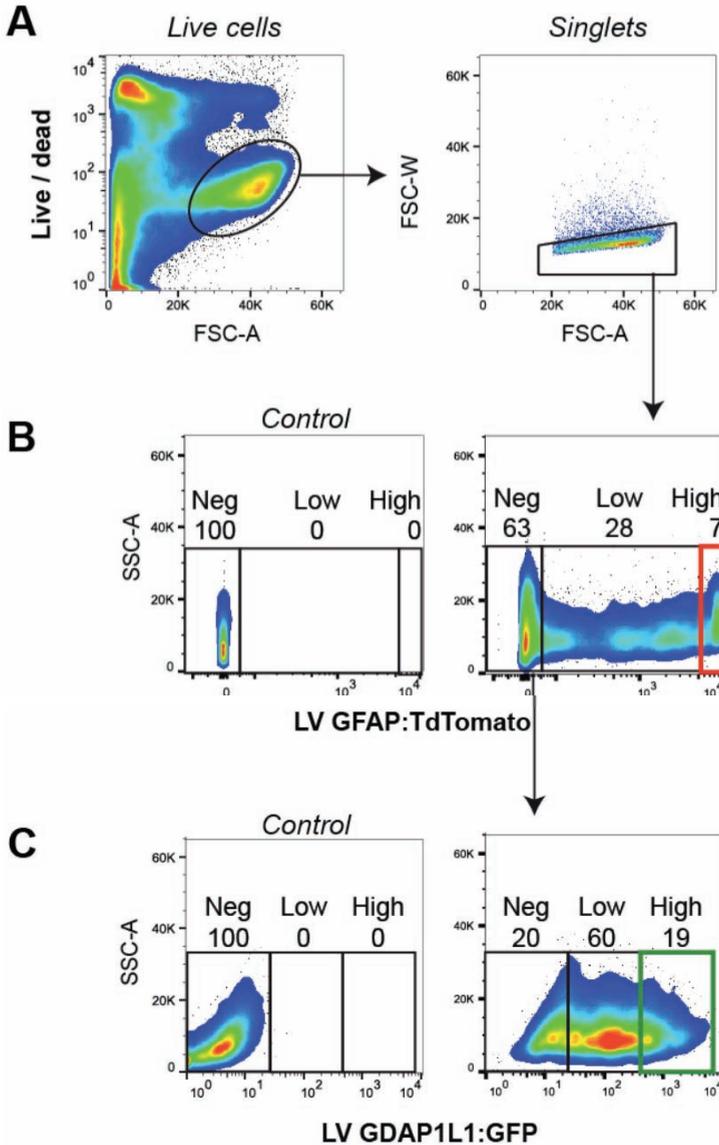
Supplementary Figure 5. (related to Figure 5). **Single-cell transcriptomes segregate functional states of differentiated neurons.** (A) PCA only on differentiated neurons (without astrocytes) shows similar cluster of highly functional neuronal states observed in PCA of Figure 5B. (B) The heat map complements Figure 5D by averaging the gene expression of single cells by neuronal AP Types or astrocytes for clarity.



Supplementary Figure 6. (related to Figure 6). **Extremely randomized tree classifier.** (A) The heatmap complements Figure 6D by showing the normalized gene expression levels averaged by AP Types. (B) Receiver Operating Characteristic (ROC) curve of the Mature/Functional vs. Immature/Non-functional neuron classifier shows trade-offs between True Positive Rate (TPR = Sensitivity, y axis) and False Positive Rate (FPR = 1-Specificity, x axis), with a good cross-validated performance: Area Under the Curve (AUC) = 0.93.



Supplementary Figure 7. (related to Figure 7). **GDAP1L1 exclusive to neural tissue *in vivo* and additional staining *in vitro*.** (A) GDAP1L1 expression is exclusive to neural tissue in human. Data source: GTEx Analysis Release V4 (dbGaP Accession phs000424.v4.p1). (B) Immunostainings of fixed human neuronal cultures confirm the translation of GDAP1L1 at the protein level in some neurons (MAP2+) but not in less mature neural progenitors (Nestin+). (C) Example of a neuron expressing high levels of GDAP1L1:GFP (lentivector) and immunostained with TUJ1.



Supplementary Figure 8. (related to Figure 7). **Live single-cell FACS sorting to select subpopulation for replating.** Live human neuronal cultures (5 weeks in BrainPhys + supplements) were infected with lentivectors GDAP1L1:GFP and GFAP:tdTomato and stained with DAPI before dissociation and FACS sorting. **(A)** Live cells were distinguished from dead cells and debris based on DAPI staining (y axis, lei graph) and size (FSC-A, FSC-W). **(B)** Glia and astrocytes were identified with the expression of tdTomato under the control of GFAP promoter. The control was a sample of cells infected with LV GDAP1L1:GFP only to account for the leak of GFP in the red channel. **(C)** The expression of GDAP1L1:GFP was measured in the cells completely negative for GFAP:tdTomato. The GFAP-negative cells expressing high levels of GFP were isolated and re-plated for patch-clamping analyses.

Supplementary Table 1. List of genes that appeared in the top lists of independent analyses for the paper (regression, machine learning, differential expression). Among the most highly ranked genes, 29 genes appeared in two or more analyses (highlighted in lightgray). Only one gene appeared in all of these independent analyses: GDAP1L1 (highlighted in darkgray). The descriptions for each gene were obtained from <https://david.ncifcrf.gov> and <http://string-db.org>.

Top genes	DAVID_GO functions	STRING_GO functions
ACOT7	acyl-CoA thioesterase 7	acyl-CoA thioesterase 7 (380 aa)
AFMID	arylformamidase	arylformamidase; Catalyzes the hydrolysis of N-formyl-L-kynurenine to L-kynurenine, the second step in the kynurenine pathway of tryptophan degradation. Kynurenine may be further oxidized to nicotinic acid, NAD(H) and NADP(H). Required for elimination of toxic metabolites (By similarity) (308 aa)
AGXT2L2	alanine-glyoxylate aminotransferase 2-like 2	alanine-glyoxylate aminotransferase 2-like 2; Catalyzes the pyridoxalphosphate-dependent breakdown of 5-phosphohydroxy-L-lysine, converting it to ammonia, inorganic phosphate and 2-aminoadipate semialdehyde (450 aa)
AHI1	Abelson helper integration site 1	Abelson helper integration site 1; Component of the tectonic-like complex, a complex localized at the transition zone of primary cilia and acting as a barrier that prevents diffusion of transmembrane proteins between the cilia and plasma membranes (By similarity) (1196 aa)
AIMP3		HCG2043275; Uncharacterized protein (151 aa)
AKR1C1	aldo-keto reductase family 1, member C1 (dihydrodiol dehydrogenase 1; 20-alpha (3-alpha)-hydroxysteroid dehydrogenase)	aldo-keto reductase family 1, member C1 (dihydrodiol dehydrogenase 1; 20-alpha (3-alpha)-hydroxysteroid dehydrogenase) (323 aa)
AKR1C2	aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III)	aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III) (323 aa)
ALDH1B1	aldehyde dehydrogenase 1 family, member B1	aldehyde dehydrogenase 1 family, member B1; ALDHs play a major role in the detoxification of alcohol-derived acetaldehyde. They are involved in the metabolism of corticosteroids, biogenic amines, neurotransmitters, and lipid peroxidation (517 aa)
ANKRD13B	ankyrin repeat domain 13B	ankyrin repeat domain 13B (626 aa)
ANKRD36C		ankyrin repeat domain 36C (1778 aa)
ANKRD39	ankyrin repeat domain 39	ankyrin repeat domain 39 (183 aa)
APOO	apolipoprotein O	apolipoprotein O; Promotes cholesterol efflux from macrophage cells. Detected in HDL, LDL and VLDL. Secreted by a microsomal triglyceride transfer protein (MTTP)-dependent mechanism, probably as a VLDL-associated protein that is subsequently transferred to HDL. May be involved in myocardium-protective mechanisms against lipid accumulation (198 aa)
ARHGEF3	Rho guanine nucleotide exchange factor (GEF) 3	Rho guanine nucleotide exchange factor (GEF) 3 (558 aa)
ATP6V1C2	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C2	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C2; Subunit of the peripheral V1 complex of vacuolar ATPase. Subunit C is necessary for the assembly of the catalytic sector of the enzyme and is likely to have a specific function in its catalytic activity. V-ATPase is responsible for acidifying a variety of intracellular compartments in eukaryotic cells (427 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
BCL10	B-cell CLL/lymphoma 10; hypothetical LOC646626	B-cell CLL/lymphoma 10; Promotes apoptosis, pro-caspase-9 maturation and activation of NF-kappa-B via NIK and IKK. May be an adapter protein between upstream TNFR1-TRADD-RIP complex and the downstream NIK-IKK-IKAP complex. Is a substrate for MALT1 (233 aa)
BEX4	brain expressed, X-linked 4	brain expressed, X-linked 4 (120 aa)
BOLA2	bolA homolog 2 (<i>E. coli</i>); bolA homolog 2B (<i>E. coli</i>)	bolA homolog 2 (<i>E. coli</i>) (152 aa)
BOLA2B	bolA homolog 2 (<i>E. coli</i>); bolA homolog 2B (<i>E. coli</i>)	bolA homolog 2B (<i>E. coli</i>) (152 aa)
CAMLG	calcium modulating ligand	calcium modulating ligand; Likely involved in the mobilization of calcium as a result of the TCR/CD3 complex interaction. Binds to cyclophilin B (296 aa)
CCDC137	coiled-coil domain containing 137	coiled-coil domain containing 137 (289 aa)
CKDN2D	cyclin-dependent kinase inhibitor 2D (p19, inhibits CDK4)	cyclin-dependent kinase inhibitor 2D (p19, inhibits CDK4); Interacts strongly with CDK4 and CDK6 and inhibits them (166 aa)
CHCHD6	coiled-coil-helix-coiled-coil-helix domain containing 6	coiled-coil-helix-coiled-coil-helix domain containing 6; Required for maintaining mitochondrial crista morphology, ATP production and oxygen consumption (235 aa)
CKMT1A	creatine kinase, mitochondrial 1A; creatine kinase, mitochondrial 1B	creatine kinase, mitochondrial 1A; Reversibly catalyzes the transfer of phosphate between ATP and various phosphogens (e.g. creatine phosphate). Creatine kinase isoenzymes play a central role in energy transduction in tissues with large, fluctuating energy demands, such as skeletal muscle, heart, brain and spermatozoa (417 aa)
CKMT1B	creatine kinase, mitochondrial 1A; creatine kinase, mitochondrial 1B	creatine kinase, mitochondrial 1B (417 aa)
CNOT4	CCR4-NOT transcription complex, subunit 4	CCR4-NOT transcription complex, subunit 4 (713 aa)
COMM1D1	copper metabolism (Murr1) domain containing 1	copper metabolism (Murr1) domain containing 1; Promotes ubiquitination of NF-kappa-B subunit RELA and its subsequent proteasomal degradation. Down-regulates NF-kappa-B activity. Down-regulates SOD1 activity by interfering with its homodimerization. Plays a role in copper ion homeostasis. Can bind one copper ion per monomer. May function to facilitate biliary copper excretion within hepatocytes (190 aa)
DEDD	death effector domain containing	death effector domain containing; A scaffold protein that directs CASP3 to certain substrates and facilitates their ordered degradation during apoptosis. May also play a role in mediating CASP3 cleavage of KRT18. Regulates degradation of intermediate filaments during apoptosis. May play a role in the general transcription machinery in the nucleus and might be an important regulator of the activity of GTF3C3. Inhibits DNA transcription <i>in vitro</i> (By similarity) (318 aa)
DSEL	dermatan sulfate epimerase-like	dermatan sulfate epimerase-like (1222 aa)
DYNC112	similar to dynein cytoplasmic 1 intermediate chain 2; dynein, cytoplasmic 1, intermediate chain 2	dynein, cytoplasmic 1, intermediate chain 2 (638 aa)
ECH1	enoyl Coenzyme A hydratase 1, peroxisomal	enoyl CoA hydratase 1, peroxisomal; Isomerization of 3-trans,5-cisdienoyl-CoA to 2-trans,4- trans-dienoyl-CoA (By similarity) (328 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
EDIL3	EGF-like repeats and discoidin I-like domains 3	EGF-like repeats and discoidin I-like domains 3; Promotes adhesion of endothelial cells through interaction with the alpha-v/beta-3 integrin receptor. Inhibits formation of vascular-like structures. May be involved in regulation of vascular morphogenesis of remodeling in embryonic development (480 aa)
EEF1A1	eukaryotic translation elongation factor 1 alpha-like 7; eukaryotic translation elongation factor 1 alpha-like 3; similar to eukaryotic translation elongation factor 1 alpha 1; eukaryotic translation elongation factor 1 alpha 1	eukaryotic translation elongation factor 1 alpha 1; This protein promotes the GTP-dependent binding of aminoacyl-tRNA to the A site of ribosomes during protein biosynthesis. With PARP1 and TXK, forms a complex that acts as a T helper 1 (Th1) cell-specific transcription factor and binds the promoter of IFN-gamma to directly regulate its transcription, and is thus involved importantly in Th1 cytokine production (462 aa)
EEF1A1P24	eukaryotic translation elongation factor 1 alpha-like 7; eukaryotic translation elongation factor 1 alpha-like 3; similar to eukaryotic translation elongation factor 1 alpha 1; eukaryotic translation elongation factor 1 alpha 1	
EEF1A1P5	eukaryotic translation elongation factor 1 alpha-like 7; eukaryotic translation elongation factor 1 alpha-like 3; similar to eukaryotic translation elongation factor 1 alpha 1; eukaryotic translation elongation factor 1 alpha 1	
EEF1A1P9	eukaryotic translation elongation factor 1 alpha-like 7; eukaryotic translation elongation factor 1 alpha-like 3; similar to eukaryotic translation elongation factor 1 alpha 1; eukaryotic translation elongation factor 1 alpha 1	
EIF2S1	eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa	eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa; Functions in the early steps of protein synthesis by forming a ternary complex with GTP and initiator tRNA. This complex binds to a 40S ribosomal subunit, followed by mRNA binding to form a 43S preinitiation complex. Junction of the 60S ribosomal subunit to form the 80S initiation complex is preceded by hydrolysis of the GTP bound to eIF-2 and release of an eIF-2-GDP binary complex. In order for eIF-2 to recycle and catalyze another round of initiation, the GDP bound to eIF-2 must exchange with GTP by way of a reaction [...] (315 aa)
EMC4		ER membrane protein complex subunit 4; May mediate antiapoptotic activity (183 aa)
ENSG00000268220		cDNA FLJ37906 fis, clone COLON2004318 (138 aa)
EPB41L1	erythrocyte membrane protein band 4.1-like 1	erythrocyte membrane protein band 4.1-like 1 (881 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
ERCC2	excision repair cross-complementing rodent repair deficiency, complementation group 2	excision repair cross-complementing rodent repair deficiency, complementation group 2; ATP-dependent 5'-3' DNA helicase, component of the core- TFIIH basal transcription factor. Involved in nucleotide excision repair (NER) of DNA by opening DNA around the damage, and in RNA transcription by RNA polymerase II by anchoring the CDK-activating kinase (CAK) complex, composed of CDK7, cyclin H and MAT1, to the core-TFIIH complex. Involved in the regulation of vitamin-D receptor activity. As part of the mitotic spindle-associated MMXD complex it plays a role in chromosome segregation. Might h [...] (760 aa)
FBLN1	fibulin 1	fibulin 1 (703 aa)
FGF14	fibroblast growth factor 14	fibroblast growth factor 14; Probably involved in nervous system development and function (252 aa)
FNDC3A	fibronectin type III domain containing 3A	fibronectin type III domain containing 3A (1198 aa)
GATC	glutamyl-tRNA(Gln) amidotransferase, subunit C homolog (bacterial)	glutamyl-tRNA(Gln) amidotransferase, subunit C; Allows the formation of correctly charged Gln-tRNA(Gln) through the transamidation of misacylated Glu-tRNA(Gln) in the mitochondria. The reaction takes place in the presence of glutamine and ATP through an activated gamma-phospho-Glu- tRNA(Gln) (136 aa)
GDAP1L1	ganglioside-induced differentiation-associated protein 1-like 1	ganglioside induced differentiation associated protein 1-like 1 (367 aa)
GNAI1	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1 (354 aa)
GNAL	guanine nucleotide binding protein (G protein), alpha activating activity polypeptide, olfactory type	guanine nucleotide binding protein (G protein), alpha activating activity polypeptide, olfactory type; Guanine nucleotide-binding proteins (G proteins) are involved as modulators or transducers in various transmembrane signaling systems. G(olf) alpha mediates signal transduction within the olfactory neuroepithelium and the basal ganglia. May be involved in some aspect of visual transduction, and in mediating the effect of one or more hormones/neurotransmitters (458 aa)
GNG4	guanine nucleotide binding protein (G protein), gamma 4	guanine nucleotide binding protein (G protein), gamma 4; Guanine nucleotide-binding proteins (G proteins) are involved as a modulator or transducer in various transmembrane signaling systems. The beta and gamma chains are required for the GTPase activity, for replacement of GDP by GTP, and for G protein- effector interaction (By similarity) (75 aa)
GPS1	G protein pathway suppressor 1	G protein pathway suppressor 1; Essential component of the COP9 signalosome complex (CSN), a complex involved in various cellular and developmental processes. The CSN complex is an essential regulator of the ubiquitin (Ubl) conjugation pathway by mediating the deneddylation of the cullin subunits of SCF-type E3 ligase complexes, leading to decrease the Ubl ligase activity of SCF-type complexes such as SCF, CSA or DDB2. The complex is also involved in phosphorylation of p53/TP53, c-jun/JUN, IkkappaBalpha/NFKBIA, ITPK1 and IRF8/ICSBP, possibly via its association with CK2 and PKD kinases. [...] (527 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
GTF2F	general transcription factor IIF, polypeptide 2, 30kDa	general transcription factor IIF, polypeptide 2, 30kDa; TFIIF is a general transcription initiation factor that binds to RNA polymerase II and helps to recruit it to the initiation complex in collaboration with TFIIB. It promotes transcription elongation. This subunit shows ATP-dependent DNA- helicase activity (249 aa)
HGS	hepatocyte growth factor-regulated tyrosine kinase substrate	hepatocyte growth factor-regulated tyrosine kinase substrate; Involved in intracellular signal transduction mediated by cytokines and growth factors. When associated with STAM, it suppresses DNA signaling upon stimulation by IL-2 and GM-CSF. Could be a direct effector of PI3-kinase in vesicular pathway via early endosomes and may regulate trafficking to early and late endosomes by recruiting clathrin. May concentrate ubiquitinated receptors within clathrincoated regions. Involved in down- regulation of receptor tyrosine kinase via multivesicular body (MVBs) when complexed with STAM [E [...] (777 aa)
HIF1AN	hypoxia inducible factor 1, alpha subunit inhibitor	hypoxia inducible factor 1, alpha subunit inhibitor; Hydroxylates HIF-1 alpha at 'Asp-803' in the C-terminal transactivation domain (CAD). Functions as an oxygen sensor and, under normoxic conditions, the hydroxylation prevents interaction of HIF-1 with transcriptional co-activators including Cbp/p300- interacting transactivator. Involved in transcriptional repression through interaction with HIF 1A, VHL and histone deacetylases. Hydroxylates specific Asn residues within ankyrin repeat domains (ARD) of NFKB1, NFKBIA, NOTCH1, ASB4, PPP1R12A and several other ARD-containing proteins. Also [...] (349 aa)
HNRNPA0	heterogeneous nuclear ribonucleoprotein A0	heterogeneous nuclear ribonucleoprotein A0; mRNA-binding component of ribonucleosomes. Specifically binds AU-rich element (ARE)-containing mRNAs. Involved in post- transcriptional regulation of cytokines mRNAs (305 aa)
IAH1	isoamyl acetate-hydrolyzing esterase 1 homolog (<i>S. cerevisiae</i>)	isoamyl acetate-hydrolyzing esterase 1 homolog (<i>S. cerevisiae</i>); Probable lipase (By similarity) (248 aa)
IFI27L2	interferon, alpha-inducible protein 27-like 2	interferon, alpha-inducible protein 27-like 2 (130 aa)
IL6ST	interleukin 6 signal transducer (gp130, oncostatin M receptor)	interleukin 6 signal transducer (gp130, oncostatin M receptor); Signal-transducing molecule. The receptor systems for IL6, LIF, OSM, CNTF, IL11, CTF1 and BSF3 can utilize gp130 for initiating signal transmission. Binds to IL6/IL6R (alpha chain) complex, resulting in the formation of high-affinity IL6 binding sites, and transduces the signal. Does not bind IL6. May have a role in embryonic development (By similarity). The type I OSM receptor is capable of transducing OSM-specific signaling events (918 aa)
INO80	INO80 homolog (<i>S. cerevisiae</i>)	INO80 homolog (<i>S. cerevisiae</i>); DNA helicase and probable main scaffold component of the chromatin remodeling INO80 complex which is involved in transcriptional regulation, DNA replication and probably DNA repair; according to PubMed-20687897 the contribution to DNA double-strand break repair appears to be largely indirect through transcriptional regulation. Recruited by YY1 to YY1-activated genes, where it acts as an essential co-activator. Binds DNA. <i>In vitro</i> , has double stranded DNA-dependent ATPase activity. Involved in UV-damage excision repair, DNA replication and chromosome segreg [...] (1556 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
INO80D	INO80 complex subunit D	INO80 complex subunit D; Putative regulatory component of the chromatin remodeling INO80 complex which is involved in transcriptional regulation, DNA replication and probably DNA repair (1027 aa)
IRGQ	immunity-related GTPase family, Q	immunity-related GTPase family, Q (623 aa)
KIAA1456		KIAA1456 (454 aa)
KIF3C	kinesin family member 3C	kinesin family member 3C; Microtubule-based anterograde translocator for membranous organelles (By similarity) (793 aa)
LOC100134230	hypothetical protein LOC100134230; similar to KIAA0454 protein; similar to phosphodiesterase 4D interacting protein isoform 2; phosphodiesterase 4D interacting protein	
LOC646626	B-cell CLL/lymphoma 10; hypothetical LOC646626	
LOC652164	hypothetical protein LOC100134230; similar to KIAA0454 protein; similar to phosphodiesterase 4D interacting protein isoform 2; phosphodiesterase 4D interacting protein	
LOC727927	hypothetical protein LOC100134230; similar to KIAA0454 protein; similar to phosphodiesterase 4D interacting protein isoform 2; phosphodiesterase 4D interacting protein	
LOC728532	similar to dynein cytoplasmic 1 intermediate chain 2; dynein, cytoplasmic 1, intermediate chain 2	
LPPR2		Lipid phosphate phosphatase-related protein type 2 (343 aa)
LRRC40	leucine rich repeat containing 40	leucine rich repeat containing 40 (602 aa)
LRRD1		leucine-rich repeats and death domain containing 1 (860 aa)
LSM7	LSM7 homolog, U6 small nuclear RNA associated (<i>S. cerevisiae</i>)	LSM7 homolog, U6 small nuclear RNA associated (<i>S. cerevisiae</i>); Binds specifically to the 3'-terminal U-tract of U6 snRNA and is probably a component of the spliceosome (103 aa)
MAP1LC3B2	microtubule-associated protein 1 light chain 3 beta 2	microtubule-associated protein 1 light chain 3 beta 2; Probably involved in formation of autophagosomal vacuoles (autophagosomes) (By similarity) (125 aa)
MDGA2	MAM domain containing glycosylphosphatidylinositol anchor 2	MAM domain containing glycosylphosphatidylinositol anchor 2; May be involved in cell-cell interactions (By similarity) (1025 aa)
METAP1	methionyl aminopeptidase 1	methionyl aminopeptidase 1; Removes the N-terminal methionine from nascent proteins. Required for normal progression through the cell cycle (386 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
MT-CO1		mitochondrially encoded cytochrome c oxidase I; Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the bimetallic center formed by heme A3 and copper B (By similarity) (513 aa)
MTSS1	metastasis suppressor 1	metastasis suppressor 1; May be related to cancer progression or tumor metastasis in a variety of organ sites, most likely through an interaction with the actin cytoskeleton (755 aa)
ND4	NADH-ubiquinone oxidoreductase chain 4L; NADH-ubiquinone oxidoreductase chain 4	
ND4L	NADH-ubiquinone oxidoreductase chain 4L; NADH-ubiquinone oxidoreductase chain 4	
NDUFA9	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 9, 39kDa	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 9, 39kDa; Accessory subunit of the mitochondrial membrane respiratory chain NADH dehydrogenase (Complex I), that is believed not to be involved in catalysis. Complex I functions in the transfer of electrons from NADH to the respiratory chain. The immediate electron acceptor for the enzyme is believed to be ubiquinone (377 aa)
NDUFAF2	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, assembly factor 2	NADH dehydrogenase (ubiquinone) complex I, assembly factor 2; Acts as a molecular chaperone for mitochondrial complex I assembly (169 aa)
NECAB1	N-terminal EF-hand calcium binding protein 1	N-terminal EF-hand calcium binding protein 1 (351 aa)
NEFL	neurofilament, light polypeptide	
NELL2	NEL-like 2 (chicken)	NEL-like 2 (chicken) (866 aa)
NF2	neurofibromin 2 (merlin)	neurofibromin 2 (merlin) (595 aa)
NKIRAS2	NFKB inhibitor interacting Ras-like 2	NFKB inhibitor interacting Ras-like 2; Atypical Ras-like protein that acts as a potent regulator of NF-kappa-B activity by preventing the degradation of NF-kappa-B inhibitor beta (NFKBIB) by most signals, explaining why NFKBIB is more resistant to degradation. May act by blocking phosphorylation of NFKBIB and nuclear localization of p65/RELA NF- kappa-B subunit. It is unclear whether it acts as a GTPase. Both GTP- and GDP-bound forms block phosphorylation of NFKBIB (By similarity) (191 aa)
NME1	non-metastatic cells 1, protein (NM23A) expressed in; NME1-NME2 readthrough transcript; non-metastatic cells 2, protein (NM23B) expressed in	NME/NM23 nucleoside diphosphate kinase 1; Major role in the synthesis of nucleoside triphosphates other than ATP. Possesses nucleoside-diphosphate kinase, serine/threonine-specific protein kinase, geranyl and farnesyl pyrophosphate kinase, histidine protein kinase and 3'-5' exonuclease activities. Involved in cell proliferation, differentiation and development, signal transduction, G protein-coupled receptor endocytosis, and gene expression. Required for neural development including neural patterning and cell fate determination (177 aa)
NME1-NME2	non-metastatic cells 1, protein (NM23A) expressed in; NME1-NME2 readthrough transcript; non-metastatic cells 2, protein (NM23B) expressed in	

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
NME2	non-metastatic cells 1, protein (NM23A) expressed in; NME1-NME2 readthrough transcript; non-metastatic cells 2, protein (NM23B) expressed in	
NUDCD1	NudC domain containing 1	NudC domain containing 1 (583 aa)
PCLO	piccolo (presynaptic cytomatrix protein)	piccolo (presynaptic cytomatrix protein) (5142 aa)
PDE4DIP	hypothetical protein LOC100134230; similar to KIAA0454 protein; similar to phosphodiesterase 4D interacting protein isoform 2; phosphodiesterase 4D interacting protein	phosphodiesterase 4D interacting protein (2362 aa)
PFDN2	prefoldin subunit 2	prefoldin subunit 2; Binds specifically to cytosolic chaperonin (c-CPN) and transfers target proteins to it. Binds to nascent polypeptide chain and promotes folding in an environment in which there are many competing pathways for nonnative proteins (154 aa)
PGM2L1	phosphoglucomutase 2-like 1	phosphoglucomutase 2-like 1; Glucose 1,6-bisphosphate synthase using 1,3- bisphosphoglycerate as a phosphate donor and a series of 1-phosphate sugars as acceptors, including glucose 1-phosphate, mannose 1-phosphate, ribose 1-phosphate and deoxyribose 1-phosphate. 5 or 6-phosphosugars are bad substrates, with the exception of glucose 6-phosphate. Also synthesizes ribose 1,5-bisphosphate. Has only low phosphopentomutase and phosphoglucomutase activities (622 aa)
PID1	phosphotyrosine interaction domain containing 1	phosphotyrosine interaction domain containing 1; Increases proliferation of preadipocytes without affecting adipocytic differentiation (248 aa)
PITRM1	pitrilysin metallopeptidase 1	pitrilysin metallopeptidase 1 (1038 aa)
PKM		pyruvate kinase, muscle (531 aa)
PLEKHB2	pleckstrin homology domain containing, family B (evectins) member 2	pleckstrin homology domain containing, family B (evectins) member 2; Involved in retrograde transport of recycling endosomes (222 aa)
POLG2	polymerase (DNA directed), gamma 2, accessory subunit	polymerase (DNA directed), gamma 2, accessory subunit; Mitochondrial polymerase processivity subunit. Stimulates the polymerase and exonuclease activities, and increases the processivity of the enzyme. Binds to ss-DNA (485 aa)
POLR1C	polymerase (RNA) I polypeptide C, 30kDa	polymerase (RNA) I polypeptide C, 30kDa; DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates. Common component of RNA polymerases I and III which synthesize ribosomal RNA precursors and small RNAs, such as 5S rRNA and tRNAs, respectively. RPAC1 is part of the Pol core element with the central large cleft and probably a clamp element that moves to open and close the cleft (By similarity) (346 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
POLR21	polymerase (RNA) II (DNA directed) polypeptide I, 14.5kDa	polymerase (RNA) II (DNA directed) polypeptide I, 14.5kDa; DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates. Component of RNA polymerase II which synthesizes mRNA precursors and many functional non-coding RNAs. Pol II is the central component of the basal RNA polymerase II transcription machinery. It is composed of mobile elements that move relative to each other. RPB9 is part of the upper jaw surrounding the central large cleft and thought to grab the incoming DNA template (By similarity) (125 aa)
PPFIA2	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 2	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 2; Alters PTPRF cellular localization and induces PTPRF clustering. May regulate the disassembly of focal adhesions. May localize receptor-like tyrosine phosphatases type 2A at specific sites on the plasma membrane, possibly regulating their interaction with the extracellular environment and their association with substrates (1257 aa)
PPM1J	protein phosphatase 1J (PP2C domain containing)	protein phosphatase, Mg ²⁺ /Mn ²⁺ dependent, 1J (505 aa)
PSMC3IP	PSMC3 interacting protein	PSMC3 interacting protein; Plays an important role in meiotic recombination. Stimulates DMC1-mediated strand exchange required for pairing homologous chromosomes during meiosis. The complex PSMC3IP/MND1 binds DNA, stimulates the recombinase activity of DMC1 as well as DMC1 D-loop formation from double-strand DNA. This complex stabilizes presynaptic RAD51 and DMC1 filaments formed on single strand DNA to capture double-strand DNA. This complex stimulates both synaptic and presynaptic critical steps in RAD51 and DMC1-promoted homologous pairing. May inhibit HIV-1 viral protein TAT activ [...] (217 aa)
PSMD12	proteasome (prosome, macropain) 26S subunit, non-ATPase, 12	proteasome (prosome, macropain) 26S subunit, non-ATPase, 12; Acts as a regulatory subunit of the 26S proteasome which is involved in the ATP-dependent degradation of ubiquitinated proteins (456 aa)
RAB12	RAB12, member RAS oncogene family	RAB12, member RAS oncogene family; The small GTPases Rab are key regulators of intracellular membrane trafficking, from the formation of transport vesicles to their fusion with membranes. Rabs cycle between an inactive GDP-bound form and an active GTP-bound form that is able to recruit to membranes different set of downstream effectors directly responsible for vesicle formation, movement, tethering and fusion. That Rab may play a role in protein transport from recycling endosomes to lysosomes regulating, for instance, the degradation of the transferrin receptor (By similarity) (244 aa)
RAB33A	RAB33A, member RAS oncogene family	RAB33A, member RAS oncogene family (237 aa)
RAB3B	RAB3B, member RAS oncogene family	RAB3B, member RAS oncogene family; Protein transport. Probably involved in vesicular traffic (By similarity) (219 aa)
RAB6B	RAB6B, member RAS oncogene family	RAB6B, member RAS oncogene family; Seems to have a role in retrograde membrane traffic at the level of the Golgi complex. May function in retrograde transport in neuronal cells (208 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
RFC5	replication factor C (activator 1) 5, 36.5kDa	replication factor C (activator 1) 5, 36.5kDa; The elongation of primed DNA templates by DNA polymerase delta and epsilon requires the action of the accessory proteins proliferating cell nuclear antigen (PCNA) and activator 1 (340 aa)
RGS9	regulator of G-protein signaling 9	regulator of G-protein signaling 9; Inhibits signal transduction by increasing the GTPase activity of G protein alpha subunits thereby driving them into their inactive GDP-bound form. Binds to G(t)-alpha. Involved in phototransduction; key element in the recovery phase of visual transduction (By similarity) (674 aa)
RNF2	ring finger protein 2	ring finger protein 2; E3 ubiquitin-protein ligase that mediates monoubiquitination of 'Lys-119' of histone H2A, thereby playing a central role in histone code and gene regulation. H2A 'Lys-119' ubiquitination gives a specific tag for epigenetic transcriptional repression and participates in X chromosome inactivation of female mammals. May be involved in the initiation of both imprinted and random X inactivation. Essential component of a Polycomb group (PcG) multiprotein PRC1-like complex, a complex class required to maintain the transcriptionally repressive state of many genes, includ [...] (336 aa)
RPL13	ribosomal protein L13 pseudogene 12; ribosomal protein L13	ribosomal protein L13 (211 aa)
RPL13A	ribosomal protein L13a pseudogene 7; ribosomal protein L13a pseudogene 5; ribosomal protein L13a pseudogene 16; ribosomal protein L13a; ribosomal protein L13a pseudogene 18	ribosomal protein L13a; Associated with ribosomes but is not required for canonical ribosome function and has extra-ribosomal functions. Component of the GAIT (gamma interferon-activated inhibitor of translation) complex which mediates interferon-gamma-induced transcript-selective translation inhibition in inflammation processes. Upon interferon-gamma activation and subsequent phosphorylation dissociates from the ribosome and assembles into the GAIT complex which binds to stem loop-containing GAIT elements in the 3'-UTR of diverse inflammatory mRNAs (such as ceruplasmin) and suppresses [...] (203 aa)
RPL13AP16	ribosomal protein L13a pseudogene 7; ribosomal protein L13a pseudogene 5; ribosomal protein L13a pseudogene 16; ribosomal protein L13a; ribosomal protein L13a pseudogene 18	
RPL13AP18	ribosomal protein L13a pseudogene 7; ribosomal protein L13a pseudogene 5; ribosomal protein L13a pseudogene 16; ribosomal protein L13a; ribosomal protein L13a pseudogene 18	
RPL13AP5	ribosomal protein L13a pseudogene 7; ribosomal protein L13a pseudogene 5; ribosomal protein L13a pseudogene 16; ribosomal protein L13a; ribosomal protein L13a pseudogene 18	
APL13AP7	ribosomal protein L13a pseudogene 7; ribosomal protein L13a pseudogene 5; ribosomal protein L13a pseudogene 16; ribosomal protein L13a; ribosomal protein L13a pseudogene 18	

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
RPL13P12	ribosomal protein L13 pseudogene 12; ribosomal protein L13	
RPL13P5	ribosomal protein L13 pseudogene 5	
RPL39	ribosomal protein L39 pseudogene 10; ribosomal protein L39 pseudogene 20; ribosomal protein L39 pseudogene 27; ribosomal protein L39; ribosomal protein L39 pseudogene 13; ribosomal protein L39 pseudogene 32	ribosomal protein L39 (51 aa)
RPL39P10	ribosomal protein L39 pseudogene 10; ribosomal protein L39 pseudogene 20; ribosomal protein L39 pseudogene 27; ribosomal protein L39; ribosomal protein L39 pseudogene 13; ribosomal protein L39 pseudogene 32	
RPL39P13	ribosomal protein L39 pseudogene 10; ribosomal protein L39 pseudogene 20; ribosomal protein L39 pseudogene 27; ribosomal protein L39; ribosomal protein L39 pseudogene 13; ribosomal protein L39 pseudogene 32	
RPL39P20	ribosomal protein L39 pseudogene 10; ribosomal protein L39 pseudogene 20; ribosomal protein L39 pseudogene 27; ribosomal protein L39; ribosomal protein L39 pseudogene 13; ribosomal protein L39 pseudogene 32	
RPL39P27	ribosomal protein L39 pseudogene 10; ribosomal protein L39 pseudogene 20; ribosomal protein L39 pseudogene 27; ribosomal protein L39; ribosomal protein L39 pseudogene 13; ribosomal protein L39 pseudogene 32	
RPL39P32	ribosomal protein L39 pseudogene 10; ribosomal protein L39 pseudogene 20; ribosomal protein L39 pseudogene 27; ribosomal protein L39; ribosomal protein L39 pseudogene 13; ribosomal protein L39 pseudogene 32	
SCHIP-1		IQCJ-SCHIP1 readthrough (563 aa)
SCN3B	sodium channel, voltage-gated, type III, beta	sodium channel, voltage-gated, type III, beta subunit; Modulates channel gating kinetics. Causes unique persistent sodium currents. Inactivates the sodium channel opening more slowly than the subunit beta-1. Its association with neurofascin may target the sodium channels to the nodes of Ranvier of developing axons and retain these channels at the nodes in mature myelinated axons (By similarity) (215 aa)
SCN9A	sodium channel, voltage-gated, type IX, alpha subunit	sodium channel, voltage-gated, type IX, alpha subunit (1977 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
SCP2	sterol carrier protein 2	sterol carrier protein 2; Mediates <i>in vitro</i> the transfer of all common phospholipids, cholesterol and gangliosides between membranes. May play a role in regulating steroidogenesis (547 aa)
SEC11C	SEC11 homolog C (<i>S. cerevisiae</i>)	SEC11 homolog C (<i>S. cerevisiae</i>); Component of the microsomal signal peptidase complex which removes signal peptides from nascent proteins as they are translocated into the lumen of the endoplasmic reticulum (By similarity) (192 aa)
SLC25A14	solute carrier family 25 (mitochondrial carrier, brain), member 14	solute carrier family 25 (mitochondrial carrier, brain), member 14; Participates in the mitochondrial proton leak measured in brain mitochondria (325 aa)
SLC25A3	solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 3	solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 3; Transport of phosphate groups from the cytosol to the mitochondrial matrix. Phosphate is co-transported with H(+). May play a role regulation of the mitochondrial permeability transition pore (mPTP) (362 aa)
SNAPC5	small nuclear RNA activating complex, polypeptide 5, 19kDa	small nuclear RNA activating complex, polypeptide 5, 19kDa; Part of the SNAPc complex required for the transcription of both RNA polymerase II and III small-nuclear RNA genes. Binds to the proximal sequence element (PSE), a non-TATA-box basal promoter element common to these 2 types of genes. Recruits TBP and BR2 to the U6 snRNA TATA box (98 aa)
SPARCL1	SPARC-like 1 (hevin)	SPARC-like 1 (hevin) (664 aa)
SPINT2	serine peptidase inhibitor, Kunitz type,	serine peptidase inhibitor, Kunitz type, 2; Inhibitor of HGF activator. Also inhibits plasmin, plasma and tissue kallikrein, and factor Xia (252 aa)
SUMF2	sulfatase modifying factor 2	sulfatase modifying factor 2 (358 aa)
TAOK3	TAO kinase 3	TAO kinase 3; Serine/threonine-protein kinase that acts as a regulator of the p38/MAPK14 stress-activated MAPK cascade and of the MAPK8/JNK cascade. Acts as an activator of the p38/MAPK14 stress- activated MAPK cascade. In response to DNA damage, involved in the G2/M transition DNA damage checkpoint by activating the p38/MAPK14 stress-activated MAPK cascade, probably by mediating phosphorylation of upstream MAP2K3 and MAP2K6 kinases. Inhibits basal activity of MAPK8/JNK cascade and diminishes its activation in response epidermal growth factor (EGF) (898 aa)
TCTEX1D2	Tctex1 domain containing 2	Tctex1 domain containing 2 (142 aa)
TIMM10B		Fracture callus 1 homolog (Rat); cDNA FLJ11268 fis, clone PLACE1009186, highly similar to Mitochondrial import inner membrane translocase subunit Tim9B; cDNA, FLJ92024, Homo sapiens fracture callus 1 homolog (rat) (FXC1), mRNA ; Component of the TIM22 complex, a complex that mediates the import and insertion of multi-pass transmembrane proteins into the mitochondrial inner membrane. The TIM22 complex forms a twinpore translocase that uses the membrane potential as the external driving force. In the TIM22 complex, it may act as a docking point for the soluble 70 kDa complex that guide [...] (103 aa)
TMEM50A	transmembrane protein 50A	transmembrane protein 50A (157 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
TMSB4XP4		thymosin beta 4, X-linked pseudogene 4 (44 aa)
TRAPPC4	trafficking protein particle complex 4	trafficking protein particle complex 4; May play a role in vesicular transport from endoplasmic reticulum to Golgi (219 aa)
TRAPPC6B	trafficking protein particle complex 6B	trafficking protein particle complex 6B; May play a role in vesicular transport from endoplasmic reticulum to Golgi (By similarity) (158 aa)
TRIM2	tripartite motif-containing 2	tripartite motif containing 2; UBE2D1-dependent E3 ubiquitin-protein ligase that mediates the ubiquitination of NEFL and of phosphorylated BCL2L11 (By similarity). Plays a neuroprotective function (By similarity). May play a role in neuronal rapid ischemic tolerance (By similarity) (771 aa)
TSTD2	chromosome 9 open reading frame 97	thiosulfate sulfurtransferase (rhodanese)-like domain containing 2 (516 aa)
TUBB4A		tubulin, beta 4A class IVa; Tubulin is the major constituent of microtubules. It binds two moles of GTP, one at an exchangeable site on the beta chain and one at a non-exchangeable site on the alpha chain (By similarity) (444 aa)
U2SURP		U2 snRNP-associated SURP domain containing (1029 aa)
UBXN1	UBX domain protein 1	UBX domain protein 1; Ubiquitin-binding protein that interacts with the BRCA1- BARD1 heterodimer, and regulates its activity. Specifically binds 'Lys-6'-linked polyubiquitin chains. Interaction with auto-ubiquitinated BRCA1, leads to inhibit the E3 ubiquitin-protein ligase activity of the BRCA1-BARD1 heterodimer. Component of a complex required to couple deglycosylation and proteasome-mediated degradation of misfolded proteins in the endoplasmic reticulum that are retrotranslocated in the cytosol (312 aa)
UQCRC1	ubiquinol-cytochrome c reductase core protein I	ubiquinol-cytochrome c reductase core protein I; This is a component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is part of the mitochondrial respiratory chain. This protein may mediate formation of the complex between cytochromes c and c1 (480 aa)
VAT1L	vesicle amine transport protein 1 homolog (<i>T. californica</i>)-like	vesicle amine transport protein 1 homolog (<i>T. californica</i>)-like (419 aa)
WSB2	WD repeat and SOCS box-containing 2	WD repeat and SOCS box containing 2; May be a substrate recognition component of a SCF-like ECS (Elongin-Cullin-SOCS-box protein) E3 ubiquitin ligase complex which mediates the ubiquitination and subsequent proteasomal degradation of target proteins (By similarity) (404 aa)
YWHAG	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide; Adapter protein implicated in the regulation of a large spectrum of both general and specialized signaling pathways. Binds to a large number of partners, usually by recognition of a phosphoserine or phosphothreonine motif. Binding generally results in the modulation of the activity of the binding partner (247 aa)
ZFH2	zinc finger homeobox 2	zinc finger homeobox 2; May be involved in transcriptional regulation (2572 aa)

Supplementary Table 1. (continued)

Top genes	DAVID_GO functions	STRING_GO functions
ZFXH4	zinc finger homeobox 4	zinc finger homeobox 4; May play a role in neural and muscle differentiation (By similarity). May be involved in transcriptional regulation (3616 aa)
ZNF587	zinc finger protein 587	zinc finger protein 587; May be involved in transcriptional regulation (575 aa)
ENSG00000104725		
ENSG00000188765		
ENSG00000196205		
ENSG00000198868		
ENSG00000215030		
ENSG00000216285		
ENSG00000219133		
ENSG00000224066		
ENSG00000224906		
ENSG00000225855		
ENSG00000229119		
ENSG00000229447		
ENSG00000229691		
ENSG00000230455		
ENSG00000230593		
ENSG00000230916		
ENSG00000230940		
ENSG00000231609		
ENSG00000233476		
ENSG00000235174		
ENSG00000236552		
ENSG00000242814		
ENSG00000244513		
ENSG00000254373		
ENSG00000259045		
ENSG00000260006		
ENSG00000261604		
ENSG00000262902		
ENSG00000263424		
ENSG00000268191		
ENSG00000268203		
ENSG00000270567		
ENSG00000270907		
ENSG20000053439		





CHAPTER 4B

Establishment of a pre-sequencing quality control pipeline for single-neuron Patch-seq experiments

Mark van den Hurk^{1,2,3}, Cedric Bardy³, Gunter Kenis^{1,2}, Jennifer Erwin³, Mark Novotny⁴, Roger S. Lasken⁴, Harry W.M. Steinbusch^{1,2}, Bart P.F. Rutten^{1,2}, Fred H. Gage³

¹ Department of Psychiatry and Neuropsychology, Division of Neuroscience, Maastricht University, Maastricht 6200 MD, The Netherlands

² School for Mental Health and Neuroscience (MHeNS), European Graduate School of Neuroscience (EURON), Maastricht University, Maastricht 6200 MD, The Netherlands

³ Sanford Consortium for Regenerative Medicine, Salk Institute for Biological Studies, La Jolla, CA 92037, USA

⁴ J. Craig Venter Institute, La Jolla, CA 92037, USA

ABSTRACT

Single-cell RNA-sequencing (scRNA-seq) has emerged as a powerful tool for stratifying the heterogeneity of cellular populations. However, the risk of including low-quality or compromised samples in downstream bioinformatics pipelines poses a significant challenge to accurate data analysis and interpretation, as distinguishing technical artifacts from intrinsic biological variability in scRNA-seq data is rather difficult. Here, we report a pre-sequencing quality control (QC) pipeline, applied to single-neuron samples captured for transcriptome analysis following electrophysiological recording (i.e., Patch-seq), that can be used to identify and filter out potentially spurious cell samples in advance of deep sequencing. By preventing the redundant, and costly, sequencing of poor-quality cDNA preps, the presented framework forms a basis for better methodological designs, which may increase the power with which altered gene expression levels can be detected between experimental samples in single-cell transcriptome analyses.

Keywords: single-cell RNA-seq, Patch-seq, pre-sequencing quality control (QC)

Recent advances in whole-transcriptome amplification and next-generation sequencing technologies have enabled efficient, nucleotide-resolution profiling of RNA isolated from individual cells. At present, the profiling of single-cell transcriptomes by RNA-sequencing (scRNA-seq) represents a highly demanded strategy to define and disentangle, in a relatively unbiased manner, the heterogeneity of seemingly identical cell types in mixed cell populations [1]. It is without doubt that a comprehensive, transcriptome-based deconvolution of complex tissue into cellular subtypes will greatly aid researchers in better defining cell identity, and understanding cellular function and behavior in healthy, developmental and disease states [1-3].

scRNA-seq data is very different from bulk RNA-seq data generated from whole tissue or cell populations in the sense that its full exploitation requires relatively novel and continuously evolving bioinformatics analysis strategies [4, 5]. In particular, a major challenge in scRNA-seq analysis is to distinguish technical noise from genuine biological variation (e.g., is a transcript absent in the cell because the gene is not transcribed at the time of capture, or could the transcript not be detected due to a stochastic, procedural loss of lowly expressed mRNAs?) [6]. scRNA-seq experiments suffer considerably from high levels of technical noise owing to the low amount of starting RNA material that requires PCR amplification [7]. To identify potential (technical) outlier samples, studies commonly evaluate various quality control (QC) metrics after sequencing, such as the rate of read mapping to the reference transcriptome, the number of detected genes and transcripts, and the count levels of common housekeeping genes that are believed to have detectable levels of expression [8]. Here, we evaluated the relationship between such *post*-sequencing QC measures and quantitative data obtained from several *pre*-sequencing QC assays used to assess sample quality prior to RNA-seq. The outcomes of this exploration enabled us to establish a pre-sequencing QC pipeline that can be used to filter out potentially spurious (i.e., poor-quality or compromised) samples in advance of sequencing (**Figure 1**).

We collected electrophysiologically and morphologically characterized single human neurons derived from embryonic and induced pluripotent stem cells (ESCs/iPSCs) for single-cell gene expression analysis with RNA-seq – a procedure nowadays referred to as “Patch-seq” [9]. Following whole-cell patch-clamping and imaging, we applied additional negative pressure to the same pipette used for electrophysiological recording to establish a stronger seal between neuron and patch electrode. Then, careful retraction of the glass microelectrode from the recording chamber bath and breaking its tip along the inside wall of an RNase-free PCR tube filled with sample buffer enabled us to isolate the neuron for whole-transcriptome analysis. Contrary to most other single-cell capture methods, which involve neurite-damaging cell dissociation or micro-aspiration of cytosolic contents [3], our method collects the entire neuron including all its processes (axon and dendrites; see also Figure 4A in CHAPTER 4A), and therefore potentially provides a

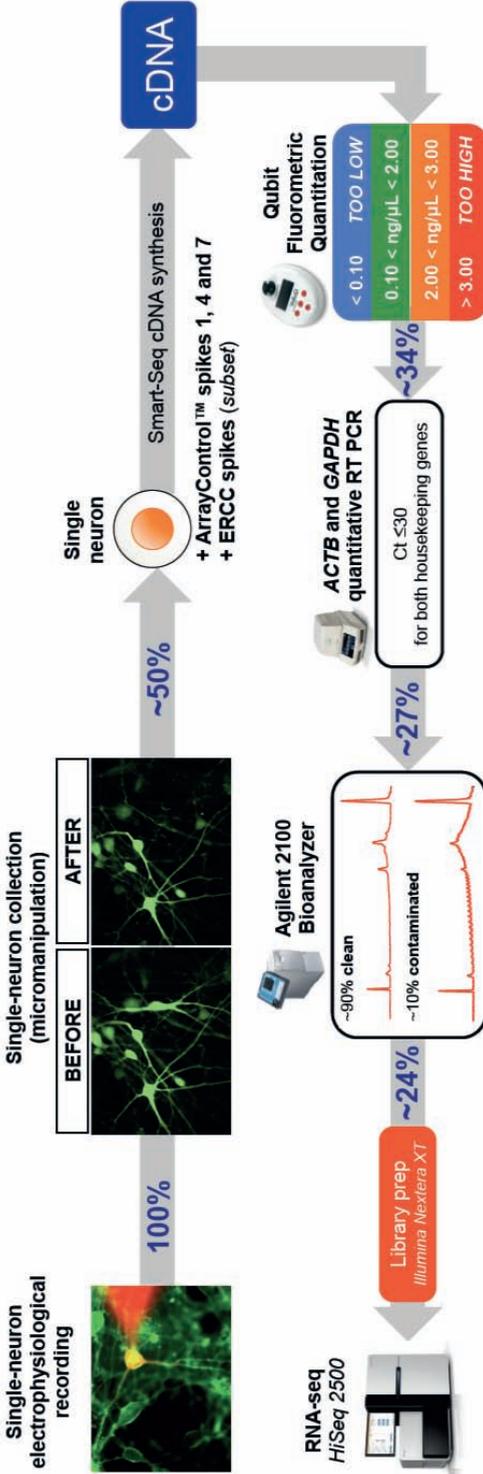


Figure 1. Pre-sequencing QC pipeline for single-neuron Patch-seq experiments. Following patch-clamp recording and morphological characterization, single neurons where collected by pipette micromanipulation. Successful neuron capture was always confirmed by microscopic qualitative evaluation (compare before and after pictures). Single-cell mRNA, along with exogenous spike-in transcripts, was then reverse-transcribed and amplified using Smart-seq technology [6]. The generated cDNA was then subjected to a series of quality control steps based on three simple assays: (i) fluorometric quantitation of cDNA yield (Qubit), (ii) expression profiling of common housekeeping genes (quantitative real-time PCR), and (iii) subjective qualitative analysis of cDNA fragment profiles (Agilent Bioanalyzer). Contaminated samples were characterized by an abnormally high yield of cDNA (Qubit >3 ng/μl) that appeared on the fragment analyzer as a broader peak with a slightly “jagged” pattern (see also Figure 2). We also identified cells not expressing either of the two housekeeping genes *GAPDH* and *ACTB* (Ct value >30 as threshold) as spurious samples, as they were typically recognized as outliers (low number of genes detected) by our post-sequencing bioinformatics QC (see text for details). Percentages in the figure indicate the success rate of sample inclusion after each respective QC step.

more accurate representation of the transcriptional program of the cell [10]. Successful removal of the neuron from the coverslip was always confirmed by microscopical evaluations, and samples were discarded from further processing if complete collection was ambiguous. This control step represents the first QC checkpoint in our pre-sequencing QC pipeline (**Figure 1**).

Following capture of a single neuron, its poly-adenylated mRNA was reverse-transcribed and amplified using Smart-seq technology [6] with minimal delay and without intermittent sample freezing to avoid RNA degradation. To assess technical performance of the reverse transcription and amplification reactions, the sample buffer used contained known amounts of exogenous reference transcripts (ERCCs and ArrayControl spike-in RNAs) [11, 12]. As revealed by fluorometric quantitation of cDNA yield (Qubit fluorometer, Invitrogen), the Smart-seq protocol typically yielded sufficient amounts of full-length cDNA for scRNA-seq. Prior to library preparation and sequencing, however, we profiled the expression of common housekeeping genes and the added spike-ins with quantitative real-time PCR to identify potentially low-quality captures and amplifications for exclusion (**Figure 1**). Although the expression levels of housekeeping genes may be considerably variable between individual cells and can even distinguish human cell types [13], we assumed that samples not expressing either of the two housekeeping genes *GAPDH* and *ACTB* (i.e., C_t value “Not Detected” or >30) are spurious, and therefore we typically excluded them from further consideration. We also subjectively assessed the quality of each cDNA sample by Agilent Bioanalyzer profiling of cDNA size distribution, which enabled us to detect and exclude contaminated samples prior to transcriptome profiling (**Figure 1**).

To get insight into the relationship between pre-sequencing QC data and post-sequencing measures of sample quality, we initially sequenced 35 libraries generated from collected single-neuron samples irrespective of whether or not these samples passed all the QC criteria of our pipeline mentioned above. These libraries were sequenced on an Illumina HiSeq platform with 100-base-pair (bp) paired-end strategy at moderate sequencing depth (~ 36.5 million (M) reads per sample on average). The mean number of reads mapped to the reference transcriptome was 20.8 M (range 0.27–43.5 M), corresponding to an efficiency of 56.9%. The wide range in read mappability pointed to the presence of a number of samples ($n = 7$) with exceptionally low mapping rate. As low read mappability is indicative of technical issues [14], we aimed to explore whether any of the quantitative pre-sequencing QC metrics was specifically associated with this post-sequencing QC variable, so that poor-quality samples could be easily identified and excluded from future sequencing runs. Notably, we found the cDNA concentration after amplification to be predictive of the read mapping rate. Specifically, 4 out of 7 samples with $<40\%$ mapped reads were characterized by abnormally high (>3 ng/ μ l) cDNA yields (**Figure 2A**), motivating us to focus future sequencing experiments on samples with a cDNA yield below

this arbitrary cut-off. Additional subjective inspection of the cDNA fragment profiles (Agilent Bioanalyzer) of the poor-mapping samples revealed that most (3 out of 4) of the samples with high (>3 ng/μl) cDNA yield were characterized by a broader than normal peak with increased cDNA amounts at lower bp sizes and a slightly “jagged” pattern indicative of PCR primer/adaptor contamination (**Figure 2B**). In contrast, samples with a high mapping rate had a shallow peak with a maximum at ~2000 bp, no jaggedness, and no low bp-sized cDNA product (**Figure 2C**). Taken together, these results indicate that both a quantitative and qualitative assessment of Smart-seq-synthesized cDNA prior to library preparation and sequencing can be used to identify, and filter out, cell samples generating low-quality data.

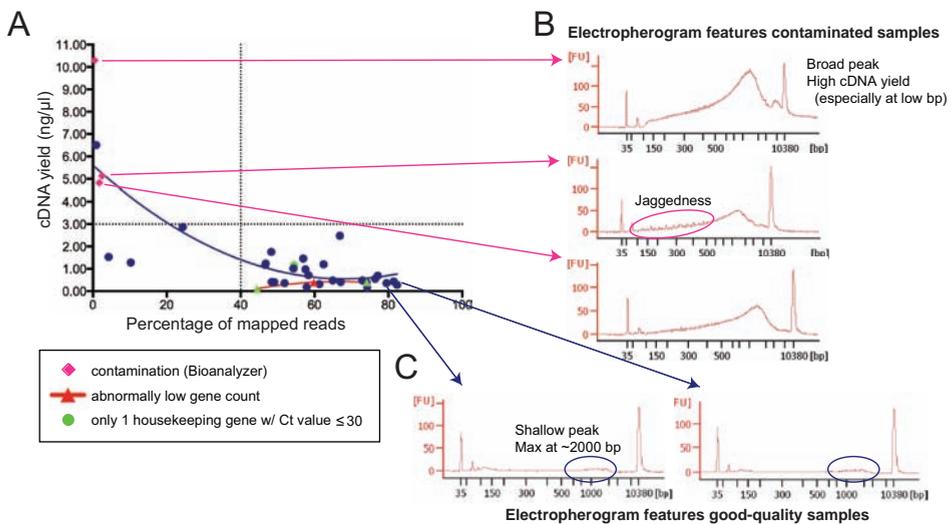


Figure 2. Contaminated samples, identified by a high cDNA yield and distinctive electropherogram features, are associated with low read mapping. (A) 4 out of 7 samples with <40% mapped reads are characterized by abnormally high cDNA yields (Qubit fluorometric quantitation >3 ng/μl). In addition, 2 out of 3 samples with a normal number of mapped reads but with a significantly reduced number of detected genes are positive for only one of two housekeeping genes (*GAPDH*, *ACTB*) whose expression we determined with PCR. **(B)** Most samples with high cDNA yield are characterized by distinctive features of the cDNA fragment profile that indicate PCR primer/adaptor contamination: a broader than normal peak with increased cDNA amounts at lower base-pair (bp) sizes and a slightly “jagged” pattern. **(C)** Samples generating high-quality data have a shallow peak with a maximum at ~2000 bp (range ~400-9000 bp), no jaggedness, and no cDNA product at lower bp sizes.

Apart from the seven samples with poor read mappability, our bioinformatics QC identified a few samples ($n = 3$) with a normal number of mapped reads, but with a significantly lower number of detected genes ($P < 0.05$). Notably, we found two out of three of these samples to be positive for only one of the two but not both housekeeping genes (*GAPDH*, *ACTB*) whose expression we profiled with PCR prior to sequencing (**Figure**

2A). Confirming the notion that expression of a few housekeeping genes is reflective of total gene expression in single cells, we found, in a larger sample of $n = 102$ cells, a highly significant negative correlation between the number of total genes detected with RNA-seq and the PCR Ct value of either housekeeping gene (*GAPDH*: $r = -0.64$, $P < 0.0001$; *ACTB*: $r = -0.54$, $P < 0.0001$) (**Figure 3A,B**). In addition, a random forest machine learning classifier trained to discriminate samples with a high number of detected genes (i.e., upper 25th percentile of the range) from all other samples (i.e., lower 75th percentile) based on various pre-sequencing QC information identified a combined measure of *GAPDH* and *ACTB* PCR expression as the top predictor of “high” versus “other” gene count (**Figure 3C-F**). The accuracy with which our classifier algorithm distinguished the two groups of samples was 86%, significantly above chance level (**Figure 3D**).

Our work demonstrates that pre-sequencing measures of cDNA quantity and quality are, to considerable extent, predictive of two post-sequencing metrics related to RNA-seq data quality [8]: (i) the rate of read mapping to the reference transcriptome, and (ii) the number of detected genes/transcripts. Based on the results of this study, we established a QC pipeline that may facilitate the identification and filtering out of contaminated and poor-quality single-neuron Patch-seq samples ahead of library preparation and sequencing (**Figure 1**). It has to be pointed out, however, that the number of genes detected in a RNA-seq experiment is related to both the quality of the library and the biological cell type [14, 15]. Jiang *et al.*'s analysis of a highly heterogeneous scRNA-seq dataset comprised of a multitude of cell types found the number of genes detected to be highly variable between different cell types [14]. Also, mRNA production has been shown to be subject to intrinsically stochastic molecular events [16, 17] and to be dependent on the cell cycle state of the cell [18], phenomena that are likely to considerably contribute to cell-to-cell variation in gene expression. Hence, substantial bias may be introduced if studies use the number of detected genes as a QC metric to select samples for biological analysis. Of note, however, Jiang *et al.* found that in a dataset composed of only one cell type, as opposed to a dataset of a mixture of cell types, samples with a significantly reduced number of genes detected had a higher probability of being technical artifacts than of being cells with variable mRNA expression [14]. Therefore, we believe that our decision to exclude samples based on abnormal Ct values for detection of housekeeping genes, is justified. Yet, it is clear from our study, and that of Jiang *et al.* [14], that scRNA-seq QC still faces many challenges. While the pre-sequencing QC pipeline presented here will reduce the chances of spurious sample inclusion, thereby reducing experimental variability and increasing the power with which altered gene expression levels are detected between conditions, future research efforts should be directed toward investigating which specific RNA-seq data properties are best reflective of sample quality.

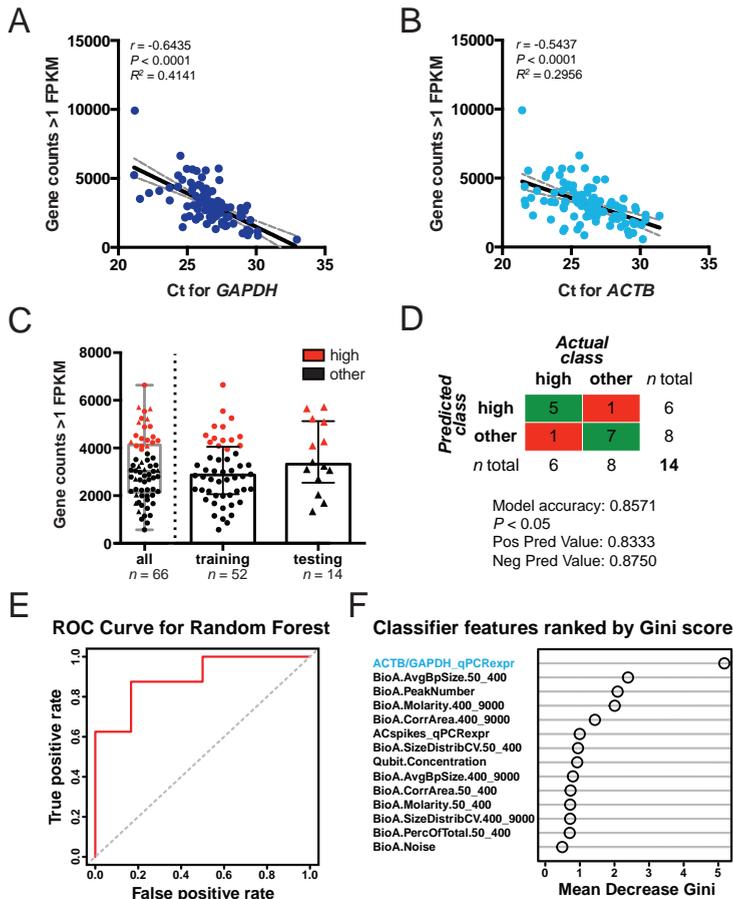


Figure 3. Expression of *GAPDH* and *ACTB* housekeeping genes is reflective of total gene expression in single cells. (A,B) Quantitative real-time PCR expression of *GAPDH* and *ACTB* is negatively correlated with the number of genes detected by RNA-seq above 1 FPKM cut-off in a large sample of $n = 102$ single cells collected for Patch-seq. (C-F) Random forest machine learning classifier algorithm accurately predicts samples with a high number of genes detected based on pre-sequencing measures (Qubit-quantified cDNA yield, PCR-measured expression of *GAPDH* and *ACTB* housekeeping genes, and quantitative information extracted from Bioanalyzer electropherogram profiles; see Supplementary Methods for details). (C) Single-cell samples for which there was data on all pre-sequencing measures ($n = 66$) were randomly divided into training ($n = 52$) and testing ($n = 14$) data sets, and dichotomized as “high” or “other” gene count based on whether the total number of genes detected by RNA-seq was among the upper 25th percentile of the range (“high”) or not (“other”). Median (interquartile range) shown for training and testing data. (D) Confusion matrix of the number of true (green) and false (red) positives and negatives of the classifier algorithm prediction performed on the testing data set. The classifier distinguished the two groups of samples with an accuracy of 86%, significantly above chance level. (E) Receiver Operating Characteristic (ROC) curve of the “high” vs. “other” classifier showing trade-offs between True Positive Rate (TPR = Sensitivity, y axis) and False Positive Rate (FPR = 1-Specificity, x axis), with good cross-validated performance: Area Under the Curve (AUC) >0.8. (F) Top predictive pre-sequencing variables selected by the machine learning classifier (ranked by Gini score). A single variable explaining most variation in PCR expression of *GAPDH* and *ACTB* housekeeping genes was identified as the top predictor of “high” vs. “other” gene count. FPKM: fragments per kilobase of exon per mapped million reads.

METHODS

Details of the methods are available in the SUPPLEMENTARY METHODS.

REFERENCES

1. Eberwine J, Sul JY, Bartfai T, Kim J. The promise of single-cell sequencing. *Nat. Methods* 11(1), 25-27 (2014).
2. Jaitin DA, Kenigsberg E, Keren-Shaul H *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* 343(6172), 776-779 (2014).
3. Stahlberg A, Kubista M, Aman P. Single-cell gene-expression profiling and its potential diagnostic applications. *Expert Rev. Mol. Diagn.* 11(7), 735-740 (2011).
4. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* 16(3), 133-145 (2015).
5. Buettner F, Natarajan KN, Casale FP *et al.* Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol.* 33(2), 155-160 (2015).
6. Ramskold D, Luo S, Wang YC *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* 30(8), 777-782 (2012).
7. Brennecke P, Anders S, Kim JK *et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* 10(11), 1093-1095 (2013).
8. Conesa A, Madrigal P, Tarazona S *et al.* A survey of best practices for RNA-seq data analysis. *Genome Biol.* 17 13 (2016).
9. Bardy C, Van Den Hurk M, Kakaradov B *et al.* Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology. *Mol. Psychiatry* (2016, in press).
10. Cajigas JJ, Tushev G, Will TJ, Tom Dieck S, Fuerst N, Schuman EM. The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* 74(3), 453-466 (2012).
11. Munro SA, Lund SP, Pine PS *et al.* Assessing technical performance in differential gene expression experiments with external spike-in RNA control ratio mixtures. *Nat. Commun.* 5 5125 (2014).
12. Jiang L, Schlesinger F, Davis CA *et al.* Synthetic spike-in standards for RNA-seq experiments. *Genome Res.* 21(9), 1543-1551 (2011).
13. Oyolu C, Zakharia F, Baker J. Distinguishing human cell types based on housekeeping gene signatures. *Stem Cells* 30(3), 580-584 (2012).
14. Jiang P, Thomson JA, Stewart R. Quality control of single-cell RNA-seq by SinQC. *Bioinformatics* 32(16), 2514-2516 (2016).
15. Kharchenko PV, Silberstein L, Scadden DT. Bayesian approach to single-cell differential expression analysis. *Nat. Methods* 11(7), 740-742 (2014).
16. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* 4(10), e309 (2006).
17. Raj A, Van Oudenaarden A. Single-molecule approaches to stochastic gene expression. *Annu. Rev. Biophys.* 38 255-270 (2009).
18. Zopf CJ, Quinn K, Zeidman J, Maheshri N. Cell-cycle dependence of transcription dominates noise in gene expression. *PLoS Comput. Biol.* 9(7), e1003161 (2013).
19. Boyer LF, Campbell B, Larkin S, Mu Y, Gage FH. Dopaminergic differentiation of human pluripotent cells. *Curr. Protoc. Stem Cell Biol.* Chapter 1 Unit1H 6 (2012).
20. Bardy C, Van Den Hurk M, Eames T *et al.* Neuronal medium that supports basic synaptic functions and activity of human neurons in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 112(20), E2725-2734 (2015).
21. Dobin A, Davis CA, Schlesinger F *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1), 15-21 (2013).
22. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31(2), 166-169 (2015).
23. Breiman L. Random forests. *Machine Learning* 45(1), 5-32 (2001).

SUPPLEMENTARY METHODS

Human Neuronal Culture

Human dermal fibroblasts were reprogrammed into pluripotent cells with the four Yamanaka factors (Oct3/4, Sox2, Klf4 and c-Myc), either in a retroviral vector or a non-integrating Sendai viral vector. Human iPSCs and embryonic stem cells (ESCs) were differentiated into neural progenitor cells (NPCs) as previously described [19]. NPCs were expanded for 3 to 5 passages (split 1:2 or 1:3 per passage) and stored at -80 °C. Then they were thawed and expanded for at least one more passage in neural progenitor medium (DMEM/F12 supplemented with FGF8, SHH, B27 and N2, without retinoic acid). Dissociated NPCs were directly plated on glass coverslips (Fisher Scientific, Cat. No. 12-545-80) coated with poly-ornithine (Sigma, Cat. No. P3655) and laminin (Invitrogen/Thermo Fisher Scientific, Cat. No. 23017-015) in 24-well plates. Twenty-four hours later, the cells were switched gradually (half medium change) to neuronal medium: BrainPhys basal [20] supplemented with 1x N2 (Gibco/Thermo Fisher Scientific, Cat No. 17502-048), 1x B27 (Gibco/Thermo Fisher Scientific, Cat No. 17504-044), brain-derived neurotrophic factor (BDNF, 20 ng/ml; PeproTech, Cat No. 450-02), glia-derived neurotrophic factor (GDNF, 20 ng/ml; PeproTech, Cat No. 450-10), ascorbic acid (AA, 200 nM; Sigma, Cat No. A0278), dibutyryl cyclic AMP (cAMP, 1mM Sigma, Cat No. D0627), and laminin (1 µg/ml; Invitrogen/Thermo Fisher Scientific, Cat No 23017-015). Half of the neuronal medium was gently replaced two to three times a week. The plates were kept in a humidified incubator at 37°C with 5% CO₂ and 21% O₂. The pH (~7.3-7.4) and the osmolarity (~300-305 mOsmol/L) of medium were maintained constant over time.

Patch Clamping, Single-Neuron Collection and cDNA Preparation

Whole-cell patch-clamping was performed as previously prescribed [9]. Following electrophysiological recording, we applied slight additional negative pressure to establish a stronger connection between cell and patch electrode, but without completely aspirating the cell in the pipette. The neuron, including its processes (axon and dendrites), was then transferred in a volume of ~2 µl of internal patch solution into a PCR tube containing 8.0 µl sample buffer by slowly retracting the patch pipette from the chamber bath and breaking the tip of the electrode along the inside wall of the tube. Sample buffer was made according to the protocol supplied by Clontech with the SMARTer™ Ultra Low RNA Kit and included both ERCC spike-ins (Thermo Fisher Scientific) and ArrayControl spikes 1, 4 and 7 (RNA standard for the Fluidigm C1) for the purpose of quality control (QC). Successful removal of the cell from the coverslip was always confirmed by DIC optics. Collected single cells were immediately spun down and put on ice for SMARTer™ cDNA synthesis following manufacturer's instructions (Clontech). Briefly, first-strand cDNA was synthesized from poly(A)⁺ RNA by incubation with 1 µl of 3' SMART CDS Primer

II A (24 μ M) for 3 min at 72°C, followed by reverse transcription in a 20- μ l final reaction volume using 200 units of SMARTScribe™ Reverse Transcriptase for 90 min at 42 °C and inactivation for 10 min at 70°C. First-strand cDNA was then purified using Agencourt AMPure XP SPRI Beads (Beckman Coulter Genomics) and amplified by long-distance PCR using the Advantage® 2 PCR Kit (Clontech) with the following PCR thermocycler program: 95°C for 1 min, 18 cycles of 95°C for 15 s, 65°C for 30 s, 68°C for 6 min, and 72°C for 10 min. PCR-amplified double-stranded (ds) cDNA was immobilized onto SPRI beads, purified by two washes in 80% ethanol, and eluted in 12 μ l of purification buffer (Clontech). The quality (Agilent 2100 Bioanalyzer High Sensitivity DNA Kit; Agilent Technologies) and quantity (Qubit dsDNA High Sensitivity Assay Kit; Invitrogen/Thermo Fisher Scientific) of each ds cDNA sample were assessed prior to library preparation.

qPCR Assays

To select the single-cell cDNA samples suitable for mRNA-seq, we determined the level of expression of *GAPDH* and *ACTB* in each sample by TaqMan real-time PCR. Samples with Ct values \leq 30 for both housekeeping genes were typically considered positive for library preparation. For each gene, duplicate 10- μ l PCR reactions were performed on an ABI Prism 7900 Sequence Detector (Applied Biosystems/Thermo Fisher Scientific) using 0.50 μ l of 1:5-diluted ds cDNA template in standard TaqMan Gene Expression Assay with FAM reporter. Real-time PCR assays for detection of the ERCCs and ArrayControl RNA spikes were performed using, respectively, standard TaqMan Gene Expression Assays and SYBR Green PCR Master Mix (Thermo Fisher Scientific) with custom primers (Fluidigm).

Illumina Transcriptome Library Preparation and Sequencing

Construction of single-cell mRNA-seq libraries was typically performed with 0.25 ng of input cDNA using the Nextera XT DNA sample prep kit (Illumina) with modified protocol. Briefly, cDNA was tagmented for 5 min at 55°C in a 5- μ l reaction containing 2.5 μ l of Tagment DNA Buffer and 1.25 μ l of Amplicon Tagment Mix; tagmentation was neutralized with 1.25 μ l of Neutralize Tagment Buffer for 5 min. Tagmented DNA was then subjected to 12-cycle PCR amplification using 3.75 μ l of Nextera PCR Master Mix and 1.25 μ l each of index 1 (i7) and index 2 (i5) library-identifying (barcoded) sequencing primers. The constructed libraries were run on a 1.5% agarose gel in Tris-borate/EDTA buffer, stained with SYBR Gold (Invitrogen), and size selected for ~300-400 or ~300-650 bp (insert size of ~165-265 or ~165-515 bp, respectively). Gel-excised library fragments were purified with the Wizard SV Gel and PCR Clean-Up System (Promega), eluted in 40 μ l of nuclease-free water, and concentrated by speedvacuum centrifugation. Each library was then quantified (Qubit dsDNA High Sensitivity Assay Kit; Invitrogen/Thermo Fisher Scientific) and examined for correct size (Agilent 2200 TapeStation High Sensitivity D1K

ScreenTape Assay; Agilent Technologies), after which equimolar amounts of uniquely barcoded libraries were pooled together and used for cluster generation and 100-bp paired-end sequencing on a HiSeq 2000 or 2500 sequencer (Illumina).

Alignment & Gene-Level Quantitation

Raw sequencing read data was mapped to the human reference transcriptome (Gencode v19) using gapped-alignment strategies. Alignment was performed by STAR (version 2.3.0) [21] followed by gene-level quantification with HTseq (version 0.6.1) [22].

Random Forest Machine Learning Classifier

A random forest classifier [23] was trained to predict from a total of 16 quantitative pre-sequencing QC features whether or not a sample was characterized by a high number of genes detected. The majority (13 out of 16) of predictive QC metrics were extracted from the cDNA sample electropherogram (Bioanalyzer) profiles, and included for both the specific (400-9000 bp) and non-specific (50-400 bp) product regions measures on the concentration/molarity, area under the curve (AUC), average bp size and coefficient of variation in size distribution, as well as more general variables related to the overall degree of noise and the number of detected peaks. Principal components analysis (PCA) was used to summarize the main components of variation in the *GAPDH* and *ACTB* C_t expression values into a single predictor. A similar predictor variable was built explaining most variation in the C_t values of our ArrayControl RNA spikes. cDNA yield (Qubit) constituted a separate predictive variable. Single-cell samples for which there was data on all these pre-sequencing measures ($n = 66$) were dichotomized based on whether the total number of genes detected by RNA-seq for that sample was among the upper 25th percentile of the range (“high gene count”) or not (“other gene count”), and were randomly subsampled into a training (80%) and testing (20%) subset for model building and cross-validation purposes, respectively. Using the random forest classifier, we achieved good validation performance with 86% generalization accuracy, as shown by a high area under the Receiver Operating Characteristic (ROC) curve (AUC >0.8, **Figure 3E**).

Statistics

GraphPad Prism (GraphPad Software Inc., Version 5.0) was used for all statistical analysis and graph presentation. Linear regression was used to correlate housekeeping gene C_t values (PCR) with the number of total genes detected (RNA-seq). To estimate the goodness of fit of the linear regression represented graphically, we reported the R^2 and the corresponding P value, which estimates how significantly the slope deviates from zero. Statistical results were considered significant for P values < 0.05.



CHAPTER 5

Identification of a co-expression gene network and transcriptional regulators for functional maturity based on electrophysiological, morphological and transcriptomic profiling of iPSC-derived single human neurons

Mark van den Hurk^{1,2}, Gunter Kenis¹, Lars Eijssen¹, Boyko Kakaradov³, Jennifer Erwin², Ruben V. Hernandez², Tameji Eames², Mark Gorris², Cynthia Marchand², Alex K. Bryant², Mariko Kellogg², Roger S. Lasken⁴, Gene W. Yeo^{3,5}, Harry W.M. Steinbusch¹, Fred H. Gage², Bart P.F. Rutten¹, Cedric Bardy²

¹ Department of Psychiatry and Neuropsychology, Division of Translational Neuroscience, Maastricht University, 6200 MD, Maastricht, The Netherlands.

² Salk Institute for Biological Studies, Sanford Consortium for Regenerative Medicine, La Jolla 92037, CA, USA.

³ Department of Cellular and Molecular Medicine, University of California San Diego (UCSD), La Jolla 92037, CA, USA.

⁴ J. Craig Venter Institute, La Jolla, 92037, CA, USA.

⁵ Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore and Molecular Engineering Laboratory, A*STAR, Singapore.

To be submitted

ABSTRACT

The brain is an immensely complex organ, consisting of hundreds of billions of cells that make over a quadrillion of connections. Even seemingly identical types of neurons exhibit a high degree of variability in behavior and function, and understanding at the molecular level why individual neurons function differently is key to identifying new disease mechanisms and potential therapeutic targets. We previously combined patch-clamp electrophysiological analysis of human embryonic and induced pluripotent stem cell (ESC/iPSC)-derived neurons with single-cell RNA-sequencing, and revealed a molecular signature identifying different functional neuronal types *in vitro*. Here, in this study, we performed further systems biology and network analysis to obtain mechanistic insights into the biology underlying different electrophysiological neuron types, which we related to neuronal maturity. Using the method of Weighted Gene Co-expression Network Analysis (WGCNA), we identified a co-expression gene module specifically associated with highly functional neurons. We revealed involvement of these genes in mitochondrial energy metabolism and cytoskeletal organization, and showed that they may play a general role in neuronal maturation. Additionally, we identified a set of 21 putative transcription factors that may be involved in regulating, or driving towards, functional neuronal maturity.

Keywords: Patch-Seq, single-neuron RNA-seq, electrophysiology, neuronal maturation, gene co-expression, WGCNA

BACKGROUND

The human brain is a vastly complex structure, consisting of a multitude of cell types with different functions, morphologies and molecular compositions. Much of current scientific research is devoted to unraveling the immense complexity of this organ, whose cellular heterogeneity and intricate cellular connections guide a broad array of complex behaviors in health and disease. To dissect the diversity of cellular types and states, single-cell transcriptome profiling approaches have proven to be extremely useful, yielding new insights that are masked in conventional bulk analyses of heterogeneous samples [1, 2]. While single-cell transcriptomics represents a powerful method to reveal molecular features associated with specific cellular phenotypes, it is technically and ethically challenging to characterize the neurophysiology and molecular profile of single neurons in the human brain. Therefore, research investigations so far with the aim of unveiling the neurophysiological properties and/or molecular heterogeneity of neuronal cells have mostly used animal model organisms [3, 4], occasionally *ex vivo* brain slices or dissociated cell suspensions harvested post-mortem from human fetal or adult brain [5-7], and more recently *in vitro* human neuronal cultures derived from fibroblasts through induced pluripotent stem cell (iPSC) or direct conversion technologies [8-13]. The ability to generate potentially unlimited quantities of live human neurons using such cell reprogramming techniques offers unprecedented avenues for the modeling of cellular types and states associated with healthy and diseased conditions.

In a previous study (CHAPTER 4A) [8], we performed electrophysiological recordings, morphology and RNA-sequencing analysis on individual human stem cell-derived neurons with the aim of delineating at the molecular level the heterogeneity of their neurophysiological properties. Examining the expression profiles of different functional neuron types, we revealed a gene expression signature unique to and predictive of highly functional electrophysiological states [8]. As our categorization of functional types of neurons followed a continuum that relates to the stage of maturity of the neurons, further bioinformatics analyses of the expression profiles of these cells using more complex systems biology and network approaches may provide comprehensive insight into the general biological mechanisms underlying neuronal development and maturation.

Specifically, in this study, we assessed co-expression of groups of genes and performed functional annotation and gene association network analysis to infer valuable knowledge on the genes' functional relationships in light of their association with neurodevelopmental maturity. Applying the method of Weighted Gene Co-expression Network Analysis (WGCNA) [14] to our single-neuron RNA-seq data, we identified a co-expression gene module specifically associated with highly functional neurons. Further characterization of this module yielded several interesting findings. First, the identified co-expression module was strongly associated with various established measures of electrophysiological

and morphological neuronal maturity. Second, functional annotation analysis revealed, among the co-expression module, an overrepresentation of genes related to the mitochondrion electron transport chain and the cytoskeleton. Third, we found our module genes to play a general role in neuronal maturity, since a substantial proportion of the module genes showed enrichment and differential expression in a previously published single-cell whole-transcriptome study [7] comparing adult (mature) versus fetal (immature) brain neurons. Fourth, through gene association network analysis, we were able to identify a set of 21 putative transcription factors that may be involved in regulating, or driving towards, the highly functional neuronal state.

RESULTS

Identification of a Co-Expression Gene Module Specific to Highly Functional Neurons

We previously defined and characterized a continuum of immature to highly functional electrophysiological states of human neurons differentiated from embryonic and induced pluripotent stem cells (ESCs/iPSCs) by combining single-neuron patch clamping, morphological and transcriptome analyses *in vitro* (CHAPTER 4A) [8]. While expression correlation, differential expression and machine learning analyses revealed a unique molecular signature predictive of highly functional neuronal states [8], we additionally wondered whether we could detect sets of highly co-expressed genes, as such genes may exhibit commonality of function or share a role in specific biological pathways. Therefore, we applied the WGCNA approach to construct a co-expression network of the 5,000 highest connected genes, and identified groups (modules) of genes with highly similar expression profiles (**Figure 1a**). The constructed network fulfilled the criterion of approximate scale-free topology (**Supplementary Figure 1**) and, after merging highly correlated ($r > 0.75$) modules, consisted of 23 detected modules, labeled by colors. Though most modules were characterized by correlation of module gene expression in a single sample, one module (darkturquoise, containing 495 genes) appeared specific to highly functional (AP Type 5) but not immature (AP Types 1, 2 and 3) neurons (**Figure 1b,c**). Indeed, calculation of the darkturquoise module eigengene expression for each sample revealed a highly significant ($P < 0.0001$) upregulation of module genes in the highly functional neuron type (**Figure 1d**). Notably, this differential expression of co-expressed module genes was not related to the ESC or iPSC line from which the neurons were derived (**Figure 1c**).

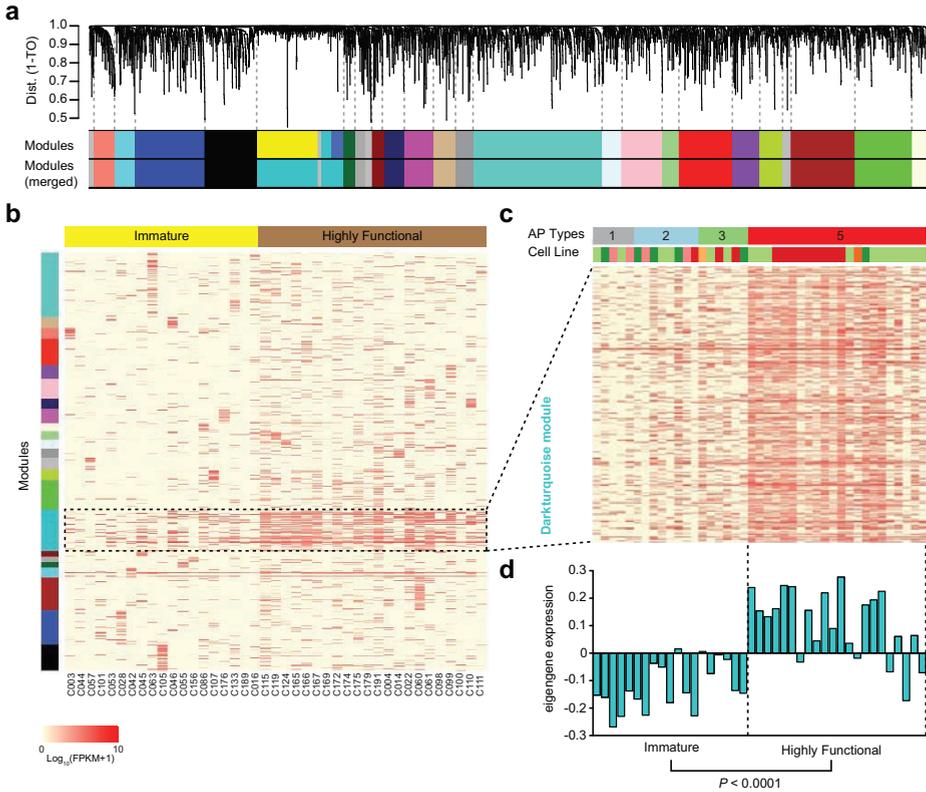


Figure 1. WGCNA identification of a gene co-expression module significantly activated in highly functional neurons. (a) Hierarchical clustering dendrogram of the 5,000 most connected genes. Modules of co-expressed genes identified by the WGCNA are labeled by colors and plotted underneath the dendrogram (top row of colors). Modules with very similar expression profiles, identified by high correlation ($r > 0.75$) of their eigengene values, were merged (bottom row of colors) prior to further analysis. (b) Heatmap plot of the expression data of the module genes (rows) for the immature ($n = 19$) and highly functional ($n = 22$) single neuron samples (columns). (c) Expression of darkturquoise module genes (495 genes in total) is positively associated with the highly functional (AP Type 5) neuronal state, and is not related to the pluripotent stem cell line from which the neurons were derived ($n = 2$ ESC lines: lightgreen, darkgreen; $n = 4$ iPSC lines: pink, yellow, orange, red). (d) Quantification of darkturquoise module eigengene expression reveals a highly significant upregulation of module genes in highly functional single neuron samples.

Correlation of Darkturquoise Module Gene Expression with Electrophysiological and Morphological Features of Neuronal Maturity

Confirming the association of the darkturquoise module with the highly functional neuronal state, we found strong correlations between the darkturquoise module's eigengene expression and several electrophysiological and morphological measures that are typical for highly functional (AP Type 5) neurons and associated with neuronal maturity (Figure 2). For instance, the firing rate ("AP frequency above -10 mV") of evoked action potentials, which represents an important measure of neuronal electrophysiological

maturation [8, 9], correlated strongly ($r = 0.73$, $P < 0.0001$) with the darkturquoise module eigengene. Of particular interest is the observation that the darkturquoise module was the only co-expression module that showed significant correlation ($r = 0.32$, $P < 0.05$) with the whole-cell capacitance, an electrophysiological estimate of intracellular volume. During neuronal differentiation and maturation, the membrane compartment and intracellular volume, and thus the capacitance, typically increase. By contrast, cell membrane resistance drops with increases in membrane area and is, therefore, typically lower in more mature neurons. The finding that the darkturquoise module eigengene was significantly anti-correlated with the membrane resistance ($r = -0.35$, $P < 0.05$), and strongly correlated to morphological measures of neuronal maturity (e.g., soma surface [$r = 0.66$, $P < 0.00001$], and dendritic arborization complexity measured as number of dendritic branch intersections with concentric circles in Scholl analysis [$r = 0.35$, $P < 0.05$]) suggests that genes of this co-expression module may play an important role in defining or regulating developmental neuronal maturation.

Potential Candidate Genes for Functional Neuronal Maturity

So far, our WGCNA analysis has revealed a strong association between the darkturquoise co-expression module and our categorization of functional types of neurons (**Figure 3a**), which reflect different stages of functional maturity of the neurons. For each of the 495 genes in this module, we determined the module membership (MM) and gene significance (GS) as the correlation of its expression value with, respectively, the darkturquoise module's eigengene and the AP Type classification. The observation that MM and GS are strongly and highly significantly correlated ($r_s = 0.6$, $P < 0.0001$; **Figure 3b**) implies that the highly connected module genes, or “hub genes”, are strongly associated with the highly functional (AP Type 5) as opposed to the immature (AP Types 1-2-3) neuronal state. Analysis of the 25 darkturquoise module members with highest MM and GS identified many genes (18/25) that we previously [8] associated specifically with highly functional neurons (**Figure 3c** and CHAPTER 4A). Importantly, *GDAP1L1*, the most important darkturquoise module gene based on combination of MM and GS, was previously recognized as the top neurophysiological gene candidate and proved useful for the isolation and enrichment of highly functional neurons *in vitro* [8].

Taken together, the results above strengthen the assertion that genes of the darkturquoise cluster mark, and potentially regulate, the functional maturity of neurons. Gene Ontology (GO) analysis of the entire module revealed enrichment for biological processes related to mitochondrion organization, adenosine triphosphate (ATP) metabolism, cellular respiration, and protein complex assembly (**Figure 3d**). The activation of a vast number of mitochondrion electron transport chain-related genes, including *ATP1B1*, *COX8A*, *NDUFB8*, *NDUFS5*, *NDUFV1*, *SLC25A4*, *UQCRC1* and *UQCRC2*, suggests an increased demand for energy by the more functional neuronal types. Additionally,

A co-expression gene network and transcriptional regulators for functional neuronal maturity

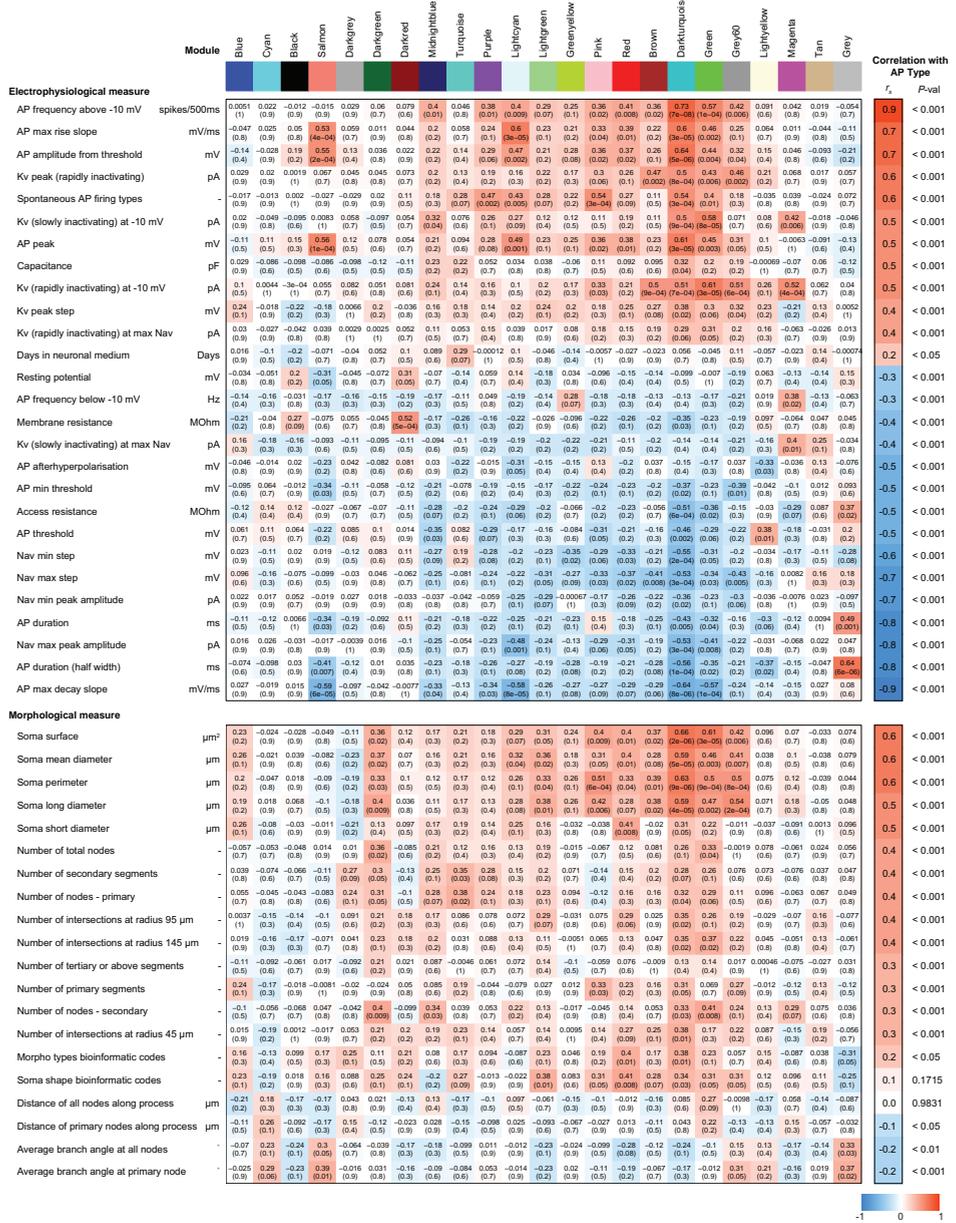
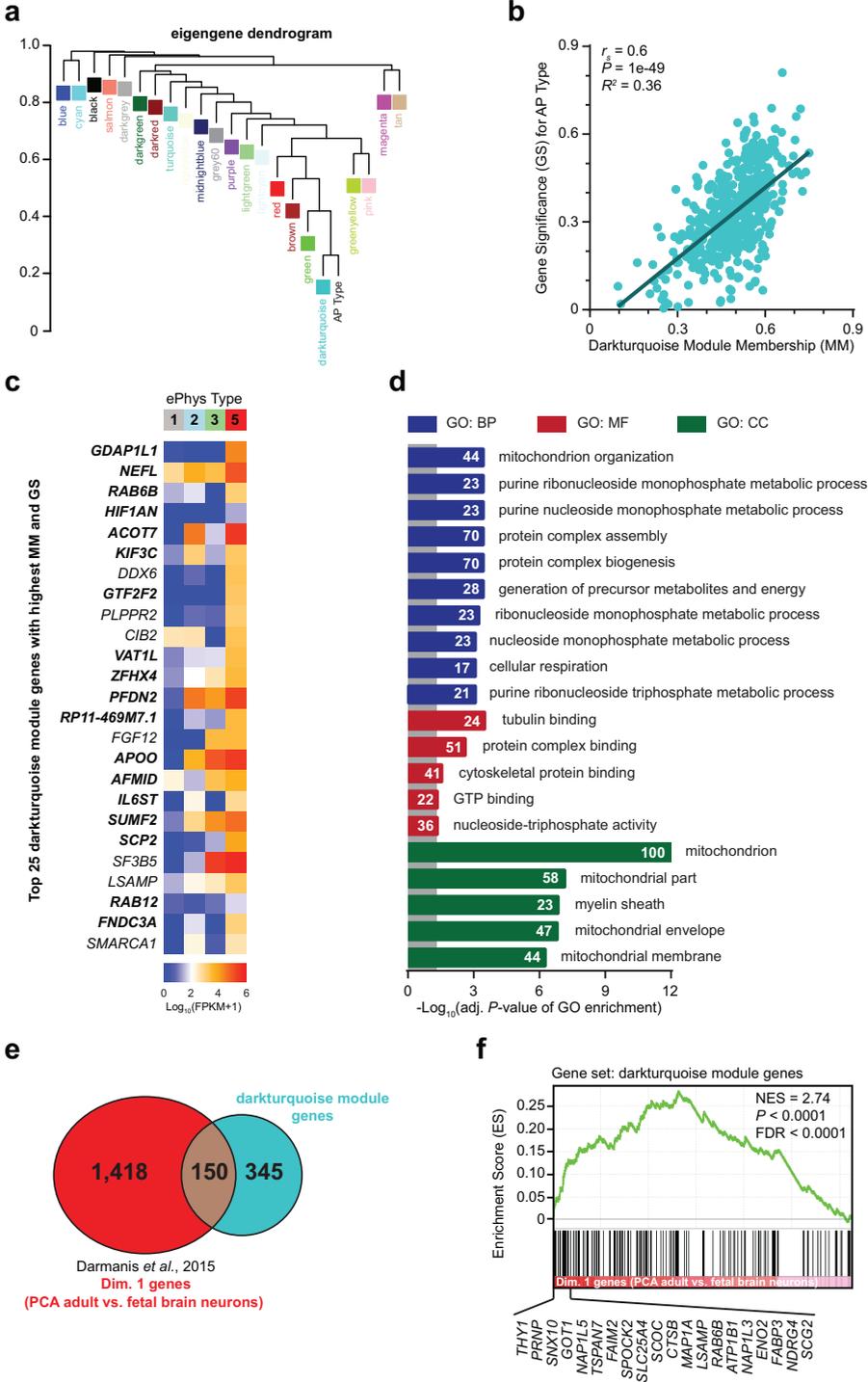


Figure 2. Module-trait correlations between WGCNA-identified modules and electrophysiological and morphological neuron features. Red color represents strong positive correlation while blue color represents strong negative correlation between the detected modules (columns) and the indicated measurement parameters (rows). White color indicates no correlation. Electrophysiological and morphological measures are ordered by decreasing Spearman's rank correlation (r_s) with the numerical classification (1-5) of AP Types (right column) [8]. The darkturquoise module showed the strongest and most significant correlation with various electrophysiological and morphological parameters typically associated with neuronal maturation (see text for details). P -values are in brackets.



← **Figure 3. Characterization of the darkturquoise WGCNA module associated with highly functional neurons.** (a) The classification of single neurons into highly functional (AP Type 5) and immature (AP Type 1-2-3) AP Types is most highly connected to the darkturquoise WGCNA module in hierarchical clustering. (b) Scatterplot of gene significance (GS) for AP Type as a function of module membership (MM) in the darkturquoise module. The strong and highly significant correlation ($r_s = 0.6$, $P < 0.0001$) between MM and GS implies that highly connected module (hub) genes are strongly associated with the highly functional neuronal state. (c) Heatmap plot of the 25 genes with highest GS and MM in the darkturquoise module. Gene names in bold font have previously been specifically associated with highly functional neuron types (Bardy *et al.*, 2016 [8]). (d) Gene Ontology (GO) analyses of the 495 darkturquoise module genes for overrepresented biological processes (BP), molecular functions (MF) and cellular components (CC). Length of bars represent Benjamini-Hochberg-corrected significance values (at $-\log_{10}$ scale) and numbers indicate number of genes annotated to GO term. Grey color indicates significance threshold. (e) 150 out of 495 (30%) darkturquoise module genes overlapped with genes that could distinguish postnatal (adult brain, mature) from prenatal (fetal brain, immature) neurons in a single-cell whole-transcriptome analysis of human cortex samples (Darmanis *et al.*, 2015 [7]). (f) Gene Set Enrichment Analysis (GSEA) revealed significant enrichment of 93 darkturquoise module genes at the top of Darmanis *et al.*'s [7] list of genes that could demarcate adult versus fetal brain neurons. Genes were ranked by their decreasing correlation with the PCA dimension explaining most variance between the adult and fetal brain neurons. The top 20 genes with highest enrichment score are listed below the enrichment plot. NES; normalized enrichment score.

we identified a subgroup of genes within the darkturquoise module that are related to cytoskeletal molecular functions (**Figure 3d**). Among these genes, *NEFL* (a major constituent of neurofilament) and *KIF3C* are particularly noteworthy since these genes have relatively high “hubness” in our darkturquoise cluster (**Figure 3c**) and are known to play a fundamental role in the regulation of axonal outgrowth [15, 16]. It may well be that the activation of these genes and the accompanying cytoskeletal rearrangements guide the maturation of neurons *in vitro* and induce the morphological properties that are characteristic of electrophysiological functional neurons (increased soma size and neurite arborization complexity [8]).

We subsequently asked whether our darkturquoise cluster of 495 putative neuronal maturation genes could be used to broadly define neuronal maturity. To this end, we performed a Gene Set Enrichment Analysis (GSEA) to test if the module was differentially expressed in a previously published single-cell RNA-seq dataset (Darmanis *et al.*, 2015) comparing neurons isolated from fetal and postnatal (adult) human cortex [7]. Interestingly, we found nearly one third (150/495) of darkturquoise module genes to be overlapping with those genes that could distinguish adult (mature) from fetal (immature) neurons (**Figure 3e**). The GSEA revealed that most of these genes (93/150) were significantly enriched at the top of Darmanis *et al.*'s [7] genes list ranked by decreasing correlation with the PCA dimension explaining most variance between adult and fetal cortical neurons ($P < 0.0001$ and $FDR < 0.0001$, Kolmogorov-Smirnoff test) (**Figure 3f**). Among the top enriched genes were several candidates known to be involved in, or related to, neuronal maturation. The candidate with the highest rank was *THY1*, a gene that encodes a glycoprotein that occurs on the surface of mature neurons [17]. *PRNP*, the second top

enriched gene, codes for prion protein (PrP^C), another cell surface glycoprotein whose expression level increases as neurons differentiate and mature [18, 19]. The enriched gene candidate *TSPAN7* is particularly noteworthy given its recognized role in the functional and morphological maturation of excitatory synapses [20]. Taken together, these results suggest an important role for darkturquoise module genes in general neuronal maturation. The 93 darkturquoise module genes significantly enriched with the maturation genes identified by Darmanis *et al.*'s [7] *ex vivo* study (**Supplementary Table 1**) might represent potentially interesting common biomarkers for human neuronal maturity.

Comprehensive Association Network Analysis of Functional Neuronal Maturity Candidate Genes

To study in more detail the relationship between genes of the darkturquoise module, we conducted a gene association network analysis using the Search Tool for the Retrieval of INteracting Genes (STRING) [21] and Cytoscape network analyzer [22]. Based on existing data on co-expression, protein interactions from curated databases, and experimental information, a gene association (protein-protein interaction) network was constructed comprising 67 interacting genes (**Figure 4**). The combination of WGCNA with gene association network construction provides important insights into the molecular system underlying the high functionality, and developmental maturity, of neuronal types. In line with the findings discussed above, GO network annotation revealed significant enrichment of genes related to the generation of energy and precursor metabolites, and the regulation of cytoskeletal organization. In addition, a collection of five genes (*YWHAG*, *EP300*, *DCC*, *THY1* and *ACTR3*) with a specific role in the regulation of neuron differentiation was identified as a distinct GO-enriched group. *EP300* encodes the histone acetyltransferase p300, which has been shown to favor differentiation and maturation of neural progenitors by acting as a transcriptional co-activator of *NEUROD1* and *BDNF* genes [23, 24]. *DCC* codes for a Netrin receptor that is required for Netrin-induced outgrowth and guidance of commissural axons [25]. The gene *YWHAG* seems to be crucially involved in cortical neuronal development as its functional ablation *in utero* has very recently been associated with morphological defects and delayed migration of developing cortical neurons [26]. *ACT3* is a member of the gene family of actin proteins, which form cytoskeletal microfilaments. Taken together, these five gene candidates seem to play an important role in neuronal developmental processes.

Identification of Putative Transcriptional Regulators of Functional Neuronal Maturity

To get insight into which transcription factors may act on top of the molecular network as putative regulators of the highly functional neuronal state, we integrated proximal transcription factor regulatory information from ENCODE into the constructed network. In

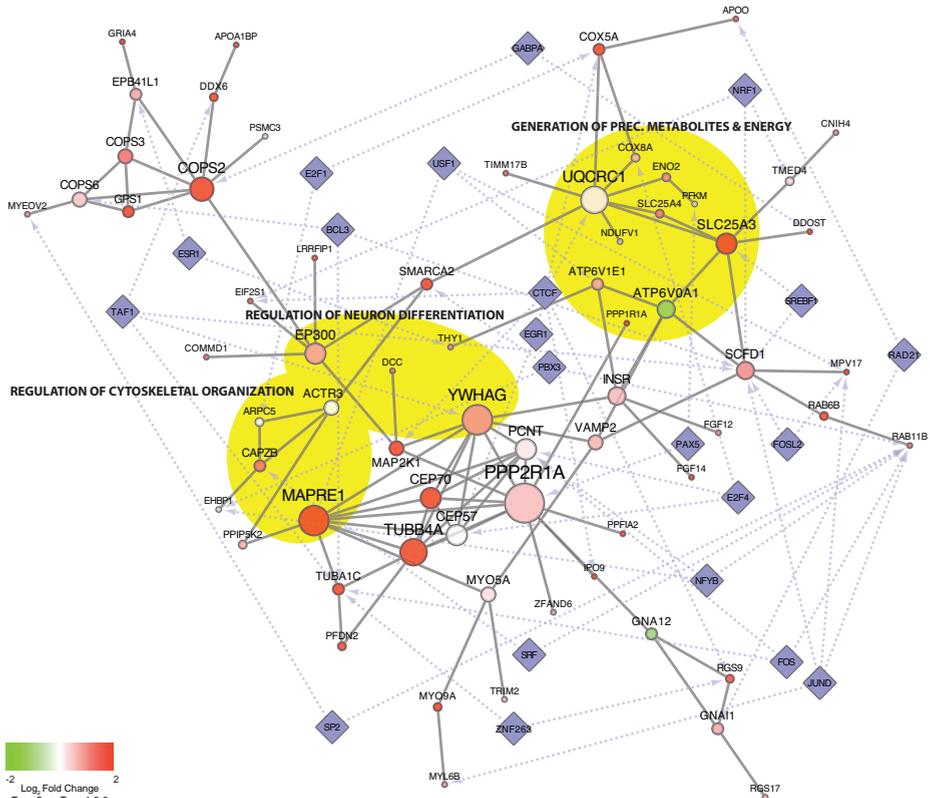


Figure 4. Gene association (protein-protein interaction) network of genes significantly associated with the highly functional neuronal state and the darkturquoise module. 271 genes with significant ($P < 0.05$) darkturquoise module membership and gene significance for AP Type were selected for gene association network construction with STRING database v10 [21] and Cytoscape 3.3.0 software [22]. Genes are represented by nodes whose color represents the fold change in expression level between highly functional and immature neurons (red color indicates upregulation while green color indicates downregulation of the gene in highly functional cells relative to immature types; see color scale). The size of the node is proportional to the degree (number of connections to other nodes). Proximal transcriptional regulators of genes in the network are shown as purple diamonds connecting to the respective genes. Nodes marked by yellow background color are members of major annotated GO biological process terms that are significantly ($P < 0.05$) overrepresented.

this analysis, we considered only those transcription factors regulating at least two gene targets. In total, we identified 21 transcription factors putatively regulating functional neuronal maturity (**Figure 4**, purple diamonds) that constitute interesting candidates for future molecular and functional analyses. Among these transcription factors, there are several that have formerly been associated with neuronal development and maturation. For instance, CTCF, the transcription factor regulating the largest number of genes (five) in our network, has a previously established role in regulating dendritic arborization and

spine formation during neuronal maturation [27]. Specifically, CTCF-deficient neurons have decreased dendritic complexity and spine density and show a reduced frequency of miniature excitatory postsynaptic currents, indicating that CTCF is important for functional neuronal development. Additionally, CTCF has been implicated in regulating neuronal diversity by controlling the expression of clustered *Pcdh* [27]. Another transcription factor, SRF, has been shown to mediate synaptic plasticity in the adult hippocampus and to be essential for specific cytoskeletal-dependent processes in the developing nervous system, including the outgrowth and guidance of neurites, and neuronal migration [28-31]. As neuronal activity-dependent expression of many immediate-early gene transcription factors is eliminated upon *SRF* gene depletion [29], the SRF transcription factor has been suggested to play a crucial role in converting synaptic activity into the structural (cytoskeletal) changes governing synaptic plasticity [32]. Our analysis of putative transcriptional regulators of functional neuronal maturity also identified several immediate early-gene transcription factors (*FOS*, *FOSL2* and *EGR1*) for which a specific role in synaptic plasticity has been previously established or proposed [33].

In addition to transcriptional regulators with an established role in the differentiation and maturation of neurons, we identified in this study several factors, including *ZNF263*, *NFYB*, *BCL3* and *RAD21*, for which a clear link with neuronal development has not been demonstrated. Future research should be directed toward investigating the possible role of these transcription factors in the context of functional neuronal maturity.

DISCUSSION

In our previous study that combined patch-clamp electrophysiology with whole-transcriptome analysis of single human stem cell-derived neurons *in vitro*, we revealed a unique molecular signature predictive of highly functional neuronal states (CHAPTER 4A) [8]. Although differential expression, expression correlation and machine learning approaches in this study identified many new neuronal genes related to electrophysiological activity [8], these analyses may not have uncovered the complete molecular programs underlying or driving electrophysiological neuronal maturity. In search for such programs, we applied in this study the method of WGCNA with the aim of detecting sets of highly co-expressed genes likely to share common function in biological processes or pathways, and that are associated with our electrophysiological phenotypes.

In summary, our WGCNA analysis revealed a module (darkturquoise, 495 genes) of highly interconnected genes whose expression was strongly upregulated in highly functional neurons, compared with immature neuron types. Correlation analyses indicated that this gene module was strongly associated with various electrophysiological and morphological measures of neuronal maturity. To gain insight into the molecular functions

of the genes in the co-expression module, we performed GO enrichment analysis, which revealed a robust activation of cytoskeletal and mitochondrion electron transport-chain related genes. Noteworthy, a recent study that also coupled patch-clamp recording with single-neuron transcriptomics found similar activation of respiratory chain machinery genes in electrophysiologically mature phenotypes of neurons, confirming the notion that high energy utilization accompanies electrophysiological maturity [9]. As cytoskeletal rearrangements, specifically microtubule and microfilament dynamics, are importantly involved in the maturation of neurons *in vitro* and *in vivo* [34, 35], we propose that the observed activation of cytoskeletal gene expression may play a role in inducing the neuromorphological features characteristic of electrophysiologically highly functional neurons (i.e., a more complex dendritic architecture and enlarged soma, see ref. [8]).

Providing evidence for the notion that genes of our identified co-expression module have a general role in neuronal maturity, we found a substantial proportion of the module (93/495 genes) to be enriched and differentially expressed in a previously published single-cell whole-transcriptome study comparing neurons isolated from adult (mature) and fetal (immature) brain [7]. Among the top enriched genes, we identified several candidates with a known function in neuronal maturation, including *THY1*, *PRNP* and *TSPAN7* (see Results for details). The 93 significantly enriched genes are summarized in **Supplementary Table 1** and may constitute generic biomarker candidates for assessing human neuronal maturity.

With the growing use of human stem cell-based models for brain-related disease modeling *in vitro*, a thorough characterization of the variance of functional states within populations of differentiated neurons has become crucially important. Human stem cell-derived neuronal progenitors mature into electrophysiologically functional neurons at different rates, and to avoid false positive findings and potential masking of disease features, it is important to compare neurons of equivalent functional maturity [8, 36]. A greater understanding of the molecular system that regulates the highly functional, or developmentally more mature, state of human neuronal types might stimulate the development of genetic engineering strategies with which one can control (or accelerate) functional maturation and potentially generate more homogeneous populations of electrophysiologically comparable neurons for research investigations. One of the most important results of our study is the identification of a set of 21 putative transcriptional regulators that may be involved in controlling (or inducing) the highly functional neuronal state by regulating cytoskeletal genes, genes related to neuron differentiation, and genes associated with the generation of energy and precursor metabolites (**Figure 4**). Several of the identified transcription factors, such as CTCF, SRF and the immediate early-gene transcription factors FOS, FOSL2 and EGR1, have a previously established role in synaptic plasticity associated with neuronal development. Other identified transcription factor genes (e.g., *ZNF263*, *NFYB*, *BCL3* and *RAD21*) represent interesting new

candidates for future research. Small molecule compounds that modulate the expression or activity of these or the aforementioned transcription factors represent promising tools for further experimental investigations.

METHODS

Human Neuronal Culture

Human dermal fibroblasts were reprogrammed into pluripotent cells with the four Yamanaka factors (Oct3/4, Sox2, Klf4 and c-Myc), either in a retroviral vector or a non-integrating Sendai viral vector. Human iPSCs and embryonic stem cells (ESCs) were differentiated into neural progenitor cells (NPCs) as previously described [37]. NPCs were expanded for 3 to 5 passages (split 1:2 or 1:3 per passage) and stored at -80 °C. Then they were thawed and expanded for at least one more passage in neural progenitor medium (DMEM/F12 supplemented with FGF8, SHH, B27 and N2, without retinoic acid). Dissociated NPCs were directly plated on glass coverslips (Fisher Scientific, Cat. No. 12-545-80) coated with poly-ornithine (Sigma, Cat. No. P3655) and laminin (Invitrogen, Cat. No. 23017-015) in 24-well plates. Twenty-four hours later, the cells were switched gradually (half medium change) to neuronal medium: BrainPhys basal [10] supplemented with 1x N2 (Gibco, Cat No. 17502-048), 1x B27 (Gibco, Cat No. 17504-044), brain-derived neurotrophic factor (BDNF, 20 ng/ml; Peprotech, Cat No. 450-02), glia-derived neurotrophic factor (GDNF, 20 ng/ml; Peprotech, Cat No. 450-10), ascorbic acid (AA, 200 nM; Sigma, Cat No. A0278), dibutyryl cyclic AMP (cAMP, 1mM Sigma, Cat No. D0627), and laminin (1 µg/ml; Invitrogen, Cat No 23017-015). Half of the neuronal medium was gently replaced two to three times a week. The plates were kept in a humidified incubator at 37°C with 5% CO₂ and 21% O₂. The pH (~7.3-7.4) and the osmolarity (~300-305 mOsmol/L) of medium were maintained constant over time.

Patch Clamping

For whole-cell patch-clamp recordings, individual coverslips were transferred into a heated recording chamber and continuously perfused (1 ml·min⁻¹) with either BrainPhys basal medium or artificial cerebrospinal fluid (ACSF) bubbled with a mixture of CO₂ (5%) and O₂ (95%) and maintained at 25°C. The composition of ACSF was adjusted to match the inorganic salt concentration and osmolarity of BrainPhys basal. ACSF contained (in mM) 121 NaCl, 4.2 KCl, 1.1 CaCl₂, 1 MgSO₄ (or 0.4 MgSO₄ and 0.3 MgCl), 29 NaHCO₃, 0.45 NaH₂PO₄·H₂O, 0.5 Na₂HPO₄ and 20 glucose (all chemicals from Sigma).

For single-cell patch-clamp electrophysiological recordings, we used a digidata 1440A/Multiclamp 700B and Clampex 10.3 (Molecular Devices). Patch electrodes were filled with internal solutions containing 130 mM K-gluconate, 6 mM KCl, 4 mM NaCl, 10

mM Na-HEPES, 0.2 mM K-EGTA; 0.3 mM GTP, 2 mM Mg-ATP, 0.2 mM cAMP, 10 mM D-glucose, 0.15% biocytin and 0.06% rhodamine. The pH and osmolarity of the internal solution were close to physiological conditions (pH 7.3, 290–300 mOsmol). Data were all corrected for liquid junction potentials (10 mV). Electrode capacitances were compensated on-line in cell-attached mode (~7 pF). Recordings were low-pass filtered at 2 kHz, digitized, and sampled at intervals of 50 ms (20 kHz). To control the quality and the stability of the recordings throughout the experiments, access resistance, capacitance and membrane resistance were continuously monitored on-line and recorded. The resistance of the patch pipettes was between 3 and 5 MOhm. The access resistance of the cells in our sample was ~40 MOhm on average. Spontaneous synaptic AMPA events were recorded at the reversal potential of Cl⁻ and could be reversibly blocked by AMPA receptor antagonist (10 μ M NBQX, Sigma Ref#N183). Spontaneous synaptic GABA events were recorded at the reversal potential of cations and could be reversibly blocked with GABA_A receptor antagonist (10 μ M SR95531, Sigma Ref#S106). Statistical analysis of electrophysiology data was assisted with Clampfit 10.3, Matlab 2014b, Igor Pro 6, Prism 6, MiniAnalysis, Excel and custom-made Python programs.

Electrophysiological Recording Protocol

The same electrophysiological protocol was applied to all neurons. Whole cell patch-clamp recordings were obtained in the absence of any receptor antagonists in ACSF or BrainPhys basal. After breaking the membrane seal, we first maintained the cell in voltage clamp (VC) at -70 mV by injecting small amounts of current when necessary (“Baseline at -70 mV”). We applied a test pulse of -5 mV to measure passive cell properties (membrane resistance, access resistance, capacitance). Voltage-dependent sodium and potassium channels properties were examined with an IV curve. Spontaneous synaptic events mediated by AMPA receptors were recorded for at least 5 min at -70 mV (close to Cl⁻ reversal potential). The nature of AMPA synaptic events was confirmed on a small subset of cells ($n = 15$, excluded from RNA-seq) by reversible blockade with AMPA receptor antagonist (10 μ M NBQX). The patch was then switched to current clamp. Small currents were injected if necessary to maintain the cell resting potential around -70 mV. Current steps (increments 2 pA, $n = 15$, 500 ms) were applied to measure the properties of evoked action potentials (APs). Next, spontaneous APs were recorded for 1 to 5 min at -70 mV and then at resting potential without injecting any current (CC 0 pA). The patch was then reversed to voltage clamp and slowly increased to the reversal potential of cations (~0 mV). Once the baseline stabilized, we recorded spontaneous GABAergic synaptic activity for 5 to 10 min. The nature of GABA synaptic events was confirmed on a small subset of cells ($n = 5$, excluded from RNA-seq) by reversible blockade with GABA_A receptor antagonist (10 μ M Gabazine).

AP Types Classification

“Type 0 cells” did not express voltage-dependent sodium currents and were excluded from analysis. “Type 1 neurons” expressed small Nav currents but were not able to fire APs above -10 mV. The arbitrary limit of -10 mV was chosen as it is close to the reversal potential of cations (0 mV). Healthy APs usually reach or overshoot the reversal potential of cations. “Type 2 neurons” fired one AP above -10 mV, which was followed by a plateau. “Type 3 neurons” also fired one AP above -10 mV and one or a few aborted spikes below -10 mV. “Type 4 neurons” fired more than one AP above -10 mV but at a frequency below 10 Hz. “Type 5 neurons” fired APs above -10 mV at 10 Hz or more. Our categorization of functional types of neurons followed a continuum that relates to the stage of maturity of the neurons. While Type 1 neurons are considered immature, Type 5 neurons are considered more mature and functionally active. Although after ~3-6 weeks of maturation we found on average ~75% of Type 4-5 neurons, we found every AP Type of cell at most differentiation time points we looked at (range of 2 weeks to 5 months after NPC maturation). This finding suggests some degree of variability in the electrophysiological maturity of neurons even within cultures of the same age. Remarkably, in our samples, the large majority of cells receiving active excitatory synapses were Type 5 neurons and we almost never found clear, spontaneously active AMPA and GABA synaptic inputs in Type 1 neurons and rarely in Type 2.

Neuromorphometry

Images of the patch-clamped cells were taken with a 40X water-immersion objective, differential interference contrast filters (all Olympus), an infrared digital camera (Rolera XR -Qimaging), and an epifluorescence mercury lamp. Mosaics of the several fields of views to cover the span of the neurites were taken and stitched later on NeuroLucida (MBF Biosciences). Before patch clamping, the targeted cells were imaged with DIC and epifluorescence (most cells patched expressed Synapsin-GFP). The entire cells were filled with a red fluorescent dye (rhodamine) and another set of images was taken rapidly after completion of the electrophysiological recordings. All the images were stitched and analyzed on NeuroLucida. The soma size and shape were measured on the DIC images before patching to avoid deformation by the patch pipette. The images from the intracellular rhodamine dye, occasionally supplemented by GFP vectors, were used to guide the neuromorphological reconstruction. For all cells, the morphology was reconstructed within a radius of 150 μm . Many cell processes reached beyond that distance but, for homogeneity of the sample, we restricted the analysis to that arbitrary radius. Occasionally, patched cells with pictures of insufficient quality were excluded from the analysis. When necessary, some processes partially hidden by the patch electrode were extrapolated.

Single-Neuron Collection and cDNA Preparation

Following electrophysiological recording, we applied slight additional negative pressure to establish a stronger connection between cell and patch electrode, but without completely aspirating the cell in the pipette. The neuron, including its processes (axon and dendrites), was then transferred in a volume of $\sim 2 \mu\text{l}$ of internal patch solution into a PCR tube containing $8.0 \mu\text{l}$ sample buffer by slowly retracting the patch pipette from the chamber bath and breaking the tip of the electrode along the inside wall of the tube. Sample buffer was made according to the protocol supplied by Clontech with the SMARTer™ Ultra Low RNA Kit and included both ERCC spike-ins (Life Technologies) and ArrayControl spikes 1, 4 and 7 (RNA standard for the Fluidigm C1) for the purpose of quality control (QC). Successful removal of the cell from the coverslip was always confirmed by DIC optics. Collected single cells were immediately spun down and put on ice for SMARTer™ cDNA synthesis following manufacturer's instructions (Clontech). Briefly, first-strand cDNA was synthesized from poly(A)+ RNA by incubation with $1 \mu\text{l}$ of 3' SMART CDS Primer II A ($24 \mu\text{M}$) for 3 min at 72°C , followed by reverse transcription in a $20\text{-}\mu\text{l}$ final reaction volume using 200 units of SMARTScribe™ Reverse Transcriptase for 90 min at 42°C and inactivation for 10 min at 70°C . First-strand cDNA was then purified using Agencourt AMPure XP SPRI Beads (Beckman Coulter Genomics) and amplified by long-distance PCR using the Advantage® 2 PCR Kit (Clontech) with the following PCR thermocycler program: 95°C for 1 min, 18 cycles of 95°C for 15 s, 65°C for 30 s, 68°C for 6 min, and 72°C for 10 min. PCR-amplified double-stranded (ds) cDNA was immobilized onto SPRI beads, purified by two washes in 80% ethanol, and eluted in $12 \mu\text{l}$ of purification buffer (Clontech). The quality (Agilent 2100 Bioanalyzer High Sensitivity DNA Kit; Agilent) and quantity (Qubit dsDNA High Sensitivity Assay Kit; Invitrogen) of each ds cDNA sample were assessed prior to library preparation.

qPCR Assays

To select the single-cell cDNA samples suitable for mRNA-seq, we determined the level of expression of *GAPDH* and *ACTB* in each sample by TaqMan real-time PCR. Samples with Ct values ≤ 30 for both housekeeping genes were typically considered positive for library preparation. For each gene, duplicate $10\text{-}\mu\text{l}$ PCR reactions were performed on an ABI Prism 7900 Sequence Detector (Applied Biosystems) using $0.50 \mu\text{l}$ of 1:5-diluted ds cDNA template in standard TaqMan Gene Expression Assay with FAM reporter. Real-time PCR assays for detection of the ERCCs and ArrayControl RNA spikes were performed using, respectively, standard TaqMan Gene Expression Assays (Life Technologies) and SYBR Green PCR Master Mix (Applied Biosystems) with custom primers (Fluidigm).

Illumina Transcriptome Library Preparation and Sequencing

Construction of single-cell mRNA-seq libraries was typically performed with 0.25 ng of input cDNA using the Nextera XT DNA sample prep kit (Illumina) with modified protocol. Briefly, cDNA was tagged for 5 min at 55°C in a 5- μ l reaction containing 2.5 μ l of Tagment DNA Buffer and 1.25 μ l of Amplicon Tagment Mix; tagmentation was neutralized with 1.25 μ l of Neutralize Tagment Buffer for 5 min. Tagmented DNA was then subjected to 12-cycle PCR amplification using 3.75 μ l of Nextera PCR Master Mix and 1.25 μ l each of index 1 (i7) and index 2 (i5) library-identifying (barcoded) sequencing primers. The constructed libraries were run on a 1.5% agarose gel in Tris-borate/EDTA buffer, stained with SYBR Gold (Invitrogen), and size selected for ~300-400 or ~300-650 bp (insert size of ~165-265 or ~165-515 bp, respectively). Gel-excised library fragments were purified with the Wizard SV Gel and PCR Clean-Up System (Promega), eluted in 40 μ l of nuclease-free water, and concentrated by speedvacuum centrifugation. Each library was then quantified (Qubit dsDNA High Sensitivity Assay Kit; Invitrogen) and examined for correct size (Agilent 2200 TapeStation High Sensitivity D1K Screen-Tape Assay; Agilent), after which equimolar amounts of uniquely barcoded libraries were pooled together and used for cluster generation and 100-bp paired-end sequencing on a HiSeq 2000 or 2500 sequencer (Illumina).

Sequencing Data Normalization and Filtering

The raw sequencing read data from $n = 56$ patched human neurons, which passed a series of QC criteria, was mapped to the human reference transcriptome (Gencode v19) using gapped-alignment strategies. Alignment was performed by STAR (version 2.3.0) [38] followed by gene-level quantification with HTseq (version 0.6.1) [39]. An initial pre-filtering step was performed to remove any genes with null expression across all samples. The count data was then corrected for sequencing depth and normalized using the regularized-logarithm (rlog) transformation from the DESeq2 package, which effectively compresses differences for the low count genes and stabilizes the variance across the mean [40].

Differential Expression Analysis

The five electrophysiological types of differentiated neurons were stratified into “immature” (Types 1-2-3), “transitional” (Type 4) and “highly functional” (Type 5) neurons. Log₂ fold changes in gene expression between groups were calculated from rlog-normalized count data with DESeq2 [40] after an additional filtering step (removal of the 35% of lowest expressed genes) that increased detection power while keeping the same experiment-wide type I error (false discovery rate) of 5%. The P -value of differential expression was adjusted using the Benjamini-Hochberg correction for multiple comparisons ($P < 0.05$ considered significant).

Weighted Gene Co-expression Network Analysis

Because of our interest to identify genes and regulators specifically associated with highly functional neurons and distinct from immature AP Types, we omitted any transitional (Type 4) cells in our co-expression gene network construction so as to have two clear, phenotypically distinct sample groups of nearly equal sample size (immature: $n = 19$; highly functional: $n = 22$). Due to computational restrictions, the normalized count data was reduced to the 10,000 most varying genes as genes that do not change expression across samples are assumed to yield little information in a co-expression network [41]. To identify groups (modules) of genes whose expression profile is highly correlated, we constructed a weighted co-expression network using the WGCNA R software package [42]. Co-expression similarity was calculated for all gene pairs as the absolute value of their expression profiles' Pearson correlation, and converted into a network adjacency measure by raising the similarity to a power β of 7, chosen based on the scale-free topology criterion (see also **Supplementary Figure 1**) [43]. Module detection and network analysis were restricted to the top 5,000 genes with highest connectivity (sum of adjacencies to all other genes). This does not result in considerable information loss as the most highly connected genes are believed to be biologically the most important [44], and genes within a particular module tend to naturally have high connectivity [45]. The topological overlap (TO) of all gene pairs, representing their relative interconnectivity, was calculated from the adjacencies and transformed into a dissimilarity measure ($=1-TO$) for use in hierarchical clustering to discriminate modules. The static tree cut method was applied to identify clusters of highly co-expressed genes (minimum module size of 50 genes) as branches of the dendrogram. All modules were assigned a color and summarized by module eigengenes (MEs) explaining the maximal amount of variance in module expression levels. The "grey" module contained all genes that could not be assigned to a co-expression module. Highly similar modules were merged if their eigengenes' correlation was greater than 0.75. MEs of all modules were Pearson correlated with the measured electrophysiological and morphological parameters to estimate module-trait relationships and extract potentially biologically relevant modules for further analysis. The module membership MM of each gene was calculated as the (Pearson) correlation between its expression value and the module eigengene, while the gene significance GS for AP Type was computed as the (Spearman) correlation of a gene's expression value with the AP Type classification. Heat maps of the modules were created via the heatmap.2 function implemented in R's gplot package, using log-normalized FPKM values of genes comprising the module. Gene Ontology (GO) enrichment analysis of module genes was performed using ToppFun (<https://toppgene.cchmc.org/>, database May 2016) to test for overrepresentation in biological process (BP), molecular function (MF) and cellular component (CC) categories. Benjamini-Hochberg adjustment of P -values was applied to

correct for multiple testing, with enriched terms being considered significant for adjusted P -values < 0.05 .

Gene Set Enrichment Analysis

GSEA [46] was performed to explore whether the darkturquoise module genes, associated with highly functional neurons, showed significant correlation with genes that could distinguish postnatal (adult brain) from prenatal (fetal brain) neurons in a single-cell transcriptome study (Darmanis *et al.*, 2015) [7] of human cortical neuron samples. Input data were (i) a list of genes from Darmanis *et al.* [7] ranked by their decreasing correlation with PCA dimension 1 explaining most variance between adult and fetal brain neurons, and (ii) the set of 495 darkturquoise module genes. The ranked gene list was analyzed for overrepresentation of the darkturquoise module genes at the top using GSEA's "pre-ranked" tool with default parameters (enrichment statistic: weighted; normalization mode: meandiv; 1000 gene-set permutations).

Gene Association Network

The interaction of module genes (proteins) was investigated with STRING database (<http://string-db.org>) v10 [21]. Genes having significant module membership and GS for ePhys type were selected as input and the subsequent gene association (protein-protein interaction) network was constructed using known data on co-expression, protein interactions from curated databases, and experimental information. Network visualization and analysis was performed with Cytoscape 3.3.0 [22]. To identify potentially important regulatory players of the network, ENCODE proximal transcription factor regulation information was integrated into the network analysis with the CyTargetLinker Cytoscape plug-in [47]. Only transcription factors regulating at least two targets were included in the network. Intra-network analysis of overrepresented GO biological process terms was performed using the BiNGO plug-in for Cytoscape [48].

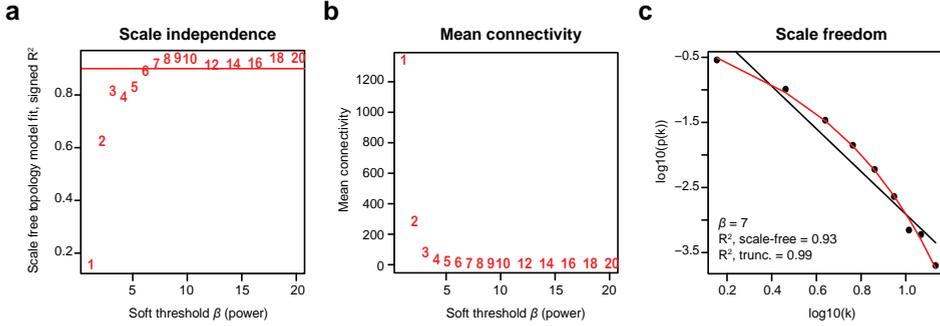
REFERENCES

1. Toriello NM, Douglas ES, Thaitrong N *et al.* Integrated microfluidic bioprocessor for single-cell gene expression analysis. *Proc. Natl. Acad. Sci. U. S. A.* 105(51), 20173-20178 (2008).
2. Patel AP, Tirosh I, Trombetta JJ *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344(6190), 1396-1401 (2014).
3. Zeisel A, Munoz-Manchado AB, Codeluppi S *et al.* Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 347(6226), 1138-1142 (2015).
4. Zhang Y, Chen K, Sloan SA *et al.* An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* 34(36), 11929-11947 (2014).
5. Moore AR, Filipovic R, Mo Z, Rasband MN, Zecevic N, Antic SD. Electrical excitability of early neurons in the human cerebral cortex during the second trimester of gestation. *Cereb. Cortex* 19(8), 1795-1805 (2009).
6. Moore AR, Zhou WL, Jakovcevski I, Zecevic N, Antic SD. Spontaneous electrical activity in the human fetal cortex in vitro. *J. Neurosci.* 31(7), 2391-2398 (2011).
7. Darmanis S, Sloan SA, Zhang Y *et al.* A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. U. S. A.* 112(23), 7285-7290 (2015).
8. Bardy C, Van Den Hurk M, Kakaradov B *et al.* Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology. *Mol. Psychiatry* (2016, in press).
9. Chen X, Zhang K, Zhou L *et al.* Coupled electrophysiological recording and single cell transcriptome analyses revealed molecular mechanisms underlying neuronal maturation. *Protein & Cell* 7(3), 175-186 (2016).
10. Bardy C, Van Den Hurk M, Eames T *et al.* Neuronal medium that supports basic synaptic functions and activity of human neurons in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 112(20), E2725-2734 (2015).
11. Nicholas CR, Chen J, Tang Y *et al.* Functional maturation of hPSC-derived forebrain interneurons requires an extended timeline and mimics human neural development. *Cell Stem Cell* 12(5), 573-586 (2013).
12. Wu H, Xu J, Pang ZP *et al.* Integrative genomic and functional analyses reveal neuronal subtype differentiation bias in human embryonic stem cell lines. *Proc. Natl. Acad. Sci. U. S. A.* 104(34), 13821-13826 (2007).
13. Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Sudhof TC, Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463(7284), 1035-1041 (2010).
14. Zhao W, Langfelder P, Fuller T, Dong J, Li A, Hovarth S. Weighted gene coexpression network analysis: state of the art. *J. Biopharm. Stat.* 20(2), 281-300 (2010).
15. Gumy LF, Chew DJ, Tortosa E *et al.* The kinesin-2 family member KIF3C regulates microtubule dynamics and is required for axon growth and regeneration. *J. Neurosci.* 33(28), 11329-11345 (2013).
16. Yuan A, Rao MV, Veeranna, Nixon RA. Neurofilaments at a glance. *J. Cell Sci.* 125(Pt 14), 3257-3263 (2012).
17. Morris R, Grosveld F. Expression of Thy-1 in the nervous system of the rat and mouse. *Immunol. Ser.* 45 121-148 (1989).
18. Peralta OA, Huckle WR, Eyestone WH. Expression and knockdown of cellular prion protein (PrP^C) in differentiating mouse embryonic stem cells. *Differentiation* 81(1), 68-77 (2011).
19. Steele AD, Emsley JG, Ozdinler PH, Lindquist S, Macklis JD. Prion protein (PrP^C) positively regulates neural precursor proliferation during developmental and adult mammalian neurogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 103(9), 3416-3421 (2006).

20. Bassani S, Passafaro M. TSPAN7: A new player in excitatory synapse maturation and function. *Bioarchitecture* 2(3), 95-97 (2012).
21. Szklarczyk D, Franceschini A, Wyder S *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447-452 (2015).
22. Shannon P, Markiel A, Ozier O *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13(11), 2498-2504 (2003).
23. Sharma A, Moore M, Marcora E *et al.* The NeuroD1/BETA2 sequences essential for insulin gene transcription colocalize with those necessary for neurogenesis and p300/CREB binding protein binding. *Mol. Cell. Biol.* 19(1), 704-713 (1999).
24. Chatterjee S, Mizar P, Cassel R *et al.* A novel activator of CBP/p300 acetyltransferases promotes neurogenesis and extends memory duration in adult mice. *J. Neurosci.* 33(26), 10698-10712 (2013).
25. Keino-Masu K, Masu M, Hinck L *et al.* Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. *Cell* 87(2), 175-185 (1996).
26. Wachi T, Cornell B, Marshall C, Zhukarev V, Baas PW, Toyo-Oka K. Ablation of the 14-3-3gamma Protein Results in Neuronal Migration Delay and Morphological Defects in the Developing Cerebral Cortex. *Dev. Neurobiol.* 76(6), 600-614 (2016).
27. Hirayama T, Tarusawa E, Yoshimura Y, Galjart N, Yagi T. CTCF is required for neural development and stochastic expression of clustered Pcdh genes in neurons. *Cell Rep.* 2(2), 345-357 (2012).
28. Knoll B, Kretz O, Fiedler C *et al.* Serum response factor controls neuronal circuit assembly in the hippocampus. *Nat. Neurosci.* 9(2), 195-204 (2006).
29. Ramanan N, Shen Y, Sarsfield S *et al.* SRF mediates activity-induced gene expression and synaptic plasticity but not neuronal viability. *Nat. Neurosci.* 8(6), 759-767 (2005).
30. Etkin A, Alarcon JM, Weisberg SP *et al.* A role in learning for SRF: deletion in the adult forebrain disrupts LTD and the formation of an immediate memory of a novel context. *Neuron* 50(1), 127-143 (2006).
31. Alberti S, Krause SM, Kretz O *et al.* Neuronal migration in the murine rostral migratory stream requires serum response factor. *Proc. Natl. Acad. Sci. U. S. A.* 102(17), 6148-6153 (2005).
32. Knoll B, Nordheim A. Functional versatility of transcription factors in the nervous system: the SRF paradigm. *Trends Neurosci.* 32(8), 432-442 (2009).
33. Lyons MR, West AE. Mechanisms of specificity in neuronal activity-regulated gene transcription. *Prog. Neurobiol.* 94(3), 259-295 (2011).
34. Hlushchenko I, Koskinen M, Hotulainen P. Dendritic spine actin dynamics in neuronal maturation and synaptic plasticity. *Cytoskeleton* doi:10.1002/cm.21280 (2016).
35. Shirao T, Gonzalez-Billault C. Actin filaments and microtubules in dendritic spines. *J. Neurochem.* 126(2), 155-164 (2013).
36. Sandoe J, Eggen K. Opportunities and challenges of pluripotent stem cell neurodegenerative disease models. *Nat. Neurosci.* 16(7), 780-789 (2013).
37. Boyer LF, Campbell B, Larkin S, Mu Y, Gage FH. Dopaminergic differentiation of human pluripotent cells. *Curr. Protoc. Stem Cell Biol.* Chapter 1 Unit1H 6 (2012).
38. Dobin A, Davis CA, Schlesinger F *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1), 15-21 (2013).
39. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31(2), 166-169 (2015).
40. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15(12), 550 (2014).

41. Kogelman LJ, Cirera S, Zhernakova DV, Fredholm M, Franke L, Kadarmideen HN. Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue RNA Sequencing in a porcine model. *BMC Med. Genomics* 7 57 (2014).
42. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9 559 (2008).
43. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* 4 Article17 (2005).
44. Jeong H, Mason SP, Barabasi AL, Oltvai ZN. Lethality and centrality in protein networks. *Nature* 411(6833), 41-42 (2001).
45. Carlson MR, Zhang B, Fang Z, Mischel PS, Horvath S, Nelson SF. Gene connectivity, function, and sequence conservation: predictions from modular yeast co-expression networks. *BMC Genomics* 7 40 (2006).
46. Subramanian A, Tamayo P, Mootha VK *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102(43), 15545-15550 (2005).
47. Kutmon M, Kelder T, Mandaviya P, Evelo CT, Coort SL. CyTargetLinker: a cytoscape app to integrate regulatory interactions in network analysis. *PLoS One* 8(12), e82160 (2013).
48. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21(16), 3448-3449 (2005).

SUPPLEMENTARY INFORMATION



Supplementary Figure 1. The constructed weighted gene co-expression network fulfills the criterion of approximate scale-free topology. (a) Plot of the scale-free topology index (R^2 , y axis) against different soft thresholding powers (β , x axis). For network construction, we chose a β value of 7, for which the R^2 crossed 0.9 for the first time and reached saturation. The higher the R^2 value, the closer the network is to an ideal scale-free system. **(b)** Mean connectivity (y axis) against the power β . Higher powers result in lower mean connectivity. **(c)** Scale-free plot for the weighted network (power $\beta = 7$).

Supplementary Table 1. Putative generic biomarkers for human neuronal maturity revealed by Gene Set Enrichment Analysis of darkturquoise module genes.

Gene Symbol	Gene Name	Rank in Gene List	Rank Metric Score	Running Enrichment Score
THY1	Thy-1 cell surface antigen	0	0.778	0.0100
PRNP	prion protein	5	0.738	0.0167
SNX10	sorting nexin 10	7	0.736	0.0255
GOT1	glutamic-oxaloacetic transaminase 1	9	0.733	0.0342
NAP1L5	nucleosome assembly protein 1 like 5	13	0.720	0.0414
TSPAN7	tetraspanin 7	18	0.708	0.0477
FAIM2	Fas apoptotic inhibitory molecule 2	30	0.686	0.0488
SPOCK2	SPARC/osteonectin, cwcv and kazal like domains proteoglycan 2	36	0.675	0.0539
SLC25A4	solute carrier family 25 member 4	38	0.672	0.0619
SCOC	short coiled-coil protein	40	0.670	0.0698
CTSB	cathepsin B	51	0.662	0.0713
MAP1A	microtubule-associated protein 1A	54	0.659	0.0784
LSAMP	limbic system-associated membrane protein	55	0.658	0.0869
RAB6B	RAB6B, member RAS oncogene family	56	0.658	0.0953
ATP1B1	ATPase Na ⁺ /K ⁺ transporting subunit beta 1	60	0.655	0.1017
NAP1L3	nucleosome assembly protein 1 like 3	63	0.654	0.1087
ENO2	enolase 2	66	0.652	0.1157
FABP3	fatty acid binding protein 3	70	0.650	0.1219
NDRG4	NDRG family member 4	78	0.645	0.1253
SCG2	secretogranin II	82	0.642	0.1314
SOD1	superoxide dismutase 1	95	0.630	0.1311
ARF3	ADP ribosylation factor 3	109	0.617	0.1299
FGF12	fibroblast growth factor 12	117	0.612	0.1328
PLEKHB2	pleckstrin homology domain containing B2	124	0.609	0.1364
ATP6V1E1	ATPase H ⁺ transporting V1 subunit E1	126	0.606	0.1435
VAMP2	vesicle-associated membrane protein 2	133	0.602	0.1471
ERLEC1	endoplasmatic reticulum lectin 1	136	0.601	0.1534
ATP6V0A1	ATPase H ⁺ transporting V0 subunit a1	141	0.600	0.1583
PTPRN	protein tyrosine phosphatase, receptor type N	157	0.595	0.1554
SNCA	synuclein alpha	167	0.593	0.1567
PCYOX1	prenylcysteine oxidase 1	178	0.591	0.1572
FGF14	fibroblast growth factor 14	183	0.590	0.1620
HPRT1	hypoxanthine phosphoribosyltransferase 1	187	0.588	0.1675
FAM162A	family with sequence similarity 162 member A	204	0.580	0.1636
GPX1	glutathione peroxidase 1	213	0.578	0.1654
AHI1	Abelson helper integration site 1	222	0.575	0.1672
SEZ6L2	seizure related 6 homolog like 2	226	0.574	0.1725

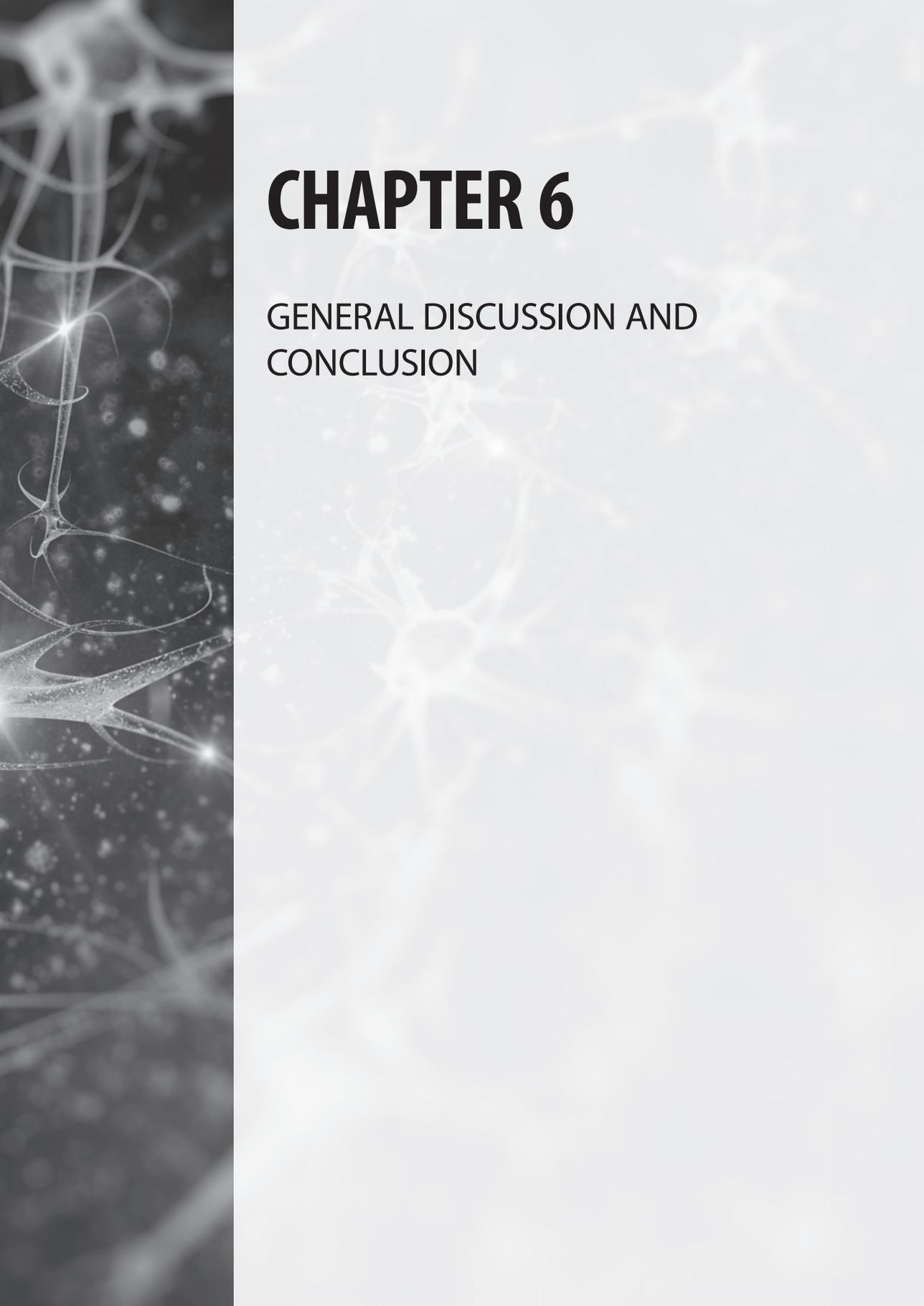
Supplementary Table 1. Putative generic biomarkers for human neuronal maturity revealed by Gene Set Enrichment Analysis of darkturquoise module genes. (continued)

Gene Symbol	Gene Name	Rank in Gene List	Rank Metric Score	Running Enrichment Score
KLC1	kinesin light chain 1	232	0.572	0.1763
EPB41L1	erythrocyte membrane protein band 4.1 like 1	241	0.569	0.1780
TAGLN3	transgelin 3	244	0.568	0.1839
PSMA5	proteasome subunit alpha type 5	265	0.562	0.1771
HIGD1A	HIG1 hypoxia inducible domain family member 1A	296	0.550	0.1630
NEGR1	neuronal growth regulator 1	299	0.549	0.1686
DNM3	dynammin 3	319	0.544	0.1623
CD47	CD47 molecule	322	0.543	0.1678
NDUFV1	NADH:ubiquinone oxidoreductase core subunit V1	324	0.543	0.1741
PFKM	phosphofructokinase, muscle	328	0.543	0.1790
BLCAP	bladder cancer associated protein	337	0.541	0.1803
MPV17	MPV17, mitochondrial inner membrane protein	341	0.539	0.1852
SUMO3	small ubiquitin-like modifier 3	352	0.536	0.1850
TCF25	transcription factor 25	355	0.536	0.1905
PPP2CA	protein phosphatase 2 catalytic subunit alpha	356	0.536	0.1974
IFI6	interferon alpha-inducible protein 6	362	0.535	0.2008
BSG	basigin (Ok blood group)	366	0.534	0.2055
ACOT7	acyl-CoA thioesterase 7	377	0.533	0.2053
SYP	synaptophysin	393	0.529	0.2016
SCN3B	sodium voltage-gated channel beta subunit 3	396	0.529	0.2070
ISCA1	iron-sulfur cluster assembly 1	398	0.528	0.2131
COX14	COX14, cytochrome c oxidase assembly factor	409	0.527	0.2128
C2orf80	chromosome 2 open reading frame 80	439	0.520	0.1991
NDUFB8	NADH:ubiquinone oxidoreductase subunit b8	441	0.520	0.2051
PRKAA2	protein kinase AMP-activated catalytic subunit alpha 2	451	0.519	0.2054
MYL6B	myosin light chain 6B	452	0.519	0.2121
COPS6	COP9 signalosome subunit 6	457	0.518	0.2159
SMARCA2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	458	0.518	0.2226
PCMT1	protein-L-isoaspartate (D-aspartate) O-methyltransferase	461	0.518	0.2279
UQCRC1	ubiquinol-cytochrome c reductase core protein I	470	0.516	0.2289
ST3GAL5	ST3 beta-galactoside alpha-2,3-sialyltransferase 5	478	0.515	0.2306
RAB18	RAB18, member RAS oncogene family	479	0.515	0.2372
TUBB4A	tubulin beta 4A class IVa	485	0.513	0.2403
NEFL	neurofilament, light polypeptide	488	0.513	0.2455
ANOS5	anoctamin 5	491	0.512	0.2507
NUDCD3	NudC domain containing 3	500	0.511	0.2516
DHRS7	dehydrogenase/reductase 7	510	0.510	0.2518

Supplementary Table 1. Putative generic biomarkers for human neuronal maturity revealed by Gene Set Enrichment Analysis of darkturquoise module genes. (continued)

Gene Symbol	Gene Name	Rank in Gene List	Rank Metric Score	Running Enrichment Score
G3BP2	G3BP stress granule assembly factor 2	516	0.508	0.2549
CRBN	cereblon	517	0.508	0.2614
PPP1R2	protein phosphatase 1 regulatory inhibitor subunit 2	541	0.504	0.2517
SBDS	SBDS ribosome assembly guanine nucleotide exchange factor	544	0.504	0.2568
ACSL3	acyl-CoA synthetase long-chain family member 3	564	0.501	0.2498
RAP1GDS1	Rap1, GTPase-GDP dissociation stimulator 1	566	0.500	0.2555
GAD1	glutamate decarboxylase 1	577	0.499	0.2549
POLR2I	RNA polymerase II subunit I	578	0.499	0.2614
NGFRAP1	nerve growth factor receptor (TNFRSF16) associated protein 1	594	0.497	0.2572
MICU1	mitochondrial calcium uptake 1	612	0.494	0.2515
TUBA1B	tubulin alpha 1b	620	0.493	0.2530
COX8A	cytochrome c oxidase subunit 8A	628	0.492	0.2544
PGM2L1	phosphoglucomutase 2 like 1	633	0.492	0.2579
GNAI1	G protein subunit alpha i1	635	0.491	0.2635
KIAA1107	KIAA1107	644	0.490	0.2642
FAM171B	family with sequence similarity 171 member B	647	0.490	0.2691
NCALD	neurocalcin delta	649	0.489	0.2747
PPP2R1A	protein phosphatase 2 scaffold subunit Aalpha	655	0.488	0.2774
SLC25A3	solute carrier family 25 member 3	656	0.488	0.2837





CHAPTER 6

GENERAL DISCUSSION AND CONCLUSION

GENERAL DISCUSSION

Recent advances in human cell reprogramming technologies have motivated a panoply of induced pluripotent stem cell (iPSC)-based *in vitro* experiments to study and reveal differences between healthy control subjects and patients afflicted with neurological or psychiatric disease [1-3]. Although iPSC technology has emerged as a powerful tool to recapitulate various cell biological and molecular features of human neuropsychiatric illnesses “in a laboratory dish”, the variability inherent to cell culturing conditions constitutes a significant roadblock to the accurate modeling and identification of disease-related phenotypes. Robust and replicable experimental measurements are needed if one wishes to enable successful translation of research findings into clinical practice and impact the prevention or treatment of human neuropsychiatric disease. Therefore, technical and biological sources of variability, which can dramatically affect comparative analyses and may obscure identification of disease-related phenotypes *in vitro* [4], should be carefully considered and controlled for, or eliminated (if possible), in any disease modeling study. In this context, the work presented in this thesis aimed to (i) create more physiological *in vitro* neuronal culture conditions and (ii) stratify the heterogeneity of functional states within populations of differentiated neurons. More realistic neuronal culturing conditions and the ability to efficiently identify and isolate neurons of a particular functional state from heterogeneous differentiating cultures will enable better comparative analyses in human stem cell models of brain disorders, and may ultimately lead to a more successful translation of research findings to the clinic.

To better understand the molecular mechanisms that control cell identity and underlie the generation of iPSCs from somatic cells, I first reviewed recent research on the transcriptional and epigenetic regulation of pluripotency induction (CHAPTER 2 and ref. [5]). The derivation of iPSCs from fibroblasts is an inefficient process, and the precise molecular mechanisms governing somatic cell reprogramming are just being revealed. From the reviewed studies, it became clear that faithful reprogramming is very dependent on various technological parameters, such as the specific composition of the cell culture environment [6] and the levels and stoichiometric ratio of reprogramming factors [7, 8]. Importantly, suboptimal reprogramming conditions *in vitro* have been reported to introduce substantial variability in iPSCs’ transcriptional, epigenetic and proteomic states, which may affect these cells’ biological functionality (e.g., developmental competence) [6-11]. We conclude that to better understand how specific reprogramming conditions impact the molecular profiles and biological function of iPSCs, future research efforts should be directed toward the parallel profiling of transcriptomes and epigenomes of single reprogramming cells [5]. The new information gained from such single-cell ‘omic approaches should enable the development of more efficient reprogramming conditions

and, ultimately, the high-fidelity generation of iPSCs with limited heterogeneity for use in research and clinical applications.

We concluded from the above that iPSC technology is still in its infancy due to low efficiency of the reprogramming process, and that numerous technical hurdles need to be overcome to attain highly efficient and accurate human disease modeling, and safely employ iPSCs for clinical/therapeutic purposes [4, 12-15]. Yet, many neuroscience laboratories throughout the world have succeeded in deriving neurons from iPSCs from patients afflicted with various neurological and psychiatric disorders, and have demonstrated the ability to recapitulate certain features of nervous system disease in a laboratory dish [16-19]. In the lab, we established midbrain-like human neuronal cultures from embryonic stem cells (ESCs) and fibroblast-derived iPSCs from four healthy individuals. When examining the calcium activity of these human stem cell-derived neurons in DMEM/F12-based neuronal culture medium, we unexpectedly found only very few neurons to be spontaneously active (CHAPTER 3A). Interestingly, many cells became active when imaging the same cultures in perfusate of artificial cerebrospinal fluid (ACSF), a neurophysiological solution commonly used for assaying neuronal functionality. This observation held even after omitting all medium supplements (growth factors, antioxidants, etcetera) and matching the inorganic salt concentration, pH and osmolarity of our ACSF to those of DMEM. However, the finding that removal of amino acids, vitamins or extra components from DMEM basal medium rescued the low calcium activity of the cells provided a first indication that DMEM contains various components that acutely interfere with neurophysiological function. Because electrical activity is the essence of nervous system function *in vivo*, this discovery suggested for the first time that current *in vitro* neuronal culture conditions may not be optimized to reflect the fundamental principles of *in vivo* brain physiology.

Hesitant to question the validity of a standard basal medium such as DMEM, which has somehow successfully been used in research laboratories for decades, we investigated the basal medium-induced impairment of neuronal activity in more detail using patch-clamping methods (CHAPTER 3B and ref. [20]). These electrophysiological analyses revealed a strong reduction in synaptic function and action potential (AP) firing of neurons in DMEM, as well as in Neurobasal (a modified version of DMEM/F12 basal that has been designed to optimize the survival of primary neurons in culture [21]), and serum. This finding, that many important neurophysiological properties are substantially altered in neuronal culture media widely used throughout the scientific community, motivated us to design a new basal neuro-medium—BrainPhys™—that is more representative of the *in vivo* brain extracellular environment and supports optimal synaptic activity and APs (CHAPTER 3B and ref. [20]). Specifically in this medium, we removed or reduced any neuroactive amino acids that could directly influence synaptic transmission, and adjusted the concentrations of inorganic salts and energetic substrates, as well as the

osmolarity, to neurophysiological levels. We found that BrainPhys™ basal medium with supplements adequately supported long-term electrical activity and survival of mature human neurons *in vitro*. Importantly, differentiation in BrainPhys™+sup medium did not affect neuronal fate specification, but it increased the proportion of human neurons that were synaptically active.

Because spontaneous and synaptic neuronal activity are essential for the formation and maintenance of proper connectivity in developing neural networks [22-26], one might expect that the improvements made in BrainPhys™ basal will indirectly augment *in vitro* neuronal development. As many neuropsychiatric disorders, such as schizophrenia and autism, are hypothesized to have a neurodevelopmental origin [27, 28], it is conceivable that studying neurons in more physiological conditions that better foster neuronal development will increase the likelihood of identifying new disease-relevant phenotypes. Furthermore, many neurological and psychiatric disorders have been associated with synaptic dysfunction and altered neuronal excitability and network activity [29-32]; hence, a failure to model such disorders in accurate physiological conditions that support synaptic functions and activity may obscure the exact mechanisms underlying their pathologies. Neural culture conditions that more realistically resemble the living brain will be more likely to faithfully recapitulate the pathophysiology as it occurs in the disease and, consequently, improve the discovery of new biomarkers and treatments for clinical practice.

In addition to rigorously controlling and optimizing cell culturing conditions, which effectively reduces technical variability and increases chances of translational success, a critical yet somewhat elusive parameter to consider in investigations of human nervous system disease with stem cell models is the natural diversity (biological variability) of the cells. Human neuronal tissue with its heterogeneous, intricate, precise and dynamic assembly of a multitude of different cell types constitutes the principal example of cellular diversity. The extraordinary variety of neuronal types was illustrated for the first time in Cajal's diligent drawings of the morphological and structural features of neurons in the late 19th century [33, 34]. Then, in the late 1970s, the advent of patch-clamp recording technology [35] further stimulated the urge among scientists to identify and classify different types of neurons of the brain based on function. It is now recognized that a thorough characterization of the functional diversity obtained with tissue culture is imperative to avoid the risk of incorrect interpretation of cellular phenotypes due to heterogeneity of cellular populations. Particularly in neuronal tissue, the rate at which neural progenitors differentiate into electrophysiologically functional neurons is significantly variable; therefore, to prevent confounding of disease-relevant discoveries and erroneous conclusions, we must aim to control for functional maturational equivalence in comparisons of diseased to healthy neuronal cells [4]. Building upon recent advances

in single-cell whole-transcriptome analysis approaches [36-46], we established a multi-dimensional characterization of single human ESC- and iPSC-derived neurons to correlate their functional properties (electrophysiology) with their morphology and mRNA expression profiles (CHAPTER 4A and ref. [47]). Our stratification of *in vitro* functional neuronal heterogeneity revealed strong correlations between neuronal electrophysiological function, morpho-structural features and gene expression. Specifically, we have identified a series of genes more abundantly expressed in highly functional types of human neurons compared with more immature electrophysiological neuronal types and astrocytes. Importantly, the identified gene expression signature enabled the prediction and isolation of highly functional neuron states with machine-learning classifiers and fluorescent markers. Taken together, this study opens a new way to predict and study neurophysiological cell types based on molecular signatures. Such a molecular approach, which is more high-throughput than patch-clamp recording and is applicable to any electrophysiological type of cell, will undoubtedly prove very efficient in large-scale analyses necessary for translational studies.

Up until now, coupling of electrophysiological recordings to gene expression profiles has been primarily dependent on single-cell reverse transcription PCR (scRT-PCR) [48-51] or cDNA microarray analysis [52] of individual patch-clamped neurons. Unfortunately, PCR and array-based methods are restricted to the examination of a limited number of pre-specified genes, and combining whole-cell electrophysiology with RNA-sequencing analysis of entire single neuron transcriptomes—as performed in our study—can achieve a much more detailed, accurate and completely unbiased classification of cellular types and states [42]. In contrast with most single-cell transcriptome approaches, which require neurite-disruptive tissue dissociation to sequester single cells or rely on patch-pipette micro-aspiration of cytosolic mRNA molecules [53], the method we developed collects and processes the entire neuron and, therefore, not only includes mRNA from the cell soma but also from distant neurites, potentially providing a more complete representation of the transcriptome [54]. In this context, our approach might prove useful for studying neurological disorders such as fragile X syndrome that are characterized by dysregulation of mRNA transport and translation in local dendrites or at distant axonal sites, processes that crucially affect synaptic function and development [55, 56].

A major challenge of single-cell transcriptome analyses is to assure that only intact and non-compromised single cells are included in bioinformatics analysis pipelines. Sample quality control (QC) is a crucially important aspect of sequencing experiments as including spurious samples may introduce unwanted variability in the data, potentially blurring or confounding biological analyses. To filter out potentially spurious captures in advance of deep sequencing, we developed a pre-sequencing QC pipeline for single-neuron Patch-seq samples (described in CHAPTER 4B, applied in CHAPTER 4A) based on four simple assays: (i) microscopic qualitative evaluation of successful neuron capture; (ii)

expression profiling of common housekeeping genes and added spike-ins (quantitative real-time PCR); (iii) fluorometric quantitation of cDNA yield (Qubit); and (iv) qualitative analysis of cDNA fragment profiles (Agilent Bioanalyzer). While our results indicated some significant correlations between pre-sequencing QC measures and post-sequencing variables that are, at least to some extent, indicative of expression data quality, future research is warranted to explore in detail what (combination of) technical and biological data features best discerns technical artefacts from intrinsic biological heterogeneity.

In the last main chapter of this thesis (CHAPTER 5), we performed Weighted Gene Co-expression Network Analysis (WGCNA) [57] of our single-neuron RNA-sequencing data with the purpose of detecting sets of co-expressed genes likely to exhibit commonality of function, and that are associated with our electrophysiological neuron types. We identified a module of 495 highly interconnected genes whose expression was significantly upregulated in highly functional neurons, compared with immature neuron types. Various electrophysiological and morphological measures of neuronal maturity were strongly associated with this gene module, and Gene Ontology analysis revealed enrichment for processes related to mitochondrial energy metabolism and cytoskeletal organization. We additionally provide evidence that the identified co-expression module may play a general role in neuronal maturation, as a substantial number of module genes was enriched in a previously published single-cell RNA-seq study that compared adult (mature) versus fetal (immature) brain neurons [58]. Our most notable finding, however, is the identification of a set of 21 putative transcriptional regulators that may be involved in defining, regulating or even inducing the highly functional neuronal state. In the light of inducibly controlling (e.g., accelerating) functional maturation with the aim of creating less heterogeneous cultures of electrophysiologically more comparable neurons, pharmacological compounds that modulate these transcription factors represent very promising tools for future experimental studies. Alternatively, our results may guide the development of genetic engineering strategies that enable manipulation of key genetic regulators of the highly functional, developmentally more mature, neuronal state.

CONCLUSION AND OUTLOOK

Many neuroscience laboratories throughout the world employ iPSC technology to derive disease-specific neurons from patients afflicted with neurological or psychiatric disorders. To ensure accurate identification of disease-associated phenotypes in iPSC-derived neurons *in vitro*, a careful consideration of both technical and biological sources of variability, which impact comparative analyses, is warranted. In the present thesis, we report the design of a new basal neuronal medium—BrainPhys™—that more closely mimics *in vivo* brain neurophysiological conditions, and better supports synaptic functions and

activity of neurons *in vitro*. Although the resulting improvement of *in vitro* physiological conditions is likely to improve the rate of translational success of stem cell-based research investigations, we should be aware of the fact that many neurological and psychiatric disorders have a developmental origin, and that fully capturing the complexity of mammalian brain development *in vitro* is a difficult, if not impossible, task. However, recent developments in three-dimensional (3D) cell culturing technologies offer great opportunities to model in a laboratory dish various of the structural and functional features that are characteristic to the human brain [59, 60]. Compared with standard monolayer (i.e., 2D) neuronal cultures, cultures of neurons grown in 3D more adequately recapitulate the intricate cell-cell interactions characteristic of *in vivo* brain anatomy. Hence, the 3D culturing of human stem cell-derived neurons, which is becoming more and more common these days, will undoubtedly enable a more precise *in vitro* investigation of nervous system function, development and disease. Important for such “organoid” culture systems is that they resemble the *in vivo* brain microenvironment as accurately as possible, and model not only the complexity of neuronal connections, but also the intricate interaction of neurons with astrocytes, which have been shown to modulate neuronal activity, synaptic transmission and synaptic plasticity at the tripartite synapse [61, 62].

Another important point to note regarding neuropsychiatric disease modeling with human stem cells is that it will be impossible to recapitulate the complete phenomenology of psychiatric illnesses in a laboratory dish. Biological findings resulting from *in vitro* culture models of psychiatric disorders represent a strong simplification of the complex biological processes underlying specific psychological functions, and their successful translation to human clinical trials constitutes a major challenge. Although a strong genetic basis exists for many neurological and psychiatric illnesses, it is now generally accepted that these disorders are the result of complex interactions between genes and the environment [63-65]. While it is not possible to accurately mimic, in the lab, the complex environment to which we humans are exposed, it would be very interesting to study the effect of specific environmental exposures (e.g., hypoxia, cortisol, cannabis, etcetera) in interaction with complete patient genomes in iPSC-based models of neuropsychiatric disease [64].

In the second part of this thesis, we detailed the results of an in-depth profiling of individual ESC- and iPSC-derived neurons at the functional, morphological and molecular level. We found strong correlations between electrophysiological function, morphology and gene expression. Although our single-neuron transcriptome analyses have revealed relevant insights into the molecular regulation of neuronal maturational diversity, I propose that future studies aimed at examining how the neuronal transcriptome regulates neuronal function account for diversity at the genomic and epigenomic levels. Single-cell transcriptome studies invariably assume that all cells have the same genome and stable epigenomes; yet, accumulating evidence suggests that neuronal diversity extends to the

genetic and epigenetic levels (reviewed in ref. [66]). In particular, variability in neuronal genomes has been shown to arise due to (sub)chromosomal alterations [67-69] or “jumping” endogenous mobile genetic elements [70-73]. As such alterations in neuronal genomes are likely to induce alterations in gene expression [70, 74], it is very possible that a neuron’s transcriptome is a reflection not only of the cell’s electrophysiological function, but also of the status of the genome at a given point in time [66]. Emerging technologies for the parallel profiling of single-cell transcriptomes and genomes [75, 76] or epigenomes [77] may help further elucidate the specific contribution of genomic and epigenomic variation to neuronal transcriptional diversity. Ultimately, such investigations will provide us with a more accurate understanding of the molecular programs driving neuronal identity and maturation.

REFERENCES

1. Dolmetsch R, Geschwind DH. The human brain in a dish: the promise of iPSC-derived neurons. *Cell* 145(6), 831-834 (2011).
2. Bellin M, Marchetto MC, Gage FH, Mummery CL. Induced pluripotent stem cells: the new patient? *Nat. Rev. Mol. Cell Biol.* 13(11), 713-726 (2012).
3. Marchetto MC, Brennand KJ, Boyer LE, Gage FH. Induced pluripotent stem cells (iPSCs) and neurological disease modeling: progress and promises. *Hum. Mol. Genet.* 20(R2), R109-115 (2011).
4. Sandoe J, Eggan K. Opportunities and challenges of pluripotent stem cell neurodegenerative disease models. *Nat. Neurosci.* 16(7), 780-789 (2013).
5. Van Den Hurk M, Kenis G, Bardy C *et al.* Transcriptional and epigenetic mechanisms of cellular reprogramming to induced pluripotency. *Epigenomics* 8(8), 1131-1149 (2016).
6. Chen J, Liu H, Liu J *et al.* H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat. Genet.* 45(1), 34-42 (2013).
7. Carey BW, Markoulaki S, Hanna JH *et al.* Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell* 9(6), 588-598 (2011).
8. Tonge PD, Corso AJ, Monetti C *et al.* Divergent reprogramming routes lead to alternative stem-cell states. *Nature* 516(7530), 192-197 (2014).
9. Hussein SM, Puri MC, Tonge PD *et al.* Genome-wide characterization of the routes to pluripotency. *Nature* 516(7530), 198-206 (2014).
10. Clancy JL, Patel HR, Hussein SM *et al.* Small RNA changes en route to distinct cellular states of induced pluripotency. *Nat. Commun.* 5 5522 (2014).
11. Benevento M, Tonge PD, Puri MC *et al.* Proteome adaptation in cell reprogramming proceeds via distinct transcriptional networks. *Nat. Commun.* 5 5613 (2014).
12. Saha K, Jaenisch R. Technical challenges in using human induced pluripotent stem cells to model disease. *Cell Stem Cell* 5(6), 584-595 (2009).
13. Lu X, Zhao T. Clinical therapy using iPSCs: hopes and challenges. *Genomics Proteomics Bioinformatics* 11(5), 294-298 (2013).
14. Lund RJ, Narva E, Lahesmaa R. Genetic and epigenetic stability of human pluripotent stem cells. *Nat. Rev. Genet.* 13(10), 732-744 (2012).
15. Cahan P, Daley GQ. Origins and implications of pluripotent stem cell variability and heterogeneity. *Nat. Rev. Mol. Cell Biol.* 14(6), 357-368 (2013).
16. Brennand KJ, Simone A, Jou J *et al.* Modelling schizophrenia using human induced pluripotent stem cells. *Nature* 473(7346), 221-225 (2011).
17. 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).
18. Marchetto MC, Carromeu C, Acab A *et al.* A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* 143(4), 527-539 (2010).
19. Nguyen HN, Byers B, Cord B *et al.* LRRK2 mutant iPSC-derived DA neurons demonstrate increased susceptibility to oxidative stress. *Cell Stem Cell* 8(3), 267-280 (2011).
20. Bardy C, Van Den Hurk M, Eames T *et al.* Neuronal medium that supports basic synaptic functions and activity of human neurons in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 112(20), E2725-2734 (2015).
21. Brewer GJ, Torricelli JR, Evege EK, Price PJ. Optimized survival of hippocampal neurons in B27-supplemented Neurobasal, a new serum-free medium combination. *J. Neurosci. Res.* 35(5), 567-576 (1993).
22. Katz LC, Shatz CJ. Synaptic activity and the construction of cortical circuits. *Science* 274(5290), 1133-1138 (1996).

23. Zhang LI, Poo MM. Electrical activity and development of neural circuits. *Nat. Neurosci.* 4 Suppl 1207-1214 (2001).
24. Ben-Ari Y. Developing networks play a similar melody. *Trends Neurosci.* 24(6), 353-360 (2001).
25. Spitzer NC. Electrical activity in early neuronal development. *Nature* 444(7120), 707-712 (2006).
26. Blankenship AG, Feller MB. Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nat. Rev. Neurosci.* 11(1), 18-29 (2010).
27. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch. Gen. Psychiatry* 44(7), 660-669 (1987).
28. Neul JL, Zoghbi HY. Rett syndrome: a prototypical neurodevelopmental disorder. *Neuroscientist* 10(2), 118-128 (2004).
29. Shankar GM, Walsh DM. Alzheimer's disease: synaptic dysfunction and A β . *Mol. Neurodegener.* 4 48 (2009).
30. Milnerwood AJ, Raymond LA. Early synaptic pathophysiology in neurodegeneration: insights from Huntington's disease. *Trends Neurosci.* 33(11), 513-523 (2010).
31. Schirinzi T, Madeo G, Martella G *et al.* Early synaptic dysfunction in Parkinson's disease: Insights from animal models. *Mov. Disord.* 31(6), 802-813 (2016).
32. Steinlein OK. Genetic mechanisms that underlie epilepsy. *Nat. Rev. Neurosci.* 5(5), 400-408 (2004).
33. Cajal SR. Textura de las circunvoluciones cerebrales de los mamíferos inferiores. Nota preventiva. *Gac. Méd. Catalana* 1 22-31 (1890).
34. Cajal SR. Sur la structure de l'écorce cérébrale de quelques mammifères. *La Cellule* 1(125-76), (1891).
35. Neher E, Sakmann B. Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature* 260(5554), 799-802 (1976).
36. Eberwine J, Sul JY, Bartfai T, Kim J. The promise of single-cell sequencing. *Nat. Methods* 11(1), 25-27 (2014).
37. Islam S, Kjallquist U, Moliner A *et al.* Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* 21(7), 1160-1167 (2011).
38. Islam S, Kjallquist U, Moliner A *et al.* Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat. Protoc.* 7(5), 813-828 (2012).
39. Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* 10(11), 1096-1098 (2013).
40. Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* 9(1), 171-181 (2014).
41. Saliba AE, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 42(14), 8845-8860 (2014).
42. Sandberg R. Entering the era of single-cell transcriptomics in biology and medicine. *Nat. Methods* 11(1), 22-24 (2014).
43. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat. Rev. Genet.* 14(9), 618-630 (2013).
44. Tang F, Barbacioru C, Nordman E *et al.* RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat. Protoc.* 5(3), 516-535 (2010).
45. Tang F, Barbacioru C, Wang Y *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6(5), 377-382 (2009).
46. Wu AR, Neff NF, Kalisky T *et al.* Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* 11(1), 41-46 (2014).
47. Bardy C, Van Den Hurk M, Kakaradov B *et al.* Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology. *Mol. Psychiatry* (2016, in press).

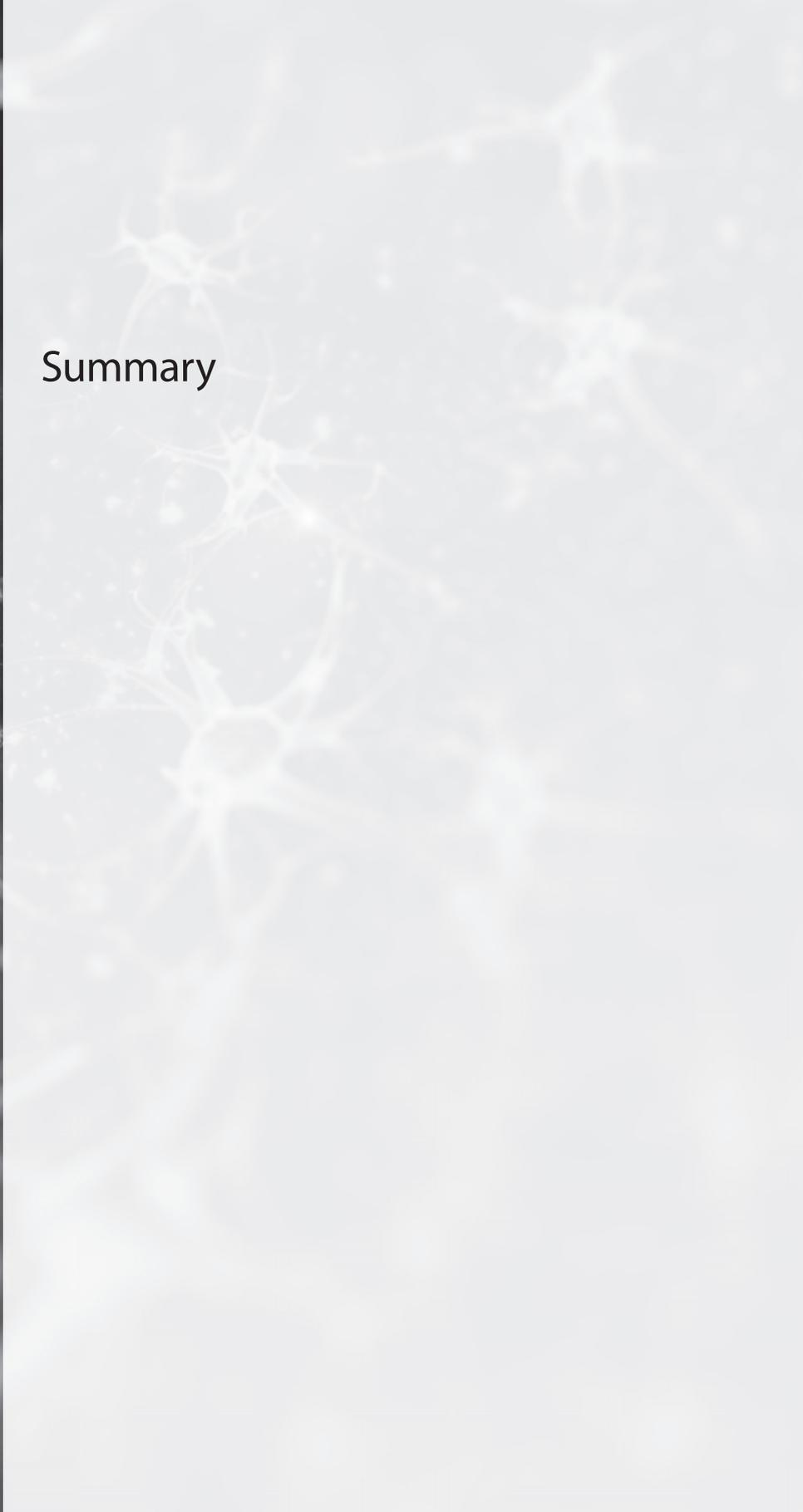
48. Belinsky GS, Rich MT, Sirois CL *et al.* Patch-clamp recordings and calcium imaging followed by single-cell PCR reveal the developmental profile of 13 genes in iPSC-derived human neurons. *Stem Cell Res.* 12(1), 101-118 (2014).
49. Sucher NJ, Deitcher DL, Baro DJ, Warrick RM, Guenther E. Genes and channels: patch/voltage-clamp analysis and single-cell RT-PCR. *Cell Tissue Res.* 302(3), 295-307 (2000).
50. Toledo-Rodriguez M, Blumenfeld B, Wu C *et al.* Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. *Cereb. Cortex* 14(12), 1310-1327 (2004).
51. Farago N, Kocsis AK, Lovas S *et al.* Digital PCR to determine the number of transcripts from single neurons after patch-clamp recording. *Biotechniques* 54(6), 327-336 (2013).
52. Subkhankulova T, Yano K, Robinson HP, Livesey FJ. Grouping and classifying electrophysiologically-defined classes of neocortical neurons by single cell, whole-genome expression profiling. *Front. Mol. Neurosci.* 3 10 (2010).
53. Stahlberg A, Kubista M, Aman P. Single-cell gene-expression profiling and its potential diagnostic applications. *Expert Rev. Mol. Diagn.* 11(7), 735-740 (2011).
54. Cajigas IJ, Tushev G, Will TJ, Tom Dieck S, Fuerst N, Schuman EM. The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* 74(3), 453-466 (2012).
55. Bassell GJ, Warren ST. Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* 60(2), 201-214 (2008).
56. Bagni C, Greenough WT. From mRNP trafficking to spine dysmorphogenesis: the roots of fragile X syndrome. *Nat. Rev. Neurosci.* 6(5), 376-387 (2005).
57. Zhao W, Langfelder P, Fuller T, Dong J, Li A, Hovarth S. Weighted gene coexpression network analysis: state of the art. *J. Biopharm. Stat.* 20(2), 281-300 (2010).
58. Darmanis S, Sloan SA, Zhang Y *et al.* A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. U. S. A.* 112(23), 7285-7290 (2015).
59. 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).
60. Lancaster MA, Renner M, Martin CA *et al.* Cerebral organoids model human brain development and microcephaly. *Nature* 501(7467), 373-379 (2013).
61. Volterra A, Meldolesi J. Astrocytes, from brain glue to communication elements: the revolution continues. *Nat. Rev. Neurosci.* 6(8), 626-640 (2005).
62. Navarrete M, Perea G, Fernandez De Sevilla D *et al.* Astrocytes mediate in vivo cholinergic-induced synaptic plasticity. *PLoS Biol.* 10(2), e1001259 (2012).
63. Landrigan PJ, Sonawane B, Butler RN, Trasande L, Callan R, Droller D. Early environmental origins of neurodegenerative disease in later life. *Environ. Health Perspect.* 113(9), 1230-1233 (2005).
64. Van Os J, Kenis G, Rutten BP. The environment and schizophrenia. *Nature* 468(7321), 203-212 (2010).
65. Sullivan PF. Puzzling over schizophrenia: schizophrenia as a pathway disease. *Nat. Med.* 18(2), 210-211 (2012).
66. Harbom LJ, Chronister WD, McConnell MJ. Single neuron transcriptome analysis can reveal more than cell type classification: Does it matter if every neuron is unique? *Bioessays* 38(2), 157-161 (2016).
67. Rehen SK, McConnell MJ, Kaushal D, Kingsbury MA, Yang AH, Chun J. Chromosomal variation in neurons of the developing and adult mammalian nervous system. *Proc. Natl. Acad. Sci. U. S. A.* 98(23), 13361-13366 (2001).
68. McConnell MJ, Lindberg MR, Brennand KJ *et al.* Mosaic copy number variation in human neurons. *Science* 342(6158), 632-637 (2013).

69. Cai X, Evrony GD, Lehmann HS *et al.* Single-cell, genome-wide sequencing identifies clonal somatic copy-number variation in the human brain. *Cell Rep.* 8(5), 1280-1289 (2014).
70. Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* 435(7044), 903-910 (2005).
71. Coufal NG, Garcia-Perez JL, Peng GE *et al.* L1 retrotransposition in human neural progenitor cells. *Nature* 460(7259), 1127-1131 (2009).
72. Evrony GD, Cai X, Lee E *et al.* Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* 151(3), 483-496 (2012).
73. Baillie JK, Barnett MW, Upton KR *et al.* Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 479(7374), 534-537 (2011).
74. Kaushal D, Contos JJ, Treuner K *et al.* Alteration of gene expression by chromosome loss in the postnatal mouse brain. *J. Neurosci.* 23(13), 5599-5606 (2003).
75. Dey SS, Kester L, Spanjaard B, Bienko M, Van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. *Nat. Biotechnol.* 33(3), 285-289 (2015).
76. Macaulay IC, Haerty W, Kumar P *et al.* G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat. Methods* 12(6), 519-522 (2015).
77. Angermueller C, Clark SJ, Lee HJ *et al.* Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* 13(3), 229-232 (2016).





Summary



There exists a tremendous degree of diversity within the neuronal population, and even morphologically similar-looking neurons can be markedly different in function. The major aim of this doctoral dissertation was to stratify the functional heterogeneity of human stem cell-derived neurons *in vitro* by investigating the molecular features that define neuronal identity in terms of specific functional and morpho-structural properties.

CHAPTER 1 is a general introduction to this thesis, outlining the scientific background and rationale to the studies presented. Specifically, in this chapter, I introduce the concept of modeling neurological and psychiatric disease features *in vitro* using patient-derived neurons, and discuss the existence of cellular heterogeneity as a major challenge to accurately interpreting findings from research investigations which employ stem cell-derived neuronal models. I then describe how the technology of single-cell transcriptome sequencing can be used to dissect the biological heterogeneity of cell types, and obtain a detailed characterization of cellular state and function. Lastly, I detail the specific aims of the experimental work conducted, and provide an outline of the structure of the chapters of this thesis.

In **CHAPTER 2**, we provide an extensive literature review on the transcriptional and epigenetic mechanisms of somatic cell reprogramming to obtain a better understanding of the complex molecular regulation of cell identity and pluripotency induction in somatic cells. The technology of induced pluripotent stem cell (iPSC) generation is quite inefficient and research studies are only beginning to elucidate the molecular mechanisms involved. From the studies reviewed, it is clear that faithful reprogramming is greatly dependent on various technical factors. Notably, non-optimal culture conditions *in vitro* can bring about considerable variation in cell-molecular profiles that could impinge on cells' biological functionality.

In **CHAPTER 3A**, we present the first experimental evidence that *in vitro* neuronal culturing conditions may not be optimized to reflect fundamental principles of *in vivo* brain physiology. Specifically, in calcium imaging experiments, we found only very few neurons to be spontaneously active when imaged in Dulbecco's Modified Eagle's Medium (DMEM), a standard basal medium routinely used in neuronal cell culture. Yet, many cells became active when imaging in a neurophysiological solution of artificial cerebrospinal fluid (ACSF) that had the exact same inorganic salt concentration, pH and osmolarity as DMEM. Providing a first indication that various components in DMEM (acutely) interfere with neuronal function, we found that removing all of the amino acids, vitamins or extra components of DMEM avoided the basal medium-induced impairment of neuronal activity.

Further in-depth electrophysiological analyses, detailed in **CHAPTER 3B**, revealed a strong impairment in synaptic function and action potential (AP) firing of neurons in DMEM, as well as in Neurobasal (a modified version of DMEM basal), and serum. The detrimental alteration of essential electrophysiological neuronal properties in these

media stimulated us to design a new basal medium that better mimics the healthy brain's *in vivo* micro-environment and adequately supports essential physiological neural activity *in vitro*. Specifically in this new medium called BrainPhys™, we excluded or decreased the levels of neuroactive amino acids, and set the concentrations of inorganic salts and energetic substrates, as well as the osmolarity, to neurophysiological conditions. Notably, BrainPhys™ basal medium, with the appropriate supplements, supported long-term electrical activity and survival of human neurons in culture, and enhanced neuronal synaptic function. We suggest that this synaptic enhancement is due to strengthening of silent synaptic contacts, rather than to the formation of new synapses.

CHAPTER 4A describes the results of a multi-dimensional characterization of single human neurons by combined electrophysiological, morphological and transcriptome (i.e., “Patch-seq”) analysis. We show that the inherent variation in basic electrophysiological profiles of human stem cell-derived neurons can be largely explained by two AP metrics (amplitude and firing rate), which together enable categorization of the neurons into five functional types that largely reflect stages of increasing neurophysiological maturity. For these functional neuronal types, we found strong correlations between their electrophysiological properties, morphological features and molecular profiles. Importantly, machine-learning strategies enabled the identification of new molecular biomarkers that efficiently predict which neurons are highly functional independently of time spent *in vitro*. As proof of concept, we selected one of the putative neurophysiological biomarkers—GDAP1L1—to effectively single out and enrich highly functional, live human neurons from heterogeneous cultures *in vitro*. This approach of predicting and studying neurophysiological cell types based on molecular signatures has higher throughput than patch clamping and may prove very useful for large-scale analyses in translational studies.

In **CHAPTER 4B**, we detail our Patch-seq pre-sequencing quality control (QC) pipeline, which we established to facilitate the identification and filtering out of poor-quality and contaminated single-neuron captures ahead of library preparation and deep sequencing. Captured neuron samples were subjected to a series of QC steps that included (i) expression profiling of common housekeeping genes (quantitative real-time PCR), (ii) fluorometric quantitation of cDNA yield (Qubit) and (iii) subjective qualitative analysis of cDNA fragment profiles (Agilent Bioanalyzer), and any sample not meeting predefined quality standards was filtered out. Our results indicated some significant correlations between these pre-sequencing QC measures and post-sequencing variables that are, at least partially, indicative of expression data quality. The presented framework thus forms a basis for better methodological designs, which may decrease variability and increase statistical power in single-cell gene expression experiments.

In **CHAPTER 5** of this dissertation, we conducted a systems-level analysis of the single-neuron transcriptome data by Weighted Gene Co-Expression Network Analysis

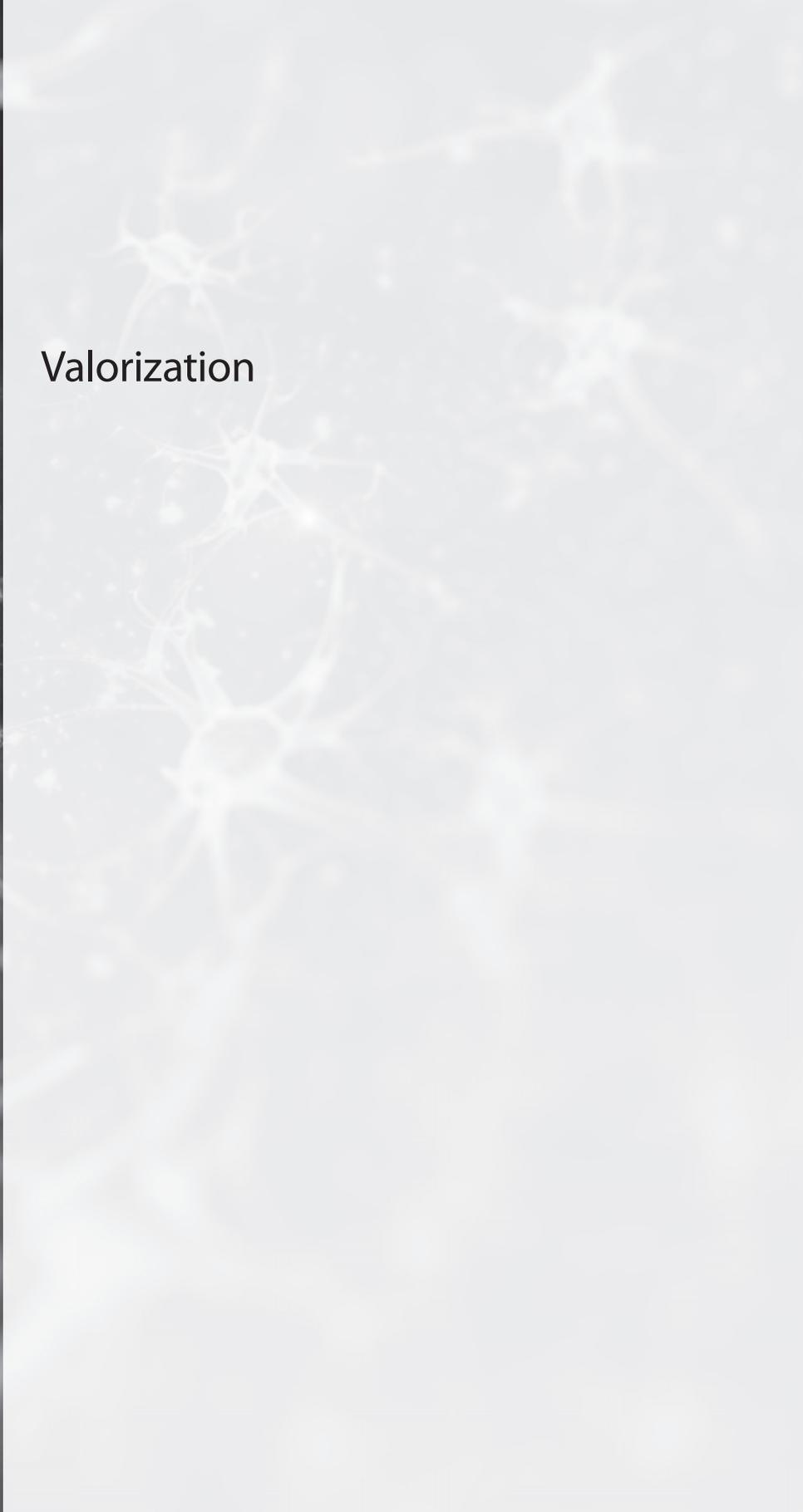
(WGCNA) to acquire further knowledge of the biology underlying the different functional neuron types. Our major finding is the identification of a cluster of 495 highly co-expressed genes that are strongly correlated with the functional properties and morphological features of highly functional human neurons. These genes were found to be involved in processes related to mitochondrial energy metabolism and cytoskeletal function, and Gene Set Enrichment Analysis (GSEA) indicated that they might play a general role in neuronal maturity. Another important result is the identification of a set of 21 putative transcriptional regulators of the highly functional, mature neuronal state. We suggest that these transcription factors represent promising targets for the development of novel chemical compounds that can accelerate the functional maturation of neurons.

The final chapter of this thesis, **CHAPTER 6**, critically discusses the experimental findings presented, and points out potential limitations and future avenues for investigation. Overall, the results of our studies provide a basis for biologically more relevant and accurate *in vitro* models of human neuronal cells derived from stem cells, which may facilitate the likelihood of translational success. The use of more physiological media to grow brain cells *in vitro*, and the stratification of functional neuronal heterogeneity to generate more homogeneous neuronal cultures, as detailed in this thesis, represent important means towards this end.





Valorization



Neurological and psychiatric disorders account for a substantial proportion of the world's total disease burden (i.e., 9% of global disability-adjusted life years (DALYs) in 2010) [1]. With the growth and changing age composition of the population, the societal and healthcare burden of mental health conditions is expected to become an even more significant problem in forthcoming years. For example, the World Health Organisation has estimated that depression will be the leading cause of lost DALYs by 2030 [2], and has calculated a 12% increase in DALYs from neurological disorders by that year from 2005 [3].

Mental disorders are often chronic and impact various (if not almost all) aspects of everyday living for patients and their families. Two major challenges that contribute to this personal burden can be identified as (i) a lack of compliance to treatments, and (ii) a huge knowledge gap between psychoeducation about the biological aspects of mental disorders as offered to patients in clinical setting, and current scientific knowledge as it is developing in the neurosciences. Overcoming this second challenge, for example by improving the information about neurobiological factors and mechanisms, may enable patients to formulate a better narrative and conceptual understanding regarding the development of their vulnerability. Knowledge development, as currently performed by studying living neurons from healthy subjects and patients “in a laboratory dish” (as presented in this thesis), may form—if interpreted with caution—a body of information on neurobiological factors and mechanisms involved in the onset and course of mental disorders, which may benefit patients and family members in better narrative and conceptual understanding on the development of their vulnerability, and may furthermore improve compliance to treatment.

Apart from the individual and societal burden that mental disorders represent for the patients and those who care for them, these brain-related disorders have a significant economic impact. The global cost of mental illnesses was estimated at US\$ 2.5 trillion in 2010, and is predicted to escalate to US\$ 6.0 trillion by 2030 [4]. It is important to point out that the true economic burden of mental health conditions may be considerably underestimated in these calculations due to the known comorbidity of these disorders with other chronic medical conditions, including cardiovascular disease [5, 6], diabetes [7], obesity [8, 9] and cancer [10], as well as with accidental death [11]. Moreover, the fact that only about one-third of total costs come from direct medical care, with the rest of expenses being incurred from lost employment or diminished work productivity due to the mortality and morbidity associated with these disorders [4], indicates that treatment options available to date are largely ineffective. Therefore, with a global surge in aging populations, the need for novel health discoveries through state-of-the-art scientific research has greatly exacerbated.

Although translational neurosciences research is relatively underfunded [12, 13], it has tremendous potential to dramatically impact the way we understand, prevent and

treat neurological and psychiatric disease. Preclinical trials with experimental animal models are essential and will continue to provide important insights into the biological mechanisms of mental health conditions; yet, pharmacological interventions found to be effective in *in vivo* animal studies almost invariably fail to translate to the human situation due to factors such as species differences, human brain complexity and patient/disease heterogeneity [14]. Most mental illnesses are complex disorders whose phenotypes can seldom be fully recapitulated by single-gene manipulation. The premise is therefore that preclinical trials with living human brain material are the most accurate and desired approach to study the precise physiological dysfunction of a neurological or psychiatric disorder, and avoid translational pitfalls. In this thesis, we describe a series of experiments that have resulted in the establishment of a biologically more relevant and accurate *in vitro* model of human neuronal cells derived from stem cells. The major implications of our findings and opportunities for valorization are detailed in this valorization addendum.

IMPROVING TRANSLATIONAL SUCCESS IN NEUROLOGY AND PSYCHIATRY: TOWARDS MORE EFFECTIVE PHARMACOLOGICAL TREATMENTS

Through recent advances in human cell reprogramming [15-18], it is now possible to generate virtually unlimited quantities of live human neuronal cells from readily accessible tissue sources, including skin cells. The ability to generate neuronal cell lines from patients afflicted with neurological and psychiatric disorders as well as from matched healthy control subjects has opened up new avenues for the modeling of both cellular and molecular pathophysiological features of brain disorders “in a laboratory dish” [19]. Research approaches using human patient-derived neuronal cultures represent a favorable intermediate between animal experiments and human clinical trials, and have vast potential to transform research and medicine by enabling drug discovery to be directly pursued on live human neuronal tissue [20-22].

Although *in vitro* human neuronal disease models derived using cell reprogramming technologies offer great promises to revolutionize medicine by facilitating the discovery of novel pharmacological interventions, it is imperative that such models are as realistic as possible to the human situation so as to optimize translational success to the clinic. Our results demonstrate that commonly used media for the differentiation and culturing of human neurons generally interfere with proper neuronal physiological function, and provide conditions that are much different from the living brain. More realistic neuronal models will be more likely to recapitulate the dysfunctional biology of neurological and psychiatric disorders, and improve the chances of discovery and translational success of novel pharmacological treatments. As detailed in this thesis (CHAPTER 3B), our design of a new neuronal medium that better resembles *in vivo* brain conditions, now

commercially available through STEMCELL Technologies Inc. as BrainPhys™, brings researchers around the world one step closer to this goal. The improvements made in this new medium, along with ongoing further developments in cell culture conditions and techniques, are expected to advance our understanding of mental health disorders and promote a more robust and accurate platform for preclinical drug development and testing.

Despite continuous advances in cell culturing methods, neuronal cultures remain inherently variable and are often characterized by a considerably heterogeneous proportion of functionally mature neurons [23, 24]. In comparing patient- and healthy control-derived neurons, we must strive to control for this source of variability by e.g. limiting cellular and molecular analyses to only mature functional neuronal cells. A major outcome of the research presented in this dissertation is the identification of a set of new live biomarkers associated with highly functional human neurons. As we demonstrate in CHAPTER 4A, these markers can be used to purify neurons of highly functional neuronal states from heterogeneously differentiating neuronal cultures. It may be expected that these and yet-to-be-discovered biomarkers of different neuronal types and states will be increasingly applied in future disease modeling experiments of brain disorders to reduce immanent phenotypic variability and streamline more accurate and robust investigations of drug development in high-throughput fashion. Without doubt, minimizing variation among neuronal populations and using more physiological neuronal models will importantly contribute to the successful translation of new discoveries into effective treatments.

PHARMACOGENOMICS APPROACHES TO PREDICTING DRUG RESPONSIVENESS: TOWARDS PERSONALIZED MEDICINE

Experiments like those described in this thesis may also offer future opportunities for personalizing drug therapy in patients. Although in this thesis we restricted our electrophysiological and molecular analyses to neurons derived from healthy donors, the research model we established based on cell reprogramming methodology can be applied to studying the molecular and neuronal dysfunction of any neurological or psychiatric disease *in vitro*. The electrophysiological and molecular properties of single patient-derived neurons can be correlated with patients' clinical and genomic profiles, an approach that may advance the discovery of disease mechanisms and potential new therapeutic targets. Additionally, the cellular effects of drugs prescribed to the patients can be studied in the patient-derived neuronal lines and correlated with observed behavioural symptoms, thus enabling a prediction of the clinical responsiveness of the patients to specific drugs to be made based on outcomes of *in vitro* cellular assays. Ultimately, such an approach that correlates cellular drug effects with clinical patient features will enable clinicians to

take a bold step toward personalized medicine by providing novel rationales for effective pharmacological treatment.

CONCLUSION

The work detailed in this thesis provides a promising framework for facilitating the discovery of novel biomarkers and treatments for various neurological and psychiatric disorders. Specifically, it is anticipated that creating more physiological *in vitro* culturing conditions and establishing more homogeneous neuronal cultures in disease modeling experiments will result in new mechanistic insights into the pathology of brain disorders and the identification of novel drug actions. The approach we outline in this thesis can be applied to model “in a laboratory dish” any brain disorder of public health importance in a more realistic and less variable manner, and improve rates of translational success. In a step towards personalized medicine, such improved *in vitro* disease models may be used in the future to study in detail the mechanisms of action of pharmaceuticals, and predict the clinical responsiveness of patients to prescribed drugs. Taken together, the framework provided by our studies offers manifold opportunities for medically relevant investigations aimed toward decreasing the personal, societal and economic burden of various neurological and psychiatric disorders. In this respect, productive collaboration of research scientists with clinical groups, biotechnology companies and industry is key to accelerating discovery and generating outcomes that benefit society.

REFERENCES

- Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010. *PLoS One* 10(2), e0116820 (2015).
- Lepine JP, Briley M. The increasing burden of depression. *Neuropsychiatr. Dis. Treat.* 7(Suppl 1), 3-7 (2011).
- World Health Organization. Neurological Disorders: Public Health Challenges. (2006).
- Bloom DE, Cafiero ET, Jané-Llopis E *et al.* The Global Economic Burden of Noncommunicable Diseases. Geneva: World Economic Forum. (2011).
- Glassman AH. Depression and cardiovascular comorbidity. *Dialogues Clin. Neurosci.* 9(1), 9-17 (2007).
- Kawachi I, Sparrow D, Vokonas PS, Weiss ST. Symptoms of anxiety and risk of coronary heart disease. The Normative Aging Study. *Circulation* 90(5), 2225-2229 (1994).
- Katon WJ. The comorbidity of diabetes mellitus and depression. *Am. J. Med.* 121(11 Suppl 2), S8-15 (2008).
- Daumit GL, Clark JM, Steinwachs DM, Graham CM, Lehman A, Ford DE. Prevalence and correlates of obesity in a community sample of individuals with severe and persistent mental illness. *J. Nerv. Ment. Dis.* 191(12), 799-805 (2003).
- Blaine B. Does depression cause obesity?: A meta-analysis of longitudinal studies of depression and weight control. *J. Health Psychol.* 13(8), 1190-1197 (2008).
- McGinty EE, Zhang Y, Guallar E *et al.* Cancer incidence in a sample of Maryland residents with serious mental illness. *Psychiatr. Serv.* 63(7), 714-717 (2012).
- Gau SS, Cheng AT. Mental illness and accidental death. Case-control psychological autopsy study. *Br. J. Psychiatry* 185 422-428 (2004).
- Sobocki P, Lekander I, Berwick S, Olesen J, Jonsson B. Resource allocation to brain research in Europe (RABRE). *Eur. J. Neurosci.* 24(10), 2691-2693 (2006).
- Sobocki P, Olesen J, Jonsson B. Brain research has high returns but Europe is lagging behind. *Eur. J. Neurol.* 14(6), 708-710 (2007).
- Dragunow M. The adult human brain in preclinical drug development. *Nat. Rev. Drug Discov.* 7(8), 659-666 (2008).
- 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).
- Yu J, Vodyanik MA, Smuga-Otto K *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318(5858), 1917-1920 (2007).
- Park IH, Zhao R, West JA *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 451(7175), 141-146 (2008).
- Lowry WE, Richter L, Yachechko R *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.* 105(8), 2883-2888 (2008).
- Dolmetsch R, Geschwind DH. The human brain in a dish: the promise of iPSC-derived neurons. *Cell* 145(6), 831-834 (2011).
- Grskovic M, Javaherian A, Strulovici B, Daley GQ. Induced pluripotent stem cells--opportunities for disease modelling and drug discovery. *Nat. Rev. Drug Discov.* 10(12), 915-929 (2011).
- Ebert AD, Svendsen CN. Human stem cells and drug screening: opportunities and challenges. *Nat. Rev. Drug Discov.* 9(5), 367-372 (2010).
- Bellin M, Marchetto MC, Gage FH, Mummery CL. Induced pluripotent stem cells: the new patient? *Nat. Rev. Mol. Cell Biol.* 13(11), 713-726 (2012).

23. Belinsky GS, Moore AR, Short SM, Rich MT, Antic SD. Physiological properties of neurons derived from human embryonic stem cells using a dibutyryl cyclic AMP-based protocol. *Stem Cells Dev.* 20(10), 1733-1746 (2011).
24. Hu BY, Weick JP, Yu J *et al.* Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc. Natl. Acad. Sci. U. S. A.* 107(9), 4335-4340 (2010).





Acknowledgments

Completing a PhD is a long and strenuous journey, which cannot be undertaken nor finished without the support and guidance of a community of friends, family, colleagues and supervisors. Therefore, I would like to take some time to express the gratitude due for the support I have received over the past four years. With so many of you assisting me throughout the years, I can only hope to acknowledge everyone involved, and that the following words will suffice in expressing my gratitude.

First and foremost, I would like to give heartfelt and warm thanks to Prof. Harry Steinbusch. Prof. Steinbusch, I am utmost grateful to have worked under your supervision and for the opportunity you provided to undertake this exciting, though challenging PhD project, in collaboration with Prof. Rusty Gage of the Salk Institute for Biological Studies. I was therefore not only able to further develop myself academically, but also personally, spending time away from home, in new places, meeting amazing people from different cultures. Your continuous support and guidance are greatly appreciated and have been essential in the successful completion of this PhD.

Next, I would like to express my sincere appreciation of Prof. Bart Rutten and Dr. Gunter Kenis. By also being my supervisors at Maastricht University, both of you dedicated a lot of time and energy to guide me through the entire process, and your academic expertise set the stage for taking the necessary and important decisions along the way. You both provided invaluable and thorough feedback regarding not only written pieces of work, but also presentations, approaches for data analysis and interpretation, without which the completed work would not have been what is portrayed in the previous sections of this thesis.

Then, I come to Prof. Fred (Rusty) Gage of the Salk Institute. Dear Rusty, I am utmost thankful for the opportunity to work in your lab at the Salk Institute in La Jolla during both my master's and PhD project (2011-2014). I felt like a kid in a candy store being in your lab, and my time at the Salk provided me with great scientific insights that I would not have gotten otherwise. In addition, I am very grateful for the welcoming environment, for interns in general, at the Salk Institute.

Dr. Cedric Bardy, thanks for being the direct supervisor of my scientific efforts at the Gage lab. You have provided me with great competences to ask and answer relevant scientific questions, and perform well-conceived experiments. Cedric, together we constituted, and continue to constitute, a great team, working together almost flawlessly under many different, sometimes stressful, circumstances. You have taught me how to most accurately analyze and interpret cellular and molecular experimental data, and motivated me to often question and test their validity, for which I am most grateful.

At this stage, I would like to thank all members of the Gage lab, as well as colleagues from the J. Craig Venter Institute, who in some way or another contributed to the success of my time spent in San Diego. Special thanks go to Lynne Moore, Dr. Ruben Hernandez, Dr. Jennifer (Jenny) Erwin, Dr. Manching (MC) Ku, Dr. Ben Lacar, Prof. Roger Lasken

and Dr. Mark Novotny. Lynne, thank you for your interest in and most appreciated help/guidance with my experiments; I am most grateful for all your technical advice. Also Ruben: many thanks for sharing your technical knowledge and scientific ideas with me. It is with pleasure that I look back on the evenings I spent with you and Terra at Liberty Station, enjoying great hamburgers and beers at Slater's 50/50 and Stone Brewery. Jenny, MC, Ben, Prof. Lasken and Mark: thank you so much for introducing me to the field of single-cell transcriptomics. Your invaluable technical expertise and advice enabled me to perform high-quality research.

My sincere gratitude is also extended to Dr. Jos Prickaerts, Dr. Daniel van den Hove and Dr. Lars Eijssen, all of whom are esteemed colleagues and advisors at Maastricht University, and who inspired me to be passionately involved in science and think critically. Jos, you sparked my interest in neuroscience and neurobiology early on. Your inspirational mentoring during my Research Master's brought lots of positive energy and enjoyment to this part of my education. Daniel, thanks for making me think critically during department lectures and meetings of our Translational Neuropsychiatry group, and for taking good care of me as a tutor of the Molecular Life Sciences Bachelor's course "Nerves and Senses". You taught me "Science is fun!" and I will continue my life on that very note. Lars, thank you for making me familiar with the world of bioinformatics; it is because of your expertise and advice that I was able to significantly reduce the time required for my analyses.

Continuing with my network at Maastricht University, I would like to thank Dr. Laurence de Nijs for her assistance with PCR-related issues, and for her helping me obtain mouse brain tissue material from the University of Liège. I also wish to express my gratitude to the administrative and financial staff of the School, as well as the technical staff of the lab. Rachele Capponi, thank you for your administrative support, especially in the last months leading up to my defense. You saved me a lot of time and effort by planning meetings and facilitating the communication necessary for timely completion of my PhD project. Marie-Thérèse Moers-Kockelmans, thank you for your administrative efforts and for trusting me with the responsibility of providing technical support during the highly interesting workshops of the School and EURON, organized with great care by yourself, Rachele, Nicole Senden and Peggy Bisschoff. Ankie Hochstenbach, I could always reach you with any finance-related questions, no matter how small or tedious; thank you for your advice. Lastly, the technicians: Hellen Steinbusch, Marjan Philippens, Denise Hermes and Barbie Machiels; although not always visibly present in finalized pieces of work, your technical support and advice was indispensable and invaluable.

Further, I would like to thank my assessment committee, consisting of Prof. Jos Kleinjans, Prof. Chris Evelo, Prof. Bert Smeets, Prof. Klaus-Peter Lesch and Dr. Laurent Nguyen, for critically evaluating my thesis and giving me the opportunity to defend my

research. Thank you, Prof. Kleinjans, for taking the role as chairman of the committee for the dissertation assessment and the defense.

Thomas Habets and Magdalena Weidner, thank you for being my paranymphs during the PhD defense ceremony. Thomas, we have known each other since kindergarten and have supported each other in many aspects of life, including our academic endeavors. Our coffee breaks with Kevin Gerritsen were the ultimate opportunity to discuss both the happy and difficult situations that any PhD student encounters on their journey. Thank you, and Kevin, for always being there. Magda, thanks for providing me with your knowledgeable and helpful input on the scientific questions that occasionally popped up in my head. I really enjoyed the lunch and coffee breaks we had together with Johanna Zöller.

Further, I would like to thank my other roommates and fellow/past PhD students within the Translational Neuropsychiatry group of our School. Marion Lévy, Artemis Iatrou, Roy Lardenoije, Dr. Ehsan Pishva, Dr. Nicole Leibold and An Ning, thank you for being on the same team, and enduring all my stories about CrossFit. I enjoyed the really good times with you guys.

Finally, all my Maastricht University colleagues that I have failed to mention yet, especially Jo Stevens, Sandra Schipper, João Casaca Carreira, Christian Bertens and Carolin Hoffmann, as well as my fellow lab members and friends at the Salk Institute—Ben Galet, Mark Gorris and Martijn Kelder: I would like to thank all of you for your continued encouragement. Specifically for my friends at Salk, I am very grateful for all the happy times that we spent together.

Continuing with the more personal part of my network, my close friends and family. I would like to express my gratitude to everyone at CrossFit Batteraof, for making my life more “constantly varied” and providing me with the necessary distraction from my hard work during the day. Then, my dear friends Christiaan Riemens and Maarten Strijbosch, with whom I enjoyed many “guy’s nights” watching great movies and eating great food.

Of course, I’d like to thank all members of my family and my mom and dad in particular. Mom and Dad, thank you for supporting me throughout my academic career. You are my rocks in an ocean of uncertainties, always steady, safe and within sight. I would not have finished this work without your continuous encouragement.

And last, but not least, my dearest Jolien. Without a doubt, the last months of the PhD project are the hardest, with finishing all the written parts of the thesis. You endured all my late nights and weekends of typing and pushed me across the finish line. The typing is over now, adventure awaits!





Curriculum Vitae

Mark van den Hurk was born on the 12th of April 1989 in Maastricht, The Netherlands. After pre-university secondary education (Sint-Maartenscollege, Maastricht), he started his Bachelor studies in Molecular Life Sciences at Maastricht University in 2007. Following the attainment of his degree in 2010, he continued with and successfully completed (2012) the two-year Research Master's in Cognitive and Clinical Neuroscience, specialization Fundamental Neuroscience. For his research internship and master thesis, he joined the Laboratory of Genetics of Prof. Fred Gage at the renowned Salk Institute for Biological Studies in La Jolla, California, United States, where he was actively involved in a project aimed at the functional characterization of human stem cell-derived neuronal networks. This work as an exchange scholar, directed by Dr. Cedric Bardy, resulted in a four-year collaborative PhD study between the translational neuropsychiatry group of the Department of Psychiatry and Neuropsychology at Maastricht University, and the Salk. His PhD research, supervised by Dr. Bardy (Salk) and Prof. Harry Steinbusch, Prof. Bart Rutten and Dr. Gunter Kenis (Maastricht University), employed an *in vitro* neuronal model derived from human stem cells to study at the single-cell molecular level human neuronal function. The results of his PhD project are described in detail in this doctoral thesis. Mark will continue his work as a postdoctoral researcher in Dr. Bardy's Laboratory for Human Neurophysiology and Genetics that has recently been established within the SAHMRI Mind & Brain theme headed by Dr. Julio Licinio in Adelaide, South Australia.





List of Publications

PUBLISHED PEER-REVIEWED SCIENTIFIC JOURNAL ARTICLES

C. Bardy, M. van den Hurk*, B. Kakaradov*, J.A. Erwin, B.N. Jaeger, R.V. Hernandez, T. Eames, A.A. Paucar, M. Gorris, C. Marchand, R. Jappelli, J. Barron, A.K. Bryant, M. Kellogg, R.S. Lasken, B.P. Rutten, H.W. Steinbusch, G.W. Yeo, F.H. Gage (*contributed equally). **Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology.** *Mol. Psychiatry* 2016; 21(11):1573-88. doi: 10.1038/mp.2016.158

M. van den Hurk, G. Kenis*, C. Bardy*, D.L. van den Hove, F.H. Gage, H.W. Steinbusch, B.P. Rutten (*contributed equally). **Transcriptional and epigenetic mechanisms of cellular reprogramming to induced pluripotency.** *Epigenomics* 2016; 8(8):1131-49. doi: 10.2217/epi-2016-0032

C. Bardy, M. van den Hurk*, T. Eames*, C. Marchand*, R.V. Hernandez*, M. Kellogg*, M. Gorris, B. Galet, V. Palomares, J. Brown, A.G. Bang, J. Mertens, L. Böhnke, L. Boyer, S. Simon, F.H. Gage (*contributed equally). **Neuronal medium that supports basic synaptic functions and activity of human neurons *in vitro*.** *Proc. Natl. Acad. Sci. U.S.A.* 2015; 112(20):E2725-34. doi: 10.1073/pnas.1504393112

C.P. Fitzsimons, E. van Bodegraven, M. Schouten, R. Lardenoije, K. Kompotis, G. Kenis, M. van den Hurk, M.P. Boks, C. Biojone, S. Joca, H.W. Steinbusch, K. Lunnon, D.F. Mastroeni, J. Mill, P.J. Lucassen, P.D. Coleman, D.L. van den Hove, B.P. Rutten. **Epigenetic regulation of adult neural stem cells: implications for Alzheimer's disease.** *Mol. Neurodegener.* 2014; 9:25. doi: 10.1186/1750-1326-9-25

IN PREPARATION

M. van den Hurk, G. Kenis, L. Eijssen, B. Kakaradov, J.A. Erwin, R.V. Hernandez, T. Eames, M. Gorris, C. Marchand, A.K. Bryant, M. Kellogg, R.S. Lasken, G.W. Yeo, H.W. Steinbusch, F.H. Gage, B.P. Rutten, C. Bardy. **Identification of a co-expression gene network and transcriptional regulators for functional maturity based on electrophysiological, morphological and transcriptomic profiling of iPSC-derived single human neurons.**





Thesis Defenses from MHeNS

THESIS DEFENSES FROM MH&NS – SCHOOL FOR MENTAL HEALTH & NEUROSCIENCE

2013

- Rob Havermans: **Bipolar disorder in daily life; Mood and cortisol responses to naturally occurring events.** Supervisor: Prof.dr. M. de Vries; Co-Supervisor: Dr. N. Nicolson.
- Véronique Moers-Hornikx: **Deep brain stimulation and the cerebellum.** Supervisors: Prof.dr. J. Vles / Prof.dr. Y. Temel; Co-Supervisor: Dr. G. Hoogland.
- Nicole Veldhorst-Janssen: **Intranasal delivery of rapid acting drugs.** Supervisors: Prof.dr. M. Marcus / Prof.dr. C. Neef; Co-Supervisor: Dr. P.H. van der Kuy.
- Stéphanie Knippenberg: **Vitamin D and Multiple Sclerosis: immunological and clinical outcome.** Supervisor: Prof.dr. J. Cohen-Tervaert; Co-Supervisors: Dr. J. Damoiseaux / Dr. Y. Bols.
- Erik D. Gommer: **Dynamic Cerebral Autoregulation: from methodology towards clinical application.** Supervisors: Prof.dr. W.H. Mess / Prof.dr. R.B. Panerai, UK; Co-Supervisor: Dr.ir. J.P.H. Reulen.
- Olga A.H. Reneerkens: **Can PDE inhibition improve cognition ? Translational insights.** Supervisor: Prof.dr. H.W.M. Steinbusch; Co-Supervisor: Dr. J. Prickaerts.
- Lyzel S. Elias-Sonnenschein: **Clinical and biomarker correlates of genetic risk factors for Alzheimer's disease.** Supervisor: Prof.dr. F.R.J. Verhey; Co-Supervisor: Dr. P.J. Visser.
- Diego F. Mastroeni: **Epigenetic Dysregulation and the Pathophysiology of of Alzheimer's Disease.** Supervisors: Prof.dr. H.W.M. Steinbusch / Prof.dr. P.D. Coleman, Sun City, Arizona; Co-Supervisors: Dr. B.P.F. Rutten / Dr. D.L.A. van den Hove.
- Leonidas Chouliaras: **Epigenetic Regulation in Aging and Alzheimer's disease: A translational perspective.** Supervisor: Prof.dr. H.W.M. Steinbusch; Co-Supervisors: Dr. B.P.F. Rutten / Dr. D.L.A. van den Hove.
- Liesbeth Knaepen: **Perinatal events and altered pain sensitivity in later life.** Supervisors: Prof.dr. E.A.J. Joosten / Prof.dr. D. Tibboel, EUR; Co-Supervisor: Dr. J. Patijn.
- Marisela Martinez-Claros: **Hippocampal plasticity and corticosterone: From dendrites to behaviour.** Supervisor: Prof.dr. H.W.M. Steinbusch; Co-Supervisors: Dr. J.L. Pawluski / Dr. J. Prickaerts.
- Marcus D. Lancé: **A circle of improvement in bleeding management: from laboratory to clinic and back.** Supervisors: Prof.dr. M.A.E. Marcu / Prof.dr. J.W.M. Heemskerk; Co-Supervisor: Dr. Y.M.C. Henskens.
- Hilde Braakman: **Imaging the brain; neuronal correlates of cognitive impairment in children with frontal lobe epilepsy.** Supervisors: Prof.dr. A.P. Aldenkamp / Prof.dr. J.S.H. Vles; Co-Supervisors: Dr.ir. W.H. Backes / Dr. P.A.M. Hofman.
- Willem H. van Zwam: **Aneurysmal subarachnoid hemorrhage: imaging strategies and cost-effectiveness aspects in diagnostic work-up and post-therapeutic follow-up.** Supervisors: Prof.dr. J.T. Wilmink / Prof.dr. J.E. Wildberger; Co-Supervisor: Dr. P.A.M. Hofman.
- Klara De Cort: **The Pathogenesis of Panic Disorder.** Supervisors: Prof.dr. I. Myin-Germeys / Prof.dr. E.J.L. Griez; Co-Supervisors: Dr. K.R.J. Schruers / Dr. I. Van Diest, Leuven.
- Kim van Wijck: **Mind the Gap; experimental studies on splanchnic hyperfusion and gastrointestinal integrity loss in man.** Supervisors: Prof.dr. W.A. Buurman / Prof.dr. C.H.C. Dejong; Co-Supervisor: Dr. K. Lenaerts.
- Yvette Roke: **Antipsychotic-induced hyperprolactinemia in children and adolescents with mainly autism spectrum disorders. Prevalence, symptoms, clinical consequences and genetic risk factors.** Supervisors: Prof.dr. P.N. van Harten / Prof.dr. J.K. Buitelaar (RUN); Co-Supervisor: Dr. A. Boot (UMCG).

- Fleur Goezinne: **Retinal detachment surgery: pre and postoperative prognostic factors.** Supervisors: Prof.dr. F. Hendrikse / Prof.dr. C.A.B. Webers; Co-Supervisor: Dr. E.C. La Heij (Amsterdam).
- Ralph L.J.G. Maassen: **The Merits of Videolaryngoscopy during Glottic Visualisation for Endotracheal Intubation.** Supervisors: Prof.dr. M. Marcus / Prof.dr. A. van Zundert (University of Queensland).
- Maria J. de Sousa Guerreiro: **The role of sensory modality in age-related distraction.** Supervisor: Prof.dr. C.M. van Heugten; Co-Supervisor: Dr. P.W.M. van Gerven.
- Ine Rayen: **Effects of developmental fluoxetine exposure on neurobehavioral outcomes.** Supervisor: Prof.dr. H.W.M. Steinbusch; Co-Supervisors: Dr. J.L. Pawluski / Dr. T.D. Charlier (Ohio University, USA).
- Nynke M.G. Bodde: **Psychogenic non-epileptic seizures; a separate disorder or part of a continuum?** Supervisors: Prof.dr. R. van Oostenbrugge / Prof.dr. K. Vonck (UZ Gent); Co-Supervisors: Dr. R. Lazeron / Dr. A. de Louw (Epilepsiecentrum Kempenhaeghe, Heeze).
- Alejandro M. Gomez: **Novel strategies for making myasthenia less gravis: targeting plasma cells and the neuromuscular junction.** Supervisor: Prof.dr. M.H. De Baets; Co-Supervisors: Dr. M. Losen / Dr. P. Martinez-Martinez.
- Mohammad S. Rahnama'i: **Prostaglandins and Phosphodiesterases in the Urinary Bladder Wall.** Supervisors: Prof.dr. Ph. Van Kerrebroeck / Prof.dr. S. de Wachter (Universiteit Antwerpen); Co-Supervisor: Dr. G. van Koeveeringe.
- Mariken B. de Koning: **Studying biomarkers in populations at genetic and clinical high risk for psychosis.** Supervisors: Prof.dr. T. Amelsvoort / Prof.dr. J. Booij (AMC).
- Fabien Boulle: **Epigenetic regulation of BDNF/TrkB signaling in the pathophysiology and treatment of mood disorders.** Supervisors: Prof.dr. H.W.M. Steinbusch / Prof.dr. L. Lanfumey (Universiteit Parijs); Co-Supervisors: Dr. D. van den Hove / Dr. G. Kenis.

2014

- Iris Nowak-Maes: **Tinnitus; assessment of quality of life & cost-effectiveness.** Supervisors: Prof.dr. M. Peters / Prof.dr. B. Kremer; Co-Supervisors: Dr. M. Joore / Dr. L. Anteunis.
- Marjolein Huijts: **Cognitive function in patients with cerebral small vessel disease.** Supervisor: Prof.dr. R.J. van Oostenbrugge; Co-Supervisors: Dr. A.A. Duits / Dr. J. Staals.
- Markus Gantert: **Fetal inflammatory injury as origin of long term disease: Lessons from animal models.** Supervisors: Prof.dr. B. Kramer / Prof.dr. L. Zimmermann; Co-Supervisor: Dr. A. Gavilanes.
- Elke Kuypers: **Fetal development after antenatal exposures: Chorioamnionitis and maternal glucocorticoids.** Supervisors: Prof.dr. B.W. Kramer / Prof.dr. H.W. Steinbusch / Prof.dr. Suhas G. Kallapur (University of Cincinnati, Ohio, USA).
- Pieter Kubben: **Ultra low-field strength intraoperative MRI for Glioblastoma Surgery.** Supervisor: Prof.dr. J.J. van Overbeeke; Co-Supervisor: Dr. H. van Santbrink.
- Laura Baijens: **Surface electrical stimulation of the neck for oropharyngeal dysphagia in Parkinson's disease: therapeutic aspects and reliability of measurement.** Supervisor: Prof.dr. B. Kremer; Co-Supervisor: Dr. R. Speyer, Townsville.
- Janneke Hoeijmakers: **Small fiber neuropathy and sodium channels; a paradigm shift.** Supervisor: Prof.dr. R.J. van Oostenbrugge; Co-Supervisors: Dr. C.G. Faber / Dr. I.S.J. Merkies.
- Stephanie Vos: **The Role of biomarkers in preclinical and prodromal Alzheimer's disease.** Supervisor: Prof. dr. F.R. Verhey; Co-Supervisor: Dr. P.J. Visser.
- Muriël Doors: **The Value of Optical Coherence Tomography in Anterior Segment Surgery.** Supervisors: Prof. dr. R.M. Nuijts / Prof.dr. C.A. Webers; Co-Supervisor: Dr. T.T.J.M. Berendschot.

- Anneke Maas: **Sleep problems in individuals with genetic disorders associated with intellectual disability.**
Supervisors: Prof.dr. I. Curfs / Prof.dr. R. Didden.
- Sebastiaan van Gorp: **Translational research on spinal cord injury and cell-based therapies; a focus on pain and sensorimotor disturbances.** Supervisors: Prof.dr. B. Joosten / Prof.dr. M. van Kleef; Co-Supervisors: Dr. J. Patijn / Dr. R. Deumens, KU Leuven.
- Andrea Sanna: **High risk newborns and brain biochemical monitoring.** Supervisor: Prof.dr. J.S.H. Vles; Co-Supervisors: Dr. D. Gazzolo, Alessandria, Italy / Dr. A.W.D. Gavilanes.
- Julie A.D.A. Dela Cruz: **Dopamine mechanisms in learning and memory: Evidence from rodent studies.** Supervisors: Prof.dr. H.W.M. Steinbusch / Prof.dr. R.J. Bodnar, New York; Co-Supervisor: Dr. B.P.F. Rutten.
- René Besseling: **Brain wiring and neuronal dynamics; advances in MR imaging of focal epilepsy.** Supervisors: Prof.dr. A.P. Aldenkamp / Prof.dr.ir. W.H. Backes; Co-Supervisor: dr. J.F.A. Jansen.
- Maria Quint-Fens: **Long-term care after stroke; development and evaluation of a long-term intervention in primary care.** Supervisors: Prof.dr. J.E.M. Metsemakers / Prof.dr. C.M. van Heugten / Prof.dr. M. Limburg, Almere; Co-Supervisor: dr. G.H.M.I. Beusmans.
- Veronique Moulaert: **Life after survival of a cardiac arrest; the heart of the matter.** Supervisors: Prof.dr. J.A. Verbunt / Prof.dr. C.M. van Heugten / Prof.dr. D.T. Wade, Oxford, UK.
- Feikje Smeets: **The hallucinatory-delusional state: a crucial connection in the psychosis symptom network.**
Supervisor: Prof.dr. J. van Os; Co-Supervisor: Dr. T. Lataster.
- Lies Clerx: **Alzheimer's disease through the MR-eye; novel diagnostic markers and the road to clinical implementation?** Supervisor: Prof.dr. F. Verhey; Co-Supervisors: Dr. P.J. Visser / P. Aalten.
- Sonny Tan: **The subthalamic nucleus in Parkinson's disease.** Supervisors: Prof.dr. Y. Temel / Prof.dr. H.W.M. Steinbusch / Prof.dr. T. Sharp, Oxford, UK / Prof.dr. V. Visser-Vandewalle, Köln.
- Koen van Boxem: **The use of pulsed radiofrequency in the management of chronic lumbosacral radicular pain.** Supervisors: Prof.dr. M. van Kleef / Prof.dr. E.A.J. Joosten; Co-Supervisor: Assoc. Prof.dr. J. van Zundert.
- Jérôme Waterval: **Hyperostosis cranialis interna.** Supervisors: Prof.dr. J.J. Manni / Prof.dr. R.J. Stokroos.
- Sylvie Kolfshoten-van der Kruijs: **Psychogenic non-epileptic seizures; the identification of neurophysiological correlates.** Supervisors: Prof.dr. A.P. Aldenkamp / Prof.dr. K.E.J. Vonck, Universiteit Gent; Co-Supervisors: Dr. J.F.A. Jansen / Dr. R.H.C. Lazon, Kempenhaeghe.
- Wouter Pluijms: **Spinal cord stimulation and pain relief in painful diabetic: polyneuropathy, a translational approach.** Supervisors: Prof.dr. M. van Kleef / Prof.dr. E.A. Joosten; Co-supervisor: Dr. C.G. Faber.
- Ron Handels: **Health technology assessment of diagnostic strategies for Alzheimer's disease.** Supervisors: Prof.dr. F.R.J. Verhey / Prof.dr. J.L. Severens (EUR); Co-Supervisor: Dr. M.A. Joore / Dr. C.A.G. Wolfs.
- Evelyn Peelen: **Regulatory T cells in the pathogenesis of Multiple Sclerosis: potential targets for vitamin D therapy.** Supervisors: Prof.dr. R.M.M. Hupperts / Prof.dr. J.W. Cohen Tervaert; Co-Supervisor: Dr. J.G.M.C. Damoiseaux / Dr. M.M.G.L. Thewissen, Diepenbeek.
- Reint Jellema: **Cell-based therapy for hypoxic-ischemic injury in the preterm brain.** Supervisors: Prof.dr. B.W.W. Kramer / Prof.dr. H.W.M. Steinbusch; Co-Supervisor: Dr. W.T.V. Germeraad / Dr. P. Andriessen, Veldhoven.
- Maria Wertli: **Prognosis of Chronic Clinical Pain Conditions: The Example of Complex Regional Pain Syndrome 1 and Low Back Pain.** Supervisors: Prof.dr. M. van Kleef; Co-Supervisor: Dr. F. Brunner, Zürich / Dr. R. Perez, VUmc.
- Dagmar Zeef: **An experimental model of Huntington's disease: Validation & Stimulation.** Supervisors: Prof. dr. Y. Temel / Prof.dr. H.W.M. Steinbusch; Co-supervisor: Dr. A. Jahanshahi.
- Jeroen Decoster: **Breaking Down Schizophrenia into phenes, genes and environment.** Supervisors: Prof.dr. I. Myin-Germeys / Prof.dr. M. De Hert, KU Leuven; Co-Supervisor: Dr. R. van Winkel.

- Eaja Anindya Sekhar Mukherjee: **Fetal Alcohol Spectrum Disorders: exploring prevention and management.** Supervisor: Prof.dr. L.M.G. Curfs; Co-Supervisor: Prof. S. Hollins, St. George's University of London, UK.
- Catherine van Zelst: **Inside out; On stereotype awareness, childhood trauma and stigma in psychosis.** Supervisors: Prof.dr. Ph. Delespaul / Prof.dr. J. van Os.
- Ibrahim Tolga Binbay: **Extended Psychosis Phenotype in the Wider Social Environment.** Supervisor: Prof.dr. J. van Os; Co-Supervisor: Dr. M. Drukker.
- Frank Van Dael: **OCD matters in psychosis.** Supervisors: Prof.dr. J. van Os / Prof.dr. I. Myin-Germeys.
- Pamela Kleikers: **NOXious oxidative stress: from head toe too and back.** Supervisors: Prof.dr. H.H.H.W. Schmidt / Prof.dr. H.W.M. Steinbusch; Co-Supervisor: Dr. B. Janssen.
- José Luis Gerardo Nava: **In vitro assay systems in the development of therapeutic interventions strategies for neuroprotection and repair.** Supervisors: Prof.dr.med. J. Weis / Prof.dr. H.W.M. Steinbusch; Co-Supervisor: Dr. G.A. Brook, RWTH Aachen.
- Eva Bollen: **Cyclic nucleotide signaling and plasticity.** Supervisors: Prof.dr. H.W.M. Steinbusch / Prof.dr. R. D'Hooge, KU Leuven; Co-Supervisor: Dr. J. Prickaerts.

2015

- Jessica A. Hartmann: **A good laugh and a long sleep; Insights from prospective and ambulatory assessments about the importance of positive affect and sleep in mental health.** Supervisor: Prof.dr. J. van Os; Co-Supervisors: C.J.P. Simons / Dr. M. Wichers.
- Bart Ament: **Frailty in old age; conceptualization and care innovations.** Supervisors: Prof.dr. G.I.J.M. Kempen / Prof.dr. F.R.J. Verhey; Co-Supervisor: Dr. M.E. de Vugt.
- Mayke Janssens: **Exploring course and outcome across the psychosis-continuum.** Supervisor: Prof.dr. I. Myin-Germeys; Co-Supervisor: Dr. T. Lataster.
- Dennis M.J. Hernau: **Dopayours is not dopamine: genetic, environmental and pathological variations in dopaminergic stress processing.** Supervisor: Prof.dr. I. Myin-Germeys; Co-Supervisors: Prof.dr. F.M. Mottaghy / Dr. D. Collip.
- Ingrid M.H. Brands: **The adaptation process after acquired brain injury Pieces of the puzzle.** Supervisors: Prof.dr. C.M. van Heugten / Prof.dr. D.T. Wade, Oxford UK; Co-Supervisors: Dr. S.Z. Stapert / Dr. S. Köhler.
- Francesco Rizzo: **Urinary and salivary S100B monitoring in high risk infants.** Supervisor: Prof.dr. J.S.H. Vles; Co-Supervisors: Dr. D. Gazzolo, Genoa, Italy / Dr. A.W.D. Gavilanes.
- Alessandro Borghesi: **Stem and Progenitor Cells in Preterm Infants: Role in the Pathogenesis and Potential for Therapy.** Supervisor: Prof.dr. L. Zimmermann; Prof.dr. B. Kramer; Co-Supervisors: Dr. D. Gazzolo, Genoa, Italy / Dr. A.W.D. Gavilanes.
- Claudia Menne-Lothmann: **Affect dynamics; A focus on genes, stress, and an opportunity for change.** Supervisor: Prof.dr. J. van Os; Co-Supervisors: Dr. M. Wichers / Dr. N. Jacobs.
- Martine van Nierop: **Surviving childhood new perspectives on the link between childhood trauma and psychosis.** Supervisors: Prof.dr. I. Myin-Germeys / Prof.dr. J. van Os; Co-Supervisor: Dr. R. van Winkel.
- Sylvia Klinkenberg: **VNS in children; more than just seizure reduction.** Supervisors: Prof.dr. J. Vles / Prof.dr. A. Aldenkamp; Co-Supervisor: Dr. H. Majoie.
- Anouk Linssen: **Considerations in designing an adult hearing screening programme.** Supervisor: Prof.dr. B. Kremer; Co-Supervisors: Dr. L. Anteunis / Dr. M. Joore.
- Janny Hof: **Hearing loss in young children; challenges in assessment and intervention.** Supervisors: Prof.dr. B. Kremer / Prof.dr. R. Stokroos / Prof.dr. P. van Dijk, RUG; Co-Supervisor: Dr. L. Antheunis.

- Kimberly Cox-Limpens: **Mechanisms of endogenous brain protection; Clues from the transcriptome.** Supervisors: Prof.dr. J. Vles / Prof.dr. L. Zimmermann; Co-Supervisor: Dr. A. Gavilanes.
- Els Vanhoutte: **Peripheral Neuropathy outcome measures; Standardisation (PeriNomS) study part 2: Getting consensus.** Supervisors: Prof.dr. C. Faber / Prof.dr. P. van Doorn; Co-Supervisor: Dr. I. Merkies, Spaarne ziekenhuis Hoofddorp.
- Mayienne Bakkers: **Small fibers, big troubles; diagnosis and implications of small fiber neuropathy.** Supervisors: Prof.dr. C. Faber / Prof.dr. M. de Baets; Co-Supervisor: Dr. I. Merkies, Spaarne ziekenhuis Hoofddorp.
- Ingrid Kramer: **Zooming into the micro-level of experience: An approach for understanding and treating psychopathology.** Supervisor: Prof.dr. J. van Os; Co-Supervisors: Dr. M. Wichers, UMC Groningen / Dr. C. Simons.
- Esther Bouman: **Risks and Benefits of Regional Anesthesia in the Perioperative Setting.** Supervisors: Prof. dr. M. van Kleef / Prof.dr. M. Marcus, HMC, Qatar / Prof.dr. E. Joosten; Co-Supervisor: Dr. H. Gramke.
- Mark Janssen: **Selective stimulation of the subthalamic nucleus in Parkinson's disease; dream or near future.** Supervisors: Prof.dr. Y. Temel / Prof.dr. V. Visser-Vandewalle, Keulen / Prof.dr. A. Benazzouz, Bordeaux, France.
- Reina de Kinderen: **Health Technology Assessment in Epilepsy; economic evaluations and preference studies.** Supervisors: Prof.dr. S. Evers / Prof.dr. A. Aldenkamp; Co-Supervisor: Dr. H. Majoie / Dr. D. Postulart, GGZ O-Brabant.
- Saskia Ebus: **Interictal epileptiform activity as a marker for clinical outcome.** Supervisors: Prof.dr. A. Aldenkamp / Prof.dr. J. Arends, TUE / Prof.dr. P. Boon, Universiteit Gent, België.
- Inge Knuts: **Experimental and clinical studies into determinants of panic severity.** Supervisor: Prof.dr. I. Myin-Germeys; Co-Supervisor: Dr. K. Schruers; Influencing panic.
- Nienke Tielemans: **Proactive coping post stroke: The Restored4Stroke Self-Management study.** Supervisors: Prof.dr. C. van Heugten / Prof.dr. J. Visser-Meily, UMC Utrecht; Co-Supervisor: Dr. V. Schepers, UMC Utrecht.
- Tom van Zundert: **Improvements Towards Safer Extraglottic Airway Devices.** Supervisors: Prof.dr. A.E.M. Marcus / Prof.dr. W. Buhre / Prof.dr. J.R. Brimacombe, Queensland, Australia / Prof.dr. C.A. Hagberg.
- Tijmen van Assen: **Anterior Cutaneous Nerve Entrapment Syndrome Epidemiology and surgical management.** Supervisors: Prof.dr. G.L. Beets / Prof.dr. M. van Kleef / Dr. R.M.H. Roumen / Dr. M.R.M. Scheltinga, MMC Veldhoven.
- Rohit Shetty: **Understanding the Clinical, Immunological and Genetic Molecular Mechanisms of Keratoconus.** Supervisors: Prof.dr. R.M.M.A. Nuijts / Prof.dr. C.A.B. Webers.
- Christine van der Leeuw: **Blood, bones and brains; peripheral biological endophenotypes and their structural cerebral correlates in psychotic disorder.** Supervisor: Prof.dr. J. van Os; Co-supervisor: Dr. M. Marcelis.
- Sanne Peeters: **The Idle Mind Never Rests; functional brain connectivity across the psychosis continuum.** Supervisor: Prof.dr. J. van Os; Co-supervisor: dr. M. Marcelis.
- Nick van Goethem: **$\alpha 7$ nicotinic acetylcholine receptors and memory processes: mechanistic and behavioral studies.** Supervisor: Prof.dr. H.W.M. Steinbusch; Co-supervisor: Dr. J. Prickaerts.
- Nicole Leibold: **A Breath of fear; a translational approach into the mechanisms of panic.** Supervisor: Prof.dr. H.W.M. Steinbusch; Co-supervisors: Dr. K.R.J. Schruers / Dr. D.L.A. van den Hove.
- Renske Hamel: **The course of mild cognitive impairment and the role of comorbidity.** Supervisor: Prof.dr. F.R.J. Verhey; Co-supervisors: Dr. I.H.G.B. Ramakers / Dr. P.J. Visser.

- Lucia Speth: **Effects of botulinum toxin A injections and bimanual task-oriented therapy on hand functions and bimanual activities in unilateral Cerebral Palsy.** Supervisors: Prof.dr. J. Vles; Prof.dr. R. Smeets; Co-supervisor: Dr. Y. Janssen-Potten, Adelante Hoensbroek.
- Yuan Tian: **The effects of Lutein on the inflammatory pathways in age-related macular degeneration (AMD).** Supervisors: Prof.dr. C. Webers; Prof.dr. A. Kijlstra, WUR; Co-supervisor: Dr. M. Spreeuwenberg; Dr. H. Tange.
- Peggy Spauwen: **Cognition and Type 2 diabetes; the interplay of risk factors.** Supervisors: Prof.dr. F. Verhey; Prof.dr. C. Stehouwer; Co-supervisor: Dr. M. van Boxtel
- Marc Hilhorst: **Crescentic glomerulonephritis in ANCA associated vasculitis.** Supervisors: Prof.dr. J. Cohen-Tervaert; Co-supervisor: Dr. P. van Paassen
- Martin Gevonden: **The odd one out: exploring the nature of the association between minority status and psychosis.** Supervisors: Prof.dr. J-P. Selten; Prof.dr. J. Booij, Uva; Prof.dr. I. Myin-Germeys
- Bart Biallosterski: **Structural and functional aspects of sensory-motor Interaction in the urinary bladder.** Supervisors: Prof.dr. Ph. Van Kerrebroeck; Prof.dr. S. De Wachter, UvAntwerpen; Co-supervisors: Dr. G. van Koeveringe; Dr. M. Rahnama'i.
- Alexandra König: **The use of information and communication technologies (ICT) for the assessment of patients with Alzheimer's Disease and related disorders.** Supervisors: prof.dr. F. Verhey; prof.dr. Ph. Robert, Nice, Fr; Co-supervisors: dr. P. Aalten; dr. R. David, Nice, Fr.
- Micheline Chenault: **Assessing Readiness for Hearing Rehabilitation.** Supervisors: prof.dr. M.P.F. Berger; prof.dr. B. Kremer; Co-supervisor: dr. L.J.C. Anteunis.
- Anand Vinekar: **Retinopathy of Prematurity. Recent advances in tele-medicine screening, risk factors and spectral domain optical coherence tomography imaging.** Supervisor: prof.dr. C.A.B. Webers; Co-supervisor: dr. N.J. Bauer
- Fleur van Dooren: **Diabetes and Depression: exploring the Interface between Pathophysiological and Psychological factors.** Supervisors: prof.dr. F.R.J. Verhey; prof.dr. J.K.L. Denollet, UvT; prof.dr. F. Pouwer, UvT; Co-supervisor: dr. M.T. Schram.
- Gabriëlla Pons van Dijk: **Taekwondo and physical fitness components in middle-aged healthy volunteers; the Sekwondo study.** Supervisors: prof.dr. J. Lodder; prof.dr. H. Kingma; Co-supervisor: dr. A.F. Lenssen.
- Yara Pujol López: **Development and psychoneuroimmunological mechanisms in depression.** Supervisor: prof.dr. H.W.M. Steinbusch; Co-supervisors: Dr. G. Kenis; Dr. D. van den Hove; Dr. Aye Mu Myint, München.
- Romina Gentier: **UBB⁴²; an important switch in the onset of Alzheimer's disease.** Supervisors: Prof. H. Steinbusch; Prof. D. Hopkins; Co-supervisor: Dr. F. van Leeuwen.
- Sanne Smeets: **Insights into insight: studies on awareness of deficits after acquired brain injury.** Supervisor: Prof. C. van Heugten; Prof. R. Ponds; Co-supervisor: Dr. I. Winkens
- Kim Beerhorst: **Bone disease in chronic epilepsy: fit for a fracture.** Supervisor: Prof. A. Aldenkamp; Prof. R. van Oostenbrugge; Co-supervisor: Dr. P. Verschuure.
- Alex Zwanenburg: **Cerebral and cardiac signal monitoring in fetal sheep with hypoxic-ischemic encephalopathy.** Supervisor: Prof. T. Delhaas; Prof. B. Kramer; Co-supervisors: Dr. T. Wolfs; Dr. P. Andriessen, MMC.
- Ismail Sinan Guloksuz: **Biological mechanisms of environmental stressors in psychiatry.** Supervisor: Prof. J. van Os; Co-supervisors: Dr. B. Rutten; Dr. M. Drukker.
- Seyed Ehsan Pishva MD: **Environmental Epigenetics in mental health and illness.** Supervisor: Prof.dr. J. van Os; Co-supervisors: Dr. B.P.F. Rutten; Dr. G. Kenis.
- Ankie Hamaekers: **Rescue ventilation using expiratory ventilation assistance; innovating while clutching at straws.** Supervisors: Prof.dr. W.F. Buhre; Prof.dr. M. van Kleef.

- Rens Evers. **22q11.2 deletion syndrome: intelligence, psychopathology and neurochemistry at adult age.** Supervisors: Prof.dr. L.M.G. Curfs; Prof.dr. T. v. Amelsvoort.
- Sarah-Anna Heschem. **Novel insights towards memory restoration.** Supervisor: Prof.dr. Y. Temel; Co-supervisor: Dr. A. Blokland; Dr. A. Jahanshahi.
- João P. da Costa Alvares Viegas Nunes. **Insulin receptor sensitization improves affective pathology in various mouse models.** Supervisor: Prof.dr. H.W.M. Steinbusch; Co-supervisors: Dr. K-P. Lesch; Dr. T. Strelakova; Dr.B.H. Cline, Oxford.
- Yanny Ying-Yee Cheng. **Clinical Outcomes After Innovative Lamellar Corneal Transplantation Surgery.** Supervisor: Prof.dr. R.M.M.A. Nuijts; Co-supervisor: Dr. J.S.A.G. Schouten.

2016

- Oliver Gerlach. **Parkinson's disease, deterioration during hospitalization.** Supervisor: Prof.dr. R. van Oostenbrugge; Co-supervisor: Dr. W. Weber.
- Remo Arts. **Intracochlear electrical stimulation to suppress tinnitus.** Supervisor: Prof.dr. R.J. Stokroos; Co-supervisor: Dr. E.L.J. Georg.
- Mitchel van Eeden. **The €-Restore4stroke study: Economic evaluation of stroke care in the Netherlands.** Supervisors: Prof.dr.mr. S.M.A.A. Evers; Prof.dr. C.M. v. Heugten; Co-supervisor: dr. G.A.P. van Mastrigt.
- Pim Klarenbeek. **Blood pressure and cerebral small vessel disease.** Supervisor: Prof.dr. R.J. van Oostenbrugge; Co-supervisor: Dr. J. Staals.
- Ramona Hohnen. **Peripheral pharmacological targets to modify bladder contractility.** Supervisor: Prof.dr. Ph.E.V. van Kerrebroeck; Co-supervisors: Dr. G.A. van Koeveringe; Dr. M.A. Sahnama'i; Dr. C. Meriaux.
- Ersoy Kocabicak. **Deep brain stimulation of the subthalamic nucleus: Clinical and scientific aspects.** Supervisors: Prof.dr. Y. Temel; Prof.dr. K. van Overbeeke; Co-supervisor: Dr. A. Jahanshahi.
- Sven Akkerman. **Temporal aspects of cyclic messenger signaling in object recognition memory; a pharmacological approach.** Supervisor: Prof.dr. H.W.M. Steinbusch; Co-supervisors: dr. J. Prickaerts; dr. A. Blokland.
- Anja Moonen. **Emotion and Cognition in Parkinson's disease; etiology and neurobiological mechanisms.** Supervisor: Prof.dr. F.R.J. Verhey; Co-supervisor: dr. A.F.G. Leentjens.
- Anna Schüth. **Three-dimensional bladder tissue morphology.** Supervisors: Prof.dr. G.A. van Koeveringe; Prof.dr. M. v. Zandvoort, Aachen; Prof.dr. Ph. V. Kerrebroeck.
- Elisabeth van der Ven. **Ethnic minority position as risk indicator for autism-spectrum and psychotic disorders.** Supervisors: Prof.dr. J.P. Selten; Prof.dr. J. van Os.
- Zuzana Kasanova. **Environmental reactivity for better or worse; The impact of stress and reward on neurochemistry, affect and behavior across the psychosis continuum.** Supervisor: Prof.dr. I. Myin-Germeys, KU Leuven/UM; Co-supervisor: dr. D. Collip.
- Danielle Lambrechts. **Ketogenic diet therapies; treatment for children and adults with refractory epilepsy.** Supervisors: Prof.dr. H.J.M. Majoie; Prof.dr. J.S.H. Vles; Prof.dr. A.P. Aldenkamp; Co-supervisor: dr. A.J.A. de Louw, Kempenhaghe, Heeze.
- Frank van Bussel. **Advanced MRI in diabetes; cerebral biomarkers of cognitive decrements.** Supervisors: Prof.dr.ir. W.H. Backes; Prof.dr. P.A.M. Hofman; Co-supervisor: dr. J.F.A. Jansen.
- Lisa Schönfeldt. **Neurostimulation to treat brain injury?** Supervisors: Prof.dr. Y. Temel; Prof.dr. S. Hendriks, Hasselt; Co-supervisor: dr. A. Jahanshahi.
- Rianne Geerlings. **Transition in patients with childhood-onset epilepsy; a long way to adulthood.** Supervisor: Prof.dr. A.P. Aldenkamp; Co-supervisors: dr. A.J.A. de Louw, dr. L.M.C. Gottmer, Kempenhaeghe.

- Nele Claes. **B cells as multifactorial players in multiple sclerosis pathogenesis: insights from therapeutics.** Supervisors: Prof.dr. V. Somers, Hasselt; Prof.dr. R. Hupperts; Co-supervisors: Prof.dr. P. Stinissen, dr. J. Fraussen, Hasselt.
- Olaf Schijns. **Epilepsy surgery and biomarkers from history to molecular imaging.** Supervisors: Prof.dr. J.J. van Overbeek; Prof.dr. H. Clustermann, Aachen; Co-supervisors: dr. G. Hoogland; dr. M.J.P. v. Kroonenburgh.
- Lizzy Boots. **Balanced and Prepared; development and evaluation of a supportive e-health intervention for caregivers of people with early-stage dementia.** Supervisors: Prof.dr. F.R.J. Verhey; Prof.dr. G.I.J.M. Kempen; Co-supervisor: dr. M.E. de Vugt.
- Wouter Donders. **Towards patient-specific (cerebro-) vascular model applications.** Supervisors: Prof.dr. T. Delhaas; Prof.dr.ir. F.N. van de Vosse, TUE; Co-supervisor: dr.ir. W. Huberts.
- Sizzle Vanterpool. **The implications of intrauterine invasion by microbes for placental Pathology and the occurrence of adverse pregnancy outcomes.** Supervisor: Prof.dr. B.W. Kramer. Co-supervisors: dr. J.V. Been, Erasmus MC Rotterdam, dr. U von Rango.
- Manuela Heins. **The Relationship between Social Adversity, Psychosis, and Depression across an Individual's Life Span.** Supervisor: Prof.dr. I. Myin-Germeys.
- Christianus van Ganzewinkel. **NEONATAL PAIN; Out of Sight, Out of Mind?** Supervisor: Prof.dr. B.W.W. Kramer; Co-supervisor: dr. P. Andriessen, MMC Veldhoven.
- Anne-Hilde Muris. **Hype or hope? Vitamin D in multiple sclerosis; A clinical and immunological perspective.** Supervisor: Prof.dr. R.M.M. Hupperts; Co-supervisor: dr. J.G.M.C. Damoiseaux.
- Gerard Bode. **The link between ceramide transporters, innate Immunity and Alzheimer's disease.** Supervisor: Prof.dr. M.H.V. de Baets; Co-supervisors: dr. P. Martinez, dr. M. Losen.
- Jo Stevens. **Advanced diagnostics and therapeutics for Alzheimer's disease.** Supervisor: Prof.dr. M. de Baets; Co-supervisors: dr. M. Losen, dr. P. Martinez-Martinez.
- Rosan Luijckx. **Stress and pain in muscles and brain; developing psychophysiological paradigms to examine stress and pain interactions.** Supervisors: Prof.dr. J.J. van Os; Prof.dr.ir. H.J. Hermens, UT; Co-supervisor: dr. R. Lousberg.

