UNIVERSITA' DEGLI STUDI DI MILANO
PhD Course in Molecular and Cellular Biology

XXIX Ciclo

Generation of authentic human neocortical neurons from induced pluripotent stem cells to investigate 7q11.23 gene dosage imbalances

Alice Maria Giani
PhD Thesis

Scientific tutors: Prof. Nenad Sestan
Prof. Marco Muzi Falconi

Academic year: 2016-2017

SSD: BIO/11

Thesis performed at:iiDepartment of Biosciences
University of Milan,
Milano 20133, Italy

Department of Neuroscience
Yale University,
New Haven 06510, CT (US)

CONTENTS

Acknowledgements

I would like to express my sincere gratitude to my thesis advisor Professor Nenad Sestan (Yale University) and Professor Muzi-Falconi (University of Milan) for their support and guidance in science and scientific career development accompanied by a constant encouragement in pursuing new ideas with independence and envisioning scientific questions with creativity. I would like to thank Professor Elena Cattaneo (University of Milan) for having offered me the opportunity to join her laboratory in the first rotation of my graduate studies and having introduced me to the stem cell filed and Professor Graziella Messina (University of Milan) for her valuable project advices in her role of third supervisor and for having offered me the possibility to enrich and broaden my background in the stem cells of different lineages during the second rotation in her laboratory.

I would like to thank all members of Sestan laboratory for their intellectual support and constructive scientific exchanges. In particular, I wish to thank Candace Bichsel for having been a mentor and having accompanied me in the journey of reaching scientific independence with constant encouragement and support and serving as an example. I would like to thank Andrew Tebbenkamp for the constructive collaboration and valuable advices in elaborating the study design of the project aimed to investigate the effects of 7q11.23 gene-dosage dysregulation, Zhen Li and Tianliuyun Gao for sharing their expertise in single cell RNA-sequencing and the other members of the laboratory stem cell team Marco Onorati and Fuchen Liu for their productive collaboration and feedbacks. I wish to thank also Mingfeng Li, Zhen Li, Gabriel Santpere, Belen Lorente-Galdos, Kyle Meyer and Ying Zhu for collaboration and discussion on data analysis and Francesca Talpo (University of Pavia) for contributing her precious expertise in electrophysiology.

I would also like to thank all members of Elena Cattaneo and Graziella Messina laboratories with whom I had the pleasure to collaborate and work side by side for providing a wonderful scientific environment.

My special thanks goes to my loving family and true friends to whom this dissertation is dedicated.

Part I: Introduction

Abstract

This research project has been aimed to investigate human neocortical development in healthy and diseased subjects by analyzing and comparing the transcriptional profiles and cellular morphologies of human neocortical cells derived from induced pluripotent stem cells (iPSCs).

Given the importance to rely on a solid and highly reproducible iPSCs-based differentiation protocol that generates authentic neocortical neurons *in vitro* with high efficiency before applying it as a model system of human neurodevelopmental disorders, in the first phase of this study we performed a comprehensive transcriptional, cellular and physiological characterization of the *in vitro* neurodevelopmental paradigm. The transcriptional dynamics regulating *in vitro* neocortical development have been investigated by performing RNA-sequencing (RNA-seq) at both population and single-cell level in combination with several bioinformatics analyses including principal component analysis (PCA), differential gene expression analysis and weighted gene co-expression network analysis (WGCNA). The transcriptional results were corroborated by the widespread positivity for a selected panel of informative cell-fate and cell-stage specific markers detected through immunocytochemistry and the physiological maturity of our iPSCs-derived neocortical neurons was further confirmed by their ability to generate action potentials, develop complex firing patterns and sustain excitatory and inhibitory spontaneous synaptic activity. Overall, these results fully validated the reproducibility of the differentiation protocol and its efficiency and reliability in generating physiologically mature authentic neocortical neurons.

Subsequently, we applied this extensively characterized neocortical differentiation paradigm to model *in vitro* two human neurodevelopmental disorders caused by symmetrical copy number variations (CNVs) of the Williams-Beuren syndrome chromosome region (WBSCR) located on the long arm (q) of chromosome 7 at position 11.23 (7q.11.23 locus). 7q11.23 CNVs are of special interest as the two disorders resulting from the deletion (Williams syndrome, WS) and duplication (7q.11.23 duplication syndrome, 7q11DUP) of this region exhibit cognitive and behavioral phenotypes marked by both similar features and symmetrically opposite traits. The association of 7q11DUP to complex neurodevelopmental disorders such as autism spectrum disorder and schizophrenia, while WS is a well-characterized syndrome without clear overlap to complex neurodevelopmental disorders make the study of this locus extremely interesting to identify the molecular mechanisms unique to each clinical condition, common to both syndromes and shared with other complex neurodevelopmental disorders. To this aim, we generated several iPSCs lines from a large cohort comprising WS individuals, 7q11DUP patients and healthy subjects and differentiated them into neocortical neurons by applying the previously in-depth characterized protocol. Having assessed the quality of our iPSCs-derived neocortical neurons, we are currently identifying neuronal subtypes specific genes and gene networks having the most statistically significant relationship to these disorders through single cell RNA-sequencing analysis. Furthermore, morphometric analysis of WS and control iPSCs-derived neocortical neurons has confirmed in humans many neuronal morphological abnormalities observed in a mouse knockout for Dnajc30, a previously uncharacterized gene contained in the 7q11.23 locus.

Chapter 1: Human neocortical development

1.1 Structural organization of the human postnatal cerebral cortex

The human cerebral cortex, also known as neocortex, is a 1-4.5 mm thick mantle of gray matter constituting the outermost surface of the cerebral hemispheres enclosing the underlying white matter [1]. The gray matter is composed of a tightly packed network of different neurons supported by a diverse range of glial cells and capillaries that constitutes up to 80% of the brain mass [2], while the white matter consists of the myelinated axons of cortical neurons projecting intra- or extra-cortically and their input afferents from subcortical structures. The human neocortex is distinctively characterized by an expansion of its surface area combined with a minor increase in its thickness compared to other mammals [3-8]. The increase of its surface area confers to the human neocortex its folded and convoluted appearance characterized by a complex pattern of grooves (sulci) and ridges (gyri), comprising two-thirds and one-third of the cerebral cortex respectively [7, 9, 10]. Gyrification appears to be a strategic evolutionary solution to minimize the enlargement of the skull while accommodating a larger number of neurons [5, 10-12].

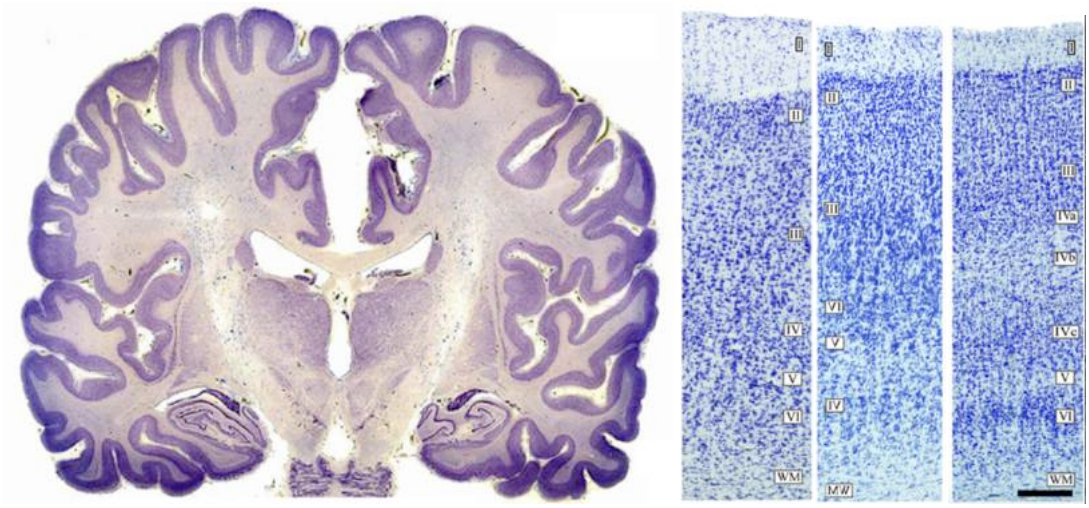


Figure 1: Human cerebral cortex and its cytoarchitectonic organization. *On the left, Nissl-staining of a coronal section of the human brain. The neocortex is the outer gyrified layer labeled in dark violet due to the high cell density while the underneath white matter displays a clearer stain as being mainly constituted by myelinated axons. Adapted with permission from http://www.brains.rad.mus.edu. On the right, three photomicrographs from 100 µm thick Nissl-stained sections showing the six cytoarchitectonically distinct horizontal laminae of respectively the human frontal (area 10), parietal (area 3b), and occipital (area 17) neocortical areas that highlight histological differences between the layers organization in distinct areas. Scale bar: 250 µm. Adapted from Ballesteros-Yánez et al. [13].*

The human cerebral cortex is estimated to consist of approximately 16.34 billions

neurons [14-16] finely interconnected through the establishment of about 164 trillion synapses [15, 17] and precisely organized in the radial and tangential planes. In the radial dimension the human neocortex is organized into six cytoarchitectonically distinct horizontal laminae, also known as cortical layers [18-23]**.** Originally seen in unstained sections of brain tissue by Gennari and Baillarger, the layers have since been distinguished on the basis of cell and fiber stainings. These six histologically different laminae are numbered in increasing order from superficial to deep with layer I constituting the surface of the cerebral hemispheres facing the pia mater and layer VI being located more ventrally in contact with the white matter. Along the tangential axis the neocortex displays a columnar organization with all the major subtypes of neocortical neurons located in layers II to VI establishing stereotypical connections in the vertical dimension as visualized in Golgi impregnated and Nissl-stained sections [24-26] and assessed through electrophysiological recordings [27]. A cortical minicolumn is the smallest level of vertical organization in the neocortex consisting of approximately 80-100 neurons [24] and it operates as a modular functional processing microcircuit. Several synaptically connected minicolumns form larger processing and functional units known as macrocolumns. Both minicolumns and macrocolumns display considerable heterogeneity between different cortical areas and even within a given macrocolumn, as they vary in size, cell composition, internal connectivity, synaptic organization [28] and perform different functions [24, 25]. At a broader level, always along the tangential axis, different cortical areas can be distinguished based on their histological structure as originally identified and classified through Nissel-staining by Broadmann. Many of the Brodmann areas originally identified on the basis of their cytoarchitecture have later been correlated to the performance of diverse cortical functions, including sensory areas that are the target of subcortical structures and are designated to process sensory information, motor areas that coordinate voluntary movements and association areas that integrate sensory information to produce more complex cognitive functions. The performance of these functions relies on specific connectivity patterns between distinct areas. Briefly, primary sensory areas send projections to a set of higher-order unimodal secondary and tertiary areas in parallel and following a hierarchical organization. In turn higher-order unimodal areas project to multimodal association areas where different sensory information are integrated. Subsequently, the axons of the cortical neurons located in the multimodal association areas project to other multimodal areas and both primary and non-primary motor areas, enabling the direction and coordination of movements and behavior.

1.2 Neurons of the cerebral cortex

The cerebral cortex is the biological substrate that enables humans to perform high cognitive functions including reasoning, perception, attention, consciousness, memory, language and refined social and emotional skills. To perform these advanced cognitive functions, the cerebral cortex is organized into a complex circuitry consisting of a myriad of diverse neurons characterized by specific transcriptional identities and epigenetic profiles, distinctive molecular and neurochemical signatures, peculiar electrophysiological properties, unique dendritic morphologies and precise axonal projections patterns [29, 30] that serve different functions. Broadly, neocortical neurons can be distinguished into two major classes according to their connectivity patterns and function: excitatory projection neurons and inhibitory interneurons [14, 31]. Excitatory

projection neurons constitute approximately 70-80% of all cortical neurons while the remaining 20-30% cortical neurons are inhibitory interneurons [32]. Inhibitory interneurons are an heterogeneous population of neurons including multiple phenotypes that share some common features while being distinctively characterized by unique morphological, physiological and neurochemical profiles [33-35]. They are generally identified by a cell soma that is smaller than most cortical pyramidal neurons, sparsely spiny or aspiny dendrites and a short axon forming local connections as their name suggests [36]. They receive both excitatory and inhibitory synapses on their soma, dendrites and axon. Through the secretion of the neurotransmitter γ-aminobutyric acid (GABA), they exert an inhibitory action on their target neurons that is crucial for the modulation of cortical output and plasticity [37, 38]. Differently from excitatory projection neurons, inhibitory interneurons are generated from neural progenitors located in the lateral, medial and caudal ganglionic eminences of the ventral telencephalon, also known as subpallium [36, 39-44]. To reach their final laminar position in the neocortex, they first migrate tangentially guided by membrane-bound and secreted factors and subsequently, once arrived in the cortical plate, they migrate radially to their final destination along with excitatory projection neurons [36, 41, 45]. On the other hand, excitatory projection neurons are generated from neocortical progenitors residing in the ventricular and subventricular zone of the dorsolateral telencephalon [30, 46, 47] and they migrate only radially along the basal process of apical radial glia to reach their laminar position following an inside-out gradient [15, 48-50]. Extending long axonal projections to other cortical areas and subcortical structures, excitatory projection neuron represent the only output system of the cerebral cortex and they constitute also its primary input system, being the major target of afferents from subcortical structures. The great majority of projection neurons are characterized by the distinctive pyramidal shape of their cell bodies from which they are named pyramidal neurons. First discovered and characterized by Ramon y Cajal at the end 19th century, pyramidal neurons have been investigated in great detail since then [51]. In the human cerebral cortex they are present in the neocortical laminae II-VI, with the exception of layer I. Pyramidal neurons display considerable variability between different cortical regions and layers [51, 52] but they share some common distinctive morphological traits and functional principles [53]. In addition to the peculiar triangular shape of their cell body, they are characterized by a single axon that extends from the base of the soma towards the white matter branching several times to reach multiple targets and by a dendritic tree that exhibits two distinct domains, the apical domain and the basal domain. The apical domain originates from the apex of the soma and the basal domain consists of multiple dendrites arising from the base of the cell body. The basal dendrites are relatively short and extend laterally from the cell soma. Instead, the apical domain is generally structured as one large apical dendrite stemming from the soma and directed toward the pia mater that bifurcates one or more times and gives rise to smaller oblique dendrites with different angular orientations before branching in a tuft of dendrites [54]. Pyramidal neurons receive different synaptic inputs at distinct cellular domains. Their soma and axon are the targets of inhibitory synapses while the dendrites receive both inhibitory inputs from local interneurons and excitatory inputs from various cortical and subcortical structures. Generally, proximal dendrites receive excitatory inputs from pyramidal neurons located in the same or adjacent areas, whereas the distal apical tuft is the target of more distant cortical areas and subcortical structures. The distinct morphologies of the dendritic tree in different locations suggests that inputs to these domains might be integrated differently [53]. Another characteristic feature shared by all pyramidal neurons is that their dendrites are covered with thousands of dendritic spines that constitute the

postsynaptic terminals of most excitatory glutamatergic synapses. Dendritic spines display a great variability in their size and shape [55] and are highly plastic [56] thus providing an anatomical substrate for memory storage and synaptic transmission [57-59]. In addition to pyramidal neurons, the neocortex comprises also a small proportion of non-pyramidal excitatory neurons that include the developmentally transient excitatory neurons of the subplate and the spiny stellate neurons of layer IV [53, 60].

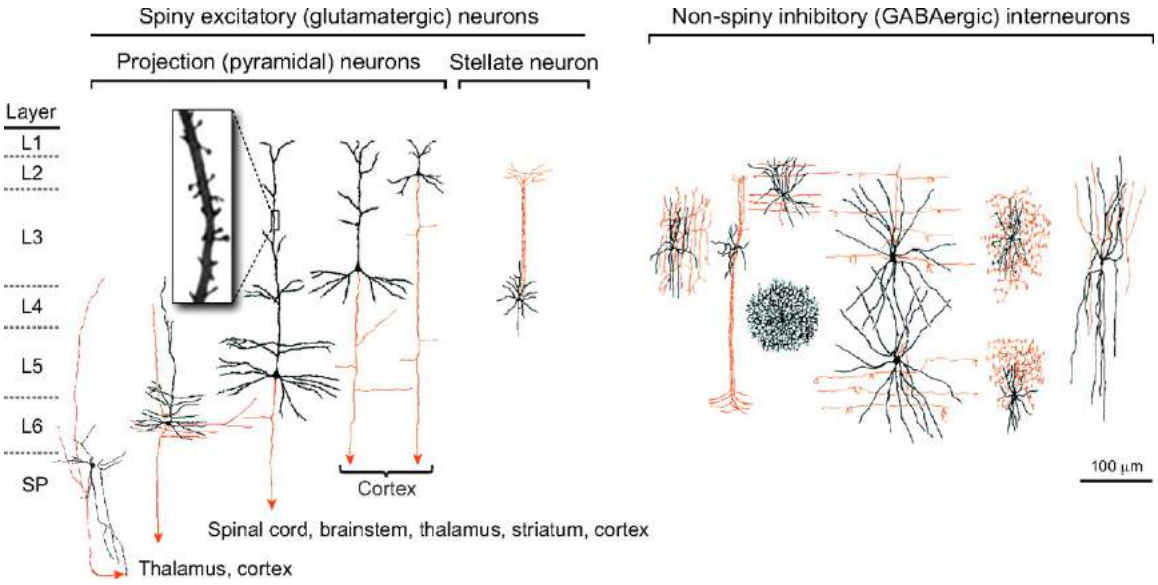


Figure 2: ***Schematic representation of the major types of excitatory and inhibitory human neocortical neurons.*** *On the left, there are pictured different subclasses of spiny neurons that secrete the neurotransmitter glutamate to mediate their excitatory action and are characterized by the presence of synaptic spines on their dendrites. A first distinction among excitatory neurons is between pyramidal projections neurons residing in cortical layers II to VI and stellate neurons mainly located in layer IV of the primary sensory areas that are endowed with short axons projecting locally. Each type of pyramidal neuron has basal and apical dendrites and an apical tuft but there are considerable differences in the morphology and connectivity patterns of pyramidal neurons belonging to different cortical layers as can be seen. On the right, it is given a panoramic of the different subtypes of aspiny inhibitory interneurons that project locally, as their name suggests, and populate selectively different laminae. Adapted from Kwan K.Y. et al. [50].*

In the cerebral cortex the transcriptional profiles and hodological properties of pyramidal neurons are strictly correlated with their laminar position as neurons belonging to a specific cortical layer share similar gene-expression and connectivity patterns. Deep layer neurons mainly project to subcortical regions forming corticofugal projections, with layer VI neurons sending their axons principally to the thalamus and layer V neurons targeting several subcortical structures, including the thalamus, striatum, spinal cord and brainstem. A minor subpopulation of deep layer neurons establishes also ipsilateral and contralateral connections with other neocortical areas [60]. Layer IV accommodates both pyramidal projection neurons and spiny stellate neurons. Neurons located in this layer

mainly receive input afferents from both inter-hemispheric cortical areas and subcortical regions, especially the thalamus, and they transmit signals to the rest of the column. On the other hand, upper layer neurons principally form corticocortical-projections that connect one region of the neocortex to another. Corticocortical-projection neurons can establish ipsilateral connections between different cortical areas of the same hemisphere or they can form commissural connections between the two opposite cerebral hemispheres by sending their axons through the corpus callosum. Generally, layer III neurons project to adjacent columns while neurons belonging to layer II target more distant cortical regions and tend to establish inter-hemispheric connections.

1.3 Neocortical neurogenesis

Neurogenesis of the human central nervous system (CNS) begins after the fusion of the neural folds leading to the formation of the neural tube and proceeds at different rates following specific temporal and spatial patterns along the rosto-caudal, dorso-ventral and medio-lateral body axes (Mora-Bermu´dez et al. 2013). The dorsolateral telencephalon that will later generate the neocortex derives from the most anterior region of the neural tube. The rostral wall of the neural tube that contours the central cavity is initially composed of a pseudostratified layer of undifferentiated neuroepithelial cells polarized along their apical-basal axis and spanning the entire width of the neuroepithelium. Neuroepithelial cells serve as stem and progenitor cells for all the neural lineage derivatives, including neurons, astrocytes and oligodendorcytes. They display typical epithelial features and are highly polarized along their apical–basal axis [61, 62] as reflected by the presence of apical-specific [63, 64] and basolateral-specific [65, 66] transmembrane proteins. At the most apical end of the lateral plasma membrane, neuroepithelial cells are interconnected by tight junctions and adherens junctions [67-70]. Their nucleus periodically moves along the apical-basal axis in phase with the cell-cycle progression, a process known as interkinetic nuclear migration (INM) and mitosis occurs near the apical plasma membrane lining the ventricular surface [71-73]. Initially, neuroepithelial cells undergo symmetrical proliferative divisions that give rise to two daughter cells identical to the mother cell with the effect of increasing the size of the progenitor pool and expanding the neuroepithelium in the lateral dimension [74].

Meanwhile, a diversified population of neurons, also known as “pioneer neurons”, positions itself between the ventricular zone (VZ) and the pial surface either migrating tangentially from the subpallium or through radial migration from the underlying ventricular zone, giving rise to a dynamic, evolving, transient developmental structure known as preplate (PP) or primordial plexiform layer [75-80]. The diverse population of cells constituting the human preplate comprise several types of reelin-expressing cells [80-83] including Cajal-Retzius cells [83, 84], predecessor cells [75, 80], different cells expressing interneuron markers [81, 82] and other neuronal subtypes [82, 85]. Among the cells expressing reelin, Cajal-Retzius cells have particular relevance as they are essential for the proper laminar positioning of the developing neocortical neurons and for the formation of early synaptic contacts [78, 82, 85]. On the other hand, predecessor cells are characterized by long horizontal processes forming an intricate network thought to serve as a scaffold for successive cell migration and axon guidance [75, 80].

As neuroepithelial progenitors proliferate and expand in number, they progressively transform into apical radial glia and change their cell division mode from symmetric proliferative to asymmetric differentiative cell division that enable to preserve self-renewal while generating more lineage-committed neural progenitors [65, 74, 86-88]. This transition marks the onset of neurogenesis and it occurs around embryonic day 33 in the lateral part of the cortical wall in humans [80, 89]. Apical radial glia are neural stem cells characterized by residual neuroepithelial features combined with newly acquired astroglial properties [61, 65, 88, 90]. As neuroepithelial progenitors, apical radial glia are highly polarized cells [63, 91] connected by adherens junctions and their cell body continues to reside in the ventricular zone (VZ). They also maintain the expression of neuroepithelial markers [87] and their nuclei undergo interkinetic nuclear movements (INM) restricted to the cell body with mitosis occuring at the apical ventricular surface [65]. In addition to these distinctive neuroepithelial features, apical radial glia express also several astroglial markers [88, 92, 93]. They are characterized by a distinctive bipolar morphology identified by an ovoid cell body in direct contact with the ventricular surface and a long thin radial process, the basal process, that spans the developing and thickening cortical wall remaining attached to the basal lamina and that will enable neuronal radial migration serving as a scaffold fiber [94, 95]. Concerning their proliferative capacity, apical radial glia are neural stem cells able to undergo multiple rounds of symmetric proliferative and asymmetric differentiative cell divisions with a progressive shift toward asymmetric differentiative cell divisions as neurogenesis progress [93, 96-98].

The products of apical radial glia asymmetric differentiative divisions range from neurons to other neural progenitors, including apical intermediate progenitors and basal progenitors, also known as intermediate progenitors [46, 87, 96-106]. As a consequence of the incessant proliferation and generation of this diverse population of neocortical progenitors and neurons, the neuroepithelium continues to enlarge both in the lateral and radial dimensions becoming a stratified tissue consisting of different cell types that are spatially segregated. The cell bodies of apical intermediate progenitors still reside in the ventricular zone maintaining direct contact with the ventricular surface and these progenitors share with apical radial glia several other features, including the expression of characteristic transcription factors like Pax6, interkinetic nuclear migration (INM) and the formation of adherens junctions. Distinctively, they downregulate radial glial genes and they undergo just one symmetric self-consuming neurogenic division [46, 99, 107]. Also their morphology differs from apical radial glia as their basal process is of variable length having lost contact with the basal lamina and it is retracted during mitotic divisions [46, 75]. While the cell bodies of neuroepithelial cells, apical radial glia and apical intermediate progenitors continue to reside in the ventricular zone (VZ), the cell bodies of the more fate-committed basal progenitors generated later during corticogenesis form a second germinal zone located directly above the ventricular zone named subventricular zone (SVZ). Interestingly, an additional type of progenitors known as subapical progenitors populate both the VZ and the SVZ. Subapical progenitors undergo mitosis at an abventricular location and display an apical process [46, 104]. In humans the subventricular zone is distinctively expanded compared to other mammals [4, 108] and further specialized into two cytoarchitectonically distinct sublayers, a smaller inner SVZ (ISVZ) and a larger outer SVZ (OSVZ) [39, 103, 109-111] separated by a band of tangentially orientated axonal fibers known as inner fiber layer (IFL) [108, 112]. The outer SVZ is further enclosed externally by an outer fiber layer (OFL). The progenitors populating the SVZ can be distinguished into three major types, basal intermediate

progenitors and basal radial glia that selectively populate the SVZ and subapical progenitors. Both types of basal progenitors have lost contact with the ventricular surface along with apicobasal polarity, they are not tight by adherent junctions [46, 98, 107, 113] and they do not undergo proper interkinetic nuclear movements. Other common features shared by both types of basal progenitors are their ability to undergo both symmetric proliferative and asymmetric differentiative cell divisions [39, 100, 101, 111, 113-115] and the expression of specific transcriptional profiles marked by the upregulation of the key transcription factors, including Pax6 [101, 103]. Distinctively, basal intermediate progenitors lack both the apical and basal processes that are retracted while they migrate to the SVZ. On the other hand, basal radial glia are a diversified group of basal progenitors that may exhibit nuclear movements preceding mitosis similar to interkinetic nuclear movements [39, 100] and they may be further distinguished into three subtypes according to the morphology of their apical and basal processes. Bipolar radial glia are characterized by a basal process attached to the basal lamina and an apical process in contact with the ventricle [101, 104], while monopolar radial glia are distinguished either by a long apical process reaching the inner SVZ or the VZ but detached from the ventricular surface [39, 101, 103, 116] or by an extended basal process that may be in contact with the pia mater [101].

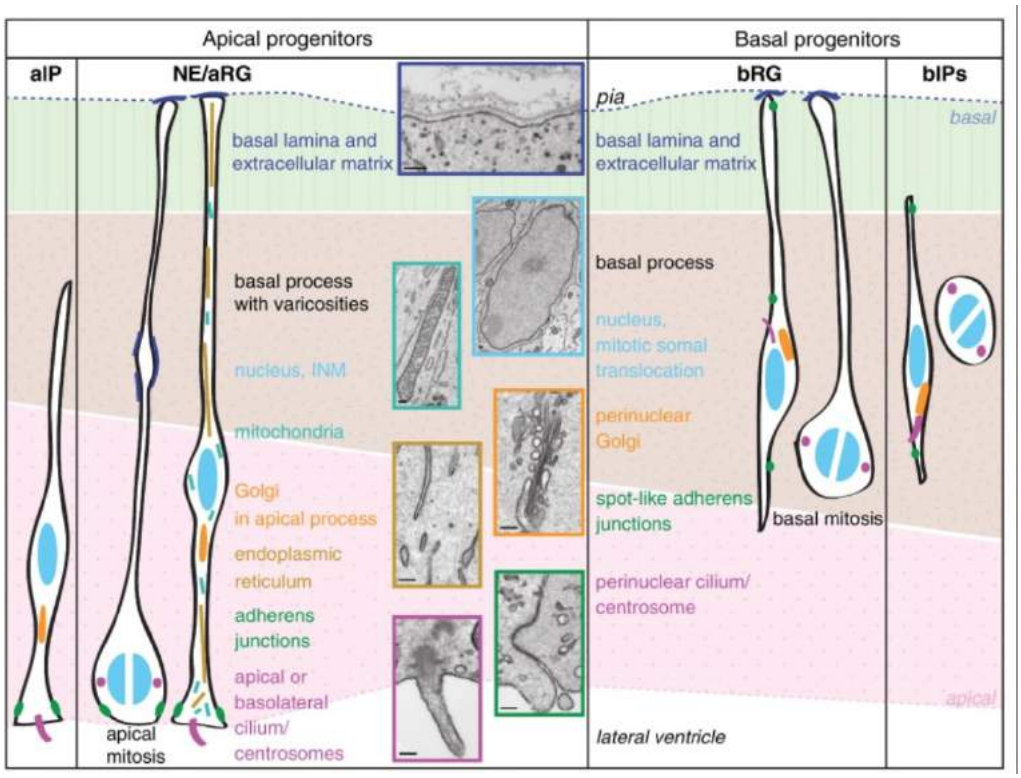


Figure 3: Schematic representation of the cell biological features of neocortical progenitor cells (NPCs). *In the ventricular zone (VZ), colored in pink, reside the cell bodies of apical intermediate progenitors (aIP) and apical radial glia (aRG) while the subventricular zone (SVZ), colored in brown, is populated by the different types of basal radial glia (bRG) and basal intermediate progenitors (bIPs). The intermediate zone (IZ) is not represented for clarity while the nascent cortical plate (CP) is marked in green.*

Reported in the schematic are also distinctive biological features of each neocortical progenitor subtype. Adapted from Wilsch-Brauninger M. et al. [47].

1.4 Neocortical migration

After generation in the germinal zones all neocortical neurons migrate to reach their final laminar position in the neocortex [94, 95, 117-119]. Inhibitory interneurons first migrate tangentially from the ganglionic eminences of the ventral telencephalon where they are born [36, 39-44] to the marginal zone of the dorsolateral telencephalon and subsequently undergo radial migration to arrive at their final laminar position in the neocortex [36, 41, 45]. On the other hand, excitatory pyramidal projection neurons are generated from the asymmetric differentiative cell divisions of neocortical progenitors located in the ventricular and subventricular zones of the dorsolateral telencephalon [46, 87, 96-106]. The first immature neurons are generated from apical radial glia cells located in the ventricular zone at 50-51 post-conceptional days, event that marks the onset of neurogenesis [15, 120]. Meanwhile, the subventricular zone (SVZ) starts to become populated with basal neocortical progenitors and it continues its expansion during early and mid-fetal development progressively becoming organized into two cytoarchitectonically distinct areas, the inner SVZ (ISVZ) and the outer SVZ (OSVZ) [4, 39, 103, 108-111, 121]. The inner subventricular zone (ISVZ) is the first subventricular zone compartment to develop followed by the enlargement of the outer SVZ (OSVZ) that becomes the major and most expanded proliferative zone with the progression of neurogenesis [39, 110, 111]. Immature excitatory neurons generated from the neocortical progenitors residing in the germinal zones migrate radially through the preplate (PP) using the thin basal process of apical radial glia cells as a scaffold fiber [94, 95] and form the nascent cortical plate (CP), a transient developmental structure that accomodates the post-migratory pyramidal neurons [50, 122, 123].

The emergence and enlargement of the cortical plate (CP) has the effect of splitting the preplate (PP) into two compartments, the marginal zone (MZ) and the subplate (SP), located respectively above and below the nascent cortical plate [15, 122, 123]. The marginal zone (MZ), that will become cortical layer 1 in the postnatal neocortex, is a transient developmental structure positioned directly below the pia mater that serves as a corridor for the tangentially migrating interneurons and is important for the proper regulation of cortical lamination as it accommodates “pioneer neurons” [75, 80], in particular reeling-expressing Cajal-Retzius cells [83, 84, 124]. As shown in studies of the mutant mouse reeler*,* the extracellular matrix protein reelin is essential for the proper regulation of radial neuronal migration and its absence cause an inversion of the characteristic “inside-out” migration gradient resulting in an inverted positioning of the neocortical layers [115, 125]. Also the subplate (SP) is a transient developmental structure located right below the cortical plate (CP) and in humans it becomes the most prominent embryonic zone of the telencephalon between 15–35 postconceptional weeks (pcw), a period known as “subplate stage” [126]. The human subplate is enriched in extracellular matrix and consist of some of the earliest born and first mature neurons of the cerebral cortex [60, 121, 126-131] along with neurons that are migrating in the radial and tangential direction to reach their final neocortical laminar position [128, 132-134] and early developing astrocytes, microglia and oligodendrocyte precursors [126]. The human subplate is also a dynamic compartment accommodating input fibers from

subcortical structures and other developing cortical areas [128-130, 135], including ipsilateral and callosal corticocortical afferents [19, 135-137], thalamocortical afferent axons [138-140] and basal forebrain cholinergic projections [141], therefore being an important station for promoting synaptogenesis [15, 128, 142], the maturation of neocortical circuits [131] and for establishing the formation of proper neocortical afferent and efferent projections [128-130, 135, 140]. As afferent and efferent projections from subcortical structures and other neocortical areas progressively develop, another transient structure arise below the subplate (SP), known as intermediate zone (IZ). The intermediate zone (IZ) contains tangentially migrating inhibitory interneurons and some sparse neurons and consists mainly of afferent and efferent axonal projections that will form the white matter of the postnatal neocortex, while neurons found in the white matter are supposed to originate from subplate neurons (SP) [60, 126].

As neurogenesis progress, different types of excitatory projection neurons are generated sequentially. Progressively, they migrate in the radial dimension from the germinal zone of origin through the intermediate zone (IZ) and the subplate zone (SP) and populate the cortical plate (CP) following a precise inside-first outside-last gradient according to which the earliest generated neurons will become the future layer 6 and the last-born neurons will constitute layer 2 [15, 29, 48-50, 77, 143, 144]. At each wave of neuronal migration, newly born neurons pass through the earlier generated deep layers before positioning at the interface between the cortical plate (CP) and the marginal zone (MZ) where they form a band in the uppermost part of the cortical plate named the dense cortical plate [15, 50].

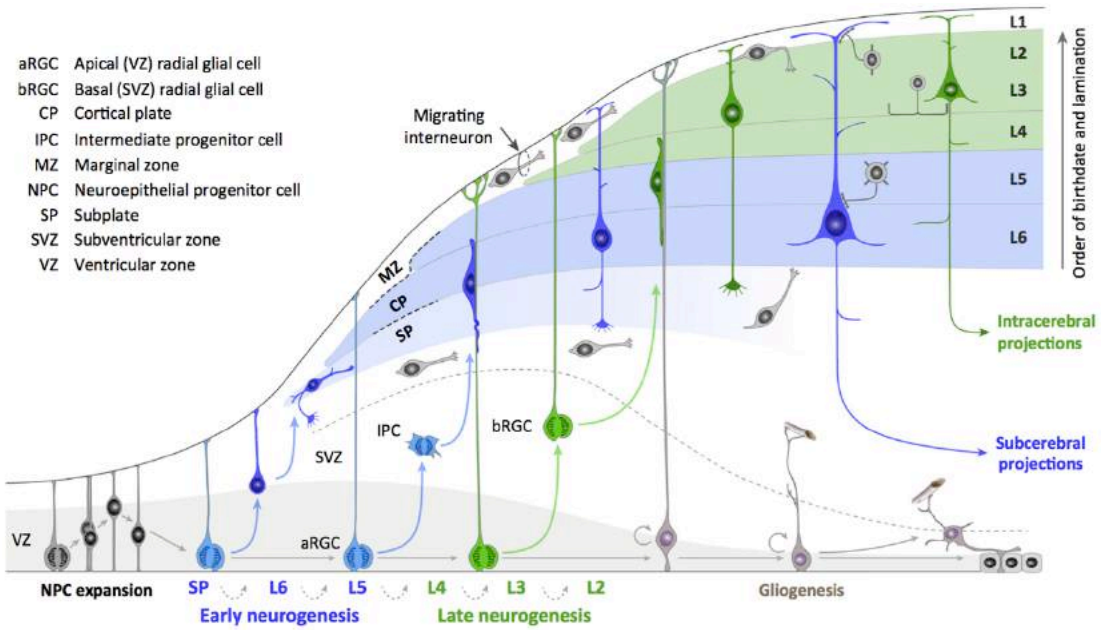


Figure 4: Schematic representation of the generation and subsequent migration of neocortical excitatory pyramidal neurons. *Neuroepithelial cells (NECs), constituting the pseudostratified neuroepithelium of the dorsolateral telencephalon, undergo interkinetic nuclear migration (INM) and divide symmetrically at the ventricular surface to increase the number of neocortical progenitors before transforming into apical radial glia cells (aRGCs), event that marks the onset of neurogenesis. Also apical radial glia cells*

(aRGCs) divide at the ventricular surface maintaining contact with the basal lamina. As neurogenesis progress and the neuroepithelium thickens, the basal plasma membrane of apical radial glia cells stretches and becomes a thin radial process, known as basal process. Apical radial glia cells undergo both symmetric and asymmetric cell divisions progressively shifting toward asymmetric differentiative cell divisions. The asymmetric differentiative cell divisions of apical radial glia cells have the combined effect of increasing their number and generating both neurons and other neocortical progenitors, including intermediate basal progenitor cells (ICP) and all the different types of basal radial glia cells (bRGC) that populate the subventricular zone (SVZ). Neurons migrate radially into the cortical plate (CP) using the apical radial glia process as a scaffold fiber and split the preplate into the marginal zone (MZ) and subplate (SP). Excitatory pyramidal neurons radial migration follows a characteristic "inside-out" gradient, with early born neurons occupying the deepest layer 6 and subsequently generated neurons giving rise to more superficial layers up to layer 2. Around the 27th postconceptional week when neurogenesis is completed, gliogenesis begins with apical radial glia cells transforming into astrocytes, oligodendrocites precursors or ependymal cells. Reported in the schematic are also the subcortical and intracortical projections characteristic of each neocortical layer. Adapted from Shibata et al. [145] and Kwan et al. [50].

Radially migrating neurons display a stereotypical bipolar morphology characterized by a leading process directed basally toward the pia mater and a lagging process pointing toward the ventricle [95, 107] and they can migrate either through somal translocation or by locomotion [97, 119, 123]. Neurons migrating by somal translocation are bipolar and to migrate they first extend and fix to the pia mater or to the marginal zone the long basally directed leading process. Subsequently, the nucleus is transported to its final laminar destination thanks to the elongations and contractions of the microtubules of the basal process, a mechanism also known as nucleokinesis [107, 123, 146, 147]. On the other hand, migration through locomotion involves the free movement of the immature neuron that climbs with its relatively short leading and trailing processes toward the cortical plate along the basal process of apical radial glia that serves as a scaffold fiber [107, 118, 119, 147].

Furthermore, radial migration of newly generated excitatory pyramidal neurons does not generally proceeds as a direct movement from the ventricle to the cortical plate but is rather a more structured process articulated into four phases [107].

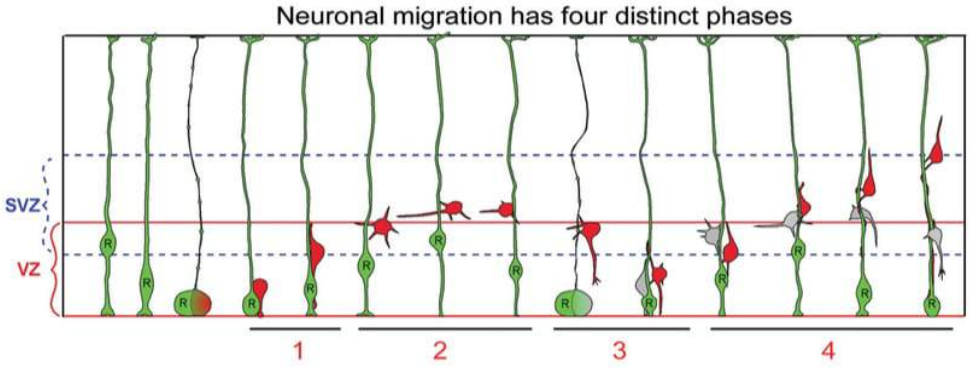


Figure 5: Schematic representation of the four migration phases undergone by newly born neocortical pyramidal neurons. *The migration of newly born pyramidal*

neurons to their final laminar destination can be divided into four phases: 1) initial radial migration, 2) arrest in the subventricular zone (SVZ), 3) retrograde migration toward the ventricle, 4) secondary radial migration. Adapted from Noctor et al. [107].

The first phase is characterized by the rapid movement of the new born neuron from the ventricular surface to the subventricular zone. Then, it temporary arrests its migration before moving backwards toward the ventricule and finally migrating to the cortical plate (Figure 4). Upon reaching their laminar position, neurons stop migrating and detach from the basal process of apical radial glia cells.

The processes of radial migration, neural fate determination and neural differentiation are finely regulated by specific gene expression patterns [60, 143, 144, 147-151] and signaling pathways [152] that mediate a broad range of cellular functions including chemoattraction and repulsion, cell adhesion, cell motility, cytoskeletal dynamics and aspects of mitotic activity. Once arrived at their final laminar position, neocortical neurons continue to differentiate, start to develop the axon and the dendrites and begin to establish synaptic connections [153, 154]. The generation of excitatory pyramidal neocortical neurons and the majority of inhibitory interneurons terminates around the 27th postconceptional week [15, 120, 155, 156] while certain subclasses of neocortical interneurons are generated also postnatally [157]. Neocortical neurogenesis is then followed by gliogenesis that is the process through which apical radial glia cells transform into glial cells [75, 158], comprising astrocytes and oligodendrocyte precursor cells. Gliogenesis starts in midgestation [15, 159] and continues postnatally [75, 160, 161]. Along with gliogenesis, also synaptogenesis, myelination, dendrites outgrowth and pruning continue in the postnatal period [15, 153, 162].

1.5 Signaling pathways regulating neocortical development

Human neocortical development is finely orchestrated by specific cell signaling pathways and transcriptional programs that act in a spatial and temporal regulated manner to coordinate all the stages of neurogenesis, including the generation of different classes of neocortical progenitors (NPCs), the balancing of their proliferation, the specification of their laminar identity, their progressive cell-fate restriction and the final differentiation toward layer-specific mature neocortical neurons. The generation of different telencephalic progenitors is articulated in a first *neural induction* phase characterized by the appearance of neuroepithelial progenitors cells lining the central cavity of the neural tube followed by a subsequent *patterning* phase during which the regional specification along both the anteroposterior (AP) and dorsoventral (DV) axes takes place resulting in the creation of spatially and molecularly distinct neural progenitor domains [163, 164]. Later with the onset of neurogenesis, it starts a protracted period of cell fate determination and terminal differentiation during which all the major types of neocortical projection neurons are generated and migrate to their final laminar position where they become fully mature neurons. During each phase of neurogenesis, many cell signaling pathways cross-talk and are integrated to direct the selection of a particular neuronal subtype identity and each signaling pathway may exert different effects as neurodevelopment progress.

Neural induction in human embryos is achieved through the inhibition of the transforming growth factor-β (TGF-β) signaling combined with the coordinated action of retinoid, fibroblast growth factor (FGF), epidermal growth factor (EGF), Wnt and sonic hedgehog (Shh) signaling pathways**.** According to the "default model" of neural induction, ectodermal cells have an autonomous tendency to differentiate into neural tissue and they acquire neural identities in the absence of cell-cell interactions and bone morphogenetic proteins (BMPs), especially BMP4, as first demonstrated by studies performed in animal models such as the frog Xenopus laevis [165-171]. These researches enabled to identify several molecules with neural inducing activity secreted by the primary organizer, including follistatin [166], chordin [167, 168] and noggin [169, 170], that compete with bone morphogenetic proteins (BMPs) for the binding to their cognate receptors, thus inhibiting the activation of the signaling cascades that induce the differentiation toward an epidermal fate [171]. Briefly, the transforming growth factor-β (TGF-β) signaling pathway is activated by the binding of a ligand belonging to the TGF-β superfamily of growth factors to its corresponding receptor [172, 173].

The TGF-β superfamily of ligands in humans consist of 30 growth factors, including 3 TGF-β isoforms, 4 activin β-chains, the protein nodal, 10 bone morphogenetic proteins (BMPs) and 11 growth and differentiation factors (GDFs) [172]. Their cognate TGF-β receptors are dimeric transmembrane serine-threonine kinases proteins that can be divided into two distinct functional classes, the type I and type II receptors encoded by seven and five genes in the human genome respectively [173]. Upon binding of the ligand, one type I and one type II receptor bridge together and form an active heteromeric receptor complex [174, 175], where the type II receptor kinase phosphorylates the type I receptor that in turn recruits and phosphorylates one of the receptor-regulated SMADs (R-SMADs), comprising SMAD1, SMAD2, SMAD3, SMAD5 and SMAD8. Phosphorylated receptor-regulated SMADs are able to form active complexes with the common mediator SMAD4 and homomeric complexes [173]. Active SMAD complexes translocate to the nucleus, where they cooperate with other transcription factors, histone-modifying enzymes and chromatin-remodelling complexes to regulate the transcription of target genes. *In vitro* studies have confirmed that similar mechanisms are also at the basis of human neural induction as it was shown that treatment of human pluripotent stem cells (hPSCs) with the recombinant protein noggin promotes the acquisition of neural stem cell identities [176-178]. Inhibition of the transforming growth factor-β (TGF-b) signaling cascade alone is not sufficient to promote neural induction in human embryos and for the adoption of neural fates and the subsequent acquisition of specific regional and cellular identities it is also required the coordinated activation of multiple other signaling pathways acting through both BMP-dependent and BMP-independent mechanisms, including retinoid, fibroblast growth factor (FGF), epidermal growth factor (EGF), Wnt and sonic hedgehog (Shh) signaling pathways.

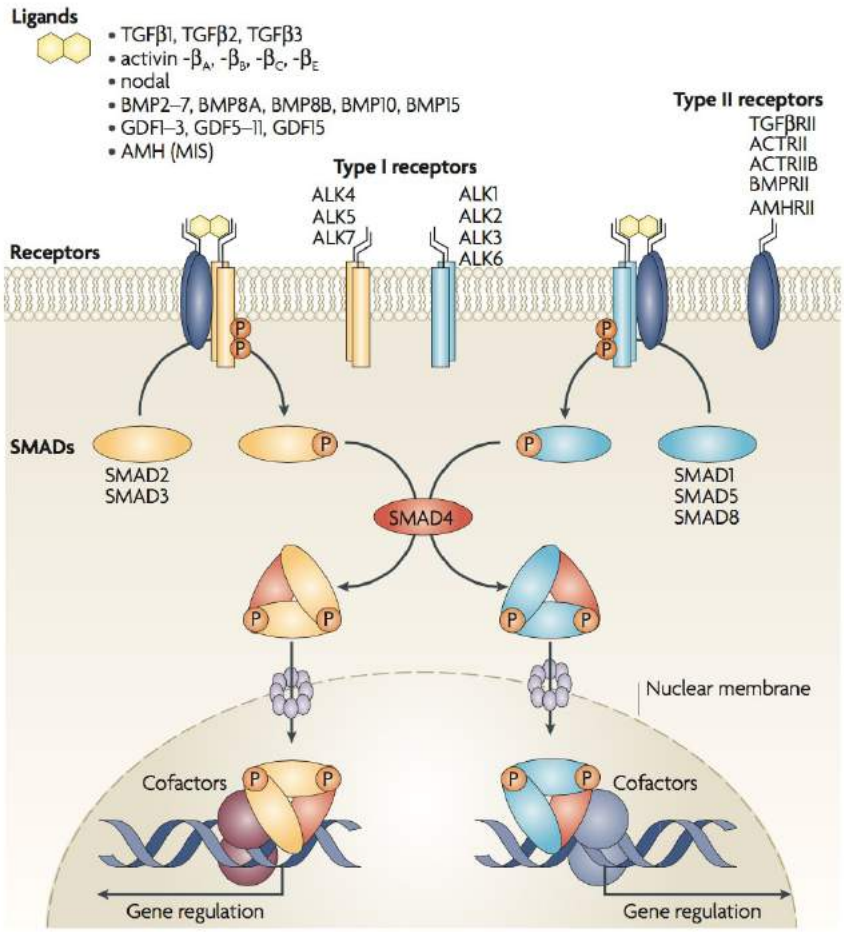


Figure 6: Schematic representation of the TGF-β signaling pathway. *In this schematic layout of the TGF-β signaling pathway are represented several human dimeric ligands belonging to the TGF-β superfamily of growth factors and their cognate type I and type II dimeric transmembrane serine-threonine kinases receptors. Binding of a TGF-β ligand to its correspondent type II receptor results in the recruitment of a type I receptor and the formation of an heteromeric receptor complex that is followed by phosphorylation of the type I receptor. As indicated, type I receptors can be subdivided into two subgroups that selectively recruit and phosphorylate different receptor-regulated SMAD proteins (R-SMADs). Type I ALK1, ALK2, ALK3 and ALK6 receptors selectively phosphorylate SMAD1, SMAD5 and SMAD8 while type I ALK4, ALK5 and ALK7 receptors specifically phosphorylate SMAD2 and SMAD3. Phosphorylated receptor-regulated SMADs form active complexes with the common mediator SMAD4. Activated SMAD complexes translocate to the nucleus where they directly regulate gene transcription. Adapted from Schmierer B. and Hill C.S. [172].*

Fibroblast growth factor (FGF) signaling is critically involved in various phases of neurogenesis during which it performs different functions [179-181], including the regulation of neural stem cells (NSCs) and neocortical progenitor cells (NPCs) proliferation and survival, the patterning and specification of their areal identity and the

orchestration of their lineage commitment and differentiation. During neural induction, studies conducted in animal models have provided evidences that FGF proteins (FGFs) antagonize BMP signaling through the phosphorylation and inhibition of the BMP effector protein Smad1 and the direct repression of BMP transcription [182, 183] while concomitantly promoting the expression of neural markers acting in combination with BMP- and WNT-antagonists [184]. Furthermore, FGF signaling acts also independently of BMP and other signaling pathways to induce the expression of neuroepithelial markers and transcription factors required for neural fate specification [185]. Before the onset of neurogenesis, the basic fibroblast growth factor (bFGF or FGF2) and FGF8 support the proliferative divisions of neuroepithelial cells [186, 187] and FGF10 contributes to the initiation of neurogenesis by promoting the maturation of symmetrically dividing neuroepithelial cells into asymmetrically dividing radial glia cells [188]. Later in neurodevelopment, especially FGF2 acts in concert with other growth factors and signaling pathways as a potent mitogenic factor that critically regulates neural stem cells (NSCs) and neural progenitor cells (NPCs) proliferation and survival as indicated by gain- and loss-of-function studies performed both *in vivo* and *in vitro* [187, 189-196] and by its high levels of expression in the whole cortical ventricular and subventricular zones before and during all the timespan of neurogenesis [187, 197-199]. Later in neurodevelopment, FGF signaling is also involved in decreasing the pace of the generation of basal progenitors from apical radial glia cells [200]. In addition to its fundamental role in the regulation of neocortical progenitors proliferation and cell-cycle exit, FGF signaling is also critical for the correct acquisition of neocortical regional and cellular identities. The patterning function of FGF signaling starts during gastrulation when FGF signaling cooperates with other signaling pathways to determine the specific positional identities of neural progenitor cells along the antero-posterior axis. Specifically, the FGF pathway together with WNT and retinoic acid (RA) signaling directs the induction and patterning of the peripheral nervous system [201, 202] and the specification of posterior neural fates as demonstrated by the induced expression of different Hox and Cdx genes in chicken and mouse neural plate explants after exposure to FGF proteins [203]. After the initial anterior-posterior patterning of the neural plate during gastrulation, FGF8 and other FGF proteins initially expressed by the rostral signaling center continue to be expressed in the telencephalic primordium during and after the formation of neural tube [204]. Especially FGF8 has a crucial role in telencephalic development as it confers and maintains the telencephalic character of the anterior neural plate by positively and negatively regulating gene expression [205-208], by modulating the expression and activity of other signaling molecules including Wnts, BMPs, and Shh [186, 209] and by promoting the survival of telencephalic progenitors [206, 210, 211]. Once the dorsal and ventral subdivisions of the telencephalic vesicles have been established, FGF8 and other FGF proteins continue to play a central role in the development of the telencephalic midline [186, 212], in the further growth of the telencephalon and in the specification of the subdivision of the dorsal cerebral cortex into multiple functional areas by regulating the regional expression of multiple transcription factors [207, 213]. Also FGF2 continues to play a critical role in directing neocortical progenitors specification that is concentration-dependent, as low concentrations of this growth factor do not affect the neuronal commitment of progenitors while the exposure to higher concentrations in combination with other extrinsic signals promotes oligodendrogliogenesis and astrogliogenesis [191, 214-216]. Later in neurodevelopment FGF signaling is also required for neuronal migration with a prominent role exerted by FGF18 [217], for the translocation of astroglial cells from the ventricular zone to the

surface of the cortex [218] and for the control of axon growth, navigation, branching and target recognition with FGF proteins acting as target-derived signals.

As other pathways, also the various intracellular FGF signaling cascades are activated by the binding of a member of the secreted fibroblast growth factor (FGF) signaling proteins to one of its corresponding FGF tyrosine kinase receptors. The human fibroblast growth factor (FGF) family comprises twenty two genes that encode secreted signaling proteins (FGFs), eighteen of which activate the signaling cascade by binding to FGF tyrosine kinase receptors. The secreted signaling FGF proteins can be classified into seven subfamilies based on their biochemical function, sequence similarities and evolutionary relationships, including five subfamilies of paracrine FGFs (also known as canonical FGFs), one subfamily of endocrine FGFs and one subfamily of intracellular FGFs [219, 220]. Canonical secreted FGF proteins are tightly bound to heparin/heparan sulfate (HS) proteoglycans (*HSPGs*) that enable to limit their diffusion through the extracellular matrix (ECM), serve as cofactors to regulate the specificity and affinity for cell surface signaling tyrosine kinase FGF receptors (FGFRs) and protect FGF proteins from degradation [221]. FGF proteins belonging to the paracrine subfamilies bind to one of the four cell surface tyrosine kinase FGF receptors (FGFRs) using heparin/HS proteoglycans as cofactors and interacting with the chaperone heat shock protein 90 (HSP90) [222, 223]. Binging of the FGF ligand induce receptor dimerization and trans-autophosphorylation of the kinase domain [224]. Activated FGF receptors phosphorylate distinct adaptor proteins that in turn start one of the four major intracellular FGF signaling cascades, comprising PI3K-AKT, RAS-MAPK, PLCγ, and STAT pathways [220, 221, 224-226]. The regulatory mechanisms of all four FGF molecular signaling cascades are based on many molecules that are themselves regulated by FGFs through positive or negative feedback loops [180, 220]. Activation of RAS-MAPK and PI3K-AKT pathways is initiated by phosphorylation of the adapter protein fibroblast growth factor receptor substrate 2a (FRS2a) that is constitutively associated with the juxtamembrane region of the FGF receptor. Phosphorylated FRS2α binds the membrane anchored adaptor protein growth factor receptor-bound 2 (GRB2) and the tyrosine phosphatase SHP2 [227, 228]. GRB2 subsequently activates the PI3K-AKT pathway and the RAS-MAPK pathway through the recruitment to the signaling complex of the GRB2 associated binding protein 1 (GAB1) [229] and the guanine nucleotide exchange factor protein son of sevenless (SOS) [227], respectively. The PI3K-AKT pathway has the downstream effect of positively and negatively modulating the activity of target molecules that regulate gene transcription through their phosphorylation by the serine/threonine-specific protein kinase AKT but there is still little evidence for the role of this signaling pathway in neural development in vivo downstream of FGF receptors [180]. Most notably ATK inhibits the activity of transcription factors as the forkhead box class transcription factor FOXO1 and cytosolic proteins as tuberous sclerosis complex 2 (TSC2) that results in the activation of other protein complexes as mTOR complex 1 [230]. Also the RAS-MAPK pathway positively regulates the expression of several target genes through the phosphorylation by the mitogen-activated protein kinase (MAPK) including members of the E26 transformation-specific (ETS) family of transcription factors as Etv4 (Pea3) and Etv5 (Erm), AP1, GATA proteins, c-*myc*, and CREB [231, 232] while it also activates negative regulators of the FGF signaling pathways such as SHP2, CBL, SPRY, SEF, MPK3 and DUSP6 [220, 233]. The MAPK/Erk pathway is particularly important in mediating the proliferative activity of neural stem cells (NES) and neocortical progenitors (NPCs), in orchestrating the acquisition of specific regional and cellular identities [233] and it also involved in regulating neuronal migration and differentiation [180].

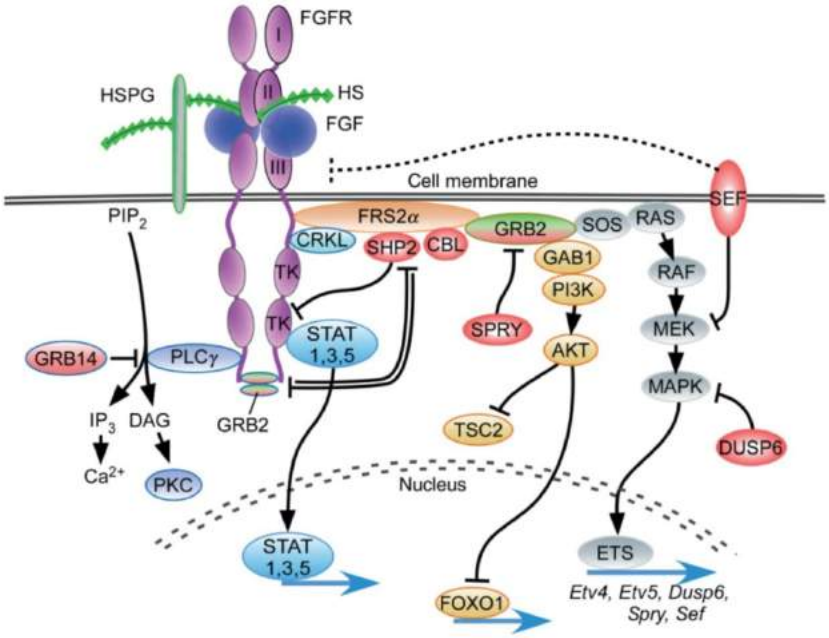


Figure 7: Schematic representation of the FGF signaling pathway. *Illustrated are the four intracellular signaling pathways activated by the binding of a canonical secreted FGF protein to its corresponding receptor (FGFR), PLCγ pathway, STAT pathway, PI3-AKT pathway and MAPK pathway. Adapted from Ornitz D.M. and Itoh N. [220].*
On the other hand, activation of PLCγ and STAT pathways begins with the phosphorylation by activated tyrosine kinase FGF receptors of member of the signal transducer and activator of transcription (STAT) protein family and membrane-associated enzyme phospholipase C g (PLCγ), respectively.

Specific tyrosine kinase sites of the FGF receptor phosphorylate selectively STAT1, 3, and 5 that translocate via an importin a/b ternary complex to the nucleus [234] where they activate gene transcription by binding to a consensus DNA-recognition motif known as gamma-activated sites (GAS) found in the promoter region of their target genes [235]. Instead, phosphorylation of PLCγ leads to the production of inositol triphosphate (IP3) and diacylglycerol (DAG) through the hydrolysis of phosphatidylinositol 4,5-bisphosphate [236]. IP3 and DAG have in turn the effects of inducing calcium ion release from intracellular stores thus increasing intracellular calcium ion levels and activating calcium/calmodulin dependent protein kinase enzymes as protein kinase C (PKC), respectively [236, 237], that have the downstream effect of stimulating neurite outgrowth [238]. The activation of PLCγ pathway is further modulated by the adaptor protein GRB14 that inhibits tyrosine phosphorylation of FGF receptors at multiple sites thus preventing PLCγ phosphorylation [239].

The epidermal growth factor (EGF) signaling is implicated in the self-renewal of neural stem cells (NSCs) and in the proliferation of neocortical progenitors as demonstrated by the elevated concentration of EGF ligands in the embryonic neocortex [240] accompanied by high levels of expression of EGF receptors in these cell types

[241, 242].

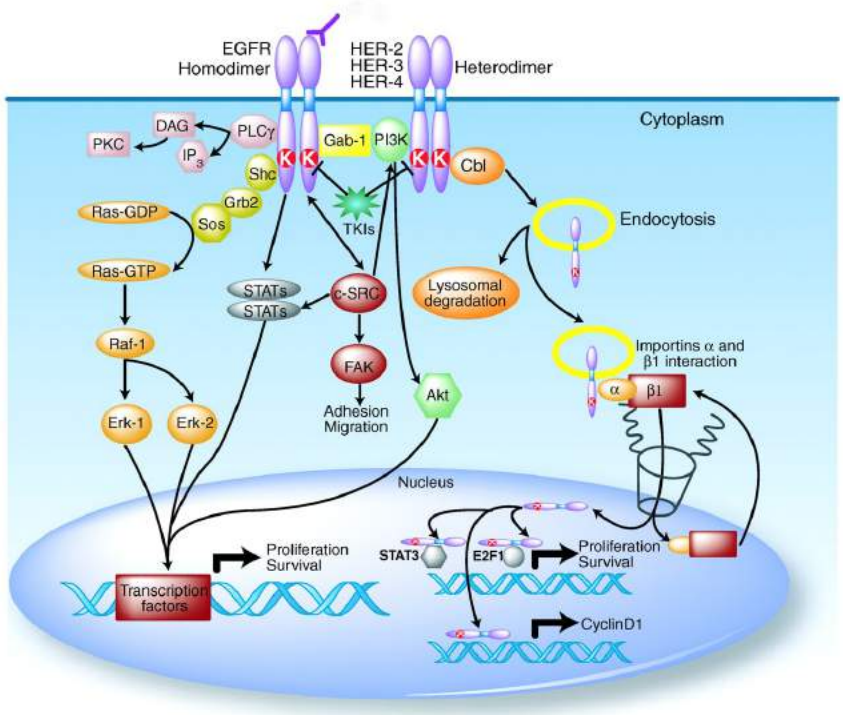


Figure 8: Schematic representation of the EGF signaling pathway. *Activation of an Erb tyrosine kinase receptor after ligand binding results in the formation of homodimeric or heterodimeric receptor complexes and the consequent phosphorylation of specific tyrosine residues located on their intracellular domains. Phosphorylation of EGF receptors may start the same four intracellular signaling cascades activated in the FGF signaling, i.e. the PLCγ, STAT, PI3K-AKT and RAS-MAPK pathway, resulting in the regulation of gene transcription. Adapted from Scaltriti M. and Baselga J. [243].*

To promote the proliferation and survival of neural stem cells (NSCs) and neocortical progenitors, EGF signaling acts in combination with FGF2 as indicated by several in vitro studies [187, 244-249] with FGF2 enhancing the responsiveness of neural progenitors to EGF, a mechanisms that might account for the synergistic activities of these two factors [250, 251]. As FGF2, also EGF influences neocortical progenitors fate decisions by supporting the maintenance of a multipotent state while concomitantly blocking neuronal and oligodendroglial differentiation [252, 253]. At later stages of neurogenesis, EGF signaling also contributes to the orchestration of neuronal migration [254] and to the determination of neocortical progenitors lineage choice and cellular identity [252, 255].

The signaling cascades triggered by the activation of the epidermal growth factor receptor (EGFR) or another of the three members belonging to the Erb family of receptor tyrosine kinases (erbB2/HER-2, erbB3/HER-3, and erbB4/HER-4) after binding of EGF, comprise the PI3K-AKT, RAS-MAPK, PLCγ, and STAT pathways [243, 256, 257], that are the same signaling cascades of the FGF signaling pathway [220, 221, 224-226]. As in the case of FGF signaling, all these four pathways have the downstream effect of regulating gene transcription through different molecular mechanisms.

Retinoic acid (RA) is mainly involved in patterning and neuronal differentiation during central nervous system development. As a patterning factor, it contributes to the regionalization along both the anteroposterior (AP) and dorsoventral axes (DV) of the neural plate and neural tube. Concerning the formation of the anteroposterior axis of the neural plate, retionic acid (RA) in combination with WNT and fibroblast growth factor (FGF) signaling pathways, is specifically responsible for the organization of the posterior regions, comprising the hindbrain and the anterior spinal cord [203, 258-261], as first indicated by studies performed in a quail model that provided evidences of absence of the posterior hindbrain and abnormalities in the anterior spinal cord in the absence of retinoic acid signaling [261-263]. These evidences were further confirmed by the discovery that retinoic acid is synthesized in the posterior mesoderm and that the retinoic acid catabolizing enzyme CYP26C1 is highly expressed in the anterior mesoderm [264]. Regarding its role in the dorsoventral patterning of the developing neural tube, retinoic acid is generated by the newly formed somites and it acts mainly in the specification the dorsoventral axis of the spinal cord regulating the acquisition of interneuron, sensory neurons and motor neuron identities and their development [265-267]. In addition to its patterning function, retinoid signaling plays also a role in the maintenance of neural stem cells and neural progenitors proliferation [268-270]. Additional researches conducted in murine and human *in vitro* models demonstrated that retinoic acid is involved the differentiation of several types of neurons and glia [271, 272]. From studies conducted in animal models it is also known that retinoid signaling is critically involved *in vivo* in the regulation of motor neurons differentiation [258, 259, 273], in the specification of motor neurons subtypes [274-276] and in their axon and neurite outgrowth [261, 277] by regulating the expression of hundreds of genes comprising transcription factors, enzymes, extracellular and structural proteins, cell-surface glycoproteins and receptors, growth factors, neurotransmitters and neuropeptide hormones [278]. Also in the adult, retinoid signaling has a role in the maintenance of motor neurons differentiated state, nerve regeneration and it is involved in neural plasticity [258, 259].

The molecular mechanisms regulating the synthesis of retinoic acid (RA) and retinoid signaling cascade are based on retinol, also known as vitamin A, that functions as both a transport form and a precursor form of retinoids [258, 279]. Retinol can not be synthesized endogenously in animals and it should be derived from the diet in the form of carotenoids and retinyl esters [258, 279]. These dietary components are stored as retinyl esters through modifications to the terminal polar end group of the molecule and serve as retinol storage supplies. Esterification of the long-chain acyl groups of retinyl esters converts them to retinol [258, 279]. To exert its transport function from the storage sites to the cells that require retinoids, retinol is released into the bloodstream and circulates bound to retinol-binding protein 4 (RBP4). The retinol-RBP4 complex is internalized into the target cells through the interaction of RBP4 with the polyprotic transmembrane receptor “stimulated by retinoic acid 6” (STRA6) [280].

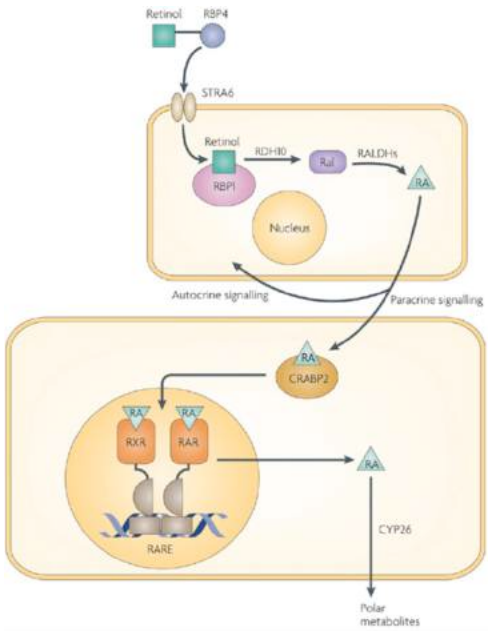


Figure 9: Schematic representation of the molecular pathways involved in the generation, catabolism and signal transduction of retinoic acid (RA). *Extracellular retinol binds to the retinol-binding protein 4 (RBP4) and the complex is internalized into the cells through the interaction of RBP4 with the polyprotic transmembrane receptor "stimulated by retinoic acid 6" (STRA6). Cytoplasmic retinol forms a complex with retinol-binding protein 1 (RBP1) and is metabolized in a two step process. First dehydrogenase 10 (RDH10) metabolizes retinol to retinaldehyde (Ral) and subsequently retinaldehyde (Ral) is metabolized to retinoic acid (RA) by retinaldehyde dehydrogenases (RALDHs). The newly synthesized retinoic acid (RA) can be released from the cytoplasm and serve as stimulus for other cells (paracrine signalling) or can directly translocate to the nucleus of the same cell where it has been produced (autocrine signaling) by forming a complex with the cellular retinoic-acid-binding protein 1 or 2 (CRABP1 or CRABP2). In the nucleus, retinoic acid (RA) binds to a transcription complex which comprise a pair of ligand-activated nuclear transcription factors belonging to two distinct classes, retinoic acid receptors (RARs) and retinoid X receptors (RXRs). The transcriptionally active complex regulates gene expression through the recognition of consensus sequences known as RA-response elements (RAREs) present in the control elements of RA-responsive genes or through other mechanisms still not fully understood. Retinoic acid (RA) catabolism is mediated by different members of the cytochrome P450 family 26 (CYP26) class of enzymes. Adapted from Maden M. (2007) [258].*

In the cytoplasm, retinol binds to the retinol-binding protein 1 (RBP1) and is metabolized to retinoic acid (RA) in a two-step process. First, dehydrogenase 10 (RDH10) converts retinol to retinaldehyde (Ral) and subsequently retinaldehyde (Ral) is metabolized to retinoic acid (RA) by retinaldehyde dehydrogenases (RALDHs) [281]. Once retinoic acid (RA) has been synthesized, it enters the nucleus assisted by the cellular retinoic-acid-binding protein 1 or 2 (CRABP1 or CRABP2) [282] and influences the transcription of target genes by binding to a heterodimeric transcription complex comprising a pair of two different ligand-activated nuclear transcription factors, the RA receptors (RARs) and the

retinoid X receptors (RXRs) [258, 259]. In humans there are three RA receptors (RARα, RARβ and RARγ) and three retinoid X receptors (RXRα, RXRβ and RXRγ), all presenting several isoforms. In addition to ligand binding, phosphorylation of these receptors and recruitment of a wide variety of co-activators or co-repressors is required for the induction or repression of gene transcription [283]. Activated heterodimeric RAR-RXR complexes regulate gene transcription through the recognition of consensus sequences known as RA-response elements (RAREs) present in the control elements of RA-responsive genes and through other mechanisms still unidentified [283, 284].

Early in development the canonical Wnt signaling controls the anterior-posterior axis formation of the neural tube through a concentration gradient characterized by higher concentrations in the posterior areas and lower concentrations in the anterior regions [285-287]. Therefore, inhibition of the canonical Wnt signaling is required for the induction of the forebrain as demonstrated in studies conducted in animal models [288-290] and human *in vitro* neural differentiation protocols [287]. Later, during the first stages of cortical development, canonical Wnt signaling supports neocortical progenitors (NPCs) proliferation, as demonstrated by gain-of-function and loss-of-function experiments in which the over-expression of a stabilized form of β-catenin resulted in the over-proliferation of neocortical progenitors [291, 292] while the conditional deletion of β-catenin leaded to the depletion of progenitor cells [292, 293]. Additional studies performed in mouse mutants have confirmed that specific Wnt genes are involved in the expansion of specific subtypes of telencephalic progenitors [294]. Concomitant to its action of stimulating the mitogenic activity of neural progenitors, Wnt signaling is also critically involved in the dorsoventral patterning of the neural tube as it contributes to cell fate specification by exerting the combined actions of promoting a dorsal telencephalic (i.e. cortical) identity while suppressing ventral telencephalic (i.e. subcortical) fates [293, 295].

Wnt signals are pleiotropic and exert multiple functions during neural induction and patterning of the forbrain. From a molecular perspective, the Wnt signaling network is extremely complex and three main signal transduction cascades can be broadly distinguished, the β-catenin dependent canonical Wnt pathway and two noncanonical β-catenin independent pathways, including the noncanonical planar cell polarity (PCP) pathway (PCP) and the noncanonical Wnt/calcium pathway. The three main Wnt signaling pathways differ with respect to the molecular cascades activated upon binding of the ligand to its corresponding receptor and for the biological function they exert. Concerning their biological function, the canonical Wnt signaling pathway mainly regulates stem cell proliferation and self-renewal, neural cell fate specification and body axis formation by modulating gene transcription [296-298] and it is the Wnt singaling cascade primarily activated during induction and patterning of the forbrain. On the other hand, the noncanonical Wnt planar cell polarity (PCP) pathway is primarily involved in controlling remodeling of the cytoscheleton, coordinated cell migration, tissue polarity and axon guidance [296, 299, 300]. Finally, the noncanonical Wnt/calcium pathway regulates intracellular calcium levels and contributes to the regulation of cell migration and axon guidance while also acting in combination with the noncanonical Wnt planar cell polarity (PCP) pathway in instructing tissue polarity [296, 299, 301]. All the three Wnt signaling pathways are activated by the binding of a Wnt protein to one of its corresponding receptors. Wnt ligands (Wnts) are cysteine-rich secreted glycoproteins encoded in humans by 19 genes [302] and their corresponding human Wnt receptors and co-receptors belong to a diversified population comprising more than 15 different

members, including Frizzled G protein-coupled receptors [303], low-density lipoprotein receptor-related protein 6 (LRP6) [304] and the closely related LRP5 [304], receptor Tyr kinase-like orphan receptor (ROR) [305], protein Tyr kinase 7 (PTK7) [306], receptor Tyr kinase (RYK) [306], muscle skeletal receptor Tyr kinase (MUSK) [306] and members of the proteoglycan families [302]. On the basis of which receptor and signaling cascade is activated within the canonical and non-canonical Wnt signaling pathways in different cellular contexts, they can be further dissected into various distinct sub-branches [299, 302, 307, 308]. In the canonical Wnt signaling pathway, the signaling cascade is activated by the concomitant binding of a Wnt glycoprotein to two cell surface proteins, a member of the Frizzled (FZD) family of serpentine G protein-coupled receptors and the low-density-lipoprotein receptor related protein 6 (LRP6) or the closely related LRP5 that act as co-receptors, resulting in the formation of the Wnt-FZD-LRP6/5 trimeric complex. The activated Frizzled receptor recruits to the plasma membrane the cytoplasmic phosphoprotein Dishevelled (Dvl) [309] which in turn recruits Axin [310-312], a key component of the multi-protein β-catenin destruction complex, including also the *adenomatous polyposis coli* gene product (APC), casein kinase Ia (CK Ia) and glycogen synthase kinase 3 (GSK3) that are two kinases with a prominent role in the phosphorylation of LRP6 or LRP5 [312-314]. Engagement of the multi-protein β-catenin destruction complex to the plasma membrane prevents cytosolic β-catenin phosphorylation by the sequential action of casein kinase 1 (CK1) and glycogen synthase kinase 3 (GSK3) and the consequent targeting of β-catenin for degradation in proteosomes via ubiquitination after having been recognized by the E3 ubiquitin ligase subunit β-Trcp [310, 315, 316]. Unphosphorylated β-catenin is then able to translocate to the nucleus where it forms a transcriptional active complex with members of the T cell factor/lymphoid enhancer factor (TCF/LEF) family of transcription factors and regulates the transcription of its target genes [310]. Also in the planar cell polarity (PCP) signaling pathway, the binding of a Wnt protein to its cognate Frizzled receptor and co-receptor results in the recruitment to the plasma membrane of the cytoplasmic phosphoprotein Dishevelled (Dvl) through the activated Frizzled receptor. Dishevelled (Dvl) in turn activates both the small GTPases Ras-related C3 botulinum toxin substrate 1 (RAC1) and Ras homolog gene family member A (RHOA) which in turn activate RHO kinase (ROCK) and JUN-N-terminal kinase (JNK), respectively. Activated RHO kinase (ROCK) leads to the actin polymerization and microtubule stabilization while activated JNK translocates to the nucleus to regulate gene transcription. Instead in the WNT/calcium pathway, WNT proteins trigger the Frizzled-mediated activation of heterotrimeric G proteins resulting in the activation of phospholipase C (PLC) which in turn stimulates diacylglycerol and inositol-1,4,5-trisphosphate ($Ins(1,4,5)P_3$) production. $Ins(1,4,5)P_3$ stimulates calcium release from intracellular stores and the activation of effectors such as calcium- and calmodulin-dependent kinase II (CAMKII), calcineurin and protein kinase C (PKC) which activate the transcriptional regulator nuclear factor associated with T cells (NFAT) that in turn translocates to the nucleus to regulate gene transcription. In addition to these three main Wnt signaling pathways, there are additional non-canonical Wnt signaling cascades triggered by the binding of WNT proteins to their corresponding Frizzled receptors and specific co-receptors that result in tissue-specific downstream events.

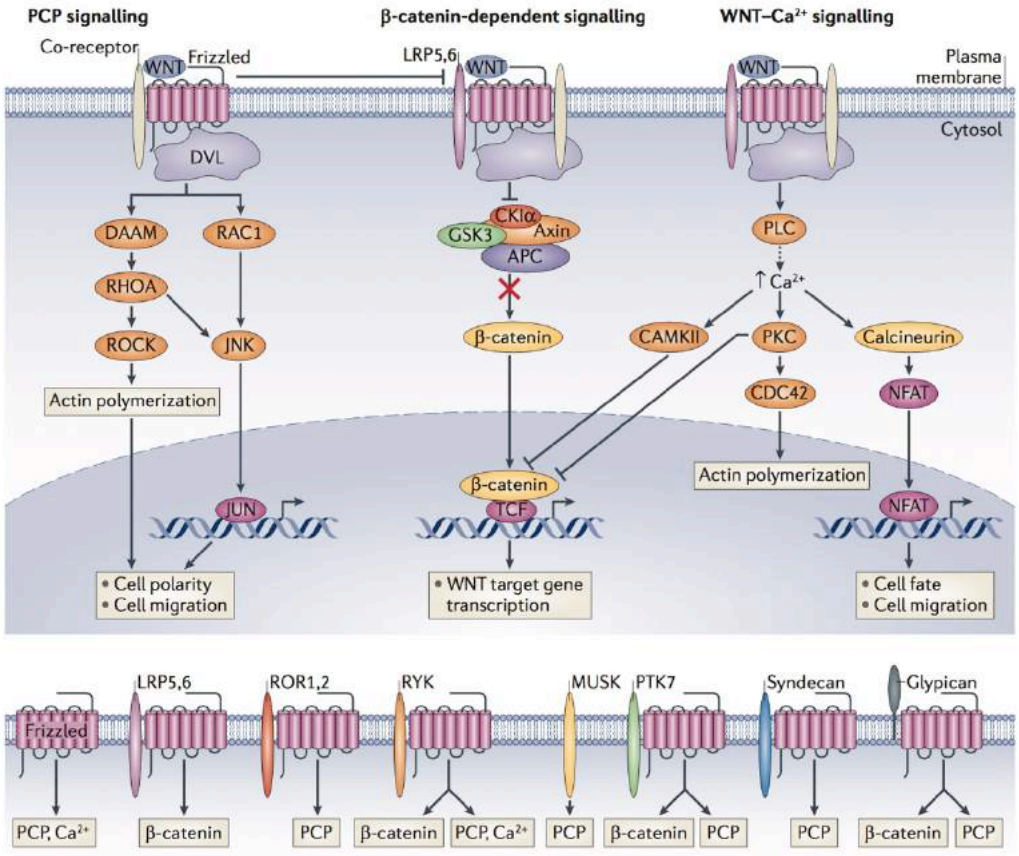


Figure 10: Simplified schematic illustrating the three main WNT signaling pathways, *the planar cell polarity (PCP) pathway, the β-catenin dependent pathway and the WNT/calcium pathway. After binding of a Wnt ligand to its cognate transmembrane receptor and co-receptor present on the cell surface, th cytoplasmic phosphoprotein Dishevelled (Dvl) is recruited to the plasma membrane and in turn activates different signaling cascades. Planar Cell Polarity (PCP) pathway: In this signaling cascade, Dishevelled (Dvl) activates the small GTPases RAC1 and RHOA, which in turn activate RHO kinase (ROCK) and JUN-N-terminal kinase (JNK), respectively, resulting in polymerization of actin and stabilization of the microtubules. β-catenin dependent pathway: In this pathway Dishevelled (Dvl) recruits to the plasma membrane and inactivates the β-catenin destruction complex comprising glycogen synthase kinase 3 (GSK3), casein kinase I a (CKIa), Axin and adenomatosis polyposis coli (APC) by binding Axin. The inactivation of the β-catenin destruction complex leads to the accumulation of β-catenin in the cytoplasm and its translocation to the nucleus, where it activates the transcription of target genes. WNT/calcium pathway: In this signaling cascade, Dishevelled (Dvl) phosphorylates calmodulin-dependent kinase II (CAMKII), protein kinase C (PKC) and calcineurin. Calmodulin-dependent kinase II (CAMKII) translocates to the nucleus where it negatively regulates β-catenin dependent gene transcription. Protein kinase C (PKC) promotes actin polymerization and calcineurin activates the nuclear factor of activated T cells (NFAT) that regulates the transcription of genes controlling cell fate and cell migration. Adapted from Niehrs C. [307].*

Sonic Hedgehog (Shh) signaling controls important developmental processes including the growth, patterning and morphogenesis of many different regions within the body plans of humans and other mammals and with respect to the central nervous system (CNS) it has a prominent role in the dorsoventral patterning of the neural tube, neural stem cells proliferation and neurons and glial cell survival. Concerning its role in the dorsoventral axis formation of the neural tube, Shh secreted from the notochord and later from the floorplate directs cell fate choices in a dose-dependent manner by inducing and maintaining the expression of key transcription factors that define ventral neural tube structures along all the anteroposterior axis of the neural tube [317-322]. As Shh acts as a ventralizing factor, it does not play a role in the patterning and acquisition of neocortical excitatory pyramidal projection neurons identities but it is the critical instructing factor in establishing the formation of the medial and lateral ganglionic eminences that will give rise to the wide variety of neocortical inhibitory interneurons [318, 320, 323, 324]. In addition to controlling cell fate choices, Shh promotes the proliferation and inhibits the differentiation of neural stem cells [325, 326] and neocortical progenitor cells (NPCs) [327, 328].

In the canonical Shh pathway, activation of Shh signal requires binding of Shh to the 12-span transmembrane receptor Patched (PTCH-1 or PTCH-2). The canonical receptor Patched inhibits downstream signaling through the physical interaction with the GPCR-like 7-span transmembrane protein Smoothened (SMO) that prevents Smoothened high expression and activity. Shh binding to Patched-1, and possibly Patched-2, results in the derepression of Smoothened (SMO) with its consequent accumulation on the cell surface membrane and phosphorylation of its cytoplasmic tail that results in the activation of the Shh downstream signalling cascade [329, 330]. In the absence of Smoothened activation, the cytoplasmic protein Suppressor of Fused (SUFU) directly binds and sequestrates in the cytoplasm members of the glioma-associated oncogene homolog (GLI) family of zinc finger transcription factors that are the terminal transcriptional effectors of Shh pathway. In vertebrates, the GLI family of zinc finger transcription factors comprises three members, GLI1, GLI2, and GLI3. Cytoplasmic sequestration of GLI transcription factors by SUFU promotes the processing and degradation of GLI proteins therefore resulting in the inhibition of the Shh signaling cascade [331, 332]. Instead, in response to activation of Shh canonical signaling pathway, GLI proteins are differentially phopshorylated and processed into transcriptional activators that translocate to the nucleus where they actively regulate gene transcription inducing the expression of Shh target genes. In addition to the classical Shh signaling cascade, there are also non-canonical pathways related to Shh signaling based either on the activation of signaling from PTCH1/SMO but independent of GLI transcription factors or on the activation of GLI transcription factors independent of Shh ligand or PTCH1/SMO [330, 332].

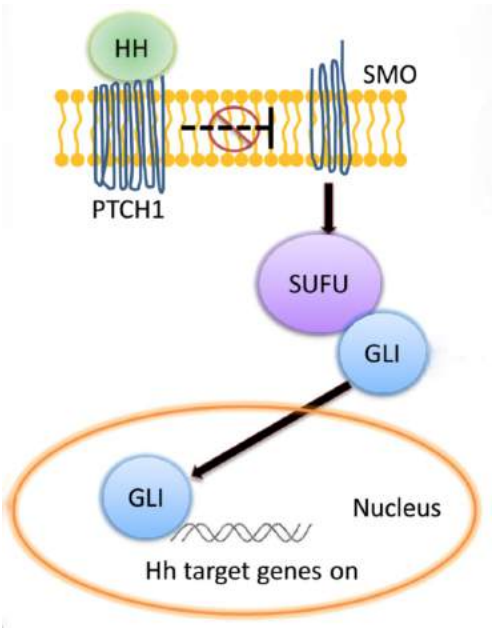


Figure 11: Simplified schematic illustrating Sonic hedgehog signaling pathway. *Binding of a member of the Hedgehog (Hh) family of proteins, including Sonic hedgehog (Shh), to its cognate 12-span transmembrane receptor Patched (PTCH1 or PTCH2) enables the derepression of the GPCR-like 7-span transmembrane protein Smoothened (SMO). Activation of Smoothened (SMO) results downstream in the activation of the cytoplasmic protein Suppressor of Fused (SUFU) that in turn enables the differential phosphorylation of GLI proteins and their processing into transcriptional activators that translocate to the nucleus where they actively regulate gene transcription. Adapted from Meng et al. (2015) [333].*

1.6 Transcriptional programs orchestrating neocortical development

The signaling pathways essential for the proper control of all the different phases of neurogenesis ultimately regulate the transcription of target genes by acting in combination with transcription factors, histone-modifying enzymes and chromatin-remodelling complexes. In the context of corticogenesis, their coordinated transcriptional regulatory effect exerted in conjunction with other molecular mechanisms contributes to direct neocortical progenitors (NPCs) cell-fate restriction and to establish the distinct laminar identities of neocortical pyramidal projection neurons [334-336]. The specific transcriptional programs and molecular mechanisms underlying neocortical pyramidal projection neurons sequential generation and regulating the acquisition of different neuronal subtypes identities have not yet been fully discovered and are currently an active area of research [62, 114, 337-341]. Although, many progresses have already been made in the identification of key transcription factors that are selectively expressed in transient cellular compartments present at different stages of embryonic and fetal neocortical development and characteristic of distinct neocortical layers. First evidences of the cellular compartment- or layer-specific expression of distinct transcription factors were gained from studies performed in murine models [163, 342-346] and these findings have been found to be highly conserved across mammals [347]. Combined with the

identification of transcription factors characterized by region- and cell type-specific expression, additional gain and loss of function studies of murine mutants have started to identify the precise roles of specific transcription factors in the determination of neocortical progenitor cells (NPCs) regional and cellular identities and their functions in regulating the progressive cell fate potential restriction from NPCs to postmitotic neurons [30, 337, 339, 341, 348, 349]. Moreover, there are evidences that specific transcription factors act also as postmitotic regulatory controls necessary to establish the precise neuronal subtype identities of newly born neurons and to further direct the postmitotic refinement and maintenance of postmitotic neurons cell-type specification [148, 151, 350-352]. Along with their role in directing the specification of cell-type identities, the activation of specific transcriptional programs and the resulting expression of particular transcription factors determine also the distinctive cellular features of different neuronal subtypes, including their morphology, electrophysiological properties and axonal connectivity [346, 353, 354]. Therefore, neuronal subtype-specific transcription factors can be used as molecular markers to identify different neuronal classes [30, 344] or they can be manipulated through genetic engineering technologies to study the mechanisms of neocortical pyramidal neurons subtype-specific development and function [346, 353, 355, 356].

Chapter 2: Pluripotent stem cells

2.1 Pluripotent stem cells

Stem cells are characterized by the properties of self-renewal, that is the ability to undergo mitosis while maintaining the same undifferentiated state of the mother cell, and their differentiation potential or potency, that is the capacity to differentiate into a range of more specialized cell types. According to their differentiation potential, stem cells can be classified as totipotent, pluripotent, multipotent, oligopotent and unipotent. In the spectrum of cell potency, totipotent stem cells represent the cells with the greatest differentiation potential as they are able to generate all the embryonic and extraembryonic tissues and to give rise to an entire organism. In the mammalian developing embryo, the zygote derived from that fusion of an oocyte and a sperm cell and the early embryonic blastomeres up to at least the 4-cell stage embryo are totipotent stem cells [357]. As embryonic development proceeds to the 8-cell stage and beyond, blastomeres gradually lose their totipotency and become pluripotent stem cells. Contrary to totipotent stem cells, pluripotent stem cells are able to differentiate into cells of the germ line and into any cell type belonging to or derived from all the three germ layers of the developing embryo but they have lost the capacity to generate extraembryonic tissues as the umbilical cord, the placenta and the amniotic membrane. As embryonic development further proceeds, stem cells maintain their ability to self-renew but their differentiation potential becomes restricted to the generation of cell types belonging to a specific lineage. Therefore, the differentiation products of multipotent stem cells are a closely related family of cells, as is the case of neural stem cells (NSC) that are able to differentiate into neurons, astrocytes and oligodendrocytes [358, 359]. Characterized by a further restriction of potency are oligopotent stem cells that, as the name suggests, are able to differentiate into only a few tissue-specific cell types as lymphoid or myeloid stem cells. At the most restricted end of the potency spectrum, there are unipotent stem cells that have the ability to produce only cells of their own type, but maintain the property of self-renewal that is required to be defined as stem cells. Examples include muscle stem cells and epidermal stem cells that are thought to be able to generate only skeletal muscle cells and keratinocytes respectively.

Special attention has been gained by pluripotent stem cells (PSCs) given their potential to differentiate into cells types belonging to the three germ layers and all their derivatives. To generate PSCs adapted to *in vitro* culture conditions and to maintain them indefinitely as stable cell lines in an undifferentiated state, different approaches have been followed. Based on their derivation method, PSCs can be classified as embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs).

Embryonic stem cells (ESCs) are derived from the inner cell mass of blastocyst stage embryos. The first embryonic stem cells were isolated from mouse embryos in 1981 by Evans M. and Kaufman M. [360] and Matin G.R. [361] independently, using different experimental approaches. The stage at which embryos were explanted, the number of cells explanted and the culture conditions in which embryonic cells were grown after explantation were essential factors that determined the success in isolating and

maintaining ESCs indefinitely in culture as stable cell lines. In the following years, derivation techniques and cell culture conditions have been adapted to generate ESCs also from other species, including non-human primates in 1995 [362] and humans in 1998 [363].

A completely different approach for generating PSCs is based on reprogramming fully differentiated somatic cells back to a pluripotent state by ectopically expressing a small set of key transcription factors. PSCs obtained using this procedure were named induced pluripotent stem cells (iPSCs) to distinguish them from PSCs of embryonic origin. Mouse adult fibroblasts were successfully converted into PSCs for the first time by Takahashi and Yamanaka in 2006 [364]. In the following years, the set of transcription factors used to achieve cell line conversion and the culture conditions adopted were adjusted to extend the reprogramming approach to other species, including humans [365-367].

Critical to the successful generation of PSCs using both generation methods are the culture conditions in which cells are grown and propagated. In fact, both ESCs and iPSCs without regard to the species of origin would spontaneously differentiate into cell types of all the three germ layers if the cell culture medium does not contain species-specific factors that support the pluripotent state. The PSCs natural tendency to differentiate into mixed cell types belonging to the three germ layers and all their derivatives follows an order that recapitulates *in vitro* the steps of early embryonic development [368-370]. Furthermore, it has been discovered that PSCs differentiation potential can be directed to produce specific cell types by exposing them and their progeny to precise molecular cues that trigger the same signaling pathways activated during *in vivo* prenatal development. These findings resulted in the blossoming of many 2D and 3D differentiation protocols for the generation of a myriad of specific cell types [371-374], including neocortical progenitors and neurons [176, 177, 375-378].

Thanks to their ability to differentiate into virtually any cell type under defined culture conditions, PSCs have attracted a lot interest. Currently, ESCs, iPSCs and their differentiation products find applications in both basic and clinical research. In basic research, PSCs have transformed the field of developmental biology providing an extremely powerful *in vitro* tool to investigate the molecular mechanisms underlying the pluripotent state and orchestrating developmental processes in health and disease. Furthermore, PSCs are also applied more broadly as disease models of pathologies that do not have a developmental origin and find applications in translational researches aimed at identifying therapeutic targets as well as in drug discovery, drug testing and drug repositioning studies. Another attractive potential application of PSCs is in regenerative medicine as specific cell types derived from PSCs can be used in cell replacement therapies [379]. Given all the fields in which PSCs are applied, technological advances in cell culture techniques and differentiation protocols are progressing at a fast pace.

2.2 Induced pluripotent stem cells

The generation of pluripotent stem cells (PSCs) starting from somatic cells was first achieved in 2006 by Takahashi and Yamanaka that were able to reprogram back into a pluripotent state mouse embryonic and adult fibroblasts [364]. These pluripotent stem cells (PSCs) generated by reprogramming fully differentiated somatic cells back to a pluripotent state were named induced pluripotent stem cells (iPSCs). The hypothesis at the basis of the generation of induced pluripotent stem cells (iPSCs) was that somatic cells can be reprogrammed back to a pluripotent state by ectopically expressing a small set of key transcription factors. This hypothesis stemmed from the convergence of three major research lines comprising the reprogramming by somatic cell nuclear transfer (SCNT), the discovery of master regulator transcription factors and the studies conducted with embryonic stem cells (ESCs). Until 1962, it was thought that cellular differentiation was a unidirectional process and that somatic cells were irreversibly committed to their cell fate and not able to be reprogrammed to a pluripotent state. John Gurdon research was able to disprove this view by experimentally proving that the nucleus taken from a differentiated frog intestinal epithelial cell and transplanted into an enucleated oocyte is capable of generating a fully functional tadpole [380]. Thus, this experiment demonstrated that oocytes contain factors that are able to reprogram a somatic nucleus and that the nucleus of a somatic cell retains the same genetic information, as early embryonic cells, required for the development of an entire organism. The procedure of transplanting the nucleus of a somatic cell into an enucleated oocyte is known as somatic cell nuclear transfer (SCNT). Following the path opened by the pioneering Gurdon experiment, SCNT has been applied for intraspecies and interspecies cloning of animals as demonstrated by the famous experiment performed by Wilmut research team of cloning Dolly the sheep [381] and by the subsequent cloning of other mammals [382-384]. In addition to somatic cell nuclear transfer (SCNT), somatic cells are able to be reprogrammed back to a pluripotent state also through the fusion with embryonic stem cells (ESCs) [385, 386]. The second stream of research that contributed to set the basis for the idea of reprogramming somatic cells through the delivery of a selected combination of transcription factors came from the realization that there are certain transcription factors able to determine and induce specific lineage identities. The first evidences were collected in 1987 when the ectopic expression of the Drosophila transcription factor Antennapedia was demonstrated to induce the formation of legs instead of antennae [387] and the forced expression of the mammalian transcription factor MyoD was proved to convert fibroblasts into myocytes [388]. The third area that contributed to create the conditions for the emergence of the iPSCs technology is the research conducted on embryonic stem cells (ESCs) that provided the necessary background knowledge about pluripotent stem cells culture conditions [363, 389]. The combination of the evidences gained from somatic nuclear transfer (SCNT) and the discovery of master regulator transcription factors able to determine and specify precise cell lineage identities led to hypothesize that “it is a combination of multiple factors in oocytes or ESCs that reprogram somatic cells back into the embryonic state and to design experiments to identify that combination” [390]. Moreover, the knowledge gained from ESCs culture conditions was essential to be able to generate and maintain induced pluripotent stem cells (iPSCs). To test the feasibility of changing the cell identity of fully differentiated somatic cells and reprogram them back to pluripotent state, Yamanaka group started to concomitantly deliver into somatic cells by

retroviral transduction a set of 24 genes known to be highly and selectively expressed in embryonic stem cells (ESCs) and to be critically involved in the maintenance of the pluripotent state. Successful induction of the pluripotent state was detected by transducing with the reprogramming factors embryonic fibroblasts derived from a transgenic mice in which the βgeo cassette, that confers resistance to the antibiotic geneticin (G418), has been inserted into a Fbx15, a gene not essential for pluripotency maintenance but highly expressed in mouse embryonic stem cells. Embryonic fibroblasts derived from the $Fbx15^{\beta geo/\beta geo}$ mice and transduced with the reprogramming factors were grown in the same culture conditions applied for culturing embryonic stem cells (ESCs) with the addition of antibiotic G418, so that only somatic G418-resistant cells successfully induced to a pluripotent state were able to survive being positively selected. Subsequently, by the process of elimination the initial set of 24 reprogramming genes was narrowed down to only 4 genes that were sufficient to effectively induce pluripotency when ectopically expressed into somatic cells. The original minimal set of reprogramming transcription factors identified by Yamanaka's group consisted of Oct4, Sox2, Klf4 and c-Myc (OSKM). One year later, in 2007 concomitantly Yamanaka research team [365, 366] and Thomson's group [367] were able to extend the reprogramming approach to generate iPSCs also from human somatic cells. The induction strategies adopted by the two groups differed with respect to the set of reprogramming genes of choice and the gene delivery method utilized for the reprogramming. Yamanaka research team was able to successfully induce adult human dermal fibroblasts into iPSCs by using the same set of transcription factors previously adopted for the reprogramming of mouse somatic cells and by optimizing the retroviral transduction method and the cell culture conditions, adapting them to human cells [365, 366]. On the other hand, Thomson's group employed a slightly different set of reprogramming genes comprising the transcription factors Oct4, Sox2 and Nanog combined with the RNA-binding protein Lin28 and adopted lentiviruses as gene delivery system [367]. The homeodomain transcription factor Oct4 (octamer-binding transcription factor 4 also known as POU5F1) belonging to the POU family [391, 392], the transcription factor Sox2 (sex determining region Y-box 2) belonging to the Sry-related HMG box family [393] and the homeobox transcription factor Nanog [394-396] have essential roles in early development as they form the core regulatory network at the basis of the pluripotent state. These three transcription factors upregulate the expression of other genes characteristic of the pluripotent state while concomitantly suppressing the expression of genes that promote differentiation and therefore their expression is crucial for the maintenance of self-renewal and pluripotency [397-401]. The other two reprogramming factors selected by Yamanaka's group, Klf4 and c-Myc, are two tumor-related genes. Klf4 (Kruppel-like factor 4) is a member of the Klf family of transcription factors and it is known to be involved in the regulation of cell proliferation, differentiation and apoptosis [402]. Concerning the pluripotent state, Klf4 contributes to the long-term maintenance of pluripotent stem cells [403] and to the activation of *Nanog* and other pluripotent stem cells specific genes through the direct repression of *p53* [404], as it has been experimentally demonstrated that the p53 protein suppresses *Nanog* and other pluripotent stem cells specific genes [405]. c-Myc (myelocytomatosis oncogene) is an oncogene that belongs to the Myc family of transcription factors and it codes for a nuclear phosphoprotein. As Klf4, also c-Myc has multiple functions and it is critically involved in controlling cell cycle progression, cell growth, cell differentiation, cellular transformation and apoptosis [406-408]. In the context of reprogramming fully differentiated somatic cells back into a pluripotent state, c-Myc protein activates the

transcription of many target genes promoting transformation and proliferation by binding enhancer box sequences (E-boxes) and by recruiting histone acetyltransferase (HAT) complexes [409-412]. The association with histone acetyltransferase (HAT) complexes may promote global histone acetylation, therefore enabling the binding of Oct3/4 and Sox2 to their specific target sites [364, 410]. Furthermore, c-Myc also directly controls DNA replication [413]. Instead, the fourth factor selected to induce pluripotency by Thomson's group in combination with Oct3/4, Sox2 and Nanog, is Lin28. Lin28 is a gene that codes for an RNA-binding protein critically involved in regulating the self-renewal of mammalian pluripotent stem cells [414, 415]. Lin28 proteins are highly expressed in the inner cell mass (ICM) and epiblast [416] where they repress *let-7* microRNAs preventing premature differentiation [417, 418] and concomitantly influence mRNA translation [414]. In the context of reprogramming, both Nanog and Lin28 are non-essential factors that enable to increase the reprogramming efficiency in human cells [367] with Lin28 specifically contributing to enhance the reprogramming efficiency by promoting an increased number of cell divisions and an accelerated cell division rate [419].

These first reprogramming protocols demonstrated the feasibility of inducing fully differentiated somatic cells of different species into a pluripotent state following a robust and simple reprogramming procedure based on the forced expression of a small set of transcription factors. However, these first reprogramming protocols were characterized by very low efficiencies, just around 0.1% of mouse fibroblasts [364] and approximately 0.1%-0.01% of human fibroblasts [365-367, 420] were successfully reprogrammed into iPSCs. In addition to the low efficiency of iPSCs generation, there were also safety concerns regarding the overexpression of the cocktail of reprogramming factors using integrating vectors as retroviruses or integrating lentiviruses, including incomplete reprogramming, immunogenicity, epigenetic changes, genetic mutations, gene insertions and the risk that using gene delivery vectors that stably integrate into the genome may activate or inactivate host genes resulting in tumors formation [421]. In a quest to improve reprogramming efficiency and the quality of iPSCs induction while avoiding the hazards due to the stable integration of transgenes into the genome of iPSCs, many researches have focused on the development of new methods for the generation of integration-free iPSCs. These new reprogramming methods can be broadly classified on the basis of the three different strategies adopted to generate transgene-free iPSCs that include the excision of the reprogramming factors, the use of alternative forms of gene delivery vectors enabling to avoid integration altogether and the direct delivery of reprogramming factors. The excision of the reprogramming factors has been achieved either using Cre-recombinase excisable lentiviruses [422] or the piggyBac transposon system [423, 424]. Another strategy pursued has been the use of alternative forms of gene delivery vectors that enable to avoid the integration of the reprogramming factors altogether. Toward this aim, several alternative viral and non-viral delivery systems have been explored, comprising non-integrating viruses as adenovirus [425], DNA expression vectors [426], RNA virus as Sendai virus [427], plasmid episomal vectors [428, 429] and minicircle DNA [430]. A third approach that has been developed consists in the direct transfection or delivery of the reprogramming factors. In this area of research, synthesized mRNA of the reprogramming factors modified to overcome innate antiviral responses have been directly transfected into somatic cells [431] or reprogramming proteins have been directly delivered to generate induced pluripotent stem cells (iPSCs) [432, 433]. In addition, reprogramming by using exclusively chemical compounds has been reported for mouse [434] but not human iPSCs induction up to now [435].

Combined with the exploration of various gene delivery vectors enabling to avoid the integration of exogenous transgenes into the generated iPSCs, several combinations of reprogramming factors and various adjustments to the cell culture conditions have been tested and adopted to expand the spectrum of the starting somatic cell types induced into a pluripotent state [436-441] and to optimize the reprogramming procedure for each different species [364, 365, 367, 442-446]. In fact, even if the reprogramming procedure is robust and simple and many different fully differentiated somatic cell types from a wide variety of species can be reprogrammed to a pluripotent state by the forced ectopic expression of the classical cocktail of reprogramming factors (OSKM), optimized combinations of reprogramming factors and adjustments of the culture conditions significantly increase the reprogramming efficiency of each gene delivery method applied to a specific somatic cell type of a definite species [447]. The main milestones achieved in the field are summarized in the timeline reported below.

Figure 11: Main milestones in iPSCs technology. *Timeline summarizing the main milestones achieved in iPSCs technology in the first decade since their initial derivation. Adapted from Chari S and Mao S. [448].*

To validate that the generated induced pluripotent stem cells (iPSCs) are indeed pluripotent stem cells (PSCs), a wide set of quality controls aimed at confirming both the functional properties and the molecular identity of iPSCs are routinely performed [449]. These screenings include the evaluation of iPSCs colony morphology, their growth in culture for several passages, the expression of key pluripotency markers characteristic also of embryonic stem cells (ESCs), the monitoring of alkaline phosphatase (AP) expression and activity, the functional assessment of their differentiation potential to

generate cells belonging to all the three germ layers either *in vivo* or *in vitro* and the ability to contribute to the generation of chimera.

The first qualitative assessment of the pluripotent identity of newly generated iPSCs is the evaluation of colony morphology that should be round shaped and densely packed with cells characterized by a high nucleus to cytoplasm ratio presenting large nucleoli and a scant cytoplasm, similarly to embryonic stem cells (ESCs). Then, selected iPSCs clones are grown in culture for several passages to assess their ability to self-renew. To confirm that candidate iPSCs display the molecular hallmarks characteristic of the pluripotent state, the expression of key pluripotency markers is monitored at protein and transcriptional level. Routinely monitored pluripotency markers include transcription factors that form the core regulatory network at the basis of the pluripotent state and surface membrane antigens. Alkaline phosphatase (AP) expression and activity, that are upregulated in pluripotent stem cells (PSCs), are also generally examined. Moreover, to better assess the specific molecular state of the newly generated iPSCs also their global DNA methylation levels may be evaluated. The gold standard to verify the differentiation potential of candidate iPSCs and their ability to generate cells belonging to all the three germ layers is the *in vivo* teratoma formation assay. This assay consists in the subcutaneous injection of candidate iPSCs into immunodeficient mice resulting one to two months after transplantation in the formation of teratomas, that are tumors containing a mixed variety of tissues belonging to all the three germ layers. The pluripotency of candidate iPSCs can also be assessed *in vitro* by performing the embryoid bodies (EBs) formation assay which consists in culturing iPSCs before as floating aggregates and subsequently replating them as a monolayer culture while using throughout the protocol a medium that enables spontaneous cell differentiation. Another key assay to validate the pluripotency of newly derived iPSCs is the generation of chimera. In this assay, candidate iPSCs are injected into the blastocysts of an animal belonging to the same or a different species and it is monitored both the ability of iPSCs to contribute to the generation of chimera and their transmission to the offspring. This assay is routinely not applicable to human iPS cell lines in compliance with ethical stem cell research guidelines. It is worth noting that it is not necessary to perform all the assays summarized above to validate the pluripotency of candidate iPSCs due to redundancy and cost limits. Therefore, only a selected combination of these assays is usually performed for routinely screening and validating newly generated iPS cell lines.

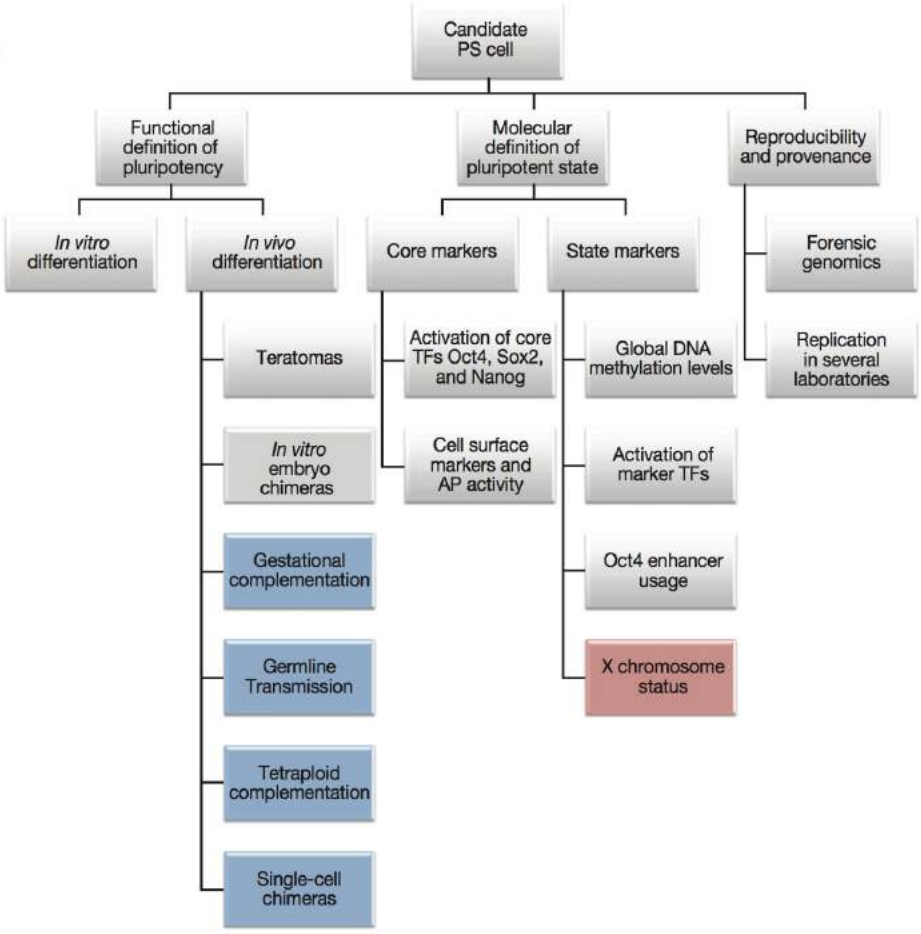


Figure 12: Schematic summarizing the control assays to validate the pluripotency of candidate pluripotent stem cells (PSCs). *Checklist summarizing the experimental controls to assess the function and the state of candidate PSCs. Evaluation of PSCs based on their functional operational definition consists of in vitro and in vivo assays aimed to assess the differentiation potential and self-renewal capacity of the cells under screening. The functional evaluation of PSCs is combined with tests aimed to confirm that candidate PSCs match the molecular definition of the pluripotent state which is assessed at the transcriptional, epigenetic and protein expression level. A third level of analysis aimed to evaluate the reproducibility of the generated PSCs would be advisable but is not always possible to achieve due to cost limits. In this schematic, the blue boxes indicate in vivo differentiation assays that can not be performed on human pluripotent stem cells in compliance with ethical stem cell research guidelines, while the red box marks the uncertain relevance of X chromosome reactivation as a criterion for assessing human ground state pluripotent stem cells, due to the lack of knowledge about the status of the X chromosome in human naive pluripotency. Adapted from De Los Angeles et al. [449].*

2.3 Pluripotent stem cells as a tool to investigate human neocortical development in health and disease

Studies of human neurodevelopment are crucial to understand the regulatory mechanisms that give rise to the peculiar structure and functions of the human brain and to investigate the molecular basis of many neurological and psychiatric disorders that have a developmental origin [450-454]. Currently, the experimental tools available for studying neocortical development include post-mortem brain tissues, animal model systems and different types of *in vitro* cellular and tissue platforms.

Post-mortem brain tissues are definitely the most physiological model system presently available to investigate human neurodevelopment. Epigenetic, transcriptional and protein synthesis investigations of these samples constitute the primary source of information about the molecular programs regulating human neocortical development and are essential to define the gold standard cell-specific signatures to which are ultimately compared data obtained from animal model systems and *in vitro* cell culture platforms, to assess their validity and relevance. Though, neocortical development researches based on human post-mortem brain tissues face several obstacles, including the difficulty in the procurement of high-quality post-mortem tissues that results in a limited number of samples available and the lack of experimental control on the developmental stage of the samples obtained that consequently translates into a constrained and suboptimal coverage of the different human neurodevelopmental periods investigated [15, 455]. A further limitation is posed by the impossibility to conduct experimental manipulations on post-mortem brain tissues which confines this experimental tool to more descriptive studies.

Many of these limitations can be overcome by conducting research using animal model systems that represent a powerful tool to investigate the molecular mechanisms, cellular functions and circuit dynamics of the neocortex, as they allow experimental manipulations and the achievement of high spatial and temporal resolution [62, 108, 111, 456]. Given these advantages, studies performed in animal model systems have provided fundamental insights into many human neurodevelopmental processes and cortical functions [4, 15, 29, 30, 336]. However, the use of animal model systems in neuroscience research is to some extent limited by the accessibility to the relevant biological targets for measurements and observation, the restricted throughput and sometimes the presence of confounding variables due to the complexity of the system. Most importantly, it has to be taken into consideration that vertebrate and invertebrate animal models are limited in precisely mimic the cellular complexity distinctive of the human brain [36, 75, 457-460] and in recapitulating certain characteristic features of human neurodevelopment and physiology due to the evolutionary distance between humans and these model organisms [4, 47, 461].

A third major category of model systems available to investigate human neocortical development consists of a wide variety of human *ex vivo* and *in vitro* tissue and cellular platforms [462]. These platforms comprise organotypic brain slices [463], primary cultures of neocortical progenitors and neurons [464], direct conversion of fully differentiated somatic cells into specific subtypes of neurons without passing through the pluripotent state, a process known as transdifferentiation or lineage reprogramming [465], and 2D or 3D cell culture platforms based on the differentiation of pluripotent stem

cells (PSCs) into neocortical progenitors [176, 177, 466] and distinct subtypes of neurons [375, 376, 467]. In this context, human PSCs-based technologies have attracted a lot of attention as their differentiation potential can be directed to generate with high efficiency the neural cell types of interest through the exposure to small molecules that trigger the same signaling pathways activated during human *in vivo* neurodevelopment. Human PSCs *in vitro* differentiation preserves the order of sequential generation of the different neocortical cell types, with deep-layer neurons developing first followed by the maturation of upper-layer neurons [375, 468]. The competence to recapitulate in a dish the main *in vivo* neurodevelopmental events and the ability to enable to reach an unprecedented spatial and temporal precision in the selection and analysis of specific developmental stages, makes human PSCs-based differentiation platforms the model system of choice to study the spatio-temporal transcriptional programs orchestrating neocortical development [468] and their disruption in disease [377, 378, 469-471]. In fact, the spatial and temporal resolution in investigating specific developmental stages offered by PSCs-based platforms is impossible to achieve using other model systems such as human brain tissues, animal models and other cell culture platforms like transdifferentiation protocols that are also able to generate the desired neural cell type populations with high efficiency but do not mimic the cell fate specification developmental trajectories. This unique advantage as a tool to model neocortical neurogenesis has fostered the development of many 2D and 3D human PSCs-based culture systems. 2D PSCs-based differentiation protocols do not recapitulate the complex neocortical cytoarchitecture, lack cell-matrix interactions and do not fully mimic all the physiological cell-cell interactions characteristic of intact human brains but their intrinsic simplicity offers several advantages shared also with other *in vitro* 2D cell culture systems, as human primary neuronal cultures and human neurons generated through lineage reprogramming. All these three 2D cell culture platforms enable experimental accessibility to measure and investigate the cellular and molecular targets of interest, allow a precise control of the extracellular environment that results in more homogenous cultures and they can be easily experimentally manipulated [462, 472]. Overall, these advantages make 2D PSCs-based differentiation protocols an highly reliable and reproducible *in vitro* model system to study neocortical development. Additional researches, aimed to bridge the gap between the cellular and tissue/organ level of structural organization and to mimic more closely *in vitro* the cytoarchitectonic complexity of the developing human brain, have been recently devoted to develop several 3D cell culture systems [462, 473, 474]. Currently, 3D cell culture systems include cortical spheroids [475] and cerebral organoids that can be either region-specific [378, 476] or able to reproduce various brain developmental regions [377]. Cortical spheroids and even more cerebral organoids recapitulate *in vitro* many biological features of the *in vivo* developing human brain as cell-matrix interactions, proper cell-cell connection patterns and the spatial organization of different tissue-specific cell types but they still lack the ability to reproduce the overall spatial organ level organization [477]. Furthermore, the increased complexity and faithfulness in mimicking the developing human brain of these 3D *in vitro* model systems is currently accompanied by an increased difficulty in accessing and measuring the biological targets of interest and by a reduced experimental control over the differentiation process that results in an enhanced heterogeneity and diminished reproducibility of the differentiation products. To overcome these hurdles and to optimize these *in vitro* technologies, many studies are focusing on applying tissue engineering approaches to 3D culture systems [473]. These approaches include attempts of integrating these platforms with 3D bio-printing methods [478] following the preliminary successes for other organs [479-481] and research plans to

provide organoids with systems mimicking the vascular network [462, 473, 482-485] and with scaffolds or beads that release patterning molecules to improve and standardize their spatial organization [486, 487]. The convergence of tissue engineering with 2D and 3D stem cells-based cultures has opened an active area of research that is already producing new technologies generally refereed to as organs on-a-chip [473, 488].

In addition to be an essential tool to investigate human neocortical development, 2D and 3D PSCs-based differentiation protocols become also a powerful tool for modeling neurodevelopmental disorders when applied to iPSCs, as iPSCs fully recapitulate the genetic background of diseased individuals and therefore they allow to bridge the gap between the genetic lesions and the clinical data by providing a model system to investigate *in vitro* pathophysiological mechanisms [469]. iPSCs conservation of the patient-specific genetic makeup not only allows to identify the molecular pathways underlying neurodevelopmental disorders and analyze the disease-related cellular phenotypes but enables also to distinguish important differences in disease states specific to individual patients, therefore increasing the degree of resolution in understanding shared and unique disease mechanisms. This property renders iPSCs also attractive to the pharmaceutical industry as their integration into drug discovery pipelines may be critical in the evaluation of potential side effects arising in humans but undetected in animal models [489, 490] and offers the promise to move toward developing patient-specific therapies, a field also known as personalized or precision medicine [491, 492].

Chapter 3: Genetic basis of neurodevelopmental disorders

3.1 Genetic architecture of neurodevelopmental disorders

Neurodevelopmental disorders comprise a wide variety of syndromes characterized by impaired development and growth of the central nervous system (CNS) that have a prenatal, infancy or early childhood origin. Their onset is generally in the first years of life even if certain pathologies such as schizophrenia emerge later, in adolescence and early adulthood. Their initial diagnosis is mainly based on clinical symptoms and for certain conditions it is also aided by the presentation of distinctive facial features. The etiology of many neurodevelopmental disorders is still unclear despite their high incidence (at least 13-15% in the US, according to [493] and NSCH data) as the majority of these pathologies involve complex traits and for only a minority of these disorders a defined causative genetic mutation could be identified. Though, the high heritability of complex neurodevelopmental disorders established by twin [494, 495] and family [496, 497] correlation studies and the identification of rare Mendelian syndromes that meet the same diagnostic criteria [498-501] have highlighted the importance of the genetic component as being causative or a predisposing risk factor acting in combination with environmental components. Consequently, these evidences have fostered the research in identifying the genetic architecture of complex polygenic neurodevelopmental disorders such as autism spectrum disorders (ASD), attention deficit hyperactivity disorder (ADHD), schizophrenia and intellectual disability (ID) and this is currently an active area of investigation [453, 502-504]. Thanks also to technological advances in genetics and genomics, several studies have already enabled the description of various mutations that lead to rare Mendelian syndromes [498-501] and have revealed many variants correlated to complex neurodevelopmental disorders [503, 505-509].

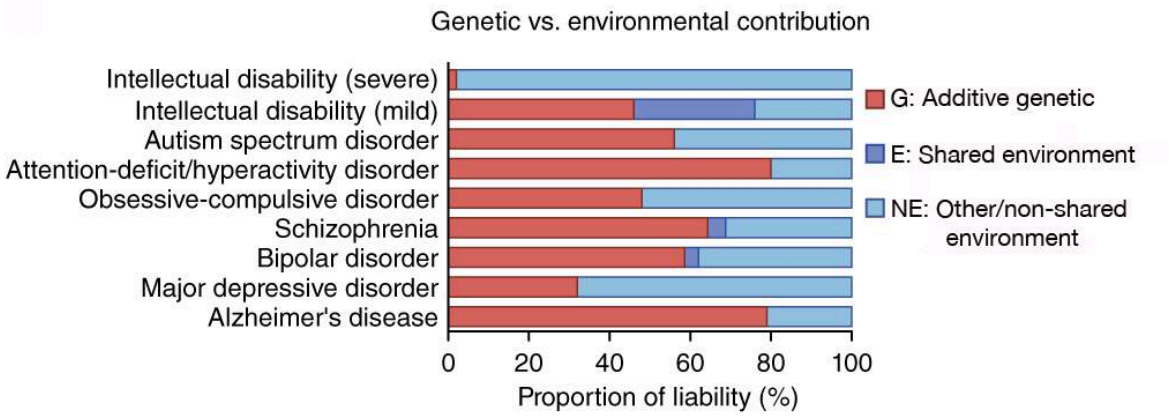


Figure 13: Estimates of the genetic and environmental contributions to several neurodevelopmental disorders. *Histograms representing the relative percentages of the additive genetic variance (G) and environmental factors, subdivided into shared environment (E) and non-shared environment (NE) variances, predisposing to different*

neuropsychiatric disorders, including neurodevelopmental disorders. Data were obtained from twin, family and population studies. Adapted from Gandal M.J. et al. [503].

The identified causal genetic variants and genetic risk factors associated to neurodevelopmental disorders can take different forms and be classified according to the type of genetic mutation, ranging from genomic structural variations affecting the chromosome structure to alterations in the nucleic acid sequence. Genomic structural variations associated to neurodevelopmental disorders include copy number variations (CNVs), in which genomic regions larger than 1 kilo base pairs are duplicated or deleted resulting in altered gene dosage characterized respectively by triploidy or haploid number of copies of the genetic elements encoded in the affected genomic loci [510, 511], and other forms of small chromosomal rearrangements (<50 kb) as inversions and translocations [503, 509]. Also alterations in the nucleotide composition have been linked to neurodevelopmental disorders pathogenesis in the form of trinucleotide repeats, insertions and deletions of bases known as indels or single base pair changes that can be further classified on the basis of their occurrence as single nucleotide polymorphisms (SNPs), when the single nucleotide change is present to some noticeable degree within a population (> 1%), or single nucleotide variations (SNVs) if the single nucleotide modification is either rare (< 1%) or common [512-514]. The genetic variants associated to neurodevelopmental disorders can also be classified on the basis of their origin as new germline mutations, also known as *de novo* mutations or their inheritance pattern, according to their frequency within a population as rare or common and depending on their effect size that reflects the degree of biological causality.

Genome wide association studies (GWAS) performed on large cohorts of healthy controls and individuals affected by complex neurodevelopmental disorders have identified that there is an inverse relationship between mutation frequency and effect size. For common alleles (allele frequency > 0.5%) the link to biological causality is more blurred and their effect size is reduced (odds ratio < 1.3) with respect to selected high-impact rare variants that are able to produce larger effects size (odds ratio typically 2-60) and more evident phenotypic mutations [503, 515]. This inverse relationship between mutation frequency and effect size is also followed by neurodevelopmental disorders characterized by a simple genetic architecture in which disease liability can be ascribed to one or few genes, as in the case of rare Mendelian syndromes and specific rare *de novo* mutations. This trend can by explained by the negative selection of high-impact mutations producing large effect size due to the reduced reproductive fitness of their phenotypes. Consequently, high-impact genetic variations usually occur *de novo* and are generally rare. This observation has lead to question the relative contributions to disease liability that can be ascribed to common and rare variants respectively and to investigate their penetrance. To this regard, two different models emphasizing the role of either common or rare mutations have been proposed. The common disease common variant (CDCV) model hypothesize that genetic susceptibility to complex neurodevelopmental disorders is ascribable to several high-frequency variants that confer moderate risk taken alone but are responsible for a large percentage of genetic predisposition in aggregate. On the other hand, the common disease rare variant (CDRV) model highlights the contributions of rare variants to disease susceptibility. Many CNVs and genome wide association studies (GWAS) aimed at testing the CDCV

and CDRV hypothesis have been performed [516-522], though their results remain inconclusive toward providing a defined answer to this debate and new perspectives and investigation approaches are emerging [503, 511, 523].

The identification of specific genotypes linked to neurodevelopmental disorders phenotypes is only a first objective of the field. Other major aims include the discovery of the molecular mechanisms underlying disease etiology, the identification of the specific disease-related abnormalities occurring at cellular, neuroanatomical and physiological level and building a framework that connects these different levels of biological investigation between each other and to the peculiar cognitive and behavioral impairments related to specific clinical phenotypes. In this larger perspective, two different approaches are followed. On one hand, the research focus is still directed toward the identification of single genotype-phenotype correlations. By following a bottom-up strategy and increasing the specificity of the genetic target under investigation, this single genotype approach potentially enables to better dissect the underlying molecular mechanisms as well as to identify cellular alterations and to draw genotype-phenotype associations [524, 525]. On the other hand, a more holistic approach based on integrative methods combined with systems biology is adopted to identify convergent pathways and common cellular targets disrupted in the pathological states [503, 504, 511, 523, 526]. The first results obtained using this strategy have started to highlight shared or disease-specific biological processes altered in several neurodevelopmental disorders and suggest that common inherited and rare *de novo* variants affect different pathways [507, 527-532].

3.2 Copy number variations contribution to neurodevelopmental disorders liability

In this research context aimed at understanding the etiology and multiple-level defects implicated in neurodevelopmental disorders and at discovering therapeutic targets building on this knowledge, the study of copy number variations (CNVs) has provided initial evidences of the relevant role played by rare mutations in disease liability and risk predisposition [533, 534]. The first recognitions that CNVs underlie specific neurodevelopmental disorders emerged from laboratory tests performed on individuals presenting clinical symptoms meeting well-defined diagnostic criteria. Laboratory tests enabled to identify the specific genomic regions deleted or duplicated in the affected subjects and therefore allowed to link definite genotypes to precise clinical syndromes, also known as genomic disorders, like Williams syndrome, Smith-Magenis syndrome and DiGeorge syndrome [535-537]. These genomic disorders are characterized by high penetrance and low incidence reflecting the rare occurrence rate of the underlying CNVs that generally arise as *de novo* mutations since they are under negative selection. The importance of CNVs in the etiology of genomic disorders characterized by a simple genetic architecture highlighted by these initial findings lead to investigate the contribution of CNVs also to complex polygenic neurodevelopmental disorders. Genome-wide studies of CNVs association with complex neurodevelopmental disorders took mainly three study design forms, including family-based analyses aimed at identifying spontaneous disease-related CNVs that occurred *de novo* in the proband, population level studies also trying to assess the overall contribution of rare CNVs to disease risk through the comparison of affected patients to healthy controls and

researches focused on delineating the association of specific genomic regions or genes to multiple disorders in a comprehensive way [538]. These studies discovered a large number of additional CNVs related to specific syndromes [539, 540]. Furthermore, they identified many recurrent CNVs that are associated with different degrees of penetrance to one or more complex neurodevelopmental disorder, as developmental delay [541], intellectual disability [542], autism spectrum disorder [518], attention deficit hyperactivity disorder [543] and schizophrenia [544]. Therefore, even if specific CNVs are rare and generally arise as *de novo* mutations, according to current estimates at least 15% of human neurodevelopmental disorders have been ascribed to rare CNVs that produce selective gene dosage imbalances [534]. The percentage of CNVs contribution to disease liability increases if taking into account even other CNVs mildly related to complex neurodevelopmental disorders.

A general trend that emerged from these studies is that clinical phenotypes are more frequently associated with deletions than with duplications. This observation might be explained by taking into consideration the mechanisms that produce the structural variation. Deletions may result from interchromosomal, interchromatid and intrachromatide nonallelic homologous recombination (NAHR) while duplications are mediated only by interchromosomal and interchromatid non-allelic homologous recombination (NAHR) [545, 546]. Therefore, deletions are more likely to occur than duplications. Furthermore, in addition to induce haploinsufficiency, deletions uncover recessive alleles of the CNV region consequently increasing the probability of the manifestation of disease-related phenotypes [534, 547]. Another observation worth noting is the high level of phenotypic variability highlighted by investigations of the contribution of specific CNVs to both genomic disorders characterized by simple genetic architectures and complex neurodevelopmental disorders. The clinical phenotypes vary widely even in neurodevelopmental disorders that have a simple genetic basis involving a specific CNV [536]. This phenotypic variability seems even increased when examining CNVs association to complex neurodevelopmental disorders. The degree of association of a precise CNV to a specific complex neurodevelopmental disorder is generally partial, meaning that the presence of the CNV only sometimes results in a specific neuropsychiatric phenotype and it is also frequent that the same CNV is associated with multiple neurodevelopmental disorders [538, 548-551]. Moreover, CNV alleles associated to neuropsychiatric disorders of developmental origin can also be expressed at a low frequency in the healthy population [550].

The high phenotypic variability of CNVs genotypes and the current lack of an integrative explanatory framework for this observed variability highlights that the contributions of specific CNVs and CNVs as a genetic variant group to human neurodevelopmental disorders liability are still mostly unexplored. Further advancement of this field is necessary given the high degree of neurodevelopmental disorders liability ascribable to CNVs [534]. Therefore, research focus is at present directed toward constructing this framework and better defining the genotype-phenotype relationship by discovering the pathways and molecular mechanisms disrupted in disease as well as by identifying the resulting cellular, physiological and neuroanatomical abnormalities produced by specific alterations in gene dosage at the level of both single genes present in the CNV and all the genes contained in the CNV combined. Studies taking into consideration also specific genetic backgrounds are pursued as well, given that the high phenotypic variability of CNVs genotypes might be at least partly ascribed to the modulatory effects

exerted by epigenetic regulation and the expression of other common and rare alleles [538, 552].

3.3 Copy number variation at the 7q11.23 locus

Copy number variation of the Williams-Beuren syndrome chromosome region (WBSCR), located on the long arm (q) of chromosome 7 at position 11.23 (7q.11.23 locus), is of special interest as the two disorders resulting from the deletion (Williams syndrome, WS) and duplication (7q.11.23 duplication syndrome, 7q11DUP) of this region exhibit cognitive and behavioral phenotypes marked by both similar features and symmetrically opposite traits [553-557]. The relatively simple and defined genetic basis underlying these disorders combined with the specificity of the observed opposite cognitive and behavioral phenotypic alterations make the study of the 7q11.23 locus particularly attractive to investigate the genetic basis of precise aspects of human cognition. Furthermore, the association of 7q11.23 duplication to complex neurodevelopmental disorders such as autism spectrum disorder and schizophrenia while 7q11.23 deletion produces a well-characterized syndrome without clear overlapping to complex neurodevelopmental disorders, render the study of this locus also extremely interesting to identify the molecular mechanisms unique to each disorder, common to both syndromes and shared with other complex neurodevelopmental disorders.

The Williams-Beuren syndrome chromosome region (WBSCR) spans 1.5 to 1.8 million base pairs (Mb) comprising 26 to 28 genes (Figure 14, panel A). Rearrangements of this region occur as it is flanked by blocks of low copy repeat sequences containing genes and pseudogenes that share high sequence homology, known as duplicons. Therefore, due to its peculiar genetic architecture, WBSCR is prone to frequent structural rearrangements resulting in CNV of this locus. In fact, the spatial closeness and the high level of sequence homology between duplicons make WBSCR vulnerable to misalignments of duplicons during meiosis and subsequent unequal crossing over resulting in abnormal recombinant products, where the WBSCR is either deleted or duplicated (Figure 14, panel C) through a mutational mechanism know as nonallelic homologous recombination (NAHR).

CNV of the 7q11.23 locus is sporadic and can occur spontaneously during meiosis on either paternally or maternally inherited chromosome 7. The length of the deletion or duplication depends on which duplicons are involved in the rearrangement. Generally, the low copy repeats sequences involved in the rearrangement are the medial and centromeric duplicons (specifically, blocks B indicated in Figure 14). Their nonallelic homologous recombination results in the typical deletion or duplication of 1.5 Mb in the WBSCR encompassing 26 to 28 genes that occurs in the majority of patients. Less frequently, duplicons misalignment gives rise to a larger deletion or duplication of 1.8 Mb comprising 28 genes [558-560].

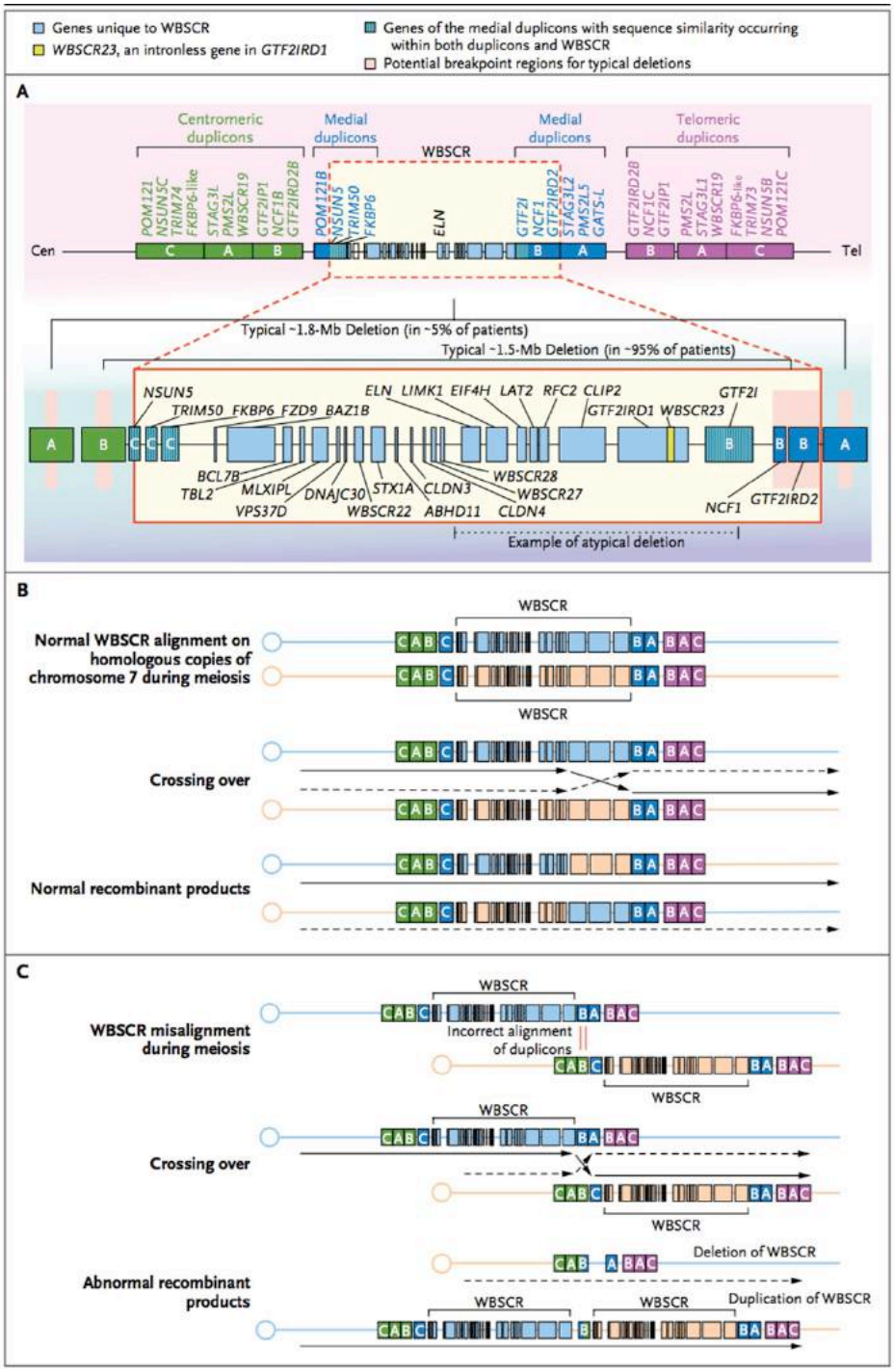


Figure 14: Schematic representation of the Williams-Beuren syndrome chromosome region and rearrangements mechanisms leading to 7q11.23 CNVs. *In panel A is reported a schematic of the Williams-Beuren syndrome chromosome region (WBSCR) highlighting the genes contained in it and the low copy repeat sequences flaking the region. In panel B and C are represented the normal and misaligned pairing of duplicons during meiosis that results after crossing over in normal recombinant products or in deletions and duplications of WBSCR, respectively. Adapted from Pober*

B.R. [536].

It is also possible that shorter or larger atypical deletions or duplications of the WBSCR arise. The effects of these atypical duplications have not yet been fully investigated while more extreme phenotypes accompanied by moderate to severe intellectual disability have been observed in individuals with deletions longer than 1.8 Mb [561]. On the other hand, deletions of the WBSCR shorter than 1.5 Mb have been associated with variable phenotypes [562-564]. The more pronounced phenotypic changes related to 7q11.23 CNV size are observed in WS with respect to 7q11DUP follow the general trend observed in many CNV studies where clinical phenotypes have been more frequently detected in cases of deletions rather than duplications.

3.4 Williams syndrome

Williams syndrome (WS) is a multisystem disorder caused by the deletion of the Williams Beuren syndrome chromosome region (WBSCR). Prevalence is estimated to range from 1:7500 to 1:20000 based on different sources [565, 566]. WS patients display peculiar phenotypes characterized by distinctive facial features, intellectual disability accompanied by psychiatric and psychopathological conditions, cardiovascular disorders, connective tissue abnormalities and endocrine system impairments. Cardiovascular abnormalities include supravalvular aortic stenosis (SVAS) present in the majority of patients and stenosis of other large and medium arteries that are due to the elastin (ELN) gene dosage imbalance, hypertension and myxomatous degeneration of aortic or mitral-valve leaflets [567, 568]. Several endocrine abnormalities are also found in WS patients including hypothyroidism, hypercalcemia and hypercalciuria [569, 570].

WS mean IQ of 50-60 indicates the presence of mild to moderate intellectual disability but WS affected individuals with normal IQs are also present [553-555]. The characteristic pattern of cognitive abilities displayed by WS patients reveals strengths in interpersonal and social skills, facial recognition, auditory memory, language performance and language-related skills above than expected for their IQ [553, 556]. These strengths are accompanied by marked weaknesses in visuospatial construction, visuomotor skills and significantly impaired performance in writing, mathematics and drawing [554, 557, 571]. This distinctive cognitive pattern is combined with unique personality traits characterized by excessive friendliness, over engaging behaviors even with strangers, lack of social inhibition and high levels of empathy [553, 572-574]. Other common behavioral problems in WS are reduced adaptive behavior than expected for IQ and difficulty with emotional regulation [575, 576]. WS is also frequently accompanied by several psychiatric and psychopathological conditions including anxiety disorders [577], specific phobias [578], attention deficit disorder [579], perseveration, restricted interests and repetitive behaviors [580].

The unique cognitive and behavioral profile that characterizes WS individuals has been associated with a peculiar pattern of cellular and neuroanatomical abnormalities. Histological analyses have highlighted differences in the global and regional neuronal organization as well as decreased neuronal size, increased packing density and coarseness abnormalities distinctive of WS brains compared to controls [581]. Anatomical examination of the central nervous system has identified a significant reduction in whole brain volume (11-13%) [582, 583] and several morphological

abnormalities in the cerebral cortex, amygdala, hippocampus, cerebellum, basal ganglia, brainstem and corpus callosum of WS patients. In WS affected individuals, the overall volume of the amygdala is increased [584] and the hippocampus has normal size but exhibit subtle shape alterations and it is characterized by an impaired metabolism [585]. It is reported that the cerebellum has normal or increased size, depending on the studies [582, 586] and at a closer inspection it reveals a distinctive increase of the neocerebellar vermal lobus area [582] while basal ganglia and brainstem display a significant reduction in size [587-589]. The volume of the corpus callosum is significantly reduced especially in the caudal part and in the splenium and its shape is more convex compared to controls [590]. Diffusion tensor imaging (DTI) studies have revealed that several major white matter tracts are compromised displaying abnormal volumes, density and integrity [591, 592]. Specifically, the right superior longitudinal fasciculus, the left fronto-occipital fasciculus, the caudate and the cingulum show compromised integrity as indicated by the increased values of fractional anisotropy, that is a measure of microstructural integrity, whereas the corticospinal tract shows decreased values [593]. Of outstanding interest is the more pronounced integrity defect revealed in the superior longitudinal fasciculus (SLF) of the right hemisphere that is associated with the dorsal stream, a pathway involved in visuospatial construction [593, 594]. This defect might be related to the visuospatial impairment characteristic of WS. Overall, lesions of these white matter tracts may constitute one of the anatomical substrates underlying the cognitive and behavioral abnormalities characteristic of WS.

The cerebral cortex of WS individuals is also characterized by multiple structural defects, including reduced volume and increased gyrification in specific regions, increased cortical complexity of specific areas and sulci abnormalities, as revealed by magnetic resonance imaging (MRI) studies. Loss of volume affects especially the occipital and superior parietal cerebral cortex while the volume of frontal lobes cortex is relatively preserved [583, 587, 595]. Cortical surface analyses have revealed decreased sulcal depth in the intraparietal and orbitofrontal sulci and a reduction in the extent of the central sulcus [596] along with a specific increased gyrification of the right parietal, right occipital and left frontal regions [597-599]. Several spatially extended cortical regions distinguished by an increased complexity and thickness have been identified [597], including delimited zones of right hemisphere perisylvian and inferior temporal cortex [598]. Furthermore, multiple sulcal abnormalities have been detected as a reduced sulcal depth in the intraparietal and orbitofrontal sulci and a reduction in the extent of the central sulcus [600].

In addition to the cellular and neuroanatomical abnormalities, functional MRI (fMRI) researches have revealed differences in neural activation between individuals affected by WS and healthy controls during tasks involving response inhibition [601], visuospatial construction [602] and auditory processing [603]. Abnormal amygdala activation in response to processing of social, emotional and fearful stimuli might be considered the neural correlate of the abnormal social cognition and emotional dysregulation distinctive of WS [572, 604, 605].

3.5 7q11.23 duplication syndrome

7q11.23 duplication syndrome (7q11DUP) is a multisystem disorder caused by the duplication of WBSCR. Prevalence is unknown and it is estimated to range from 1:7500 to 1:20000 based on prevalence estimates of WS [606, 607]. Compared to WS, 7q11DUP generally exhibit milder and highly variable phenotypes characterized by minor facial features, neurological problems, psychiatric conditions, cardiovascular disorders, motor speech disorders such as dysarthria and dyspraxia and additional phonologic problems [608]. Other frequent impairments observed in a subset of patients include joint laxity, genitourinary tract abnormalities and gastrointestinal problems comprising chronic constipation and feeding difficulties as well as chronic otitis, diastema, strabismus and hearing loss [606, 608]. It is worth noting that facial features presented by 7qDUP individuals are opposite with respect to the distinctive facies of WS patients.

Among the impairments observed in 7q11DUP, neurological problems and psychiatric conditions have special relevance as they are the predominant phenotypic trait and they are observed to some extent in the majority of patients. Neurological problems include developmental coordination disorder, involuntary movements, balance problems, hypotonia, high pain tolerance, hydrocephalus and seizures [606, 608-610]. The median IQ of 7q11DUP is in the low average but it can range widely from severe mental retardation to normal and above average scores, so intellectual disability is diagnosed just in a subset of patients [611]. The pattern of cognitive abilities presented by 7q11DUP affected individuals reveals preserved weaknesses and strengths in specific abilities. Spatial abilities and nonverbal reasoning are preserved at all ages while a decrease in verbal abilities over lifetime has been detected, especially affecting expressive language, accompanied by weak reading and mathematics performance that are at the bottom of the average or low average range [611, 612]. This pattern of cognitive abilities display some symmetrically opposite strengths and weaknesses compared to WS where language skills are preserved but visuospatial abilities are severely impaired. In contrast some weaknesses, as below average performance in mathematics, are shared by both disorders. In the majority of 7q11DUP patients adaptive behavior is limited as in WS and maladaptive behavior scores range from high to normal [611]. 7q11DUP is also frequently accompanied by several psychiatric and psychopathological conditions, including anxiety disorders, attention deficit hyperactivity disorder (ADHD), social and specific phobias, that are conditions also associated with WS [611]. In addition, 7q11DUP patients frequently display also disruptive behavior disorder (DBD), oppositional defiant disorder (ODD) and social phobias [611]. Furthermore, 7qDUP has also been associated with autism spectrum disorders [518, 608, 611, 613] and schizophrenia [560].

The identification of the neural correlates underlying the cognitive and behavioral profiles characterizing this syndrome is under study. As 7q11.23 duplication syndrome was identified later since it is characterized by milder phenotypes and occurs less frequently than WS, the number of studies documenting the cellular, anatomical and physiological alterations distinctive of this disorder are reduced and further researches are necessary

to better delineate them. Up to now, anatomical examination of the central nervous system has identified a significant increase in total brain volume [614] and several morphological abnormalities in the cerebral cortex, amygdala, hippocampus, cerebellum, basal ganglia and corpus callosum [608]. The increase in total brain volume observed in 7q11.23 duplication is frequently accompanied by a dilatation of the lateral ventricles, a condition known as ventriculomegaly, and it is in contrast with the decreased whole brain volume observed in WS patients [582, 583, 610]. Amygdala and hippocampus volumes are smaller compared to WS and 7q.11.23 duplication patients display cerebellar vermis hypoplasia [608] while WS affected individuals have a cerebellum of normal or increased size compared to controls [586]. Furthermore, putamen and caudate volumes are larger in 7q11.23 duplication with respect to both WS and controls [608]. In some patients the volume of the corpus callosum is reduced [608] as in WS and the white matter seems decreased, immature, underdeveloped and shows hyperintensities [608, 610], a sign of white matter lesion, but further studies are necessary to confirm and refine these findings. The cerebral cortex in some cases of 7q11.23 duplication is also characterized by multiple structural defects [606], including increased thickness, simplified gyral pattern and cortical dysplasia of the left temporal lobe [609, 610]. This anatomical defect is particularly interesting as it might explain the language impairments usually associated with this syndrome.

In addition to cellular and neuroanatomical abnormalities, few preliminary functional MRI (fMRI) studies have started to assess the neural activation profiles peculiar to 7q11.23 duplication syndrome highlighting a dysfunction of the limbic system characterized by hypofunctional amygdala, cingulum and orbital frontal cortex [614].

3.6 Analysis of the genotype-phenotype correlation in Williams syndrome and 7q11.23 duplication syndrome

Many researches aimed to unravel the link between 7q11.23 CNV genotypes and clinical phenotypes and the relationship of specific genes contained within the WBSCR to their cellular and neuroanatomical correlates have already been performed [470, 608-610, 615-621]. Study designs vary widely including: 1) functional and neuroanatomical examinations of individuals affected by 7q11.23 CNV, 2) behavioral, neuroanatomical and cellular studies using animal models systems engineered for specific genes contained in the WBSCR, 3) transcriptional analyses and cell morphology characterizations of iPSCs-derived neurons generated from affected individuals or engineered through CRISPR/Cas9 genome editing technology.

The results obtained from functional and neuroanatomical examinations of individuals affected by 7q11.23 CNVs constitute the majority of the current clinical knowledge of these disorders and studies of individuals affected by atypical WBSCR deletion started to provide some insights into the contribution of specific genes to the neurocognitive profile that characterizes WS [615, 616].

Several mouse models harboring either the complete [617], partial [618] or only a single gene [619, 620, 622, 623] deletion and single gene duplication [624] of WBSCR have been generated and characterized. The mouse model with complete deletion of WBSCR recapitulates the main cognitive, behavioral and physical defects observed in individuals

affected by WS [617]. Partial deletion and single gene knock-in and knockout mouse models mimic specific aspects of WS and 7q11DUP and therefore helped to define the contribution of single genes to the clinical phenotype [618-620, 622, 623]. Among the genes contained in WBSCR investigated up to now, it was found that Gtf2i, Gtf2ird1 and Limk1 are related to neurological and cognitive phenotypes, confirming the findings obtained from studies of atypical WS cases [615, 616].

Few studies have also started to explore the effects of 7q11.23 dosage imbalance on the transcriptome of iPSCs generated from affected individuals [470] and iPSCs-derived neurons [470, 621]. Transcriptional profiling of iPSCs generated from WS and 7q11DUP patients revealed deregulation of disease-relevant pathways already in the pluripotent state [470]. Specific subsets of disease-relevant transcriptional deregulations already present in the pluripotent state resulted amplified upon iPSCs differentiation toward different lineages, including neocortical progenitors (NPCs) [470]. Gene ontology (GO) enrichment analysis of Williams syndrome iPSCs-derived NPCs and neocortical neurons identified deregulated disease-relevant biological processes, including ‘cell maturation’, ‘cell adhesion’, ‘axon guidance’ categories, along with downregulation of the canonical Wnt signaling pathway [621]. Furthermore, the comparison between the transcriptional profiles of WS, 7q11DUP and healthy controls iPSCs and iPSCs-derived neuronal cells began to identify symmetrically opposite gene expression signatures that may be ascribed to the gene dosage imbalance along with some candidate genes responsible for specific disease-related traits [470]. These insightful studies have just started to identify the impact of 7q11.23 dosage imbalance on the overall transcriptome of some neocortical cell types. Further researches are necessary to better compare the transcriptional profiles of WS and 7q11DUP, to expand the neocortical cell types examined and to gain insights into the molecular mechanisms underlying these syndromes that are of interest for both basic research and to discover therapeutic targets, as currently therapeutic options are mainly restricted to behavioral interventions and symptomatic treatments.

Part II: Results

Chapter 4: Induced pluripotent stem cells as a model system to investigate human neocortical development and neurodevelopmental disorders

4.1 Derivation of human induced pluripotent stem cell lines

In this study 13 human induced pluripotent stem cells (iPSCs) lines have been generated from a cohort comprising six controls, four individuals affected by typical Williams-Beuren syndrome (WS) and three individuals affected by 7q11.23 duplication syndrome (7q11DUP). Among the healthy subjects chosen as controls, two are unaffected relatives of individuals presenting with typical Williams-Beuren syndrome and the other four are unrelated individuals. The four iPSCs control lines derived from unrelated individuals have been used in the first phase of this study to perform a comprehensive transcriptional, cellular and physiological characterization of the *in vitro* neocortical differentiation protocol. Having established the reliability and quality of our protocol, this differentiation strategy was later adopted as a model system to investigate how 7q11.23 gene-dosage imbalances impact neocortical neurons development by applying it also to iPSCs lines derived from WS patients, 7q11DUP individuals and the additional healthy controls. General information concerning the iPSCs lines generated in this research and their applications are reported in Table 1.

Condition	Sample ID	Sex	iPSCs clones	Brain matched	Application
Williams syndrome (WS)	WS 361	F	iPSC#8	no	disease modeling
	WS 728	M	iPSC#13	no	disease modeling
	WS 1574	M	iPSC#27	yes	disease modeling
	WS 5437	M	iPSC#36	yes	disease modeling
Controls	WS 361-1*	F	iPSC#20	no	disease modeling
	WS 728-1*	F	iPSC#2	no	disease modeling
	HSB 286	M	iPSC#7	yes	protocol validation
			iPSC#20		
	HSB 292	M	iPSC#7	yes	protocol validation
			iPSC#38		
	HSB 311	F	iPSC#1	yes	protocol validation disease modeling
			iPSC#36		
	HSB 314	M	iPSC#43	yes	protocol validation
			iPSC#47		
7q11.23 duplication syndrome (7qDUP)	7q11DUP 242	F	iPSC#11	no	disease modeling
	7q11DUP 314	M	iPSC#6	no	disease modeling
	7q11DUP 809	M	iPSC#12	no	disease modeling

Table 1: Summary table reporting information about the 13 human iPSCs lines included in this study.

The availability of brain tissues matched to the four unrelated control iPSCs lines were particularly valuable as their transcriptional analysis and comparison to iPSCs-derived neocortical neurons offered the possibility to better assess the quality of the neural differentiation protocol to which all iPSCs were subjected.

To induce pluripotency, a non-integrating reprogramming method based on the transfection of episomal plasmid vectors slightly modified from Okita et al. [429, 625] has been chosen. In this protocol reprogramming has been achieved through nucleofection of four episomal plasmids encoding respectively human OCT3/4 and short harpin RNA against p53, human SOX2 and KLF4, human L-MYC and LIN28, EGFP. Transcription factors OCT3/4 and SOX2 are master regulators of the transcriptional network responsible of controlling and maintaining the pluripotent state [398, 399]. These two transcription factors belong to the original reprogramming cocktail employed in the Yamanaka experiment [364], along with the transcription factor KLF4 known for contributing to the long-term maintenance of the pluripotent state [403]. To improve the reprogramming efficiency in human iPSCs generation, the fourth classical transcription factor delivered to induce pluripotency, the oncogene c-Myc, has been substituted by another member of the Myc family of transcription factors, L-Myc [626]. Always having in mind to further enhance the reprogramming efficiency in human cells, the RNA-binding protein LIN28 was added to the reprogramming cocktail [367] and the tumor suppressor protein p53 was silenced through short hairpin RNA targeting [626]. Along with the plasmids containing the set of reprogramming factors of choice, an episomal vector encoding the enhanced green fluorescent protein (EGFP) was cotransfected into somatic cells to monitor its fluorescence signal as a means of reprogramming quality control. More specifically, EGFP fluorescence was at first monitored during the early phases of the reprogramming to evaluate nucleofection efficiency and in the following weeks, before colonies picking, it was monitored to check the progression of pluripotency induction.

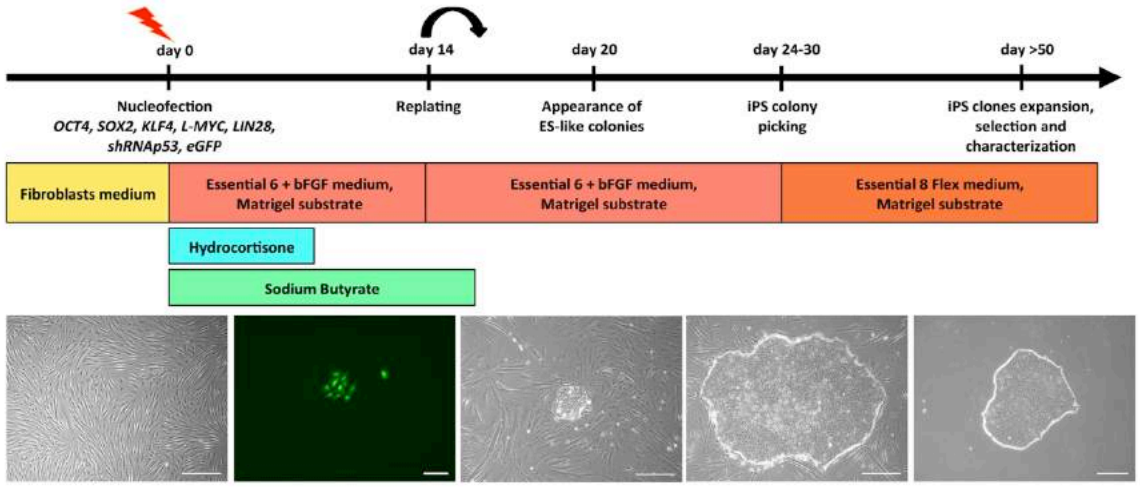


Figure 15: Schematic of the reprogramming protocol. *In this schematic representation are highlighted the main events and phenotypic changes that occur during the episomal vector-based reprogramming protocol along with specifics of the culture conditions and representative pictures of the main reprogramming stages.*

At the moment of colony picking, along with colonies morphology the absence of fluorescence expression was taken as an indicator that the iPSCs colonies selected for further expansion no longer retained the episomal plasmids. Reprogramming experiments were continued only if at least 20% of the nucleofected cells were EGFP-positive 4-6 days post-transfection. The fluorescence signal reached a pick approximately 2 weeks after nucleofection and started to decrease thereafter, indicating the progressive reduced persistence of episonal plasmids in transfected cells and finally their complete absence as expected from this non-integrating gene delivery system. By the time iPSCs colonies were ready to be manually picked, one of the parameters on which clonal selection was based was the non-detectability of the fluorescence signal through eye inspection in the colony of choice.

All reprogramming experiments performed in this study started from primary human fibroblasts cultures that were derived from skin or meninges biopsies through manual dissection and were subsequently cultured in fibroblasts medium (see Materials and Methods). After expansion, fibroblasts were nucleofected with the four episomal plasmids described above. Transfected cells were cultured until confluent for approximately 2 weeks and then replated at a lower density to avoid overconfluence and to enable clonal generation of iPSCs colonies. During these first stages of the reprogramming, transfected cells were cultured on matrigel-coated plates [627] in Essential 6 medium [628], a commercial chemically defined and xeno-free medium produced according to cGMP requirements and optimized to support the reprogramming of somatic cells. To optimize the reprogramming of human cells, Essential 6 medium was supplemented with basic fibroblast growth factor (bFGF) in combination with sodium butyrate to improve the reprogramming efficiency and hydrocortisone to stimulate transfected fibroblasts proliferation. Hydrocortisone was administrated until cell density reached 50% of confluence and sodium butyrate was withdrawn from the medium 2-4 days after replating while bFGF was consistently supplied at the same concentration. Under these culture conditions, iPSCs colonies intermingled with fibroblasts started to appear approximately 20 days after transfection and in the following 5-10 days promising individual iPSCs colonies were selected and manually picked for further clonal expansion. Clone selection was based on the lack of expression of EGFP fluorescence signal in the colony and on the visual evaluation of colony and cell morphology. Tightly packed colonies with round defined borders and composed of cells with a high nucleus-to-cytoplasm ratio were selected for further clonal expansion. Manually picked colonies were seeded on matrigel-coated plates and cultured in Essential 8 Flex medium [628-630] for long-term maintenance and cryopreservation. Also Essential 8 Flex medium is a commercial chemically defined, xeno-free and serum-free medium that differs from the reprogramming medium for the addition of TGFβ1 and for the substitution of key unstable heat-sensitive components with optimized analogs that have an extended activity, thus removing the requirement of changing the medium daily. Routine passaging of iPSCs colonies has been performed using EDTA when cells reached approximately 85% confluence [631]. EDTA-based splitting enables to routinely passage iPSCs colonies as small aggregates. As this passaging method is E-cadherin dependent, it just relies on chemically loosening direct cell-cell and cell-substrate adhesion without the necessity of enzyme neutralization, centrifugation or drug administration and therefore it allows to maximize cell survival. Furthermore, as EDTA treatment harvests iPSCs colonies faster than more differentiated cells, it also enables

differential dissociation and intrinsically sustains the purity of the culture, thus reducing the opportunities for contamination since manual cleaning is rarely required.

4.2 Pluripotency validation of human induced pluripotent stem cell lines

Before starting the quality controls aimed at validating pluripotency, iPSCs clones were cultured until at least passage 5 after manual picking. The rationale behind this choice was to allow enough time for eliminating plasmid persistence from iPSCs clones while starting to assess their pluripotency at low passages to both minimize the potential onset of chromosomal abnormalities due to long-term culture and to optimize time and costs. The first screening step has been inspecting several clones for each human iPSCs line for integration of exogenous reprogramming factors and positively selecting only transgene-free clones for further analysis (Figure 16 B-C). To assess whether episomal plasmids were integrated into the genome of individual iPSCs clones, their DNA was extracted and PCR was performed using a primer pair designed on the *EBNA-1* sequence derived from Epstein-Barr virus that is present in all four nucleofected plasmids (Figure 16A). Then, karyotype analysis was performed to rule out the presence of chromosomal abnormalities and it enabled to confirm that all selected iPSCs clones had a normal karyotype (Figure 17 F,M,T). To confirm that iPSCs clones met the standard criteria to assess pluripotency, the expression of a panel of selected classical nuclear and surface pluripotency markers was checked through immunofluorescence. All iPSCs colonies examined were found positive for the transcription factors critically involved in the maintenance of the pluripotent state OCT4, SOX2, Nanog as well as for the mambrane antigens selectively expressed on the surface of pluripotent stem cells TRA-1-60 and the stage-specific embryonic antigen-4 (SSEA-4) (Figure 17: A-E Williams syndrome, H-L healthy controls, O-S 7q-microduplication syndrome). The combined expression of these markers was a first evidence that all human iPSCs clones generated are indeed pluripotent. Furthermore, to test the differentiation potential of selected iPSCs clones *in vivo*, teratoma formation assay was performed. Briefly, iPSCs cells were injected subcutaneously into immunodeficient SCID mice and after approximately 2 months all mice developed detectable tumors. Histological examination through hematoxylin and eosin staining confirmed that the extracted tumors were teratomas and contained tissues of all the three germ layers (Figure 18 G,N,U). The combination of the results obtained from this panel of assays enabled to determine that all selected iPSCs clones are pluripotent (Figure 17). Further prove of the differentiation potential of selected iPSCs clones into all the three primary germ layers also *in vitro* came from the analysis of the products of embryoid bodies formation assay that was performed on a subset of iPSCs clones generated from control lines (Figure 18). In this standard pluripotency test, iPSCs were cultured in suspension as three-dimensional aggregates for one week in a basic differentiation medium containing fetal bovine serum (FBS) and subsequently replated and grown in the same medium as a monolayer culture for two-three additional weeks prior to fixation and immunofluorescence staining for key markers of the three germ layers.

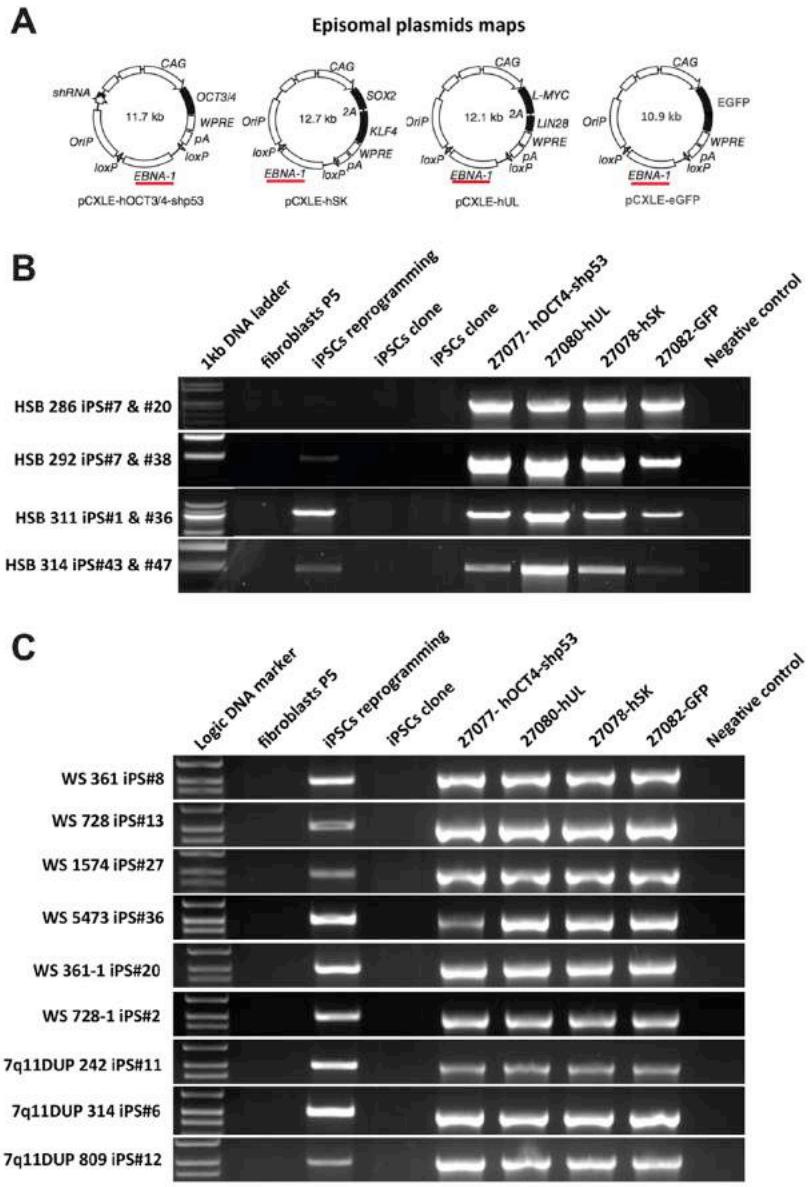


Figure 16: Assessment that the episomal plasmids are non-integrated into the genome of the generated iPSCs lines. *A. Plasmid maps of the four episomal vectors adopted to generate iPSCs. Underlined in red is the EBNA-1 region common to all vectors onto which the primers employed in the PCR were designed on. B-C. PCR results of the integration check demonstrating that the selected human iPSCs clones are transgene-free. Untransfected fibroblasts and water have been used as negative controls while cell samples collected during the early stages of the reprogramming protocol and the four reprogramming episomal plasmids have been employed as positive controls. In B are reported the results for the four human iPSCs lines derived from unrelated individuals that were adopted to extensively characterize the neocortical differentiation protocol while in C are reported the results for the nine human iPSCs lines generated from a cohort including individuals affected by typical Williams-Beuren syndrome (WS 361, WS 728, WS 1574, WS 5437), unaffected healthy relatives (WS*

361-1, WS 728-1) and 7q11.23 duplication syndrome patients (7q11DUP 242, 7q11DUP 314, 7q11DUP 809).

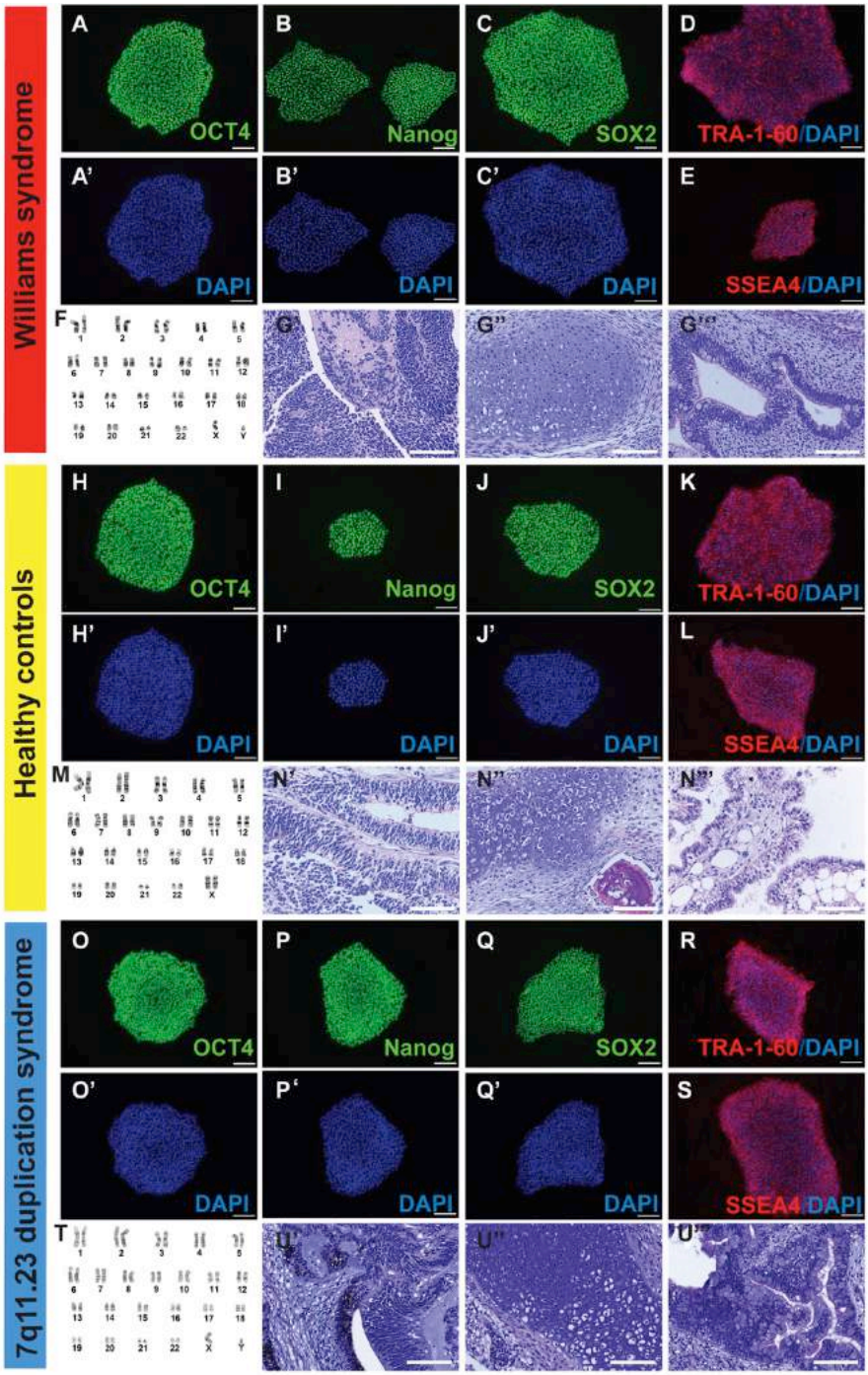


Figure 17: Human iPSCs lines characterization*. Representative images of the pluripotency validation controls performed on human iPSCs lines derived from Williams syndrome affected individuals (top panel: A-G), healthy controls (middle panel: H-N) and 7q-microduplication patients (bottom panel: O-U). A-E, H-L, O-S. Representative*

immunofluorescence stainings for the nuclear pluripotency markers OCT4, Nanog, SOX2 and for the surface markers SSEA-4 and TRA-1-60. Selected clones of all the human iPSCs lines examined presented a strong expression of all these markers. Scale bar: 100 μm. F,M,T. Representative karyotype analysis results indicating a normal karyotype in all the selected iPSCs clones. G,N,U. Representative hematoxylin and eosin stainings of iPSC-derived teratomas demonstating the presence of tissues belonging to the three germ layers, ectoderm (G', N', U'), mesoderm (G'', N'', U''), and endoderm (G''', N''', U''') in all the selected iPSCs clones. Scale bar: 100 μm.

Expression of two key cytoskeletal proteins including the type VI intermediate filament (IF) protein Nestin that is a maker of neural stem cells and neural progenitors and the microtubule protein β–III tubulin that is a marker of early postmitotic neurons confirmed that all human iPSCs lines tested were able to differentiate into ectodermal cells. Mesoderm differentiation potential was assessed monitoring the expression of the ckytoskeletal protein alpha smooth muscle actin (αSMA), a marker of myofibroblasts. Finally, evidence that the generated human iPSCs were able to form also endodermal tissues came from the positivity for both the plasma protein alpha-fetoprotein (AFP) and the transcription factor hepatocyte nuclear factor 3-beta (HNF3β, encoded by the FOXA2 gene) that are expressed in the developing liver.

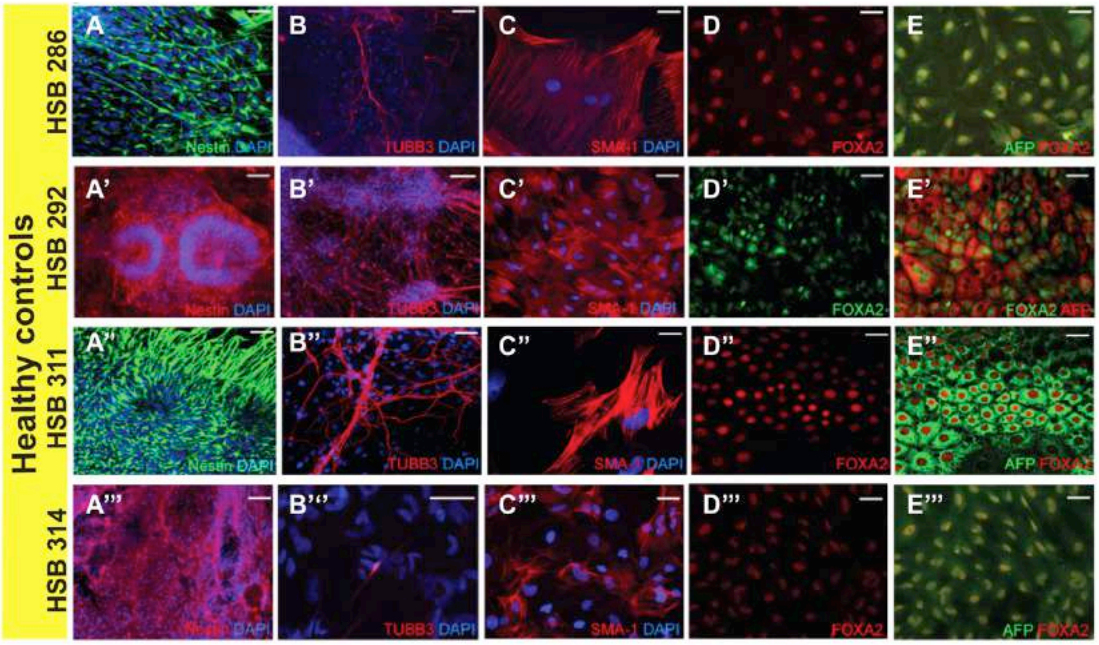


Figure 18: Embryoid bodies formation assay. *Embryoid bodies (EBs) formation assay was applied to a subset of iPSCs clones derived from to assess the differentiation potential of the generated human iPSCs also in vitro. Immunocytochemical stainings of differentiated embryoid bodies with antibodies selectively marking cells derived from the three germ layers. A-B. Nestin and β-III tubulin expression demonstrates differentiation toward ectodermal fates. C. αSMA positivity indicates mesodermal differentiation. D-E. The expression alpha fetoprotein (AFP) and FOXA2 confirmed the presence of endodermal cells. Scale bar: 50 μm.*

Chapter 5. Extensive characterization of a pluripotent stem cells based differentiation protocol for the *in vitro* generation of authentic human neocortical neurons

5.1 Pluripotent stem cells based directed neocortical differentiation protocol

In this study we applied a pluripotent stem cells (PSCs) based directed neocortical differentiation protocol that recapitulates *in vitro* the main *in vivo* neurodevelopmental events. The feeder-free differentiation protocol adopted comprise two phases, a first neural induction phase at the end of which PSCs become telencephalic neuroepithelial cells followed by a prolonged terminal differentiation phase that allows patterned neuroepithelial cells to acquire full maturation. The small molecules based strategy employed integrates several previously published neural induction [176] and neocortical differentiation protocols [375, 376] in a new synthesis to generate the major subtypes of human neocortical progenitors cells and neurons with high efficiency and in less proportions glial cells. All the stages of the differentiation protocol are performed in adherent monolayer culture conditions and using chemically defined media. Neural induction was achieved by adopting a dual SMAD inhibition strategy combined with Wnt signaling inhibition. Dual inhibition of SMAD signaling to induce neural conversion was first employed by Studer's group [176] and since then it has become a milestone small molecules-based strategy to convert PSCs into neuroepithelial cells with high efficiency in just 11 days. The rational behind the dual SMAD inhibition strategy is based on the knowledge that ectodermal cells have an autonomous tendency to acquire neural identities when the transforming growth factor-β (TGF-β) signaling pathway is not activated [165-171, 466, 630]. Therefore, neural conversion was achieved exploiting the synergistic action of two small molecules that act on different branches of the transforming growth factor-β (TGF-β) signaling pathway by competing with TGF-β ligands for binding to their cognate receptors [176]. Specifically, the small molecule LDN-193189 is used as an inhibitor of bone morphogenetic protein (BMP) type I receptors ALK2 and ALK3, that in their active form are responsible for the recruitment and subsequent phosphorylation of SMAD1, SMAD5 and SMAD8 [173, 632, 633]. Inhibition of BMP-mediated TGF-β signaling prevents differentiation toward trophectoderm and ectodermal fates [176]. Moreover, inhibition of TGF-β/Activin/NODAL pathway is achieved through SB-431542 administration. SB-431542 is a small molecule that selectively inhibits bone morphogenetic protein (BMP) type I receptors ALK4, ALK5 and ALK7 [633-635]. In their active form these type I receptors recruit and activate SMAD2 and SMAD3 through phosphorylation [172, 173]. Inhibition of the TGF-β/Activin/NODAL pathway promotes the exit from the pluripotent state by reducing Nanog expression levels [636] and suppresses differentiation toward the mesodermal lineage [176]. To increase the generation efficiency of neuroepithelial cells patterned toward telencephalic fates [287] and further promote the disruption of the pluripotent state [637], iPSCs were also exposed to Wnt signaling inhibition, that was accomplished by administrating XAV939 during the neural induction phase. This chemical compound inhibits Wnt/β-

catenin-mediated transcription by repressing tankyrase 1 [638]. The combined exposure to these three chemical compounds results in a highly efficient generation of telencephalic neuroepithelial cells competent to form neural rosettes, that are two dimensional *in vitro* radial arrangements of columnar cells mimicking *in vitro* the developing neural tube. This highly homogenous cell population is then replated at the end of the neural induction phase and grown in a neuronal medium only supplemented with brain derived neurotrophic factor (BDNF) until all the major cell types populations of neocortical neurons are generated and have achieved functional maturity. Interestingly, the differentiation protocol mimics *in vitro* the sequential generation of human neocortical pyramidal neurons, with deep layer neurons being generated first followed by the appearance and maturation of upper layer neurons. Therefore, at the end of differentiation protocol after 80 days of *in vitro* differentiation, both mature deep layer and upper layer pyramidal neurons are produced accompanied by inhibitory interneurons and minor populations of glial cells. This protracted time of *in vitro* differentiation closely recapitulates the timespan of human cortical neurogenesis that is thought to last for approximately 100 days [115].

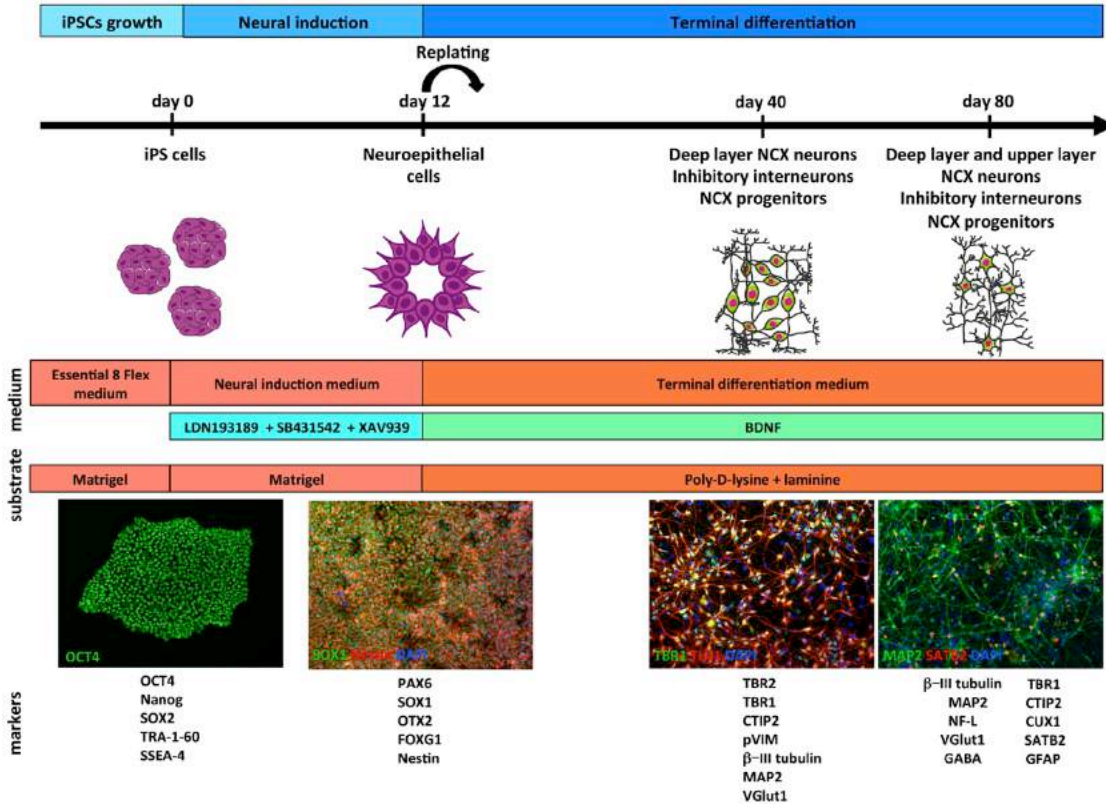


Figure 19: Schematic of the neocortical differentiation protocol adopted in this study. *The main cell types emerging during in vitro neurogenesis are represented. The media and substrates used in each differentiation phase are reported below along with representative pictures and specifics of the informative cell fate and cell-stage specific markers monitored through immunocytochemistry at each time point to confirm the generation of the expected cell types. The differentiation protocol applied in this study integrates several previously published neocortical differentiation protocols [176, 375, 376] in an optimized synthesis.*

This differentiation paradigm has been applied in the first part of this study to four different human induced pluripotent stem cells (iPSCs) lines to assess its reproducibility and ability to generate authentic human neocortical progenitors cells and neurons *in vitro* with high efficiency. In the second phase of this study, this differentiation strategy has been applied to iPSCs lines generated from patients affected by either deletion or duplication of the Williams-Beuren syndrome chromosome region (WBSCR) located on the long arm (q) of chromosome 7 at position 11.23 (7q.11.23 locus) and additional family-matched control lines to model *in vitro* these two human neurodevelopmental disorders.

5.2 Integrative analysis of iPSCs-derived neuronal populations obtained at different *in vitro* neurodevelopmental stages

Given the importance to rely on a solid and highly reproducible differentiation protocol that generates authentic neocortical neurons *in vitro* with high efficiency before applying it as model system to investigate the molecular mechanisms underlying the two syndromes resulting from 7q11.23 CNVs, we decided to perform a comprehensive characterization of the *in vitro* neurodevelopmental process. This extensive analysis has been performed on four control human induced pluripotent stem cells (iPSCs) lines.

Neocortical neurons have been generated from two separate clones of each iPSCs line to better evaluate the reproducibility of the differentiation paradigm and to finely discriminate the degree of variability due to inter-line and intra-line differences. General information about iPSCs lines and clones are reported in Table 2. On these eight human iPSCs clones we performed an integrative analysis of the iPSCs-derived neuronal populations obtained at four different time points of *in vitro* neurogenesis, when specific cell types are expected to appear during the differentiation paradigm:
i) day 0, pluripotent stem cells,
ii) day 10-14, neuroepithelial cells,
iii) day 36-42, neocortical progenitor cells intermingled with deep layer (cortical layers V/VI) pyramidal neurons and inhibitory interneurons,
iv) day 80-83, deep and upper layer (cortical layers II/III) pyramidal neurons and inhibitory interneurons accompanied by neocortical progenitor cells and minor populations of glial cells.

The transcriptional dynamics regulating *in vitro* neocortical differentiation have been investigated by performing RNA-sequencing (RNA-seq) at both population and single-cell level. In order to assess how well our *in vitro* iPSCs-based system was able to reproduce the essential features of *in vivo* neurodevelopment and to determine to which in vivo neurodevelopmental period were related the different *in vitro* differentiation stages, population level transcriptional data obtained from iPSCs-derived neuronal cells collected at different time points have been compared to the transcriptional profiles of donor-matched and unmatched brains included in the reference brain's transcriptome atlas BrainSpan, generated in our laboratory. The presence of specific cell types was also validated by monitoring the expression of a panel of informative cell fate and cell-stage specific markers through immunocytochemistry. Moreover, the functional maturity of iPSCs-derived neocortical neurons obtained at the end of the differentiation protocol

was determined by monitoring their firing properties and spontaneous synaptic activity as well as their passive properties through whole-cell patch clamp recordings. Overall, this integrative panel of analyses enabled an in-depth characterization of the transcriptional, cellular and physiological signatures of the neuronal populations obtained at different *in vitro* differentiation stages, fully validating the reproducibility of the optimized neocortical differentiation protocol of choice and its efficiency and reliability in generating all the major subtypes of neocortical progenitors cells, pyramidal neurons and inhibitory interneurons with high efficiency.

5.2.1 Transcriptional analysis of human induced pluripotent stem cells-derived neuronal cells

RNA-sequencing was performed on a total of 35 cell samples collected at four different time points of *in vitro* neocortical differentiation and on 4 additional matched brain specimens. Single or replicate differentiations were conducted on two distinct clones for each of the four distinct human iPSCs lines and one specific area of each matched brain was selected for sequencing. Details providing general information of the samples are reported in Table 2. Total RNA was extracted using RNeasy Plus Mini kit and 500 ng of each sample were employed to prepare cDNA libraries with Illumina TruSeq Stranded Total RNA LT Sample Prep Kit. cDNA libraries were provided to the Yale Center for Genomic Analysis where paired-end RNA sequencing was performed on Illumina Hi Seq2000 to an average depth of 100M reads per sample. Raw sequencing data were first processed with FastQC to have an initial assessment of data quality and identified low quality reads were filtered out before proceeding to sequence alignment against GRCh38/hg38, using the Spliced Transcripts Alignment to a Reference (STAR) software [639]. RSEQtools, SAMtools and UCSC applications were employed to convert sequence alignment data in BAM format to MRF format and to complete a wide variety of tasks including quantification of expression values, identification of transcriptional active regions, manipulation of gene annotation sets, visualization of the mapped reads and generation of signal tracks. Subsequent data processing with conditional quantile normalization (CQN) and ComBat packages in R enabled to correct for gene length, gene GC content, sequencing site and batch effects.

Sample ID	Sex	Age	iPSCs clone	Number of replicate differentiations	Matched brain
HSB 286	M	20 pcw	iPS#7	2	Parietal CX
			iPS#20	1	
HSB 292	M	13 pcw	iPS#7	1	Parietal CX
			iPS#38	1	
HSB 311	F	8 pcw	iPS#1	1	Dorsal telencephalon
			iPS#36	1	
HSB 314	M	8 pcw	iPS#43	1	Dorsal telencephalon
			iPS#47	1	

Table 2: Summary table reporting the information about the 35 cell samples comprising iPSCs and iPSCs-derived neuronal cells along with the specifications of the respective four matched-brains used for RNA-sequencing.

To examine the contribution of different variables to the global transcriptional dynamics orchestrating *in vitro* neocortical differentiation, we performed principal component analysis (PCA) including data from iPSCs, their neural derivatives and selected informative human prenatal brain tissue specimens comprised in the reference brain's transcriptome atlas BrainSpan, generated in the laboratory (www.brainspan.org). Principal component analysis (PCA) is a data reduction technique that enables to determine the variables having the biggest effect on variability in data. In our case PCA shows that samples cluster mainly on the basis of their *in vitro* differentiation stage (Figure 20).

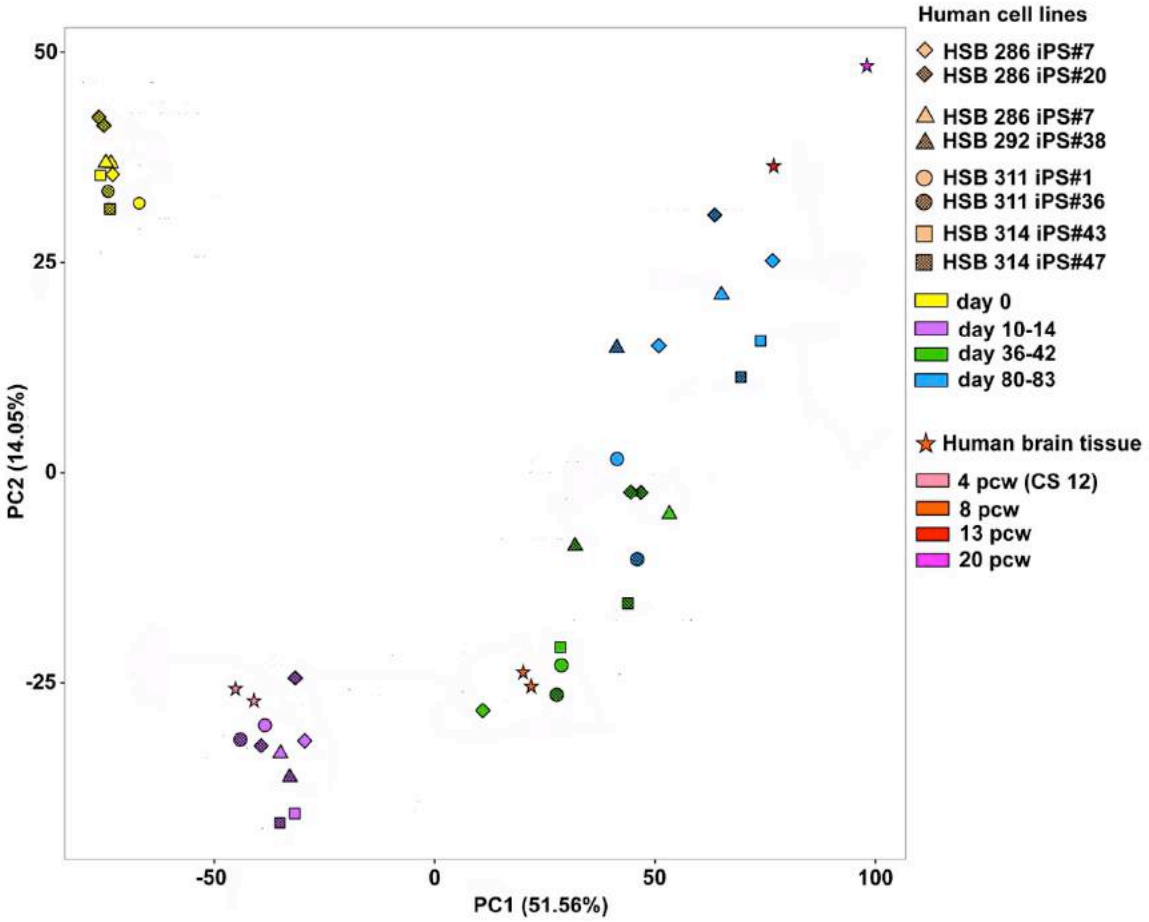


Figure 20: Principal component analysis (PCA) of iPSCs and iPSCs-derived neuronal cells samples. *PCA demonstrates that cell samples cluster on the basis of their in vitro developmental stage and enables to assess to which in vivo neurodevelopmental period are related the different in vitro differentiation stages. The percentage of data variability accounted for by each principal component is reported in parenthesis.*

This finding has been further confirmed by covariate analysis on the principal components (PCs) which indicates that the differentiation stage contributed more to the global differences in gene expression than any other sample-dependent and

sequencing-dependent tested variable (Figure 21). The sample-dependent variables examined included sex, age, differentiation stage while the sequencing-dependent variables investigated comprised flow cell, read length, raw reads, uniq percent, uniq reads, rRNA and tRNA counts.

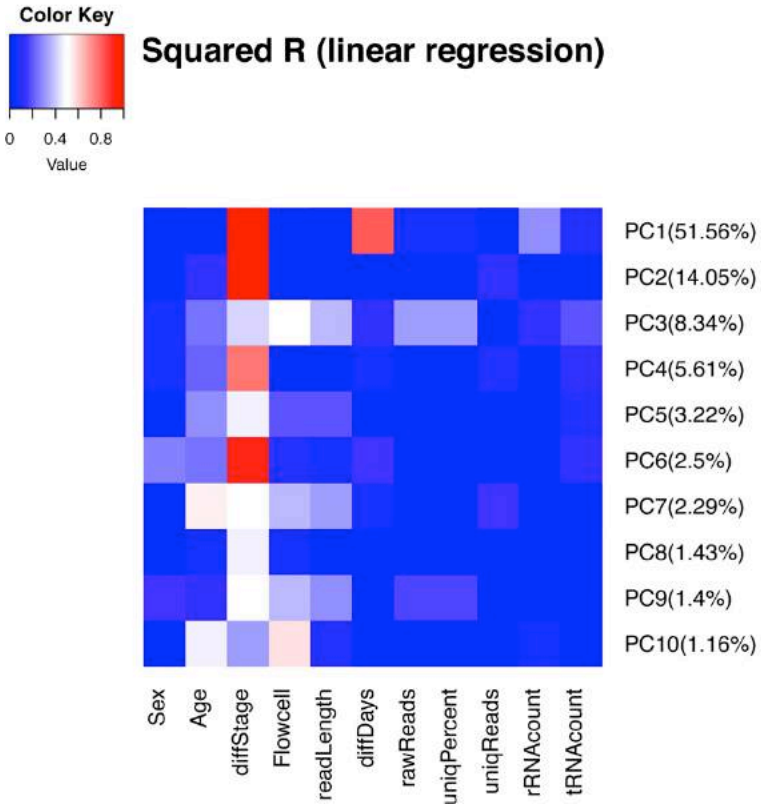


Figure 21: Heatmap of the covariate analysis performed on the principal components (PC) related to sample-dependent and sequencing-dependent variables. *In parenthesis are reported the percentages of data variability accounted for by each principal component.*

PCA results also highlighted that at all the time points analyzed the contributions to data variability due to inter-line and intra-line differences are comparable. This finding indicates that the differences detected between samples can be ascribed to the intrinsic biological variability of how each single iPSCs clone responds to the differentiation paradigm rather than to a diversity between different iPSCs lines. More specifically, considering each *in vitro* differentiation time point separately, samples collected at day 0 and at day 10-14 form more compact clusters while samples collected at day 36-42 and day 80-83 are less clustered together and intermingled in some cases (Figure 20). This observation can be explained by the different differentiations rates at which *in vitro* neurons are maturing. The more advanced along the differentiation timeline is the time point analyzed, the more apparent becomes the intrinsic biological variability of how

distinct iPSCs clones respond to the differentiation protocol, with some iPSCs clones being able to differentiate more prematurely than others.

The inclusion in the analysis of also human prenatal brain tissue specimens enabled to assess how well the transcriptional profiles of our iPSCs-derived neuronal cells reflected those of neurons obtained from human brain samples and to infer to which in vivo neurodevelopmental period were related the different *in vitro* differentiation stages. As clearly shown in the PCA plot and further confirmed by correlation analysis, iPSCs-derived neuronal cells generated at all the different time points clustered closely and had a high correlation with human prenatal brain-derived neurons (Figure 20, Figure 22). Specifically, iPSCs-derived neuronal cells obtained at the end of the neural induction phase show a high correlation with human embryonic forebrain samples (4 pcw, Carnegie stage 12) while iPSCs-derived neurons generated at intermediate stages of the differentiation protocol are highly correlated with human early fetal dorsal telencephalon specimens (8 pcw) (Figure 22). In agreement with the progression of in vitro neurogenesis, iPSCs-derived neocortical neurons obtained at the end of the differentiation paradigm show the highest correlation with human early mid fetal parietal cortex tissues (13 pcw and 16-17 pcw) (Figure 22).

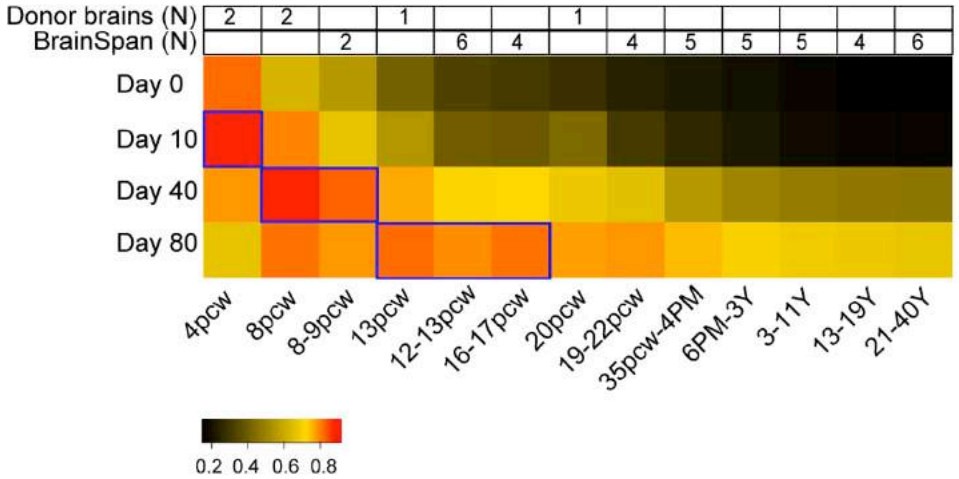


Figure 22: Correlation analysis of the transcriptomes of human iPSCs, iPSCs-derived neuronal cells and human brain tissue samples. *This analysis reveals the correlation between the different in vitro differentiation stages and in vivo neurodevelopmental periods. The numbers of donor-matched and additional BrainSpan brain tissue specimens included in the analysis are reported in the top legend.*

To better investigate the transcriptional changes occurring along *in vitro* neocortical development, we performed differential gene expression analysis. This analysis identified a total number of 28094 differentially expressed genes (DEGs) between the four different *in vitro* differentiation stages investigated. The specific number of DEGs detected performing pairwise comparisons between the four selected time points are summarized in Table 3.

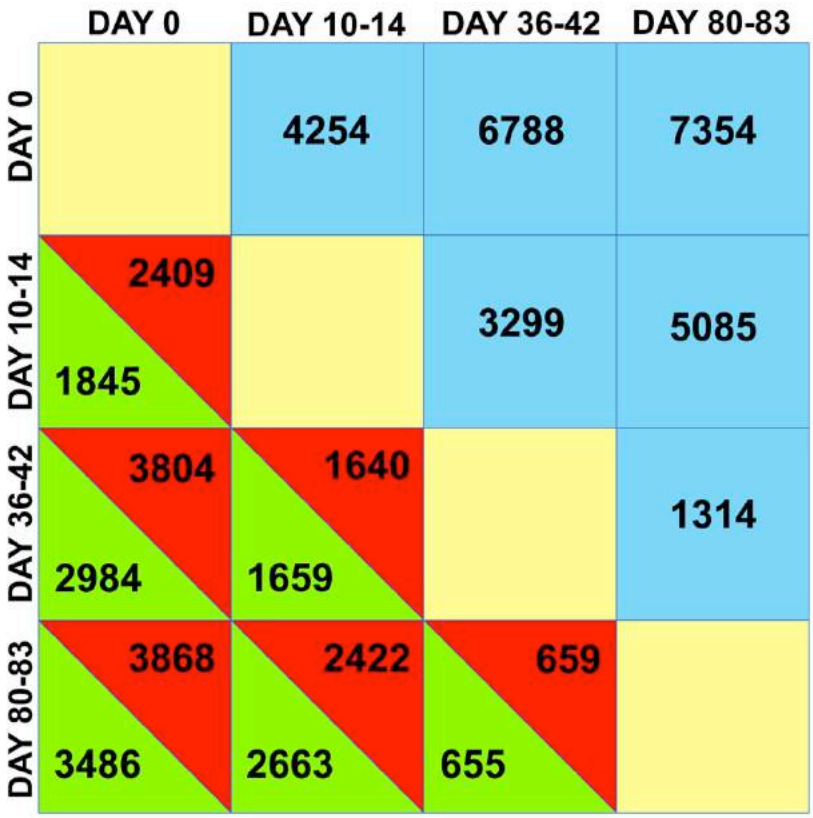


Table 3: Table summarizing the number of differentially expressed genes (DEGs). *DEGs were identified by performing pairwise comparisons between the four different time points of in vitro differentiation investigated and always comparing more differentiated in vitro stages to less differentiated ones. In light blue squares are indicated the total number of DEGs while in red and in green are reported the total numbers of upregulated and downregulated DEGs respectively. Upregulated and downregulated DEGs denote the genes that have increased or decreased expression with respect to the most differentiated in vitro stage of each comparison.*

GO enrichment analysis performed on the identified DEGs enabled to gain the first insights about *in vitro* neocortical differentiation progression and the cell types generated at different time points of *in vitro* neurogenesis. Top-ranking biological process categories for the 2409 genes upregulated at the end of the neural induction phase with respect to the pluripotent state included “nervous system development” (GO:0007399), “neural stem cell population maintenance” (GO:0097150), “Notch signaling pathway” (GO:0007219), “regulation of cell-cell adhesion” (GO:0022407) and “central nervous system development” (GO:0007417), indicating the correct conversion of iPSCs into neuroepithelial stem cells after 10-14 days of *in vitro* differentiation. Progressing along *in vitro* development, the most enriched biological process categories for the 1640 DEGs identified through the comparison of intermediate *in vitro* time points to earlier differentiation stages comprised again terms related to “nervous system development” (GO:0007399) and “central nervous system development” (GO:0007417) which confirm the neural lineage commitment of differentiating cells as well as terms related to “neuron migration” (GO:0001764),“axonogenesis” (GO:0007409), “axon guidance”

(GO:0007411), “chemical synaptic transmission” (GO:0007268), “neurotransmitter secretion” (GO:0007269), “regulation of glutamatergic synaptic transmission” (GO:0051966), “glutamate secretion” (GO:0014047) and “calcium ion-regulated exocytosis of neurotransmitter“ (GO:0048791) indicating the successful generation of functional pyramidal neurons starting from intermediate differentiation time points. Moreover, GO enrichment analysis performed on the 659 upregulated DEGs at the end of the *in vitro* paradigm with respect to intermediate differentiation stages highlighted as enriched categories many of those already identified in the previous comparison between the *in vitro* differentiation temporal windows of 36-42 days and 10-14 days confirming the further maturation of iPSCs-derived neocortical neurons. The acquisition of more functionally developed neuronal phenotypes with respect to previously examined time points was corroborated by the enrichment in many biological process categories related to ion transmembrane transport (GO:0006813, GO:0034220, GO:0035725, GO:1902476, GO:0071805, GO:0010107) and synaptic transmission (GO:0007215, GO:0048791) combined with cellular component terms associated to voltage-gated ion channels (GO:1901387, GO:1905152, GO:0008076, GO:1901387, GO:0032414) and synapses formation (GO:0014069, GO:0042734, GO:0030666, GO:0008021, GO:0045211, GO:0048786). GO enrichment analysis of DEGs detected in the other pairwise comparisons across the different differentiation time points also identified many categories indicating the successful generation of neocortical neurons starting from intermediate differentiation stages that further reached morphological and functional maturity by the end of the *in vitro* protocol.

As many of the DEGs identified through pairwise comparisons are overlapping, we also examined their distribution to better determine how they were specifically related to *in vitro* neurodevelopmental dynamics. We first explored the distribution of DEGs comparing the three representative differentiation time points to the pluripotent state (day 0) (Figure 23 A-A’). Then, we also investigated DEGs distribution between different *in vitro* differentiation time points, always comparing more differentiated *in vitro* stages to less differentiated ones (Figure 23 B-B’). GO enrichment analysis and KEGG pathway analysis were performed also on all the distinct gene lists generated by exploring DEGs distributions across the different time points. Top-ranking GO biological process categories upregulated in the pluripotent state included many terms related to cell cycle and mitosis (GO:0000083, GO:0006260, GO:0000278, GO:0007080, GO:0051726, GO:0000070) indicating the active proliferation of iPSCs. In agreement with the expected generation of neuroepithelial cells at the end of the neural induction phase, the most enriched GO categories detected for genes upregulated in samples collected at 10-14 days of *in vitro* differentiation included terms related to Notch signaling (GO:0007219), cell fate commitment (GO:0045165) and specifically telencephalon development (GO:0021537) combined with terms indicating the negative regulation of alternative cell lineage identities (GO:0042664, GO:2000691, GO:0032331, GO:1903444, GO:1904761). Among GO categories most enriched for DEGs upregulated at intermediate differentiation stages, we detected terms associated to positive regulation of cell migration (GO: 0030335, GO:0001764), cell differentiation and acquisition of neuronal identities (GO:0030154, GO:0030182, GO:0045666, GO:0007268, GO:0007411) in addition to terms related to central nervous system development (GO:0007417, GO:0007420, GO:0009953, GO:0048663) suggesting the successful generation of postmitotic neocortical neurons after 36-42 days of *in vitro* differentiation. Moreover, GO enrichment analysis performed on DEGs upregulated at the end of the differentiation process identified several terms related to synapse

maturation and functioning (GO:0060074, GO:0090126, GO:0090127, GO:0007268, GO:0098916) as well as electrophysiological activity (GO:1904457, GO:2000463, GO:0060079, GO:0042391, GO:0051969) further indicating the presence in culture of functionally mature neurons.

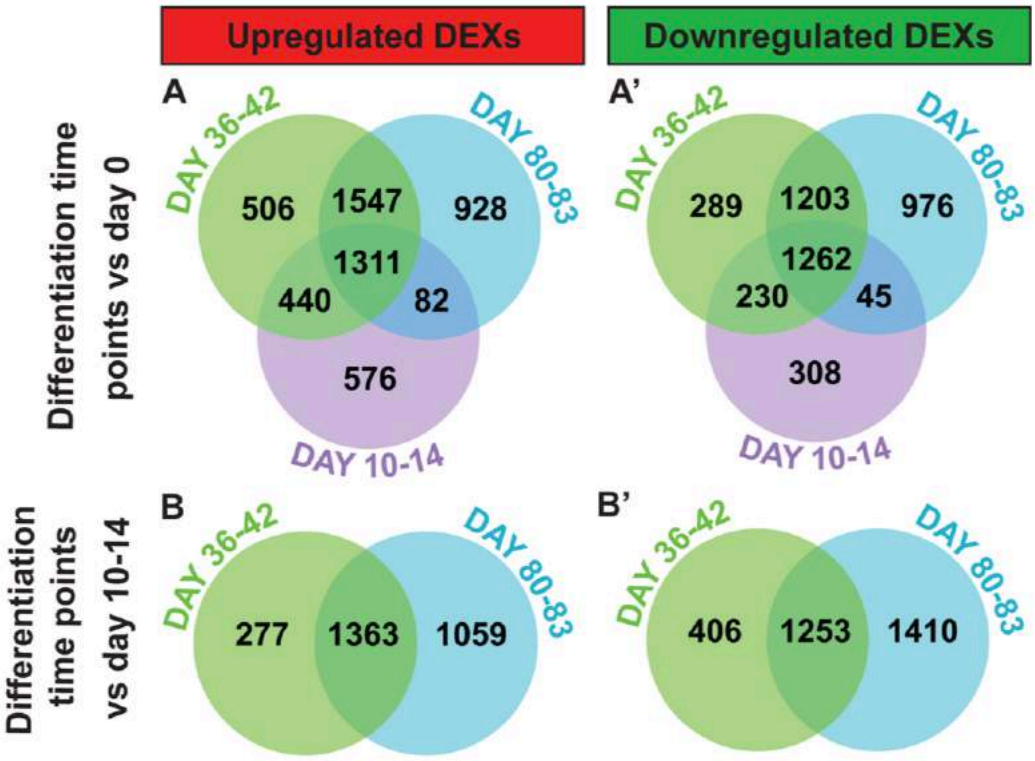


Figure 23: Number and distribution of DEGs among the different in vitro differentiation stages. *A-A'. Distribution of upregulated and downregulated DEGs identified by comparing the three representative differentiation time points to the pluripotent state. B-B'. Distribution of upregulated and downregulated DEGs identified by comparing time points at which neuronal populations have been generated (day 36-42 and day 80-83) to the end of the neural induction phase (day 10-14).*

We next examined the functional organization of the transcriptional changes underlying *in vitro* differentiation by performing weighted gene co-expression network analysis (WGCNA) [640]. WGCNA is a correlation network methodology that enables to describe the relationship between gene transcripts by constructing co-expression networks where each gene expression profile corresponds to a node and the edges represent the degree of correlation between different gene expression profiles calculated through pairwise comparisons. Gene transcripts that present similar expression patterns form clusters of highly interconnected genes known as co-expression modules. The gene transcripts that display the highest connectivity among a module are known as hub genes and the first principal component of each module is the module eigengene whose expression pattern is considered representative of the transcriptional profile of the module. In our study WGCNA identified 46 modules (Figure 24).

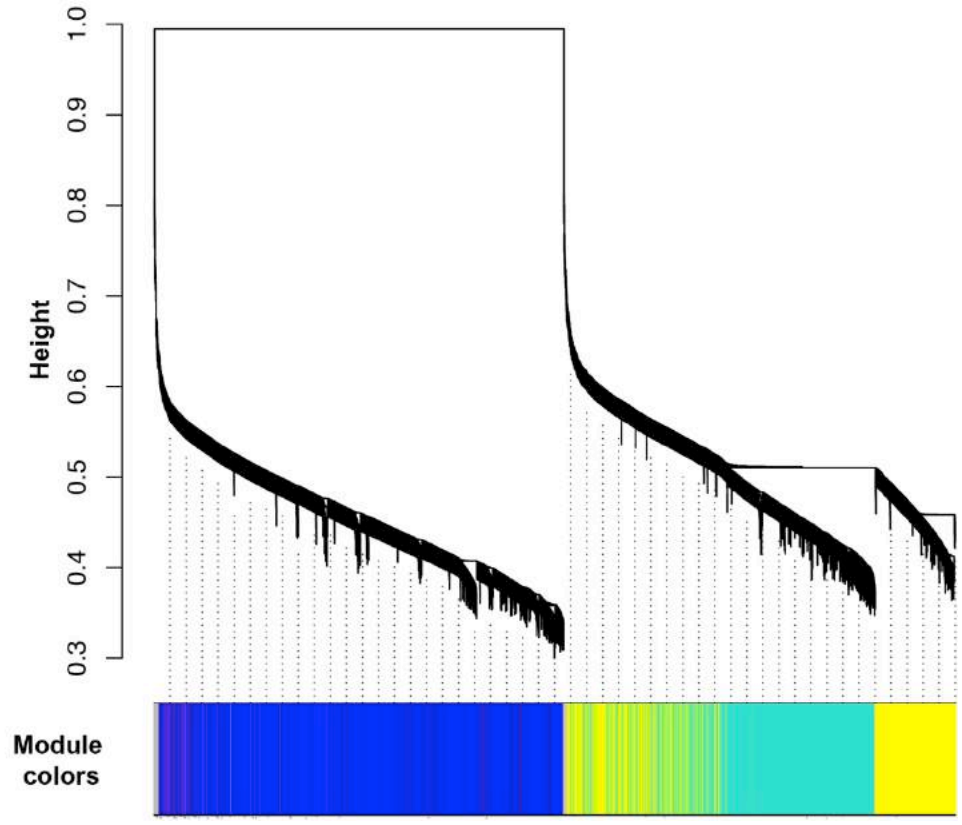


Figure 24: WGCNA cluster dendrogram on all 35 iPSCs and iPSCs-derived neuronal cells samples groups genes in 46 distinct modules. *Module assignments are denoted in the bottom color bar.*

Inspection of modules eigengenes expression enabled to establish the potential relation of distinct modules to specific *in vitro* developmental stages and developmental trajectories. Remarkably, genes comprised in module 4 were highly expressed in the pluripotent state and their expression decreased with the progressive acquisition of neocortical identities (Figure 25). Members of module 7 were characterized by a peak of expression at the end of the neural induction phase (Figure 26) while genes composing module M14 were upregulated during intermediate differentiation stages displaying an enhanced expression in samples collected at day 10-14 and day 36-42 (Figure 27). Instead, module 2 genes displayed an increased expression throughout the differentiation paradigm reaching a peak at the end of the differentiation protocol (Figure 28). Furthermore, genes constituting module 3 were upregulated beginning from day 36-42 of differentiation when deep layer pyramidal neurons and inhibitory interneurons start to be generated *in vitro* (Figure 29) and module 24 genes were selectively upregulated at the end of the differentiation paradigm, when more mature iPSCs-derived neocortical pyramidal neurons belonging to both deep and upper layers and inhibitory interneurons are produced (Figure 30). To gain a better understanding of the functional relevance of these selected modules, we performed also on them GO enrichment analysis and KEGG pathway analysis.

In agreement with the peak of expression of module 4 eigengene in the pluripotent state and its progressively decreased expression at later neocortical differentiation stages, GO enrichment analysis revealed that top-ranking biological process and molecular function categories associated to this module were mainly related to mitosis, cell cycle and cell proliferation (Figure 25 C).

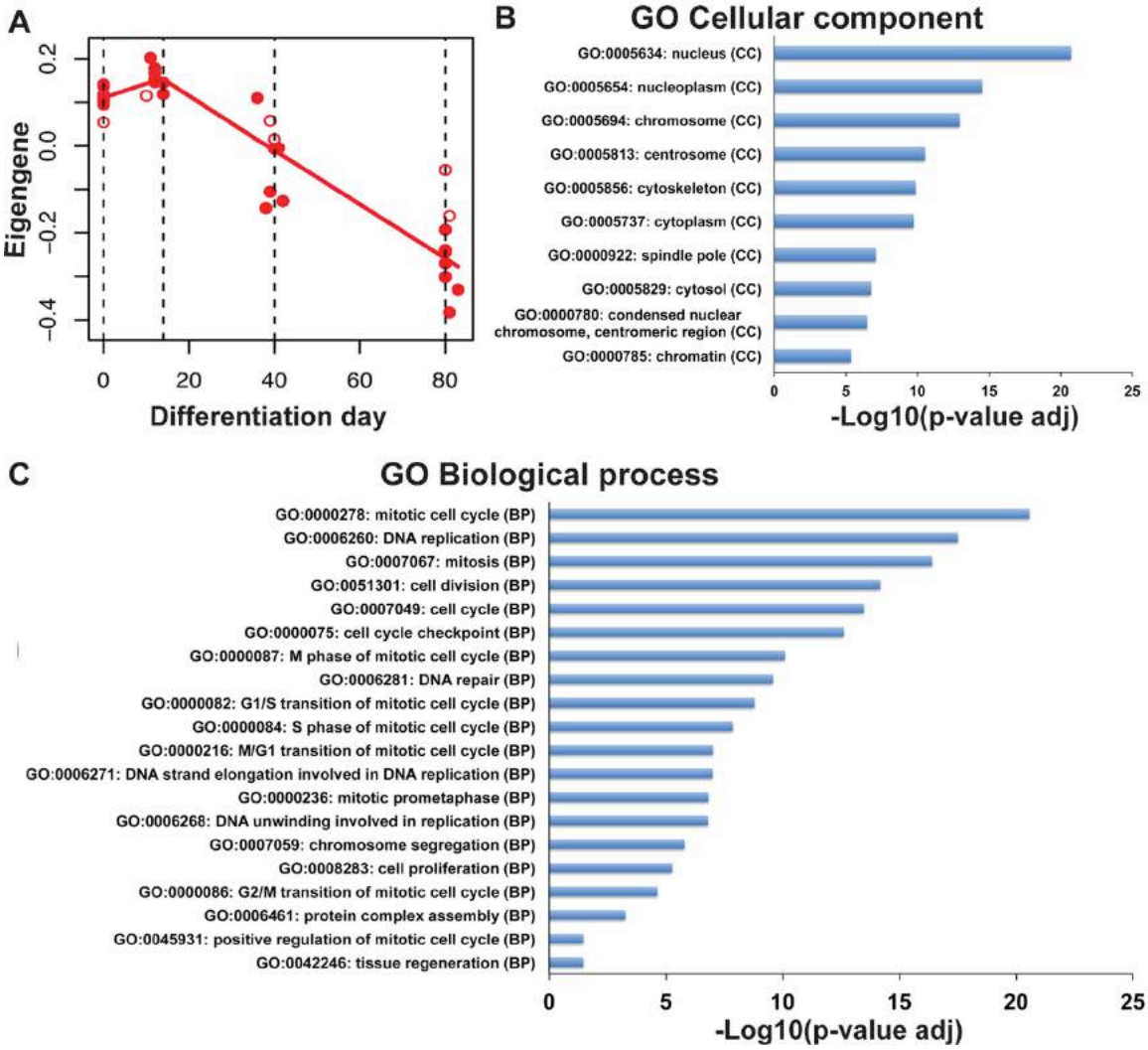


Figure 25: Analysis of module 4 identified through WGCNA. *A. Module 4 eigengene expression at the four in vitro differentiation time points analyzed. Eigengene expression for each sample is reported as dot and the line represents the developmental trajectory based on its mean expression values at each time point. B. Top enrichments for GO cellular component terms. C. Top enrichments for GO biological process categories.*

These findings were also confirmed by the inspection of enriched GO cellular components categories that were mainly related to the nucleus and chromosomes (Figure 25 B). KEGG pathway analysis further corroborated these results detecting pathways as cell cycle (KEGG: 04110), DNA replication (KEGG: 03030), oocyte meiosis (KEGG 04114), progesterone-mediated oocyte maturation (KEGG: 04914) and p53

signaling pathway (KEGG: 04115). Overall, the enrichment in GO categories and pathways related to mitosis, cell cycle and cell proliferation associated to this module reflects the active proliferation that distinguishes the pluripotent state which is increasingly reduced with the progressive acquisition of mature neuronal phenotypes along the *in vitro* differentiation paradigm.

On the other hand, consistent with the peaked expression of its eigengene at the end of the neural induction phase (Figure 26 A), module 7 most enriched GO biological process terms were related to cell maturation, cell differentiation and the acquisition of telencephalic identities (Figure 26 C). Analysis of GO molecular function terms detected as top-ranking several categories related to sequence-specific DNA binding RNA polymerase II transcription factor activity, DNA binding and protein binding suggesting the occurrence of transcriptional modifications involved in cell fate specification. In agreement with this findings, the most enriched GO cellular component categories included terms related to the nucleus, intracellular part and cell membrane as expected given the rapid cell identity transition that neuroepithelial cells are undergoing at this early neocortical differentiation stage (Figure 26 B).

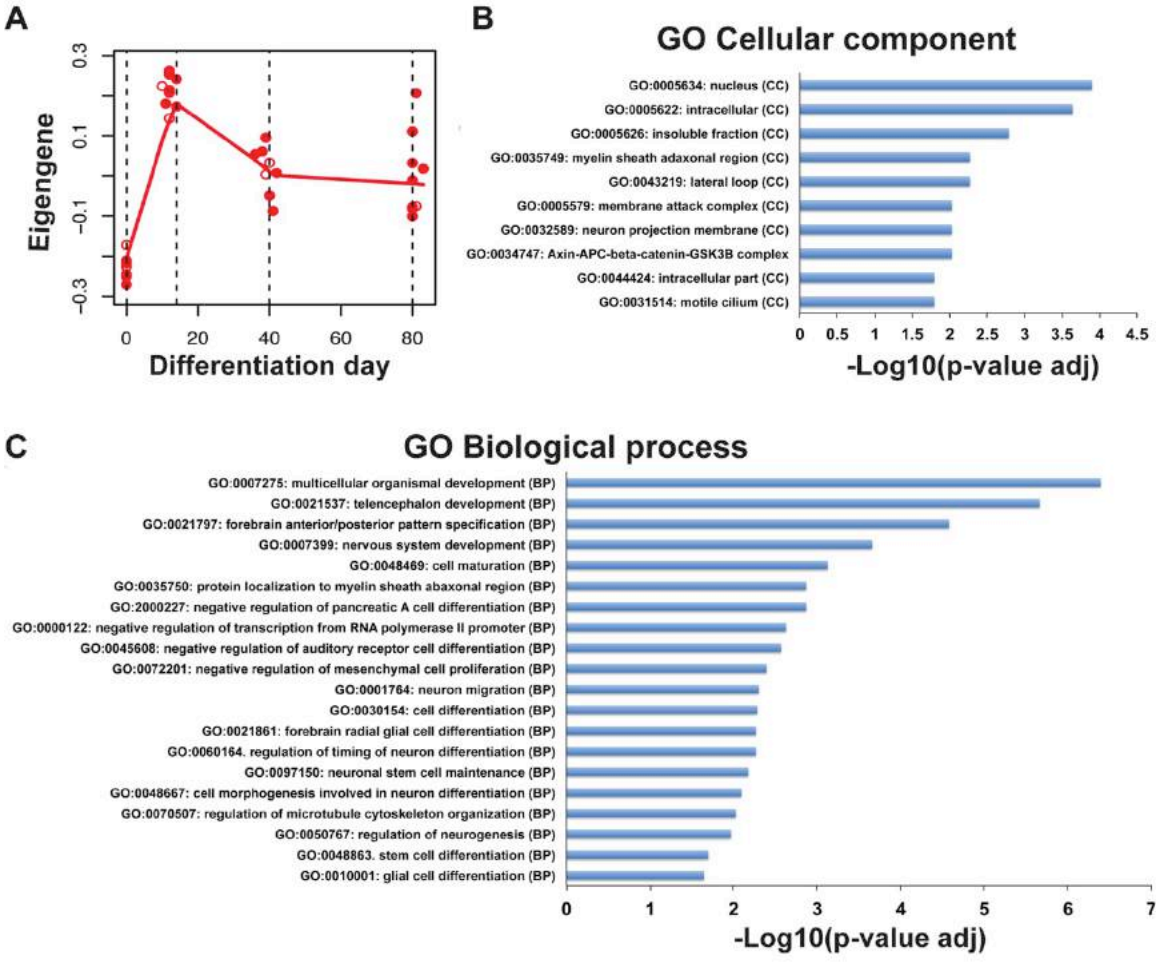


Figure 26: Analysis of module 7 identified through WGCNA. *A. Module 7 eigengene expression at the four in vitro differentiation time points analyzed. Eigengene expression*

for each sample is reported as dot and the line represents the developmental trajectory based on the mean expression values at each time point. B. Top enrichments for GO cellular component terms. C. Top enrichments for GO biological process categories.

Less selectively than transcripts composing module 7, transcripts comprised in module 14 displayed an increased expression at intermediated stages of *in vitro* differentiation as evident from the inspection of the developmental trajectory followed by this module's eigengene (Figure 27 A). Enriched GO biological processes associated to this module comprised categories related to the regulation of neurogenesis, neural plate anterior/posterior regionalization, neural stem cell maintenance, maintenance of apical/basal cell polarity, cell-cell signaling and cell-cell adhesion (Figure 27 C). Moreover, the detection of enriched GO biological process terms related to mitosis and regulation of the cell cycle further denoted the presence of actively proliferating cells committed to the neural lineage (Figure 27 C).

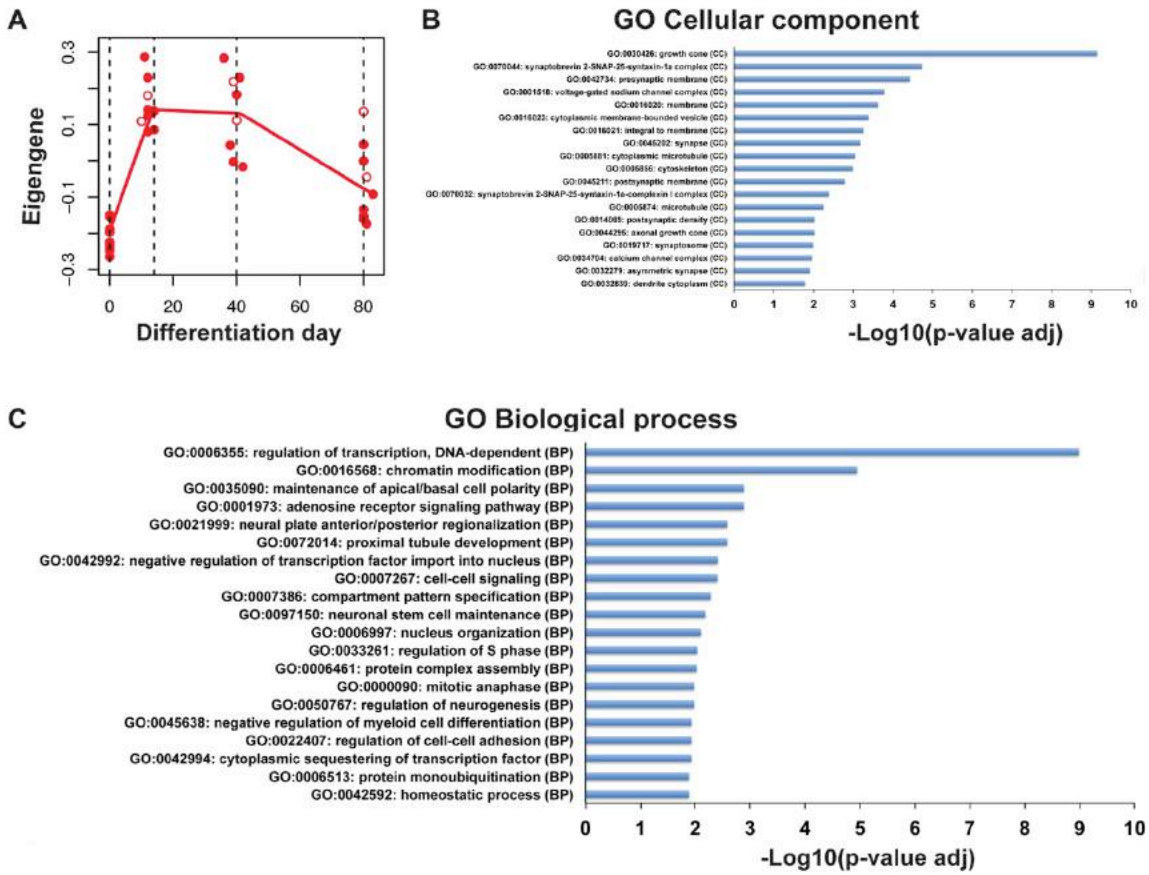


Figure 27: Analysis of module 14 identified through WGCNA. *A. Module 14 eigengene expression at the four in vitro differentiation time points analyzed. Eigengene expression for each sample is reported as dot and the line represents the developmental trajectory based on the mean expression values at each time point. B. Top enrichments for GO cellular component terms. C. Top enrichments for GO biological process categories.*

The enrichment in molecular function terms associated to the regulation of transcription (GO: 0008134, GO: 0001191), chromatin modification (GO: 0003682) and DNA binding (GO: 0003677) further suggests that iPSCs-derived neuronal cells present in culture at these intermediate *in vitro* differentiation stages are actively transitioning to different cellular identities. Enriched GO cellular component terms including nucleus, chromosome and cytoplasm were also in agreement with the other GO analyses findings (Figure 27 B). The presence in culture of proliferating neural stem cells and neocortical progenitors was further corroborated by the enrichment for GO molecular function categories related to Notch binding (GO: 0005112) and the detection of Notch signaling pathway (KEGG: 04330) through KEGG pathway analysis. In addition to neocortical progenitor cells, the generation of neocortical neurons at intermediate stages of *in vitro* differentiation was indicated by the enrichment in many cellular component categories related to synaptogenesis (GO:0070044, GO:0042734, GO:0045202, GO:0045211, GO:0070032, GO:0014069, GO:0019717, GO:0032279), voltage gated-channels (GO:0001518, GO:0034704), axonal growth (GO:0030462, GO:0044295) and dendrite development (GO:0032839) (Figure 27 B) and further supported by the identification of axon guidance (KEGG:04360) through KEGG pathway analysis. Overall, these data are in agreement with the expected acquisition of neuroepithelial stem cells identities by the majority of cells present in culture at the end of the neural induction phase and with the generation of neocortical progenitor cells intermingled with deep layer pyramidal neurons and inhibitory interneurons in the temporal window comprised between 36 and 42 days of *in vitro* neurogenesis.

Then, GO and KEGG pathway enrichment analyses were performed on transcripts belonging to modules 2, 3 and 24 to evaluate the transcriptional changes associated to the acquisition of postmitotic neocortical neurons identities, as these modules display respectively an increased expression throughout the differentiation paradigm (Figure 28 A), starting from intermediated stages of *in vitro* differentiation (Figure 29 A) and selectively at the end of the differentiation protocol (Figure 30 A) as indicated by the developmental trajectories followed by their eigengenes.

Module 2 most enriched GO biological process, cellular component and molecular function categories were mainly related to synaptogenesis and physiological maturation and comprised also terms indicating the acquisition of complex neuronal morphologies (Figure 28 B, C). The acquisition of functionally mature neuronal phenotypes was revealed by the enrichment in many GO cellular component terms related to voltage-gated ion channels and synaptic components (Figure 28 B). These findings were corroborated by the considerable number of top-ranking GO molecular function categories associated to ion channel activity (GO: 0005216), voltage-gated ion channel activity (GO: 0005244, GO: 0005242, GO: 0015269), transmitter-gated ion channel activity (GO: 0022824) and glutamate receptor activity (GO:0008066) which suggest that iPSCs-derived neuronal cells progressively become functionally mature and express the necessary ion channels to generate action potentials (APs). In addition, the presence of many top-ranking GO biological process categories associated to synaptogenesis, synaptic transmission and the structural and functional development of the presynaptic and postsynaptic terminals (Figure 28 C) further confirmed that iPSCs-derived neocortical cells increasingly acquire physiological maturity with the progression of *in vitro* differentiation and are able to form functional networks. In minor proportions, GO biological process terms associated to axonogenesis and dendritic spine morphogenesis

(Figure 28 C) accompanied by top-ranking GO cellular component categories related to the axonal growth cone, dendrite cytoplasm and microtubules (Figure 28 B) and GO molecular function terms indicating cytoskeletal rearrangements (GO: 00030276, GO: 0008017, GO: 0051010), suggested neuronal morphological development along our *in vitro* differentiation paradigm.

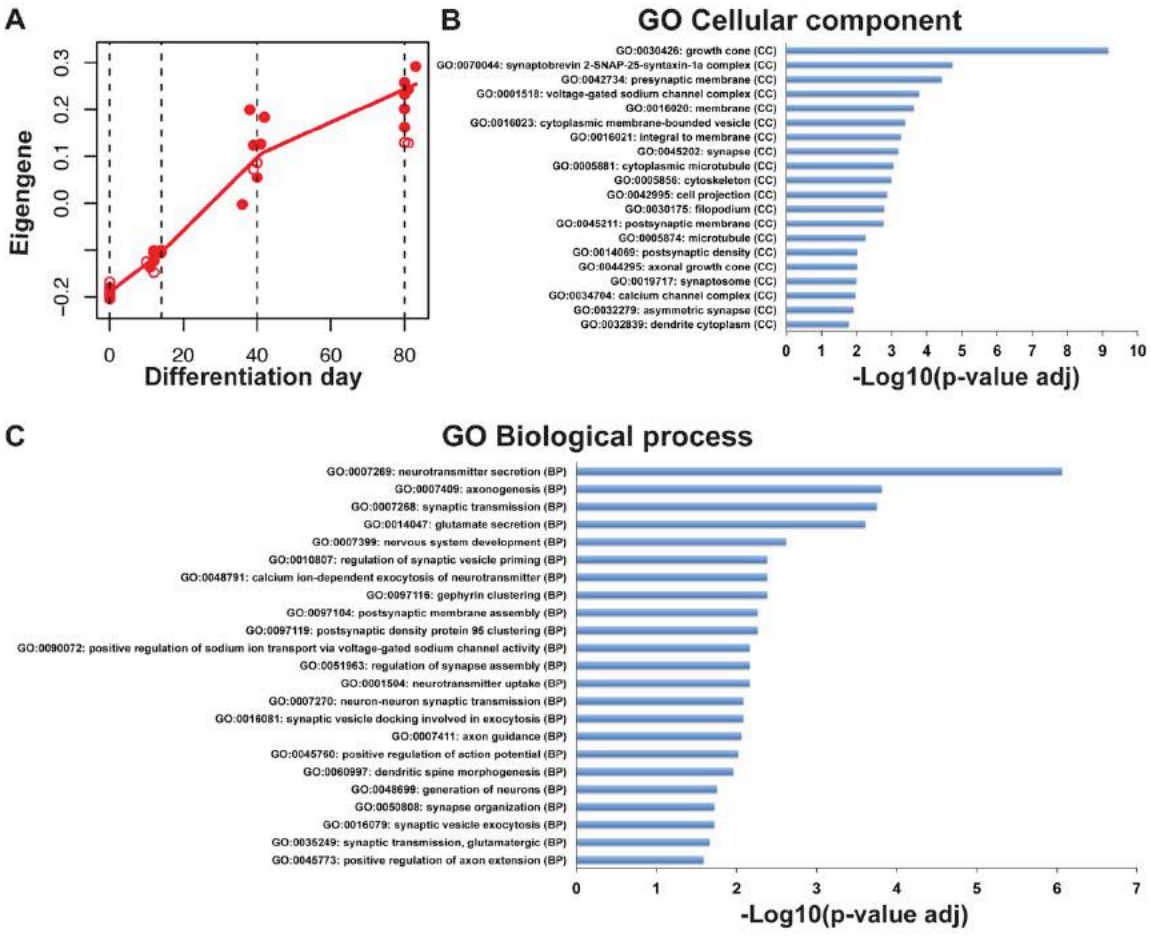


Figure 28: Analysis of module 2 identified through WGCNA. *A. Module 2 eigengene expression at the four in vitro differentiation time points analyzed. Eigengene expression for each sample is reported as dot and the line represents the developmental trajectory based on the mean expression values at each time point. B. Top enrichments for GO cellular component. C. Top enrichments for GO biological process categories.*

Enriched GO categories associated to module 3 were broadly overlapping to top-ranking GO terms previously identified analyzing module 2. Though, complementary to module 2 major focus on synaptogenesis and physiological maturation, the majority of enriched GO biological process, molecular function and cellular component terms linked to module 3 indicated neocortical neurons development and the acquisition of complex neuronal morphologies with less categories related to the development of functional maturity (Figure 29 B, C).

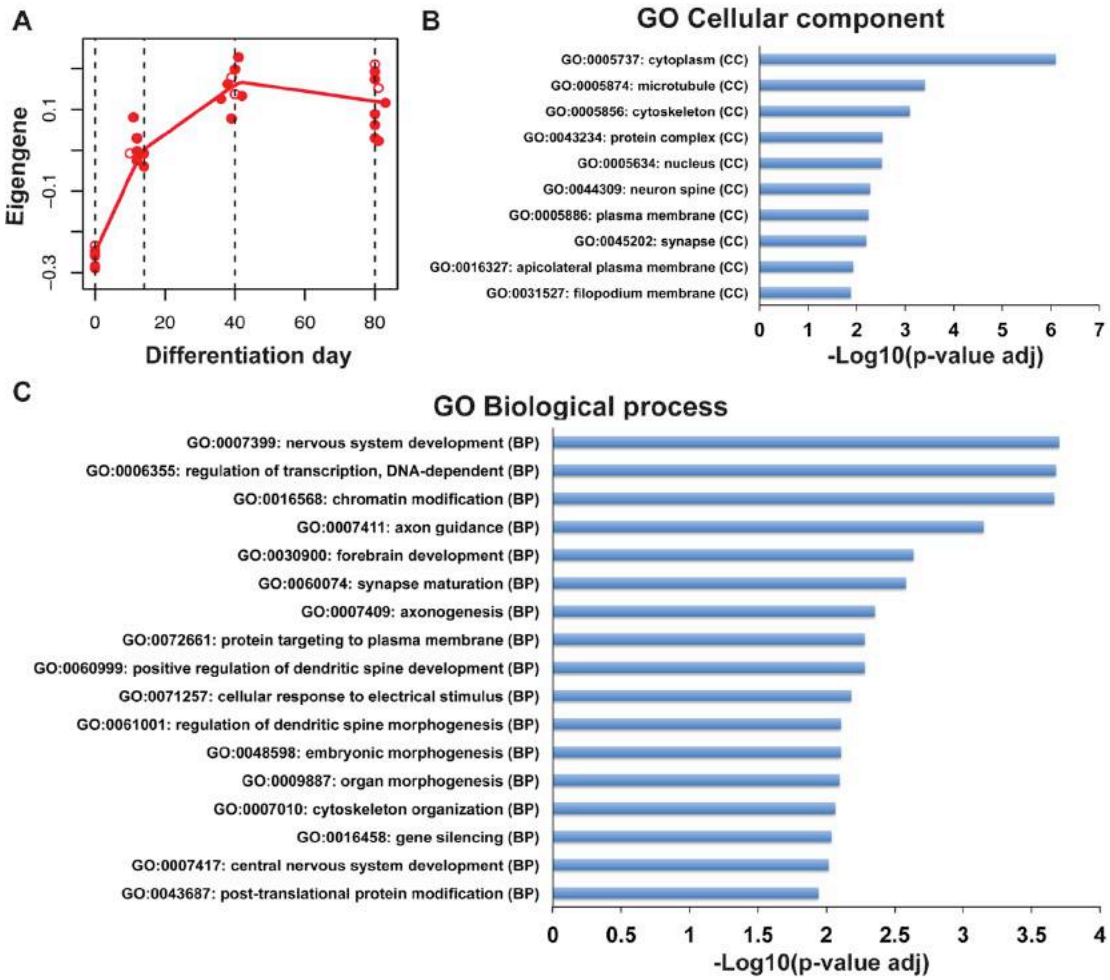


Figure 29: Analysis of module 3 identified through WGCNA. *A. Module 3 eigengene expression at the four in vitro differentiation time points analyzed. Eigengene expression for each sample is reported as dot and the line represents the developmental trajectory based on the mean expression values at each time point. B. Top enrichments for GO cellular component terms. C. Top enrichments for GO biological process categories.*

In agreement with its characteristic developmental trajectory that reaches a plateau starting from intermediate stages of *in vitro* neurogenesis (Figure 29 A), many top-ranking GO biological process terms were related to neocortical neurons development as “embryonic morphogenesis”, “nervous system”, “central nervous system” and “forebrain development” (Figure 29 C). Moreover, enrichment in many GO biological process categories related to axonogenesis, cytoskeleton organization and dendritic spine development and morphogenesis, constitute evidences of the generation of neocortical neurons with distinct neuronal morphologies (Figure 29 C) starting from intermediate differentiation stages. These results are further supported by the enrichment in GO cellular component terms as “neurons spine”, “microtubule” and “cytoskeleton” (Figure 29 B). Interestingly, GO molecular function categories associated to this module are mainly related to histone and chromatin modifications and

transcription factor binding further suggesting cell fate specification in the terminal differentiation phase.

In agreement with the peaked expression of its eigengene at the end of the differentiation paradigm (Figure 30 A), also module 24 most enriched GO biological process, cellular component and molecular function categories comprised several terms related to neuronal development, acquisition of complex and mature neuronal morphologies, synaptogenesis and voltage-gated ion channels expression and function (Figure 30 B, C).

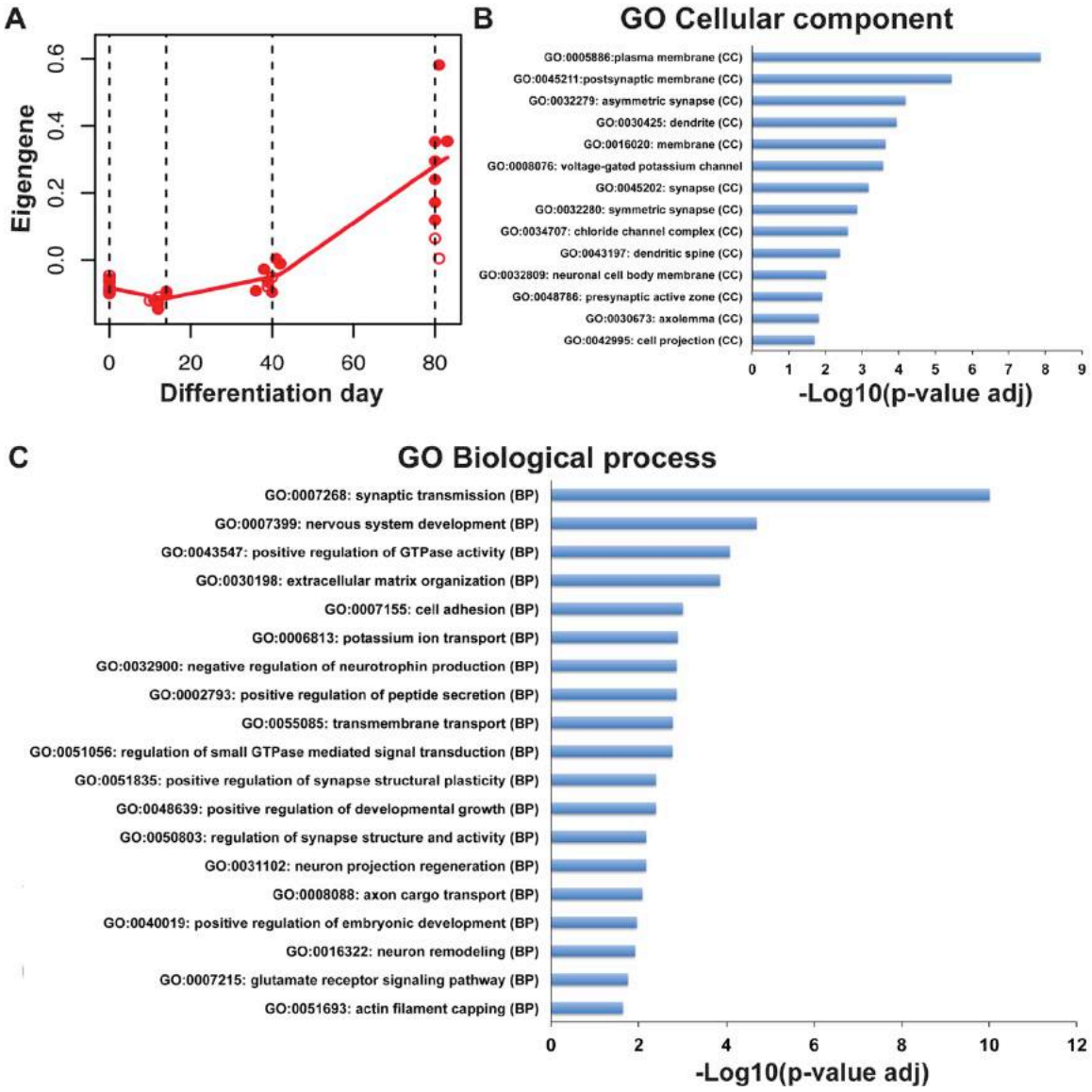


Figure 30: Analysis of module 24 identified through WGCNA. *A. Module 24 eigengene expression at the four in vitro differentiation time points analyzed. Eigengene expression for each sample is reported as dot and the line represents the developmental trajectory based on the mean expression values at each time point. B. Top enrichments*

for GO cellular component terms. C. Top enrichments for GO biological process categories.

Top-ranking GO biological process terms associated to this module as “nervous system development”, “neuron projection regeneration”, “neurons remodeling”, “positive regulation of developmental growth” and “embryonic development” (Figure 30 C) broadly indicate that *in vitro* neocortical neurogenesis continues throughout the differentiation paradigm culminating at its end as expected. Other more specific enriched GO cellular component categories as “dendritic spine” and “axolemma” (Figure 30 B) suggest the acquisition of mature postmitotic neuronal morphologies after 80 days of *in vitro* differentiation. Neuronal morphological development is also accompanied by the acquisition of functionally mature neuronal phenotypes as revealed by the enrichment in many GO cellular component terms related to synaptogenesis and voltage-gated ion channels also in this module (Figure 30 B). These findings are corroborated by the identified top-ranking GO molecular function categories comprising ion channel activity (GO: 0005216), voltage-gated ion channel activity (GO: 0005244, GO: 0005242, GO: 0015269) and transmitter-gated ion channel activity (GO: 0022824) which further suggest that iPSCs-derived neocortical neurons obtained at the end of the *in vitro* differentiation paradigm express the necessary ion channels to generate action potentials (APs). In addition, enriched GO molecular function terms broadly related to receptor activity (GO:004872) and more specifically highlighting glutamatergic and GABAergic receptor activity (GO:0004890) combined with several top-ranking GO cellular component categories related to synapse maturation (Figure 30 B) and GO biological processes terms associated to neurotransmitter secretion and the regulation of synapse structure, activity and plasticity (Figure 30 C) suggest that by the end of the differentiation protocol the synaptic terminals of iPSCs-derived neocortical neurons are sufficiently mature to enable neurons communication with each other and consequently the formation functional networks. In agreement with results obtained from GO enrichment analysis, the physiological maturity reached by iPSCs-derived neocortical neurons after 80 days of *in vitro* differentiation was also suggested by the enrichment for pathways as neuroactive ligand-receptor interaction (KEGG:04080), glutamatergic synapse (KEGG:04724) and long-term potentiation (KEGG:04720) identified through KEGG pathway analysis.

5.2.2 Transcriptional analysis of human induced pluripotent stem cells-derived neuronal cells at single cell level

To identify the distinct neuronal populations obtained at different differentiation stages,and further characterize their transcriptional profiles at single cell resolution, we performed single-cell RNA-sequencing on a total number of 1632 human iPSCs-derived neuronal cells. Neuronal cells were collected at the same three different *in vitro* differentiation stages previously investigated also through RNA-sequencing at population level (Table 4). Single cells were sequenced on Illumina HiSeq 2000 in single-read mode performing 100 cycles. Raw data were then processed with FastQC to have an initial assessment of data quality and mapped to the genome. In addition to FastQC, as part of the quality control (QC) criteria we checked the reads aligned to exonic, intronic and intergenic regions and removed cells with less than 30% aligned

reads to coding sequences in the human genome (Figure 31 A). Next, we checked the reads distribution across different chromosomes and discarded from further analysis cells with more than 10% unique reads mapped to mitochondrial chromosome (Figure 31 B). We further checked gene coverage uniformity as it is considered an indicator of RNA degradation and chose those cells with fold change of average reads coverage within gene less than 2. The number of cells that passed our quality controls and were included in subsequent analyses are also reported in Table 4.

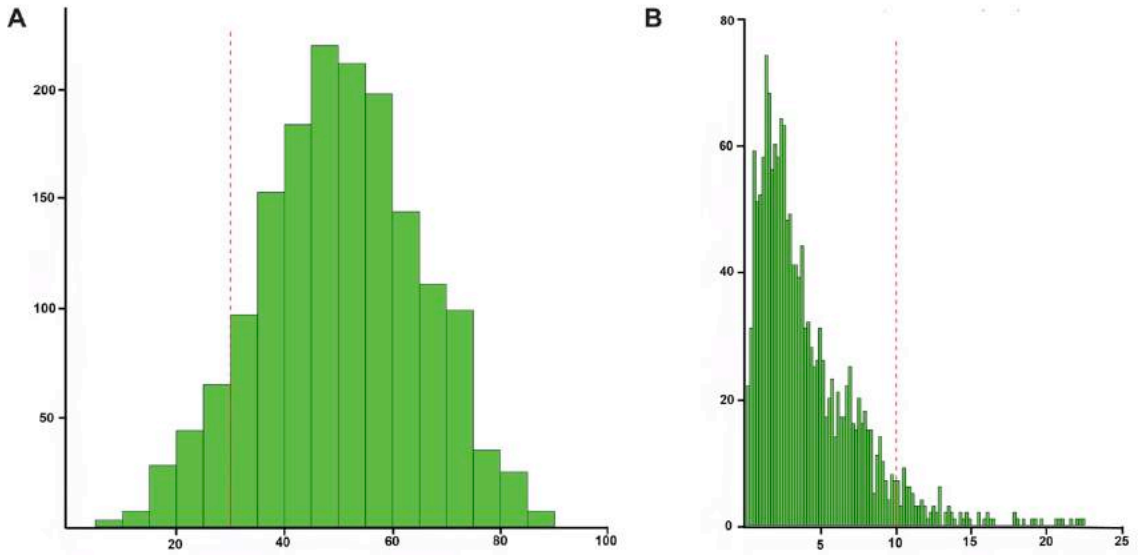


Figure 31. Quality control criteria applied to remove low quality single cells from further analysis. *A. Distribution of the coding sequences (cds) mapping ratio for all iPSCs-derived neuronal cells. Cells with less than 30% of aligned reads mapped to coding sequences (cds) were discarded from further analysis. B. Distribution of the mitochondrial chromosome (ChrM) mapping ratio for all iPSCs-derived neuronal cells. If more than 10% of all aligned reads were mapped to mitochondria chromosome (ChrM), the cell was removed from further analysis. Thresholds are reported as red dotted lines.*

Samples	D10-14	D36-42	D80-D83	Total
Human iPSCs derived neurons before QC	480	480	672	1632
Human iPSCs derived neurons after QC	356	297	429	1082

Table 4. Number of single cells sequenced for human iPSCs-derived neuronal cells generated at different stages of the in vitro neocortical differentiation protocol*.*

To identify the full cohort of iPSCs-derived neuronal cells generated at different *in vitro* differentiation stages, we performed clustering analysis using *pcaReduce* package in R which adopts an agglomerative unsupervised clustering approach that integrates principal components analysis (PCA) and hierarchical clustering [641]. Our clustering analysis identified 21 transcriptionally distinct neuronal populations (Figure 32). The majority of cells displayed transcriptional profiles of a single type, with few evidences of intermediate cell types.

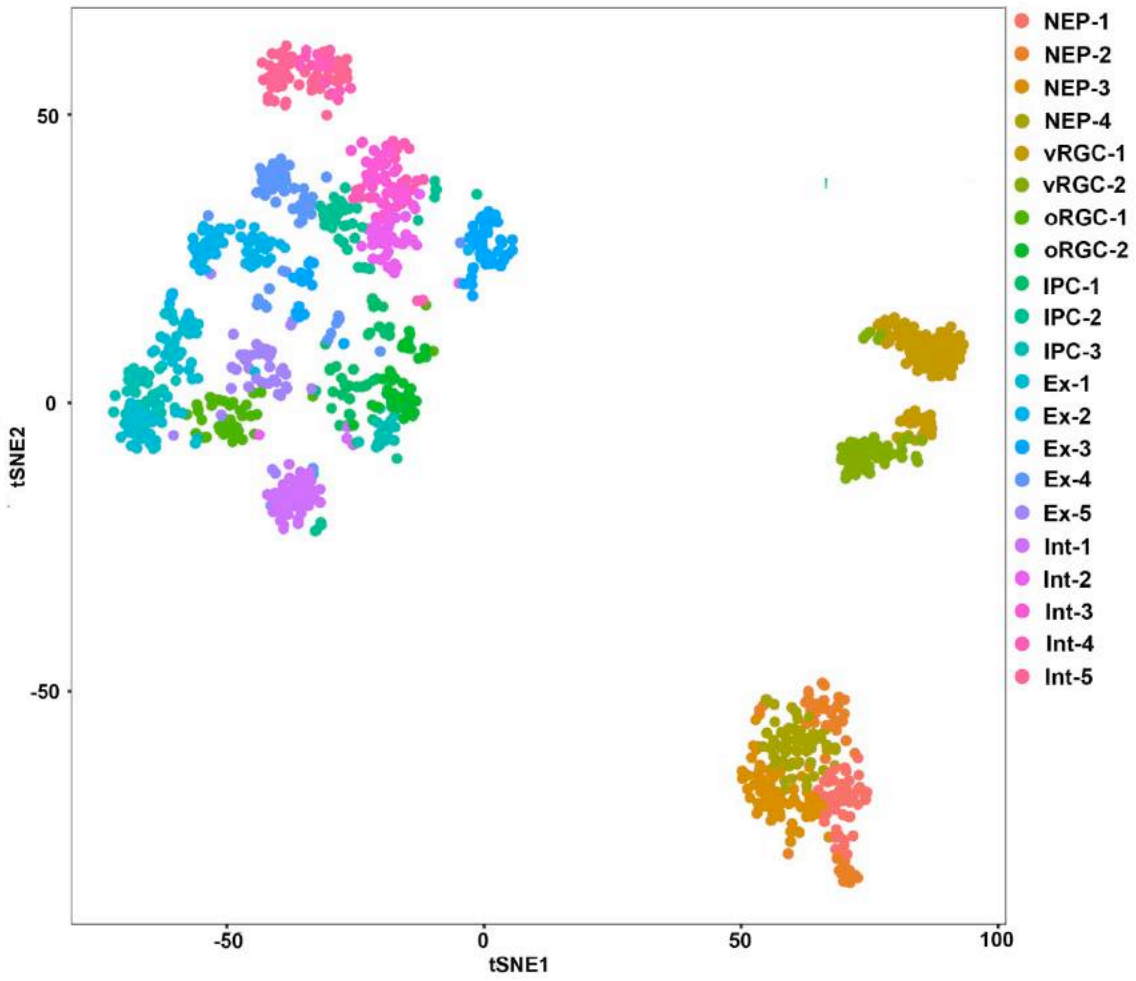


Figure 32: Visualization of single-cell clusters using t-SNE. *Clustering analysis performed on 1082 iPSCs-derived neuronal cells collected at three informative in vitro differentiation stages identified 21 cell clusters. Individual points correspond to single cells and each point is colored according to cell cluster membership. Cell clusters were determined using pcaReduce package in R and t-SNE representation was used only for visualization.*

We assigned our 21 clusters to 6 main cellular identities by inspecting the expression of known markers which resulted in the identification of 4 neuroepithelial progenitors (NEPs) subtypes, 3 apical radial glial cells residing in the ventricular zone (vRGCs) subtypes, 2 outer radial glial cells (oRGCs) subtypes, 3 intermediate progenitor cells (IPCs) subtypes, 5 excitatory pyramidal neurons (Ex) subtypes and 5 inhibitor interneurons (Int) subtypes (Figure 33). As cell classes, NEPs and vRGCs were distinguished by an enhanced expression in known neural stem cells marker genes as NES, PAX6, SOX2, PROM1, GLI3, HES1, LRIG1, PALLD, FABP7, GJA1. This increased expression of neural stem cells genes was accompanied by a pronounced downregulation of genes associated to more mature neocortical progenitors and neuronal phenotypes as the pan-neuronal markers DCX, MAP2, SOX4, SOX11, STMN2 and neuronal classes or neuronal subtypes specific genes as BCL11A, BCL11B,

MEF2C, GAD1,GAD2, SST and NPY (Figure 33). oRGCs were characterized by an increased expression of subtype-specific genes as LIFR, PTPRZ1, SLC1A3, TNC, FAM107A and other neural stem cells markers common also to NEPs and vRGCs as GLI3, PAX6, SOX2 or to IPCs as VIM (Figure 33). IPCs were identified by transcriptional signatures distinguished by the combined expression of transcripts highly expressed also in oRGCs as VIM and shared with other neuronal classes as DCX, SOX4, SOX11, BCL11A and STMN2 (Figure 33). Both excitatory pyramidal neurons and inhibitory interneurons were characterized by an increased expression in pan-neuronal markers as MAP2 accompanied by a decreased expression in neural stem cells markers as NES, PROM1, GLI3, GJA1 and oRGCs markers as LIFR, PTPRZ1 and SLC1A3. Further analyses aimed to assigning precise cellular identities to each specific excitatory and inhibitory neuronal subtype identified through our unsupervised clustering approach are currently ongoing.

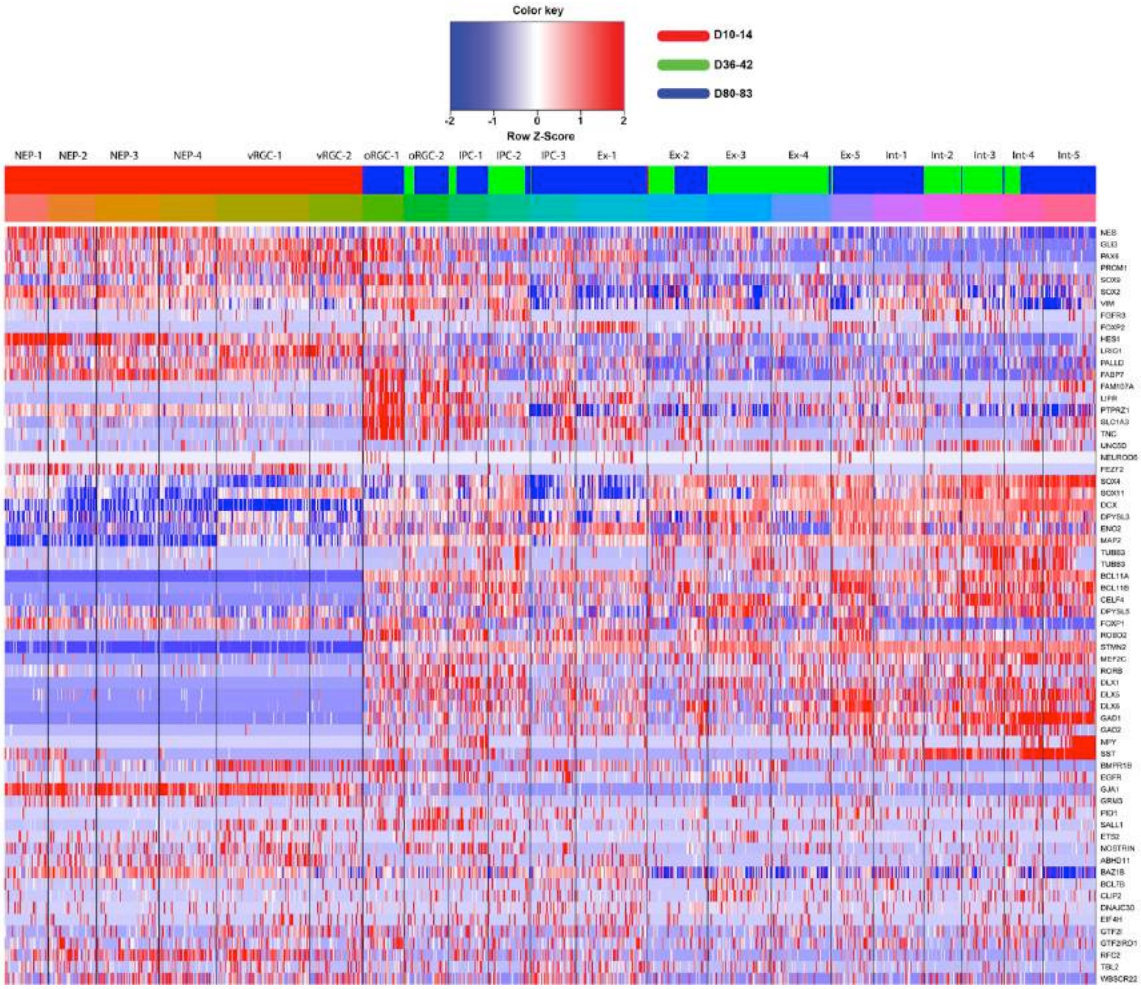


Figure 33: Subtype-specific gene expression profiling at single-cell resolution. *Representative heatmap of some cell-type specific differentially expressed genes. Analysis of the gene expression signatures distinctive of each identified cell cluster allowed to assign our 21 clusters to 6 main cellular identities comprising neuroepithelial progenitors (NEPs), apical radial glial cells residing in the ventricular zone (vRGCs), outer radial glia (oRGCs), intermediate progenitor cells (IPCs), excitatory pyramidal*

neurons (Ex) and inhibitory interneurons (Int). Further analysis aimed to assign even more precise cellular identities to each cell cluster are currently ongoing. In the first top color bar are denoted the differentiation stages at which each cell cluster was detected while in the second top bar are indicated the colors assigned to each cell cluster.

To confirm that our differentiation protocol properly recapitulates *in vitro* human neocortical neurogenesis respecting the sequential generation of neocortical pyramidal neurons, we examined the neuronal populations generated at different differentiation stages.

At the end of the neural induction phase we detected mainly NEPs and vRGCs with the only exception of one excitatory neuron (Figure 34 A).

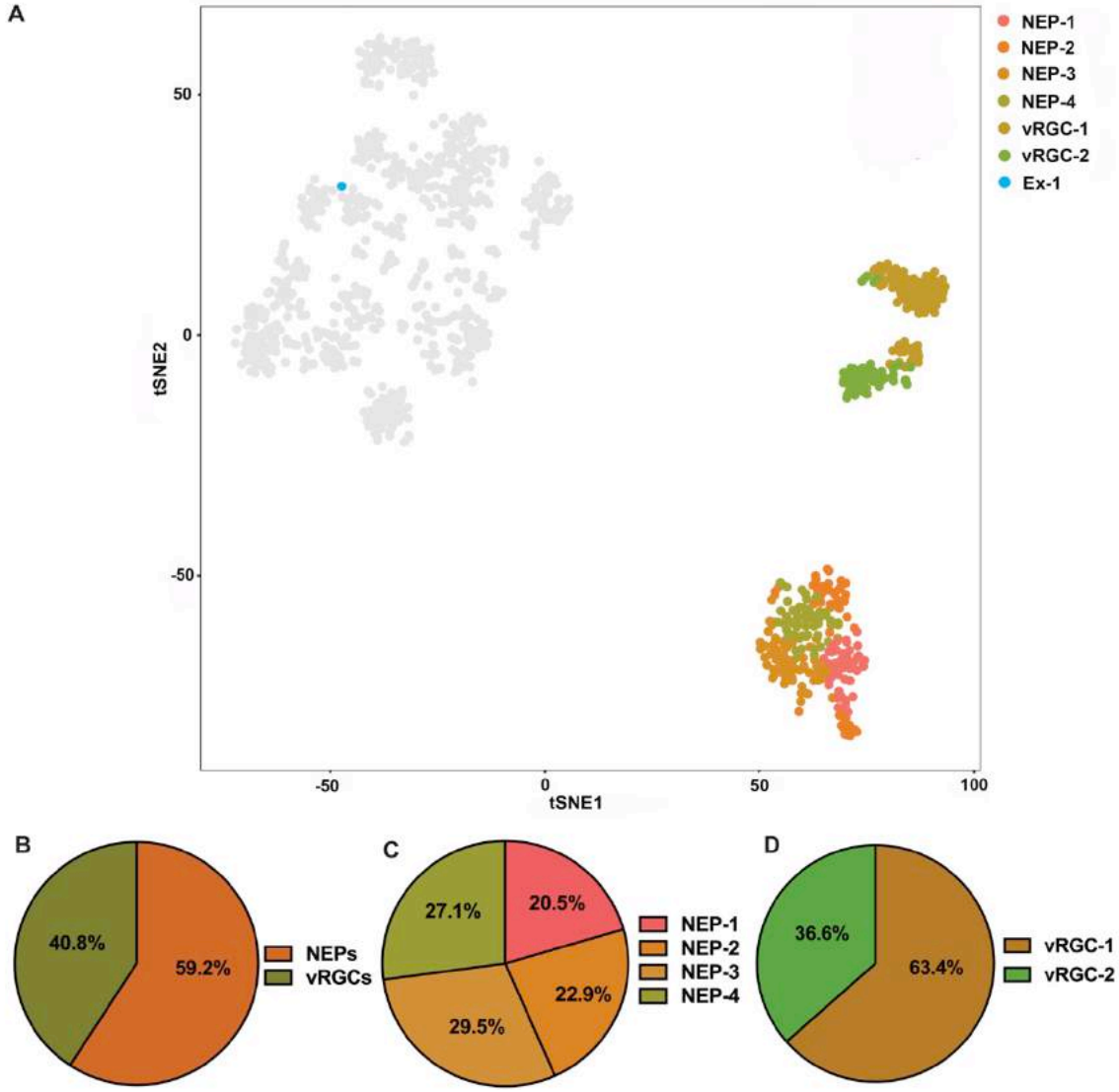


Figure 34: Clustering analysis of iPSCs-derived neuronal cells obtained at the end of the neural induction phase. *A. Clustering analysis performed on iPSCs-derived neuronal cells after 10-14 days of in vitro differentiation identified 7 cell clusters. In the t-*

SNE plot used to visualize the results of the clustering analysis individual points correspond to single cells and each point is colored according to cell cluster membership. B. Percentages of neuroepithelial progenitros (NEPs) and apical radial glial cells residing in the ventricular zone (vRGCs). C. Percentages of the 4 different NEPs subtypes. D. Percentages of the 2 different vRGCs subtypes. The presence of NEPs and vRGCs confirms the correct neural lineage commitment and the onset of in vitro neurogenesis at this differentiation stage.

Both NEPs and vRGCs were considerably present in culture (Figure 34 B) and formed two well-separated major clusters. Both cell types were further subdivided into cell subtypes that were also nearly equally represented (Figure 34 C, D). These findings indicated the correct neural lineage commitment in agreement with transcriptional data obtained from RNA-sequencing at population level. The presence of vRGCs further proves that *in vitro* neurogenesis has already started at this differentiation stage. Moreover, the identification of well-defined cell clusters assigned to two major cell identities indicates also the reproducibility of the initial phase of our differentiation strategy as neuronal cells obtained from all the 9 neocortical differentiations performed were combined in the clustering analysis.

Analysis of the cell populations present in culture at intermediate differentiation stages revealed that mainly excitatory pyramidal neurons and inhibitory interneurons were generated starting from intermediate differentiation stages (Figure 35 A, B). Specifically, 4 distinct subtypes of excitatory pyramidal neurons and 3 different subtypes of neocortical inhibitory interneurons were identified (Figure 35 A). Ex-3 and Ex-4 were the predominant excitatory neuronal subtypes generated (Figure 36 D) while Int-2 and Int-3 were the prevailing neocortical inhibitory interneurons derived (Figure 35 E), as evident from the inspection of the percentages of each cell subtype relative to the total number of excitatory and inhibitory neurons respectively. In addition to neuronal populations, we detected also the presence of later-born neocortical progenitor cells (NPCs) populating in vivo the subventricular zone, including 3 distinct intermediate progenitor cells (IPCs) subtypes and 1 outer radial glial cells (oRGCs) cluster (Figure 35 A). IPC-2 was the main neocortical progenitor population present in culture at this differentiation stage (Figure 35 C). Furthermore, it is worth noting the complete absence in our single cell dataset of NEPs and vRGCs starting from intermediate differentiation stages. These findings combined confirmed the correct differentiation progression and the successful generation of neocortical neurons after approximately 40 days of *in vitro* differentiation accompanied by later-born NPCs.

Analysis of the iPSCs-derived neuronal populations obtained at the end of the differentiation protocol confirmed the presence in culture of excitatory pyramidal neurons, inhibitory interneurons and late-born neocortical progenitors also at this time point (Figure 36 A, B). A finer inspection of the cell subtypes belonging to these three major classes indicated that additional excitatory pyramidal neurons (Ex-1) and inhibitory interneurons subtypes (Int-1, Int-5) were generated after protracted *in vitro* differentiation with respect to earlier time points. Interestingly, these neuronal subtypes were the most prominent neuronal subtypes derived at the end of our differentiation paradigm in addition to subtype 5 excitatory pyramidal neurons that were generated also at intermediate differentiation stages but with low efficiency (Figure 36 C-D). The number of Ex-2 and Int-4 populations remained similar at both intermediate and final differentiation

stages (Figure 36 C, D; Figure 36 C, D). Consistently, at this time point the proportions of previously detected neuronal subtypes, including inhibitory interneurons subtypes Int-2, Int-3 and excitatory pyramidal neurons Ex-4 were markedly reduced compared to previous differentiation stages, while excitatory pyramidal neurons Ex-3 were completely absent (Figure 36 C, D). Also at this differentiation stage, neocortical neurons were intermingled with all the later-born neocortical progenitors detected at intermediate differentiation stages and with an additional oRGCs subtype (oRGC-1). As for neuronal subtypes, also the proportions of neocortical progenitor cell subtypes detected at intermediate and final differentiation stages were different. IPC-1 and IPC-3 were prevailing at the end of the differentiation paradigm while IPC-2 that were predominant at intermediate differentiation stages resulted considerably decreased (Figure 35 E, Figure 36 E). The presence in culture of additional excitatory and inhibitory neocortical neurons subtypes and previously detected neuronal classes still combined with late-born neocortical neurons progenitors further proves the generation of a more complete set of neocortical neuronal populations by the end of our differentiation paradigm (Figure 36).

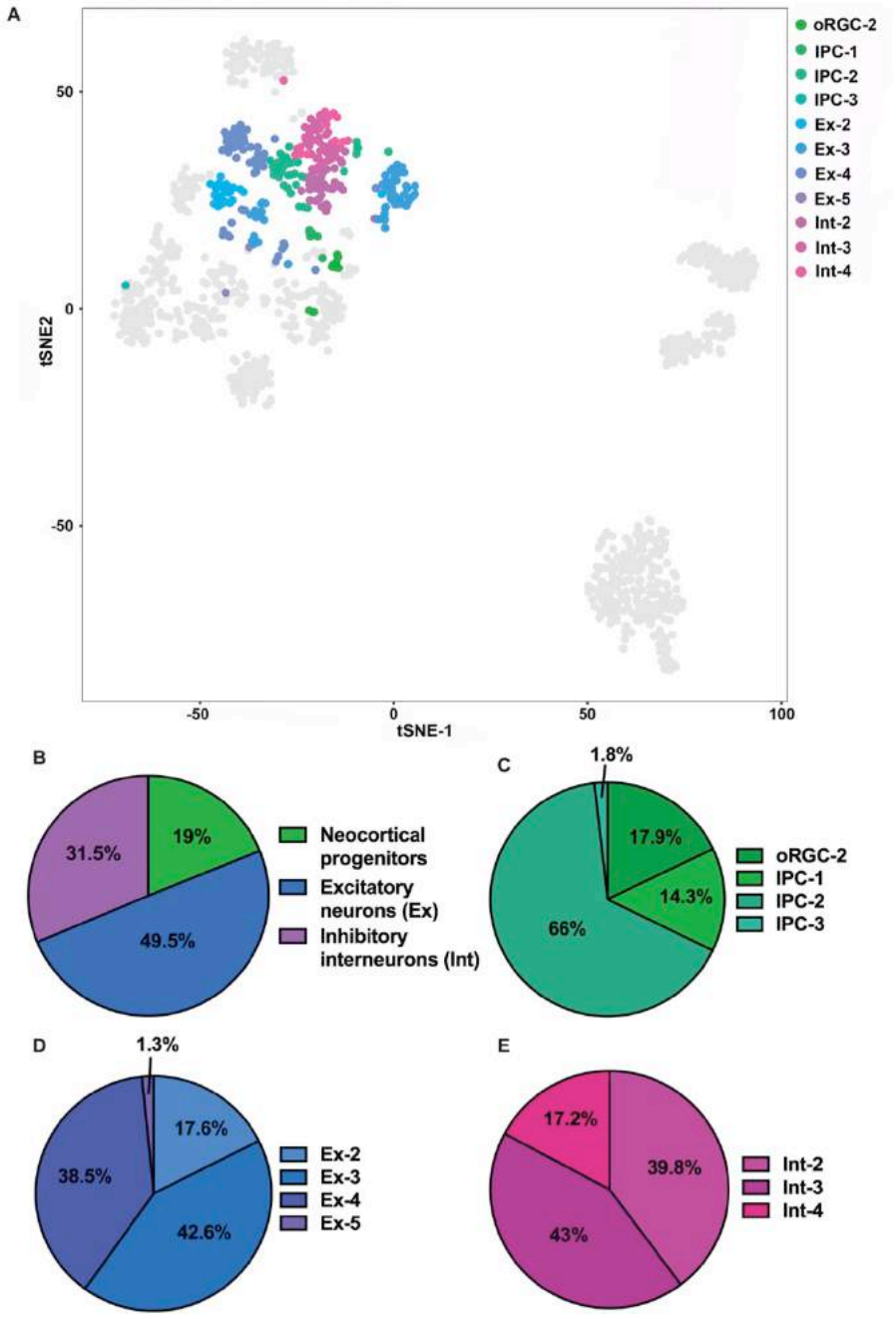


Figure 35: Clustering analysis of iPSCs-derived neuronal cells obtained at intermediate differentiation stages*. A. Clustering analysis performed on iPSCs-derived neuronal cells after 36-42 days of in vitro differentiation identified 11 cell clusters. In the t-SNE plot used to visualize the results of the clustering analysis individual points correspond to single cells and each point is colored according to cell cluster membership. B. Percentages of the three main cell types identified through our clustering analysis in this temporal window, i.e. neocortical progenitor cells (NPCs), excitatory pyramidal neurons (Ex) and inhibitory interneurons (Int). C. Percentages of the 4 different NPCs subtypes. D. Percentages of the 4 distinct pyramidal neurons subtypes. E. Percentages of the 3 different inhibitory interneurons subtypes.*

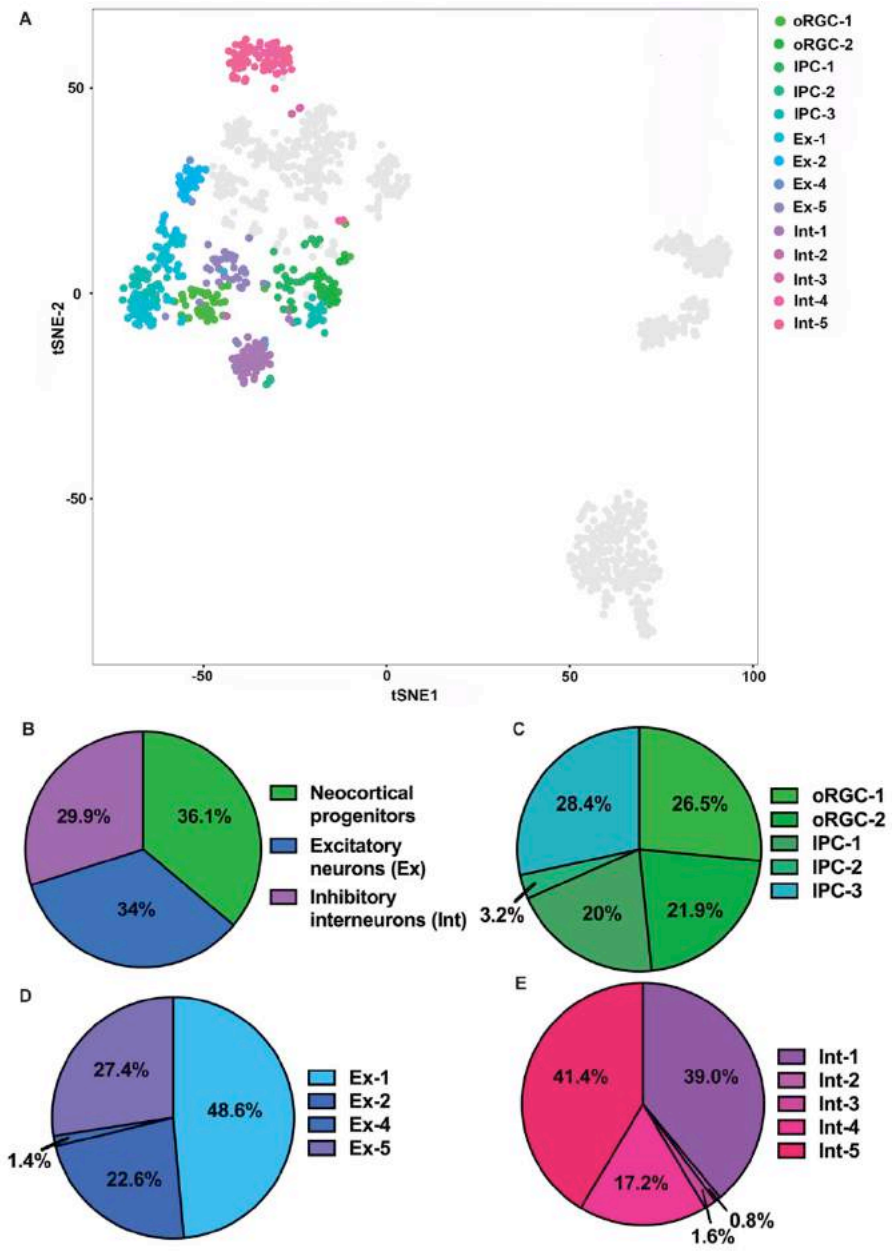
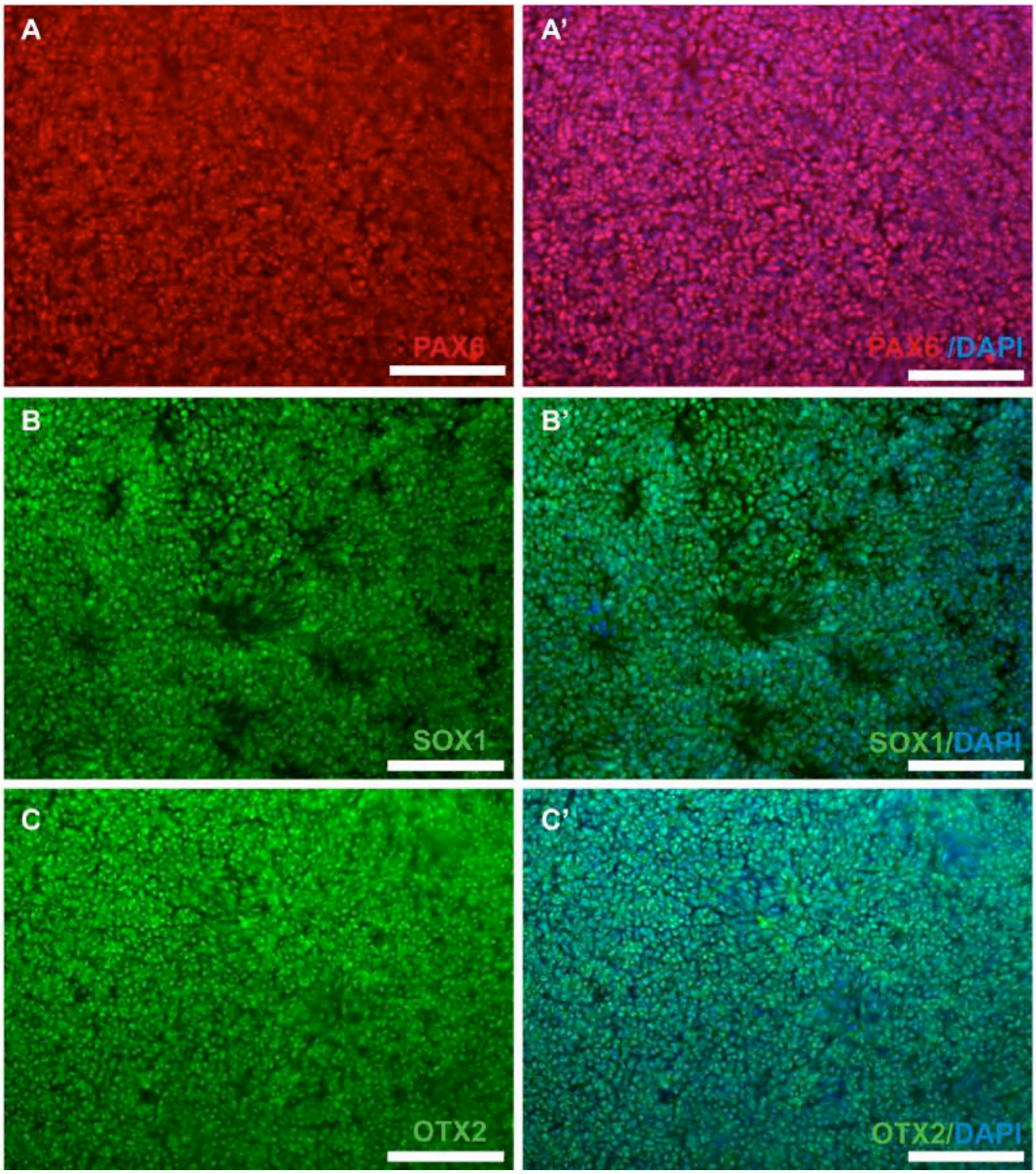


Figure 36: Clustering analysis of iPSCs-derived neuronal cells obtained at the end of the differentiation protocol. *A. Clustering analysis performed on iPSCs-derived neuronal cells after 80-83 days of in vitro differentiation identified 14 cell clusters. In the t-SNE plot used to visualize the results of the clustering analysis individual points correspond to single cells and each point is colored according to cell cluster membership. B. Percentages of the three main cell types identified through our clustering analysis in this temporal window, i.e. neocortical progenitor cells (NPCs), excitatory pyramidal neurons (Ex) and inhibitory interneurons (Int). C. Percentages of the 5 different NPCs subtypes. D. Percentages of the 4 distinct pyramidal neurons subtypes. E. Percentages of the 5 different inhibitory interneurons subtypes.*

5.2.3 Immunocytochemical validation of neuronal cell types identities

The transcriptional results were confirmed by monitoring the expression of a panel of informative cell fate and cell-stage specific markers through immunocytochemistry. At day 12 that corresponds to the end of the neural induction phase, the majority of cells derived from all the eight different iPSCs clones examined had acquired a neuroepithelial identity as demonstrated by the widespread positivity for several nuclear and cytoskeletal early neuroectodermal markers including the transcription factors PAX6, SOX1, OTX2 [642] and the type VI intermediate filament protein Nestin (Figure 37 A-D'). Furthermore, the cell fate commitment to telencephalic identities was confirmed by the positivity of the majority of cells for the transcription factor forkhead box protein G1 (FOXG1) (Figure 37 E-E').



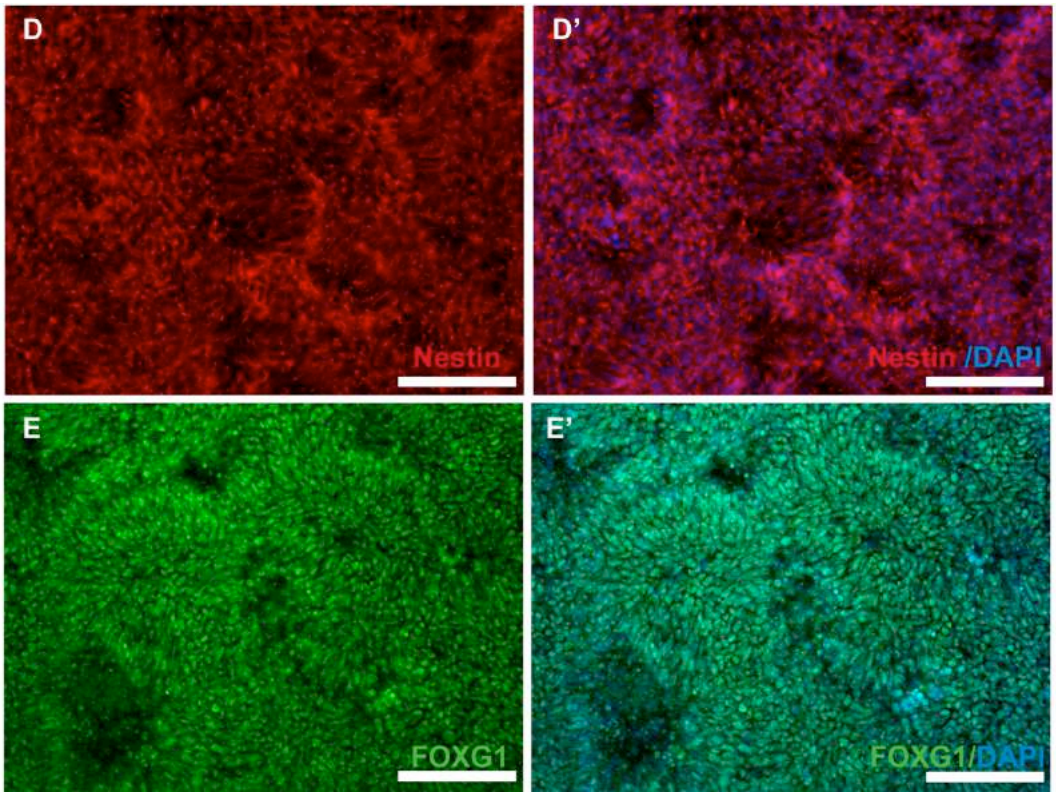


Figure 37: iPSCs-derived neuronal cells acquire telencephalic identities at the end of the neural induction phase. *Immunocytochemical analysis for a selected panel of neuroepithelial and telencephalic markers confirms the transcriptional results indicating the successful conversion of iPSCs into neuroepithelial cells after 12 days of in vitro neocortical differentiation. A-D'. Representative pictures indicating that the majority of cells in culture are positive for the neuroecotdermal markers PAX6, SOX1, OTX2 and Nestin E-E'. FOXG1 widespread positivity confirms their telencephalic commitment. Furthermore, as visibile in stainings for SOX1 (B-B'), Nestin (D-D') and FOXG1 (E-E') neural rosettes structures, that are two-dimensional radial arrangements of columnar cells mimicking in vitro the developing neural tube, are beginning to form. Scale bar: 100 μm.*

At intermediate *in vitro* neurodevelopmental time points comprised in the temporal window between day 36 and day 42 of differentiation, we assessed that iPSCs-derived neuronal cells had become either neocortical intermediate progenitor cells or had already acquired postmitotic neurons identities as suggested by the analysis of their transcriptional profiles. In order to assess the presence in culture of neocortical progenitor cells, we monitored the expression of the transcription factor T-box brain protein 2 (TBR2) which is selectively upregulated after the transition from radial glia to intermediate progenitor and subsequently downregulated in the transition from intermediate progenitor to postmitotic neuron [643]. In conjunction with TBR2 we also monitored the expression of the cytoskeletal protein Nestin that is expressed by mitotic neuronal cells. To validate the transcriptional signatures that indicate excitatory pyramidal neurons generation at these time points, we evaluated the expression of the cytoskeletal protein microtubule-associated protein 2 (MAP2) which is a postmitotic pan-neuronal marker and of the vesicular glutamate transporter 1 (VGLUT1) that is a protein mainly expressed on the membrane of synaptic vesicles involved in glutamate transport.

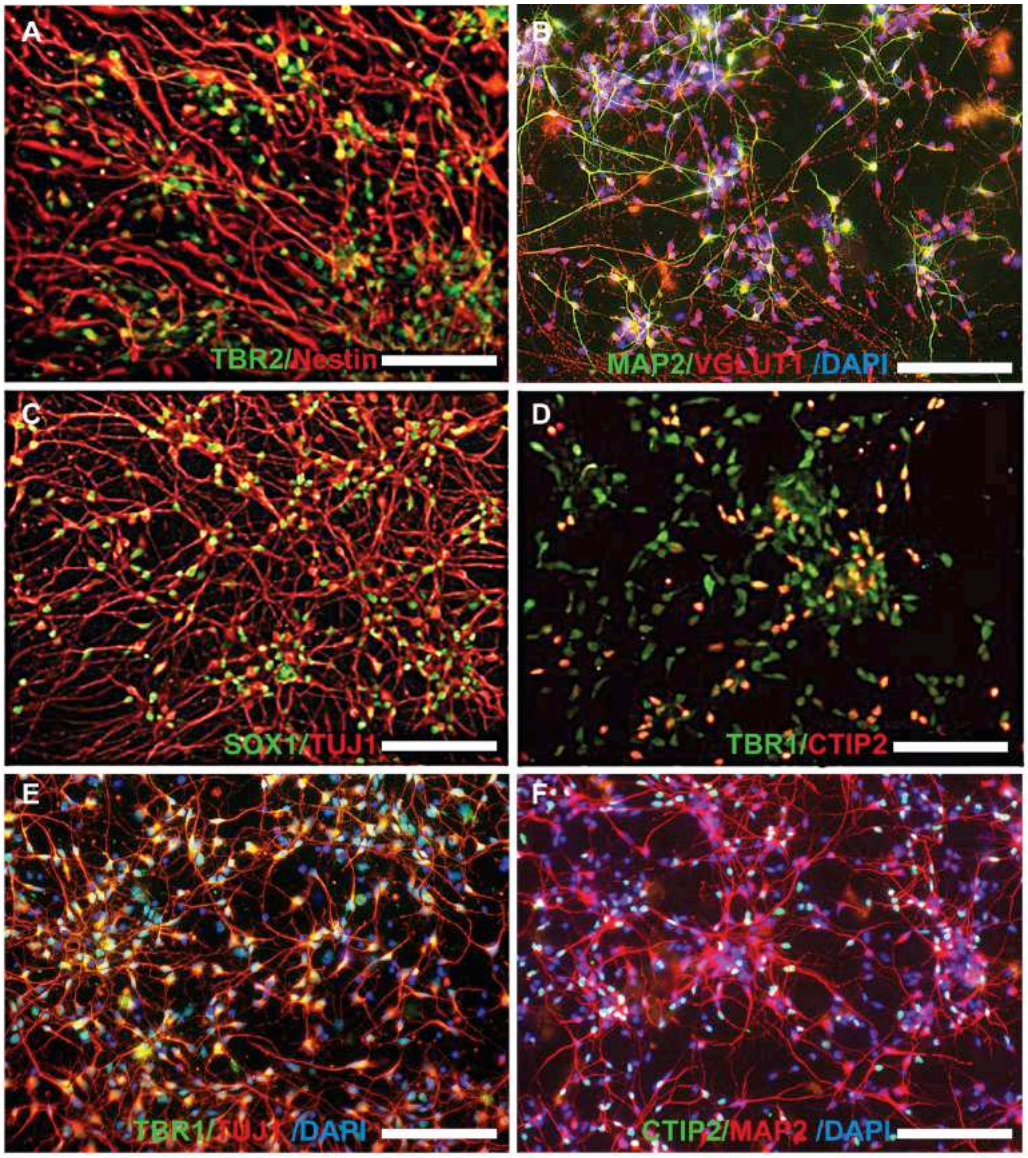


Figure 38: Neocortical progenitor cells and deep layer postmitotic pyramidal neurons are present in culture after 40 days of in vitro differentiation. *A. Positivity for TBR2 and Nestin indicates the presence in culture of neocortical progenitor cells. B. Excitatory pyramidal neurons identity is confirmed by VGLUT1 positivity combined with MAP2 expression. C. Positivity for SOX1 and βIII-tubulin further confirms the generation of immature neurons. D-F. The deep layer laminar identity of the postmitotic neurons generated at intermediate stages of in vitro differentiation has been validated by monitoring the expression of the layer specific transcription factors TBR1 and CTIP2. D. Representative image of TBR1 and CTIP2 costaining, indicating that the expression of these two transcription factors is mainly segregated. E. TBR1 and βIII-tubulin positivity confirms the generation of layer 6 corticothalamic projection neurons. F. CTIP2 and MAP2 positivity substantiates the presence in culture of layer 5 subcortical projection neurons. Scale bar: 100 μm.*

We assessed the expression of the SRY-Box 1 transcription factor (SOX1) also at this later time points as in addition to maintain the progenitor cell state during the early stages of neocortical differentiation its expression is preserved during later stages of neurogenesis when its function switches to promote neural lineage differentiation [644]. In this context SOX1 expression was monitored in conjunction with βIII-tubulin that is a microtubule element expressed by both immature and postmitotic neurons to be able to detect also the presence of neuronal populations that are not yet postmitotic and consequently are not labeled by MAP2 staining. In order to further validate the laminar identity of the glutamatergic neurons generated in this intermediate differentiation time window, we inspected the expression of two deep layer specific transcription factors, TBR1 and CTIP2 that were used as markers of layer 6 [645] and layer 5 [350] neurons respectively. Our stainings validated the presence of all the above mentioned markers confirming the generation at intermediate *in vitro* differentiation stages of neocortical progenitor cells combined with TBR1-positive layer 6 corticothalamic projection neurons and CTIP2-positive layer 5 subcortical projection neurons in agreement with transcriptional data (Figure 38).

At the end of the differentiation protocol the highly efficient generation of pyramidal neurons populating both deep and upper neocortical layers and inhibitory interneurons was confirmed by monitoring the expression of pan-neuronal and neuronal subtype- or layer-specific markers. As pan-neuronal markers we employed the cytoskeletal proteins already used in the immunocytochemical analysis performed at intermediate differentiation time points, including βIII-tubulin to detect the generation of both immature and postmitotic neurons and MAP2 to selectively assess the presence of postmitotic neurons. In addition we investigated the expression of the type IV intermediate filament neurofilament light (NF-L) that forms heteropolymers composing the axoskeleton. In order to confirm the generation of excitatory pyramidal neurons after 80 days of *in vitro* differentiation, we continued to monitor the expression of VGLUT1 in conjunction with MAP2 (Figure 39 A). To validate the presence in culture of inhibitory interneurons we performed stainings for the inhibitory neurotransmitter γ-aminobutyric acid (GABA) in conjuction with MAP2 (Figure 39 B). The generation of minor populations of glial cells was verified by the positivity for the cytoskeletal intermediate filament protein glial fibrillary acidic protein (GFAP) (Figure 39 B). To further assess the presence of pyramidal neurons belonging to both deep and upper neocortical layers, in addition to evaluating the expression of the transcription factors TBR1 and CTIP2 as markers of layer 6 and layer 5 neurons respectively, we performed stainings for the transcription factors cut like homeobox 1 (CUX1) which is expressed by layer 2-4 neurons [646] and SATB homeobox 2 (SATB2) that indicates the presence of mainly upper layers callosal projection neurons [356]. Positivity for all these markers indicated the efficient generation of pyramidal projection neurons belonging to both deep and upper neocortical layers and inhibitory interneurons accompanied by small proportions of glial cells by the end of the differentiation protocol (Figure 39).

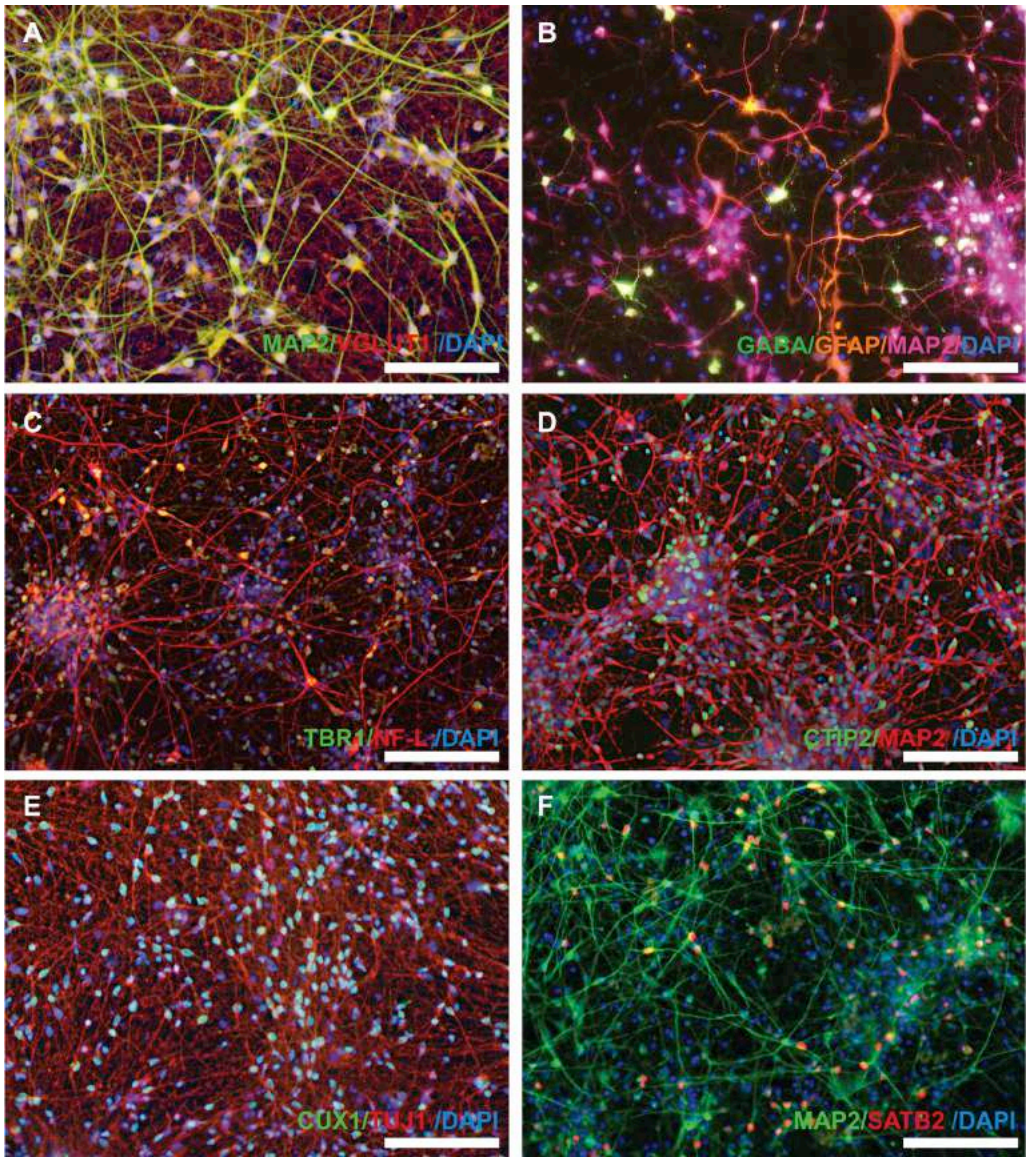


Figure 39: Both deep and upper layer neocortical pyramidal neurons accompanied by inhibitory interneurons are generated with high efficiency at the end of the differentiation protocol. *A. The widespread positivity for VGLUT1 and MAP2 confirms the successful generation of iPSCs-derived excitatory pyramidal neurons. B. The presence of cells positive for the inhibitory neurotransmitter GABA in conjunction with MAP2 substantiates the derivation of inhibitory interneurons. Furthermore, GFAP expression indicates that also minor populations of glial cells are derived. C. The expression of the layer specific transcription factor TBR1 combined with neurofilament light (NF-L) confirms the generation of layer 6 corticothalamic projection neurons. D. Positivity for CTIP2 and MAP2 indicates the presence in culture of layer 5 subcortical projection neurons. E. Expression of CUX1 and βIII-tubulin (TUJ1) demonstrates the successful generation of upper layer neurons. F. SATB2 and MAP2 positivity denotes that also callosal projection neurons are generated with high efficiency at the end of the differentiation protocol. Scale bar: 100 μm.*

In addition to confirm the successful derivation of neocortical neurons, inspection of the cytoskeletal markers reveals that iPSCs-derived neurons had acquired full morphological maturity after 80 days of *in vitro* differentiation. Stainings for βIII-tubulin (TUJ1) and MAP2 highlighted neurons forming highly interconnected dendritic trees and axons projecting to distant targets (Figure 39 A,B,D-F) also organized in extended axon bundles as apparent from neurofilament light staining (Figure 39 C). Both the presence of the major neocortical neurons subtypes obtained at the end of *in vitro* directed neurogenesis detected through immunofluorescence analysis and their morphological maturity enabled to confirm the transcriptional signatures identified through RNA-sequecing performed at both population and single-cell level.

5.2.4 Functional properties of iPSCs-derived neocortical neurons

An essential property that characterizes fully differentiated neurons is their physiological maturity. GO enrichment analysis performed on DEGs that have an increased expression at the end of the *in vitro* differentiation paradigm and selected WGCNA modules characterized by an increased expression throughout the differentiation paradigm revealed an enrichment in biological process categories related to voltage-gated sodium channel activity, positive regulation of action potential, synaptic transmission and neurotransmitter secretion (Figure 40 A) as well as in cellular component terms associated to voltage-gated ion channels and synaptogenesis (Figure 40 B).

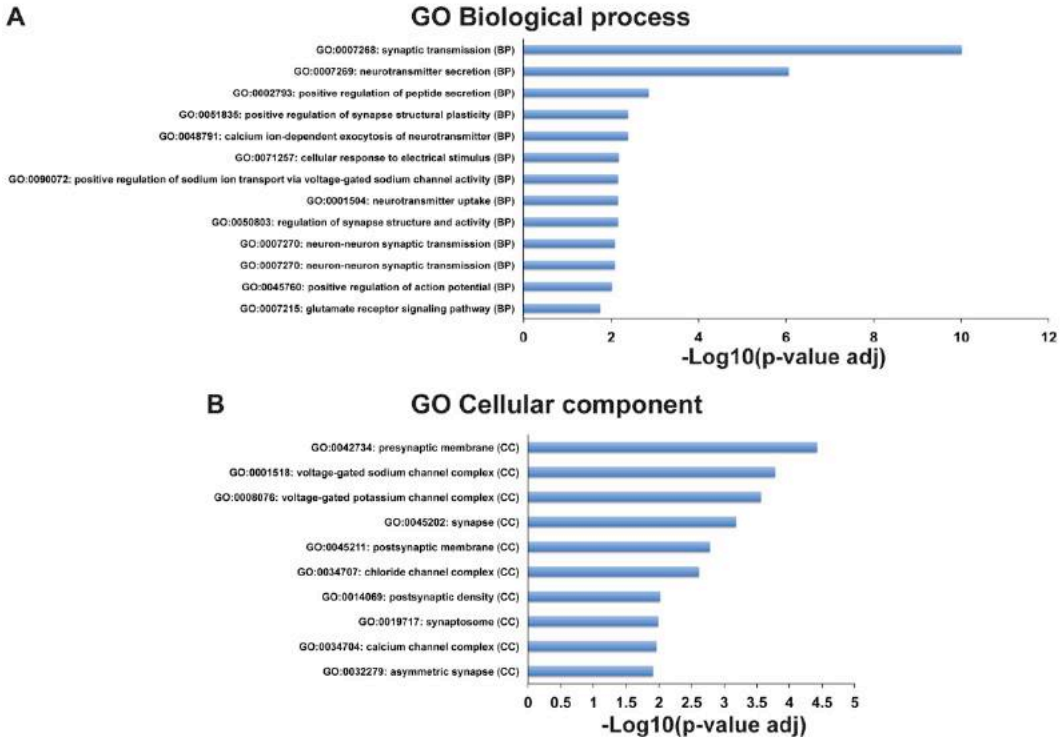


Figure 40: Top-ranking GO categories enriched at the end of the differentiation protocol suggesting the physiological maturity of iPSCs-derived neocortical

neurons. *A. Top most-specific enrichments for GO biological process categories detected at the end of the in vitro differentiation paradigm. B. Top-ranking enrichments for cellular component terms.*

The increased expression of transcripts coding for voltage-dependent ion channels identified at the end of the differentiation protocol suggests that iPSCs-derived neocortical neurons might be able to generate action potentials and sustain repetitive firing. Furthermore, the progressive enriched expression of synaptic transmission and neurotransmitter secretion categories suggests that the neocortical neurons generated *in vitro* are able to communicate with each other and form functional networks.

To directly evaluate the functional maturity of iPSCs-derived neocortical neurons generated at the end of the *in vitro* differentiation protocol, their electrophysiological properties have been investigated through whole-cell patch-clamp recordings performed both in current- and voltage-clamp mode. All analyses have been conducted in collaboration with Francesca Talpo. At first the percentage of physiologically mature neurons has been determined. The ability of neurons to initiate at least one action potential (AP) in response to the injection of progressively increasing steps of depolarizing currents has been assumed as a marker of functional maturation. The analysis revealed that the majority of cultured neocortical neurons are functionally mature (Figure 41 A). Then, the passive properties of both functionally mature (n=23) and immature (n=12) neurons have been examined, including the membrane capacitance (C_m), the input resistance (R_{in}) and the resting membrane potential (V_r) (Figure 41 B,C,D). In this context, the membrane capacitance is related to the size of cell, the input resistance is inversely proportional to the number of channels expressed by the cell while the value of the resting membrane potential can be considered an indicator of the relative proportions of the different ion channels expressed by the cell as it is determined by the concentration gradients of ions across the plasma membrane and by the membrane permeability to each type of ion. Physiologically mature iPSCs-derived neocortical neurons are characterized by a significant increase in their membrane capacitance (C_m = 32.9 ± 3.4 pF) compared to physiologically immature neurons (C_m = 15.2 ± 1.7 pF). The significantly increased membrane capacitance of neurons able to generate APs compared to immature cells indicates enlarged soma that is a distinctive trait of fully developed postmitotic neurons. The significantly decreased input resistance of fully differentiated iPSCs-derived neocortical neurons (R_{in}= 0.6 ± 0.09 GΩ) reflects the enhanced expression of ion channels in their plasma membrane with respect to more immature neuronal cells (R_{in}= 1.8 ± 0.7 GΩ). Moreover, the significantly more hyperpolarized resting membrane potential of mature iPSCs-derived neocortical neurons (V_r = -48.2 ± 1.3 mV) compared to more immature neuronal cells (V_r = -43.6 ± 1.8 mV) suggests that during *in vitro* differentiation iPSCs-derived neurons progressively acquire an ion-channels expression profile similar to that presented by fully developed postnatal neurons.

To gain more insights into the level of functional maturation reached by iPSCs-derived neurons, their passive properties have been compared to the limited reference literature values available obtained from primary cultures of human fetal neocortical neurons [647] and human fetal [648] and postnatal [649] brain slices.

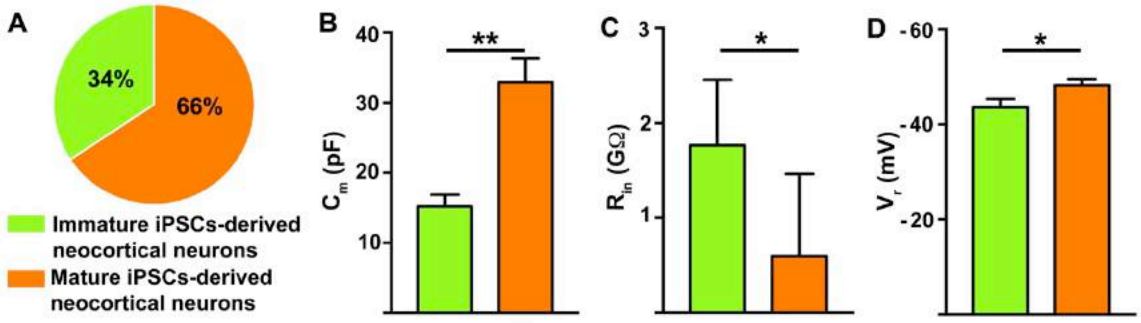


Figure 41: Passive properties of physiologically immature and mature iPSCs-derived neocortical neurons. *A. Total percentage of cells able and unable to fire action potentials when excited with suprathreshold depolarizing steps of current (n = 35). Cells unable to fire action potentials were considered still undifferentiated and immature (n = 12) while neurons able to initiate at least one action potential in response to depolarizing stimuli were considered physiologically mature (n = 23) and further analyzed. B-D. Bar plots representing the passive properties recorded from both immature and functionally mature iPSCs-derived neurons indicating that the membrane capacitance, input resistance and resting membrane potential are significantly different between iPSC–-derived immature ad mature neocortical neurons (C_m: p = 0.0009; R_{in}: p = 0.03 ; V_r: p = 0.04). Data are expressed as mean ± SEM and compared using unpaired t-test, * p < 0.05, ** p < 0.005.*

The passive properties of physiologically mature iPSCs-derived neurons resulted to be similar to those acquired from reference human fetal neocortical primary cultures derived from 7 to 9 postconceptional weeks (pcw) brain tissues and cultured for more than 20 days in vitro (DIV) prior to perform electrophysiological recordings (C_m = 36.3 ± 2.7 pF; R_{in} = 0.9 ± 0.1 GΩ; V_r= −46.3 ± 2.8 mV) [647]. Furthermore, physiologically mature iPSCs-derived neurons are characterized by a resting membrane potential similar to that recorded from the cortical plate (CP) of human fetal brain slices at 16 to 22 gestational weeks (V_r = -49.8 ± 1.3 mV) [648] but still significantly more depolarized with respect to the resting membrane potential distinctive of postnatal neurons (V_r = -75.7 ± 1.2 mV) [649]. Overall, the comparison of the passive properties characterizing *in vitro* generated neocortical neurons with reference literature values indicates that iPSCs-derived neurons present passive properties values similar to human neocortical fetal primary cultures. Moreover, it seems that iPSCs-derived neurons are characterized by passive properties also similar to human fetal neurons recorded from brain slices but still not reach the full functional maturity that identifies human postnatal neurons in brain slices. Given the differences of *in vitro* culture environment between cell culture and brain slices, the results obtained from these last comparisons should be interpreted just as a further indication of the level of functional development of iPSCs-derived neocortical neurons.

Then, the firing properties of iPSCs-derived neocortical neurons have been investigated in further detail. As the ability to generate APs and organizing them into specific firing patterns is mainly related to the activation of voltage-dependent inward Na^+ currents and outward K^+ currents, the total inward and outward ion currents that flow through the membrane at different potentials have been recorded from both physiologically mature

and immature neurons. Transient inward Na^+ current mediated by voltage-gate sodium channels is characterized by a fast kinetics while outward K^+ current due to voltage-gated delayed rectifier potassium channels is marked by slow activation kinetics. Inward sodium currents and outward potassium currents are well expressed by mature iPSCs–derived neurons (Figure 42 A; n=23), while they are almost absent in immature neurons (n=12) as expected. In accordance, quantification of the maximum sodium current density measured at -20mV shows a significant difference between mature neurons (-77.0 ± 12.8 pA/pF) and immature neurons (-3.2 ± 0.6 pA/pF). The total absence of voltage-dependent inward currents in immature neurons indicates that their inability to generate APs is due to the lack of expression of voltage-gated channels.

To better assess the level of functional development of mature iPSCs-derived neocortical neurons (n = 23), they have been classified into three categories on the basis of the spike amplitude and the number of spikes they were able to elicit when excited with suprathreshold depolarizing steps of current (Figure 42 D): i) iPSCs-derived mature neurons only able to generate spikes that did not reached the overshoot threshold of 0 mV were classified as cells able to initiate abortive spikes, ii) neurons competent to give rise to only a single spike that surpassed the overshoot threshold of 0 mV were classified as single spike neurons, iii) cells able to fire trains of APs were categorized as repetitive firing neurons. More than 80% of mature iPSCs-derived neurons were able to elicit at least one action potential that reaches the overshoot and the majority of them was competent to generate trains of action potentials (56.5%, n = 13/23) organized into characteristic repetitive firing patterns (Figure 42 B).

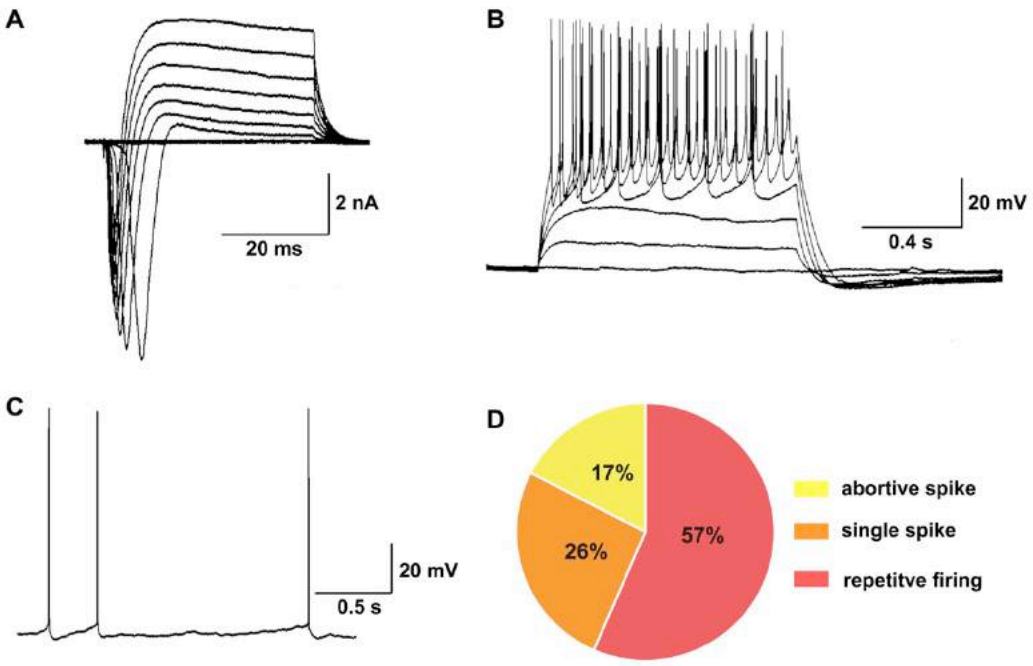


Figure 42: Firing properties of iPSCs-derived mature neocortical neurons after 80 or more days of in vitro differentiation. *A. Sample traces of total inward and outward currents generated in response to a series of depolarizing voltage steps. B. Sample trace of repetitive firing behavior in response to steps of current injection. C. Representative trace of spontaneous action potentials. All neurons able to generate at*

least one AP when stimulated with a suprathreshold depolarizing steps of current were also able to generate spontaneous APs (n=8). D. Percentages of cells subdivided according to their firing properties elicited in response to steps of current injections (n=23; abortive spike 17%, n=4; single spike 26%, n=6; repetitive firing 57%, n=13).

Also the percentages of neurons classified according to their firing properties have been compared to literature values from primary cultures of human fetal neocortical neurons [647] and human fetal [648] and postnatal [649] brain slices, using the same references previously employed in the passive properties comparison. In human neocortical primary cultures derived from 7 to 9 postconceptional weeks (pcw) old fetal tissues recorded after having grown at least 20 days in vitro, 81% of neurons (n = 22/27) were able to generated single APs and 14.8% of cells (n = 4/27) displayed repetitive firing. Only approximately one-third of the cells (n = 33/107) were able to initiate APs in recordings made from the cortical plate (CP) of 16 to 22 gestational weeks (gw) brain slices, where 15.9% of the neurons were able to initiate at least an abortive AP (n = 17/107) and 14.9 % of the cells were competent to elicit a single AP (n = 16/107). Additionally, 14.3% of CP cells at 20 and 22 gw (n = 8/56) produced repetitive firing of APs while all postnatal neurons recorded from human brain slices were able to generate a large number of APs (100%, n = 82). The comparison of these reference values with our recordings indicates that mature iPSC-derived neurons are less synchronized in their functional development compared to primary neocortical cultures derived from 7 to 9 pcw tissues, as indicated by the percentages of neurons competent to fire at least one AP (iPSCs-derived mature neurons = 82.6%; primary cultures = 96.3%), but mature iPSCs-derived neurons seem even more physiologically evolved with respect to primary cultures as suggested by the increased number of neurons able to sustain repetitive firing (iPSCs-derived mature neurons = 56.5%, primary cultures = 14.8%). Mature iPSCs-derived neurons seem also more physiologically developed than CP neurons up to 22 gw (single AP: iPSCs-derived neurons = 26.1%, CP neurons = 14.9 %; repetitive firing: iPSCs-derived neurons = 56.5%, CP neurons = 14.3%) according to brain slice recordings [648] but have not yet acquired the full functional maturity of postnatal neurons (repetitive firing = 100%) [649].
In addition to monitoring the firing properties elicited in response to injections of suprathreshold depolarizing steps of current, also the ability to generate spontaneous APs has been evaluated in a random subgroup of neurons included in the previous analysis (n = 8), excluding cells that were able to initiate only an abortive spike. Monitoring the ability to generate spontaneous APs, revealed that all neurons able to generate at least one AP when stimulated with a suprathreshold depolarizing steps of current were also able to generate spontaneous APs, a finding that further suggests their physiological maturity (Figure 42 C). Finally, the characteristics of the APs in the form of spike amplitude and spike duration have been analyzed to assess the morphology of the generated APs and compare it to literature reference values.
In iPSCs-derived neocortical neurons, spontaneous APs and APs generated in response to injections of suprathreshold steps of currents were characterized by similar values and consequently data were combined and treated as only one dataset (Table 5). The APs generated by iPSC-derived neocortical neurons are characterized by significantly lower spikes amplitudes and significantly higher spike durations compared to gold-standard reference values obtained from human postnatal neocortical neurons brain slices [649], as assessed through one sample t-test. This finding indicates that *in vitro* derived neocortical neurons are physiologically active but the morphology of their APs is still immature, an expected finding in agreement with all the other electrophysiological

measurements performed and the transcriptional profiles obtained from iPSCs-derived neocortical neurons.

APs parameters	iPSC-derived neocortical neurons	Postnatal neocortical neurons
Spike amplitude	65.3 ± 4.1 mV (n = 23)	94.3 ± 1.5 mV (n = 43)
Spike duration	3.40 ± 0.36 ms (n = 23)	2.1 ± 0.08 (n = 25)

Table 5: Parameters characterizing the action potentials (APs) generated by iPSCs-derived neocortical neurons and postnatal neocortical neurons. *In parenthesis are reported the number of cells analyzed. Mean values of APs spike amplitude and duration of postnatal neocortical neurons were obtained from [649].*

To further assess the functional maturity of our iPSCs-derived neurons generated after 80 days of *in vitro* differentiation, both the excitatory and inhibitory spontaneous activity in the form of spontaneous excitatory postsynaptic currents (sEPSCs) and spontaneous inhibitory postsynaptic currents (sIPSCs) have been investigated (Figure 43 A, B). Synaptic spontaneous events are due to the combined effects of both the spontaneous stochastic release of synaptic vesicles and the release induced by spontaneous action potentials. Therefore, monitoring the spontaneous activity provides information about the ability of the recorded postsynaptic neurons to express functional receptors, the competence of presynaptic neurons to generate action potentials and the extent of connectivity between different cells, overall enabling to evaluate the functional maturity of the cultured neurons and the comprehensive development of the neuronal network. The majority of mature neurons presented sEPSCs (69%) and a substantial subgroup of these neurons presented also sIPSCs (38%) (Figure 43 C) confirming the generation of a functional network of pyramidal neurons and inhibitory interneurons at the end of the differentiation paradigm in agreement with transcriptional data (Figure 43 D).

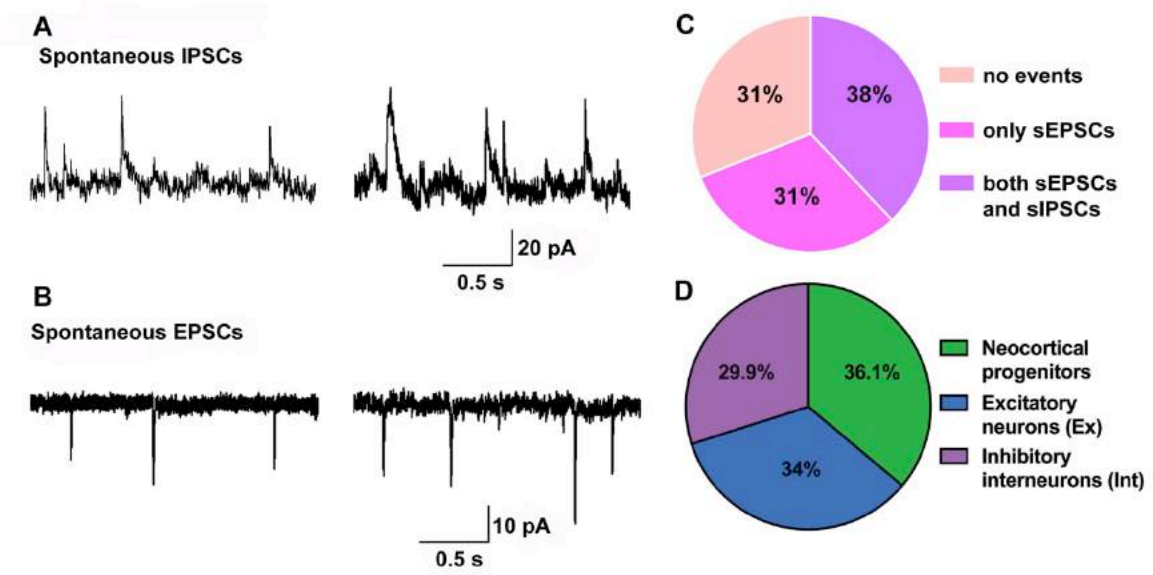


Figure 43: Spontaneous synaptic activity of iPSCs-derived mature neocortical neurons. *After 80 or more days of in vitro differentiation also the spontaneous synaptic activity of iPSCs-dervied neocortical neurons has been monitored (n=16). A-B. Sample traces of inhibitory (sIPSCs) and excitatory spontaneous postsynaptic currents*

(sEPSCs) acquired in voltage-clamp mode, performing the recordings at 0 mV and -70 mV respectively. The inhibitory or excitatory nature of the events was confirmed by applying at the end of the recordings the AMPA receptor antagonist NBQX (10 μM) or the $GABA_A$ receptor antagonist bicuculline (10 μM) that abolish respectively the sEPSCs and the sIPSCs. C. Percentages of recorded neurons that presented excitatory and inhibitory spontaneous synaptic events (no events 31%, n=5; only sEPSCs 31%, n=5; both sEPSCs and sIPSCs 38%, n=6). D. Percentages of the three main cell types identified through our clustering analysis of single cell RNA-sequencing data obtained from iPSCs-derived neocortical cells generated at the end of the differentiation paradigm, i.e. neocortical progenitor cells (NPCs), excitatory pyramidal neurons (Ex) and inhibitory interneurons (Int).

Chapter 6: Human neocortical neurons derived from induced pluripotent stem cells as a model system to investigate the effects of 7q11.23 gene-dosage dependent dysregulation

Copy number variations (CNVs) are major contributors to disorders of the central nervous system including intellectual disability, autism spectrum disorder (ASD), and schizophrenia (SCZ). To understand the pathophysiology of these disorders and develop effective therapeutic interventions, it is crucial both to determine the biological pathways and functional networks onto which multiple disease-associated genes converge and to identify critical genes within a CNV locus whose dosage imbalances drive specific neuronal phenotypes.

The human spatiotemporal expression patterns of the 26-28 protein coding genes comprised in the 7q11.23 CNV have not been analyzed in great detail. Therefore, we first performed system-level analyses utilizing human brain transcriptomic and protein-protein interaction data to investigate the molecular coordination and spatio-temporal expression specificity of genes comprised in this locus. Next, top functions for the identified modules of highly co-expressed genes were determined performing Gene Ontology enrichment analysis. Taken together these analyses provided evidence for prenatal gene regulation, postnatal synaptic transmission and mitochondrial respiration as potential contributors to 7q11.23 phenotypes.

Building on this knowledge, two different lines of research were pursued. One was aimed to identify and functionally validate underexplored, yet central genes driving neurobiological processes uniquely affected in 7q11.23 syndromes. Pursuing single genotype-phenotype correlations is a promising strategy not only to understand disease mechanisms but also to make inroads toward developing novel treatments. Thus, considering the processes highlighted by our system biology analyses prospectively as therapeutic targets, the postnatal enrichment for mitochondrial GO terms, suggesting a role for mitochondrial function in the pathogenesis of 7q11.23 syndromes, seemed an extremely interesting finding. Therefore, the genes driving these processes were identified and among them we decided to focus on unraveling the neurobiology of the previously uncharacterized gene DNAJC30. In this context, iPSCs-derived neocortical neurons, obtained using the previously extensively characterized differentiation protocol, and brain tissue samples were essential to demonstrate the involvement of mitochondrial dysfunction and the role played by DNAJC30 in mediating Williams syndrome, confirming in humans the morphological and functional alterations observed in a mouse knockout for Dnajc30.

The goal of the other project undertaken was to provide new insights into the spatiotemporal activity of genes comprised in 7q11.23 CNV and to investigate in which cell types and neurodevelopmental periods 7q11.23 gene dosage imbalances exert major effects resulting in altered cellular phenotypes and functional deficits. To achieve this aim, single-cell RNA-sequencing has been performed on samples of iPSCs-derived neocortical neurons obtained from all the cell lines included in the cohort under study using 10X Genomics platform. Analysis of single-cell RNA-sequencing data and functional validation of the identified biological targets are still ongoing.

6.1 Integrated Co-Expression and PPI Analyses Identify Three 7q11.23 Gene Clusters with Distinct Functions

Given the paucity of information about the human spatiotemporal expression patterns of protein coding genes within 7q11.23 CNV, systems-level analyses utilizing brain gene expression and protein-protein interaction datasets have been performed as a first step to identify underexplored critical genes contributing to the altered neuronal phenotypes characteristic of 7q11.23 syndromes. To investigate the correlation in the expression patterns of the genes comprised in 7q11.23 CNV during human brain development, we analyzed RNA-seq data from the publicly available BrainSpan resource generated in the laboratory. This transcriptomic resource includes data obtained from 607 high-quality, histologically verified tissue samples representing sixteen regions (comprising 11 areas of the neocortex, hippocampus, amygdala, striatum, mediodorsal nucleus of thalamus, and cerebellar cortex) systematically dissected from 41 brains (18 females and 23 males) of clinically and histopathologically unremarkable donors of different ancestry, ranging in age from 5 postconceptional weeks (pcw) to 40 years. For further information see the BrainSpan resource, www.brainspan.org. First, pairwise co-expression among all 7q11.23 genes was calculated, followed by hierarchical clustering. By visual inspection of the resulting dendrogram and heatmap, three clusters were identified (Figure 44 A). As evident by examining the position of genes belonging to the different clusters, the observed co-expression is unlikely explained by genomic adjacency given that the genes comprised in each cluster are scattered across the 7q11.23 locus (Figure 44 B).

To determine if genes in each cluster were merely co-expressed or represented a deeper molecular coordination with interacting proteins, the co-expression of each cluster with nearest neighbor interacting proteins was also explored. This analysis was performed using as a backbone interactome the publicly available biological General Repository for Interaction Datasets (BioGRID) protein-protein interaction (PPI) network, which provides a comprehensive collection of non-redundant physical protein-protein interactions [650]. In order to tailor the BioGRID PPI network to human brain development, the interactome space was confined based on the availability and the abundance of the BrainSpan RNA-seq dataset. All the nodes, i.e. proteins and corresponding genes, not profiled in the BrainSpan RNA-seq dataset were removed from the interactome. The analysis was further refined by removing genes with noise expression levels by grouping as expressed vs. non-expressed, using a mixture model. This resulted in a more brain-specific interactome network of 12,367 nodes and 159,353 edges. Within the resulting interactome network, the pairwise co-expression of nodes significantly decayed over the shortest path distance between the nodes, consistent with several reports showing that the proximity of genes in the interactome network is indicative of the functional similarity of genes.

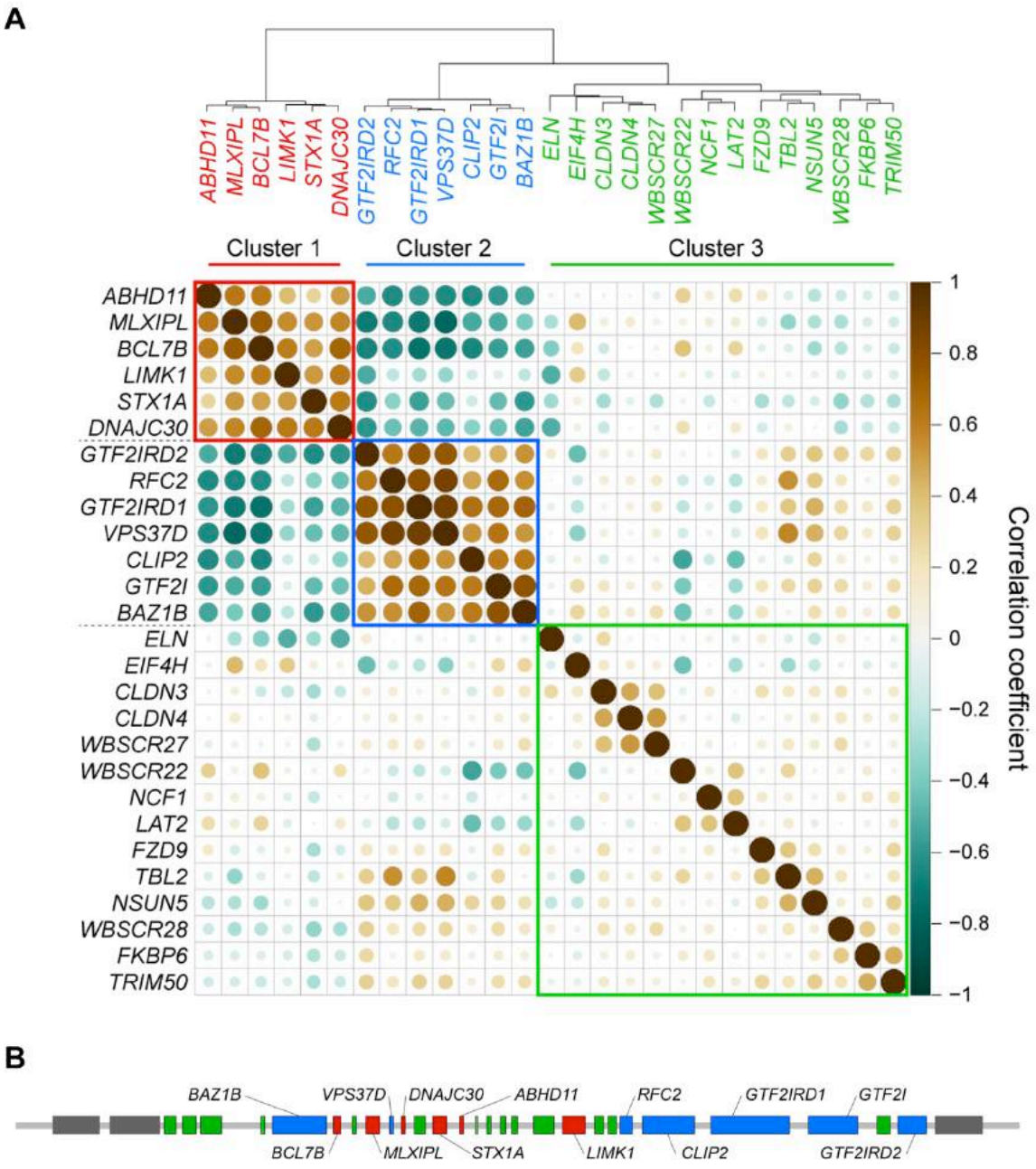


Figure 44: Analysis of pairwise co-expression among all 7q11.23 genes highlights three clusters that are independent of genomic adjacency. *A. Dendrogram and heatmap of brain transcriptional similarities for all 7q11.23 genes identify three clusters. B. Schematic representation of the genes comprised in the 7q11.23 locus color coded by cluster, showing that cluster assignment does not depend on genomic location. Gene names are reported for cluster 1 and cluster 2.*

No proteins encoded in 7q11.23 are known to directly interact with each other, but by extending the PPI network to second and third order interacting proteins, a high co-expression for clusters 1 and 2 was still observed, while cluster 3 and intercluster interacting proteins were characterized by much lower co-expression values (Figure 45 A). To further assess the molecular coordination of groups of genes belonging to a cluster, the significance of the observed mean pairwise co-expression of each gene set was tested against an empirical null expectation. To draw a null expected mean co-expression for a set of genes in a cluster, we randomly sampled a set of matched number of genes having the same number of protein interacting partners from the interactome network and averaged the absolute value of the pairwise co-expression coefficients. These data indicate that the mean pairwise co-expression of cluster 1 and cluster 2 is higher than that of randomly sampled genes across the genome with matched number of interacting proteins, a further evidence that the molecular coordination of each of these two clusters could not be fortuitous (Figure 45 B). On the other hand, the mean pairwise co-expression of cluster 3 resulted lower compared to the empirical null expectation, a finding that further confirms the results obtained from the pairwise co-expression and protein-protein interaction network analyses (Figure 45 B).

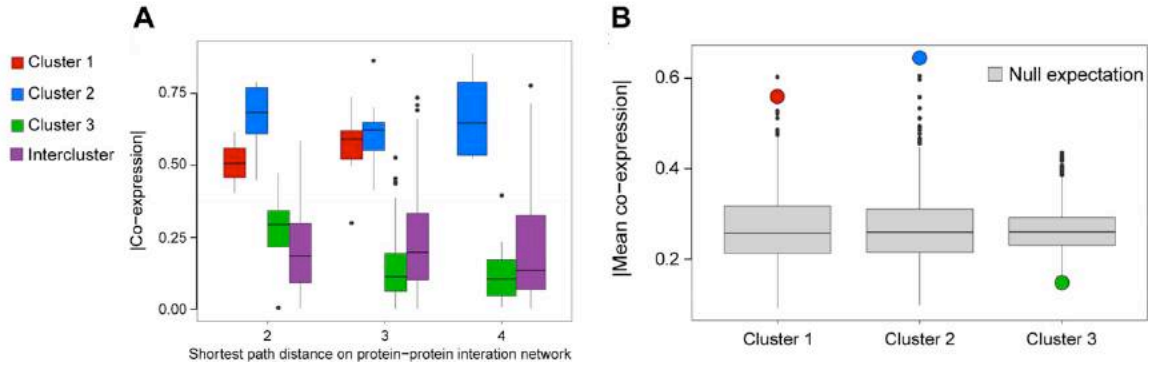


Figure 45: Co-expression analysis of all genes in each cluster and their protein-protein interaction (PPI) network highlights a deeper molecular coordination of clusters 1 and 2. *A. Co-expression of all genes in each cluster and their nearest neighbors on the BioGRID PPI network. Data are plotted as mean ±SEM and outliers as points. All of the cluster 1 proteins are connected by a distance of 3 on the PPI network. B. Boxplots illustrating that the mean co-expression for cluster 1 and cluster 2 are higher than predicted based on randomly selected genes having identical numbers of protein-protein interactors (null expectation), while cluster 3 is characterized by much lower co-expression.*

Next, it was investigated more specifically where the expression of genes within these clusters might have more pronounced effects. The relative tissue-specific gene expression enrichment between groups of genes was examined using the Genotype-Tissue Expression (GTEx) project RNA-Seq dataset (GTEx Consortium, https://gtexportal.org/home). This transcriptomic resource comprise RNA-seq data obtained from 11688 high quality tissue samples derived from 714 distinct donors (470 males and 244 females) representing 53 different tissues, including 11 brain regions. Gene expression analysis performed across several non-central nervous system (CNS) tissues revealed tissue-specific variations in cluster expression (Figure 46 A).

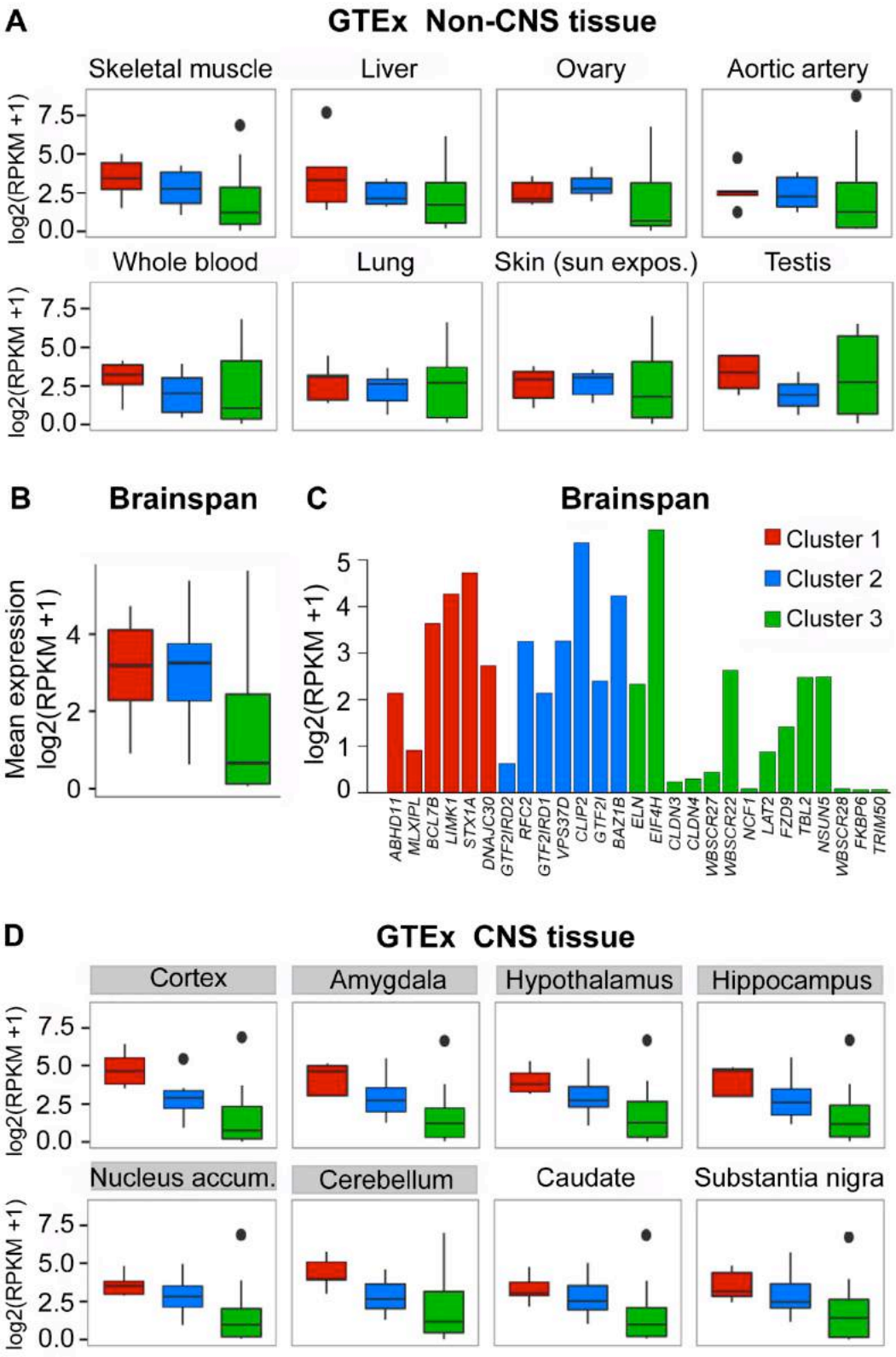


Figure 46: Spatial expression analysis of 7q11.23 gene clusters highlights that cluster 1 and cluster 2 exhibit brain-enriched expression. *A. GTEx-generated mean*

expression levels for each cluster examined across many non-central nervous system (CNS) tissues identifies tissue-specific variations in cluster expression that did not reach statistical significance. Data are plotted as ±SEM and outliers as points. B. Boxplots representing the mean expression levels averaged by cluster based on BrainSpan RNA-seq data. C. BrainSpan-generated expression levels for each gene comprised in 7q11.23 CNV. D. Mean expression levels for each cluster examined across several CNS tissues calculated from GTEx dataset. Data are plotted as ±SEM and outliers as points.

In the CNS, cluster 1 and cluster 2 are enriched over cluster 3, as indicated by gene expression analyses performed using both BrainSpan and GTEx datasets (Figure 46 B-D). The whole brain enriched expression of clusters 1 and 2 with respect to cluster 3 is not driven by specific genes but rather results from a more coordinated increased expression of the majority of genes comprised in a cluster, as evident from boxplots representing mean cluster expression (Figure 46 B) and the analysis of single gene expression (Figure 46 C) of BrainSpan data. Applying a cutoff of log2 (RPKM+1) ≥1 to this dataset showed brain expression of five of six genes from cluster 1 and six of seven genes from cluster 2, while only six of fourteen genes from cluster 3 were expressed above the threshold (Figure 46 C). Moreover, we started to gain a finer spatial resolution by analyzing mean cluster expression across several adult brain regions using GTEx dataset. This analysis revealed that clusters 1 and 2 are significantly enriched over cluster 3 in the neocortex, amygdala, hypothalamus, hippocampus, nucleus accumbens and cerebellum, while in the caudate and substantia nigra the observed increased expression of clusters 1 and 2 compared to cluster 3 did not reached statistical significance (Figure 46 D).

To further determine the spatiotemporal expression of each cluster, BrainSpan RNA-seq data were subdivided into 36 spatio-temporal intervals. Specifically, gene expression data were organized into four spatial groups, comprising the neocortex (NCX) (R1), hippocampus and amygdala (R2), striatum and thalamus (R3), cerebellum (R4) and nine time windows, to look for cluster expression specificity. Then, the specificity of gene expression of all genes in the interactome network was evaluated using gene Enrichment profiler, which provides a quantitative measure of up-down regulation in a given spatio-temporal interval with respect to all the rest intervals. The specificity scores were standardized with mean=0 and standard deviation=1 in each spatio-temporal interval. Therefore, the resulting gene-expression specificity score of each gene provides a deviation (Z-score) from the normalized specificity score distribution in a given spatio-temporal interval. This analysis revealed that cluster 2 is enriched across all brain regions prenatally while cluster 1 is selectively enriched in the neocortex postnatally (Figure 47 A). In addition to its reduced expression in all brain regions and across all the time windows examined, cluster 3 does not show any clear expression patter (Figure 47 A). While the magnitude of this result is less after repeating this analysis including also PPI partners, the same expression trend for each cluster remains (Figure 47 B). This result is consistent with the spatio-temporal enrichment observed previously [651], with slight differences in our results likely explained by the use of gene clusters, modified spatial groups, and specific interacting proteins.

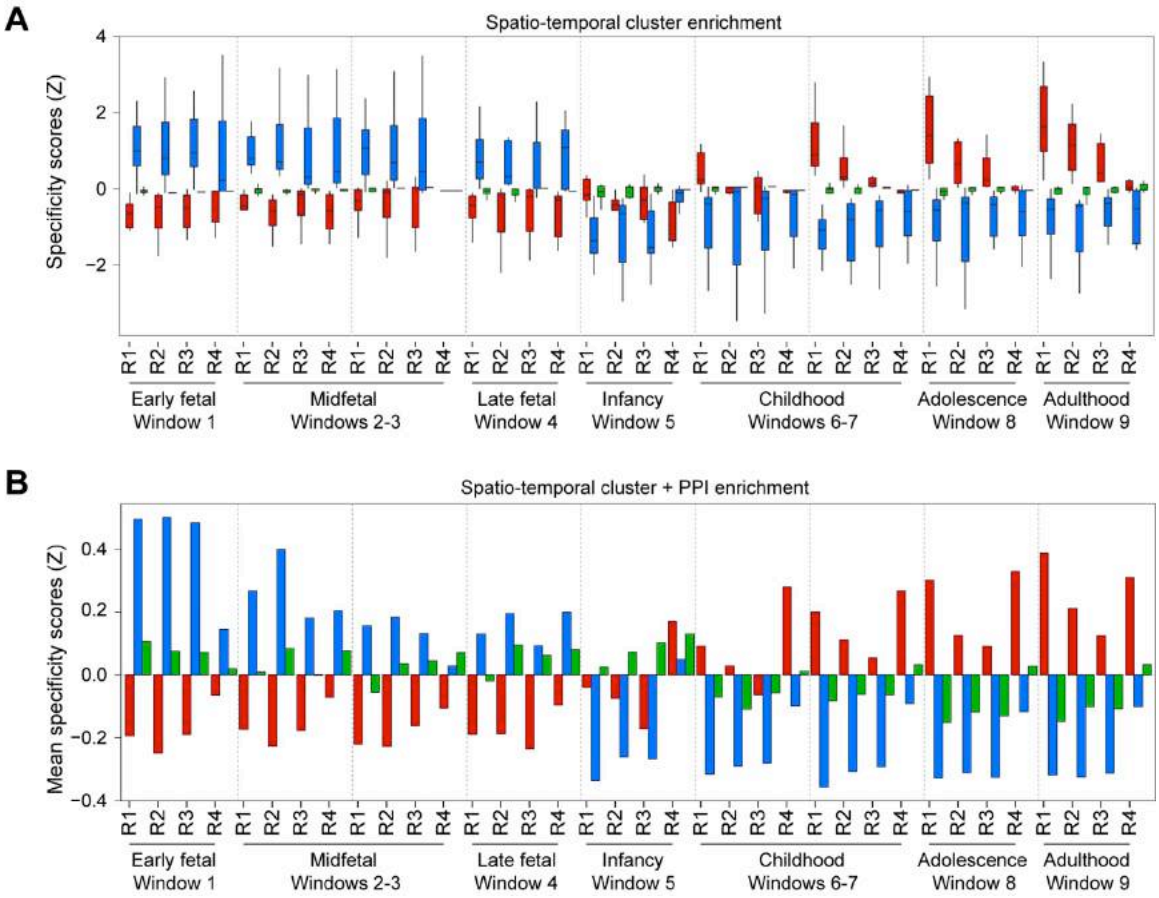


Figure 47: Spatiotemporal analysis reveals that cluster 1 is selectively enriched in the neocortex postnatally while cluster 2 is enriched across all brain regions prenatally. *A. Spatio-temporal specificity of expression using genes in clusters. B. The same analysis was repeated including first-degree interacting proteins (PPI). In both graphs, specificity scores for each cluster are represented using the same color code as previously, i.e. red for cluster 1, light blue for cluster 2 and green for cluster 3.*

We next focused on clusters 1 and 2 given their enriched expression in the human brain. To investigate the molecular functions within these two clusters, the top 1000 co-expressed genes for each gene in a cluster were identified and Gene Ontology (GO) enrichment analysis was performed for each set of genes. The lists of GO terms for each gene belonging to a cluster were then combined and summarized using REVIGO [652] to identify enriched GO terms for each cluster. For cluster 2, we found enrichment for gene regulation and mitosis processes. These processes were specifically driven by *BAZ1B*, *RFC2*, and *GTF2I* (Figure 48 A). These findings are in agreement with previous reports of transcription factor *GTF2I* has been reported to regulate some of the cognitive defects and separation anxiety in WS and ASD [608, 624]. By contrast, enriched processes for cluster 1 were synaptic transmission and mitochondrial respiration, the latter being a novel potential mediator of 7q11.23 phenotypes (Figure 48 B).

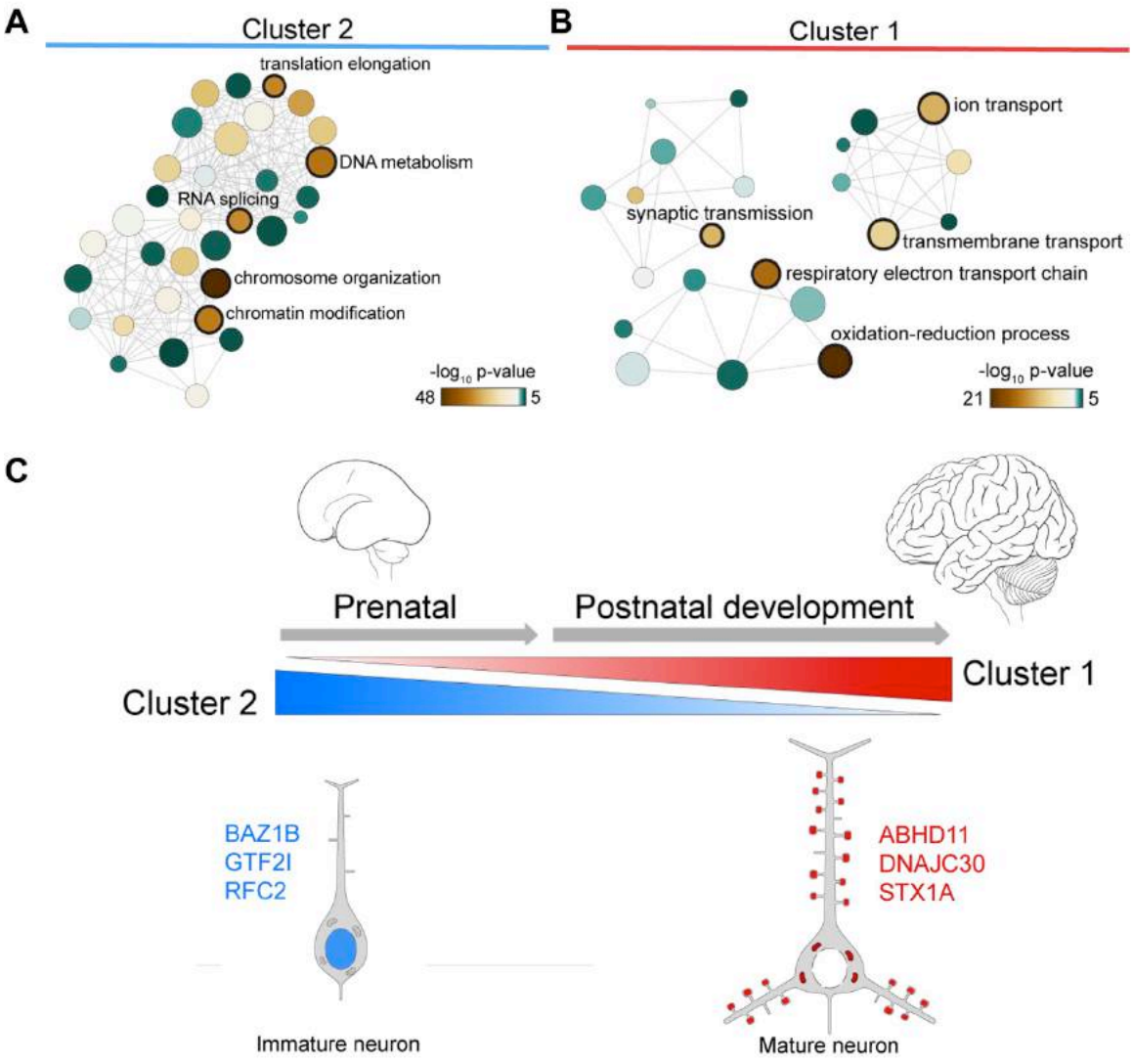


Figure 48: Top Functions of 7q11.23 brain-expressed genes are nuclear, synaptic, and mitochondrial. *A. Gene Ontology enrichment analysis performed on the top 1000 co-expressed genes for each gene in clusters 2. B. Gene Ontology analysis of the top 1000 co-expressed genes for each gene in clusters 1. Larger nodes indicate higher term frequency. C. Summary schematic showing the spatio-temporal contributions of cluster 1 and cluster 2 genes to neurodevelopmental processes.*

The identified genes driving the enrichment for these processes were *STX1A*, *ABHD11*, and *DNAJC30*. This was consistent using either BrainSpan's RNA-seq dataset or an independent exon array dataset from developmental human brains [148]. This confluence of synaptic transmission and mitochondrial respiration ontologies is consistent with the high metabolic demand and necessarily coordinated calcium flux needed for proper neuronal excitation [653]. Taken together, these analyses provide evidence for prenatal gene regulation, postnatal synaptic transmission and mitochondrial respiration as potential contributors to 7q11.23 phenotypes (Figure 48).

6.2 Spatiotemporal expression and cellular localization of DNAJC30

Considering these processes prospectively as therapeutic targets, prenatal gene regulation is not ideal given the several layers of upstream regulation and downstream targets. Furthermore, given that most 7q11.23 cases happen *de novo* and are not detected until after birth, that therapeutic option is impractical. These considerations lead to focus on genes comprised in cluster 1, as they exert their biggest effects in the neocortex after birth, when therapeutic options are more tractable following a diagnosis. In addition to synaptic transmission, GO enrichment analyses also identified mitochondrial processes as potentially novel mediators to 7q11.23 syndromes. Among the genes associated with mitochondria GO terms, it seemed interesting to investigate the previously uncharacterized gene DNAJC30 (*DnaJ (Hsp40) Homolog, Subfamily C, Member 30*) which belongs to the J-domain containing family of proteins that have a vast array of functions [654]. Therefore, DNAJC30 spatiotemporal expression was examined in more detail. Analysis of BrainSpan RNA-seq data proved that DNAJC30 mRNA is expressed throughout the developing human brain with its levels steadily increasing as development proceeds (Figure 49 A-B), consistent with postnatal enrichment of cluster 1 genes in the neocortex (Figure 47 A-B). The developmental increase of DNAJC30 protein was confirmed by immunoblotting lysates from human superior temporal cortex (STC, Brodmann area 22) (Figure 49 C-D), using that area because it was previously shown to have reduced asymmetric folding in WS patients [655]. The onset of appreciable levels of DNAJC30 protein was found during the perinatal period, which coincides with the peak period of neocortical dendritogenesis and synaptogenesis [15] (Figure 49 E).

To identify neocortical cells that express *DNAJC30*, we performed *in situ* hybridization in the human superior temporal cortex (STC) and detected mRNA in virtually all excitatory projection pyramidal neurons and many interneurons, with the strongest signal noticeably in large pyramidal neurons of layer (L) 3C (Figure 50 A). The increased DNAJC30 signal in large layer 3C pyramidal neurons of the STC was also confirmed using immunohistochemistry (Figure 50 B). In contrast, DNAJC30 strongest immunoreactivity in the human primary motor cortex was detected in large pyramidal neurons of layer (L) 5B (Betz cells) (Figure 50 C), but not layer 3 pyramidal neurons, which are smaller in size compared to pyramidal neurons in human STC and other association areas [656, 657]. Interestingly, Betz cells are among the largest neurons in the nervous system [656] and have the longest axons in the body, which extend to the spinal cord and form the corticospinal tract, which is altered in Williams syndrome [591]. Additionally, pyramidal neurons in L3C of the association areas, as STC, have the second largest somata in the NCX and form either ipsilateral or contralateral cortico-cortical projections, the latter of which comprise the anterior commissure or the corpus callosum, both of which are smaller in size in Williams syndrome [591, 593, 656]. A similar laminar expression pattern was also observed in the adult neocortex of rhesus macaque (*Macaca mulatta*) (Figure 50 D). Co-immunofluorescent staining of macaque STC showed that DNAJC30 is present in virtually all neurogranin-(NRGN) positive pyramidal neurons (Figure 50 G), as well as glutamic acid decarboxylase 1-(GAD1) positive inhibitory neurons (Figure 50 H). In contrast, there was a paucity of DNAJC30 signal in GFAP-positive neocortical astrocytes or CNP-positive oligodendrocytes, suggesting that DNAJC30 is not expressed in these cells or is below our level of detection (Figure 50 I-L).

Next, high-resolution immunogold electron microscopy (EM) was performed to precisely identify the subcellular localization of neocortical DNAJC30. In both dendrites and somata of pyramidal neurons, gold particles labeled the inner mitochondrial membrane (Figure 51 D-F), building upon recent high throughput studies that detected DNAJC30 in the mitochondrial proteome [658, 659] and supporting our hypothesis that mitochondria malfunction may contribute to 7q11.23 disorders.

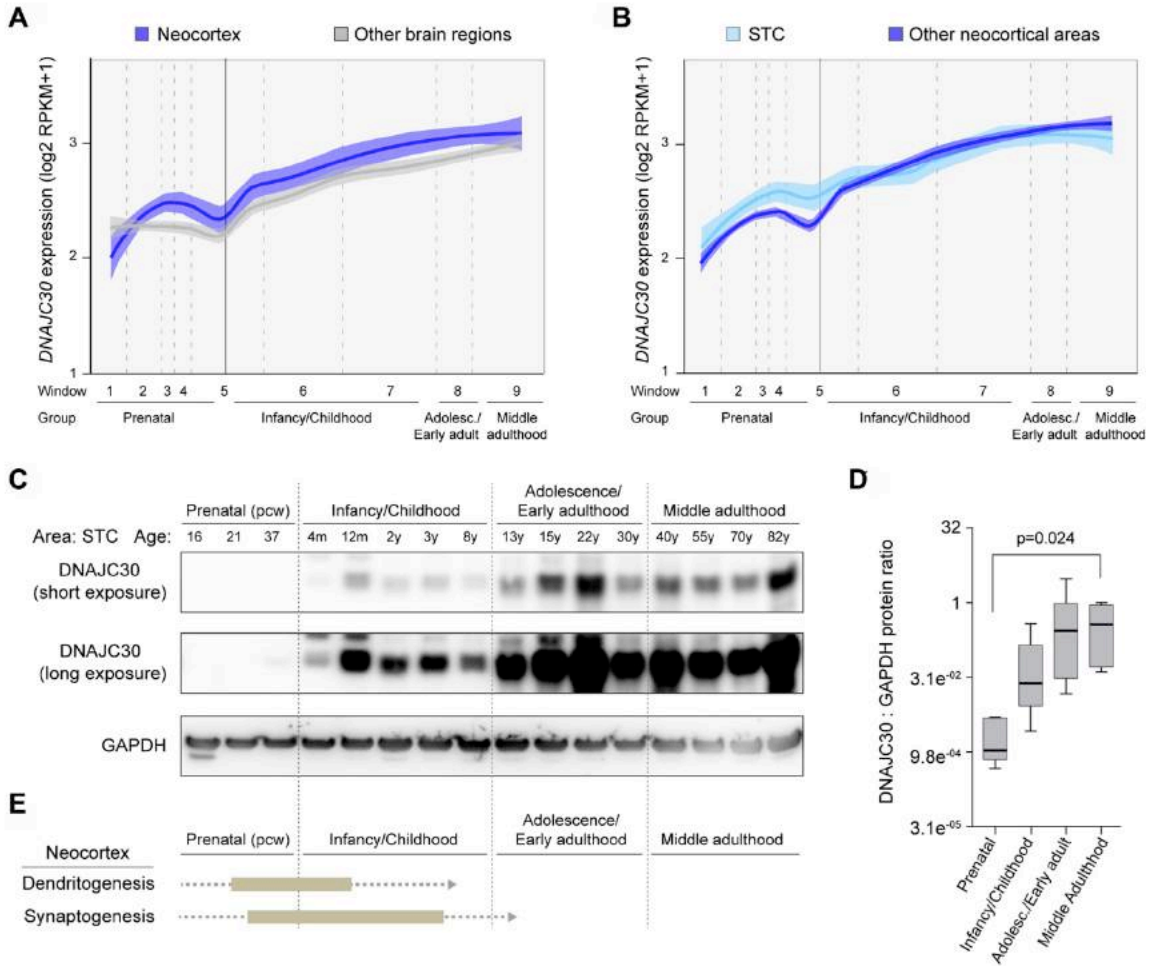


Figure 49: DNAJC30 expression increases during human postnatal brain development in neocortical neurons. *A. BrainSpan-generated expression of DNAJC30 in the neocortex compared to other brain regions. B. BrainSpan-generated expression of DNAJC30 in the superior temporal cortex (STC) compared to other neocortical areas. C. Western blots from human superior temporal cortex (STC) lysates, monitoring DNAJC30 expression from prenatal periods to late adulthood. Legend: pcw = post-conceptional weeks, m = months, y = years. D. Quantification of Western blot data reported in C, plotted as 5th-95th percentile. E. Schematic of the neurodevelopmental processes overlapping with DNAJC30 expression.*

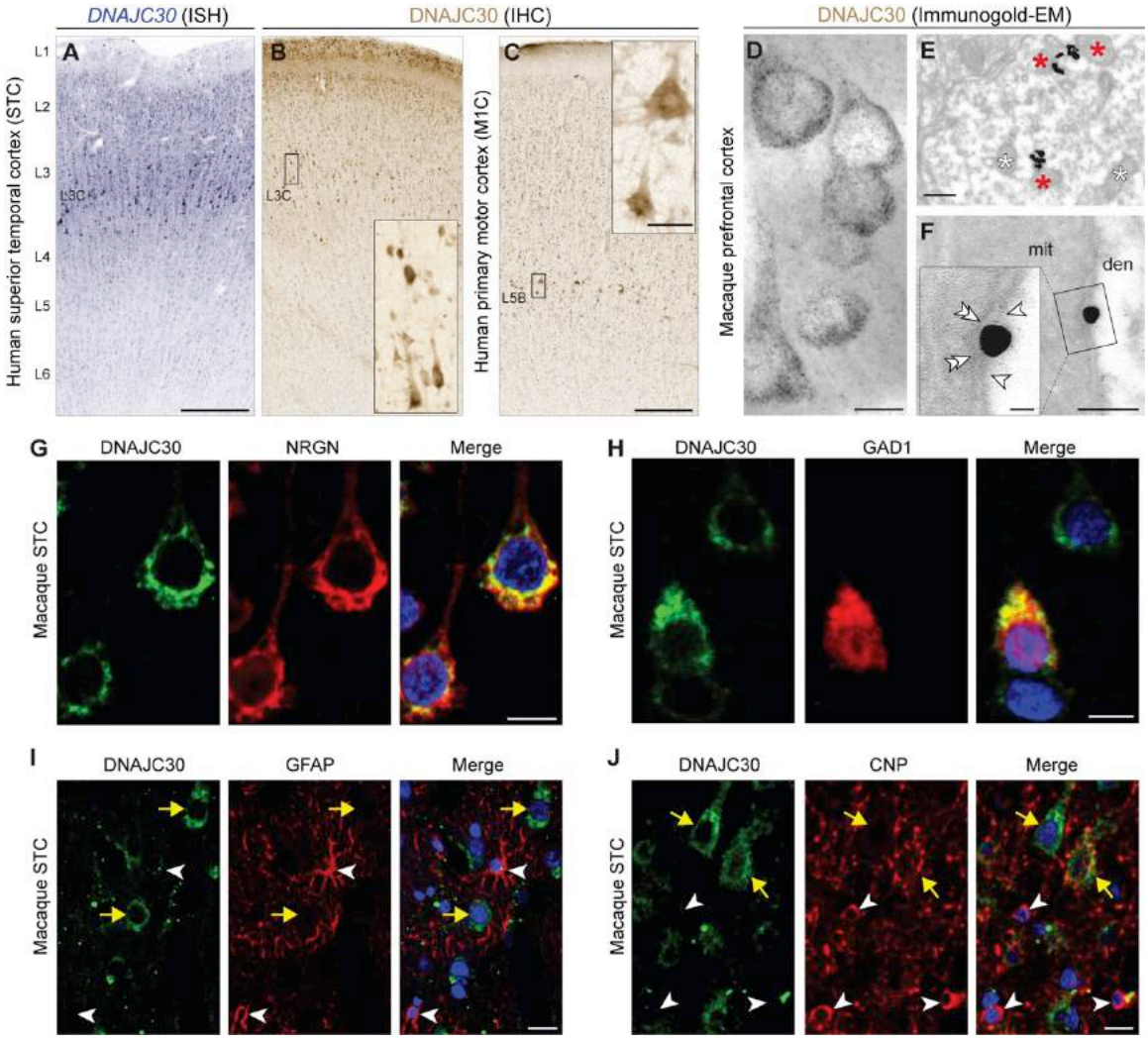


Figure 50: Anatomical and cellular localization of DNAJC30. *A. In situ hybridization for DNAJC30 in human superior temporal cortex (STC), scale bar: 250 µm. B. Human STC immunolabeled against DNAJC30 and visualized with DAB, scale bar: 500 µm. C. Human primary motor cortex (M1C) immunolabeled against DNAJC30 visualized with DAB, scale bar: 500 µm (inset scale bar: 50 µm). D-E. High-resolution immunoglod electron microscopy images of macaque dorsolateral prefrontal cortex labeled using anti-DNAJC30 and silver-enhanced 1.4nm gold. Red and white asterisks denote DNAJC30-reactive and unlabeled mitochondria, respectively. Scale bar D: 10 µm, scale bar E: 200 nm. F. High power electron micrograph captures DNAJC30 in the inner mitochondrial membrane (double arrowheads) where the outer mitochondrial membrane is fractured (arrowheads). Abbreviations: mit, mitochondrion; den, dendrite. Scale bar: 100nm (inset scale bar: 20nm). G. Co-immunofluorescence in macaque STC shows DNAJC30 is expressed in neurogranin-(NRGN) positive excitatory neurons, scale bar: 10 µm. H. Co-immunofluorescence in macaque STC shows DNAJC30 is expressed also in glutamic acid decarboxylase-(GAD1) positive inhibitory neurons, scale bar: 10 µm. I. No appreciable DNAJC30 signal was detected in GFAP-positive astrocytes, scale bar: 20µm. J. No appreciable DNAJC30 signal was detected in CNP-positive*

oligodendrocytes, scale bar: 20μm. In I-J, yellow arrows indicate (pyramidal) neurons, and white arrowheads indicate non-neuronal cell types.

Overall, these results confirmed that DNAJC30 expression in the human neocortex begins around birth steadily increasing with age and is restricted to neurons. Because of its dynamic expression during brain development, its selective neuronal expression in the neocortex and unknown function, combined with the motivation to identify unexplored mechanisms of 7q11.23 syndromes, the neurobiology of *DNAJC30* was investigated in greater detail.

6.3 Immunocytochemical validation of iPSCs-derived neocortical neurons to investigate 7q11.23 dosage-dependent dysregulation

In addition to the four human iPSCs control lines used also in the previous study aimed at validating the reproducibility and efficiency of the neocortical differentiation protocol in generating authentic neuronal populations at different *in vitro* differentiation stages, one clone for each of the nine iPSCs lines comprising 4 lines derived from patients affected by Williams syndrome (WS), 3 lines generated from 7q11.23 duplication syndrome individuals (7q11DUP) and 2 lines obtained from Williams syndrome patients unaffected relatives have been differentiated into neocortical neurons by applying the chemically defined monolayer-based differentiation protocol extensively characterized in the laboratory. To confirm the successful generation of iPSCs-derived pyramidal neurons populating both deep and upper neocortical layers accompanied by inhibitory interneurons at the end the differentiation paradigm, we have monitored through immunocytochemistry the expression of the same pan-neuronal and layer-specific markers adopted in the previous comprehensive protocol validation study. Also in this research, we examined the expression of the vesicular glutamate transporter 1 (VGLUT1) to verify the generation of excitatory pyramidal neurons after 80 days of *in vitro* differentiation (Figure 51 C,C',C'') and the inhibitory neurotransmitter γ-aminobutyric acid (GABA) to validate the presence of inhibitory interneurons (Figure 51 F,F',F''). As pan-neuronal markers we selected three different cytoskeletal proteins including the microtubule element βIII-tubulin that is expressed in both immature and postmitotic neurons, neurofilament light (NF-L) which is expressed during axonal growth and the microtubule-associated protein 2 (MAP2) that is selectively expressed by postmitotic neurons. To discriminate the presence of neocortical neurons belonging to different layers, we evaluated the expression of layer specific transcription factors including TBR1 to detect the presence of layer 6 corticothalamic projection neurons (Figure 51 A,A',A''), CTIP2 to identify layer 5 subcerebral projection neurons (Figure 51 B,B'B''), CUX1 to distinguish upper layers neurons (Figure 51 D,D',D'') and SATB2 to assess the presence of callosal projection neurons (Figure 51 E,E',E''). Furthermore, we investigated the presence in culture of minor populations of glial cells by examining the expression of the intermediate filament protein glial fibrillary acidic protein (GFAP) (Figure 51 F,F',F''). All the 13 cell lines examined were found positive for all the above mentioned markers, indicating that disease-related genotypes did not completely impaired the differentiation toward any of the major neocortical neuronal subtypes investigated (Figure 51).

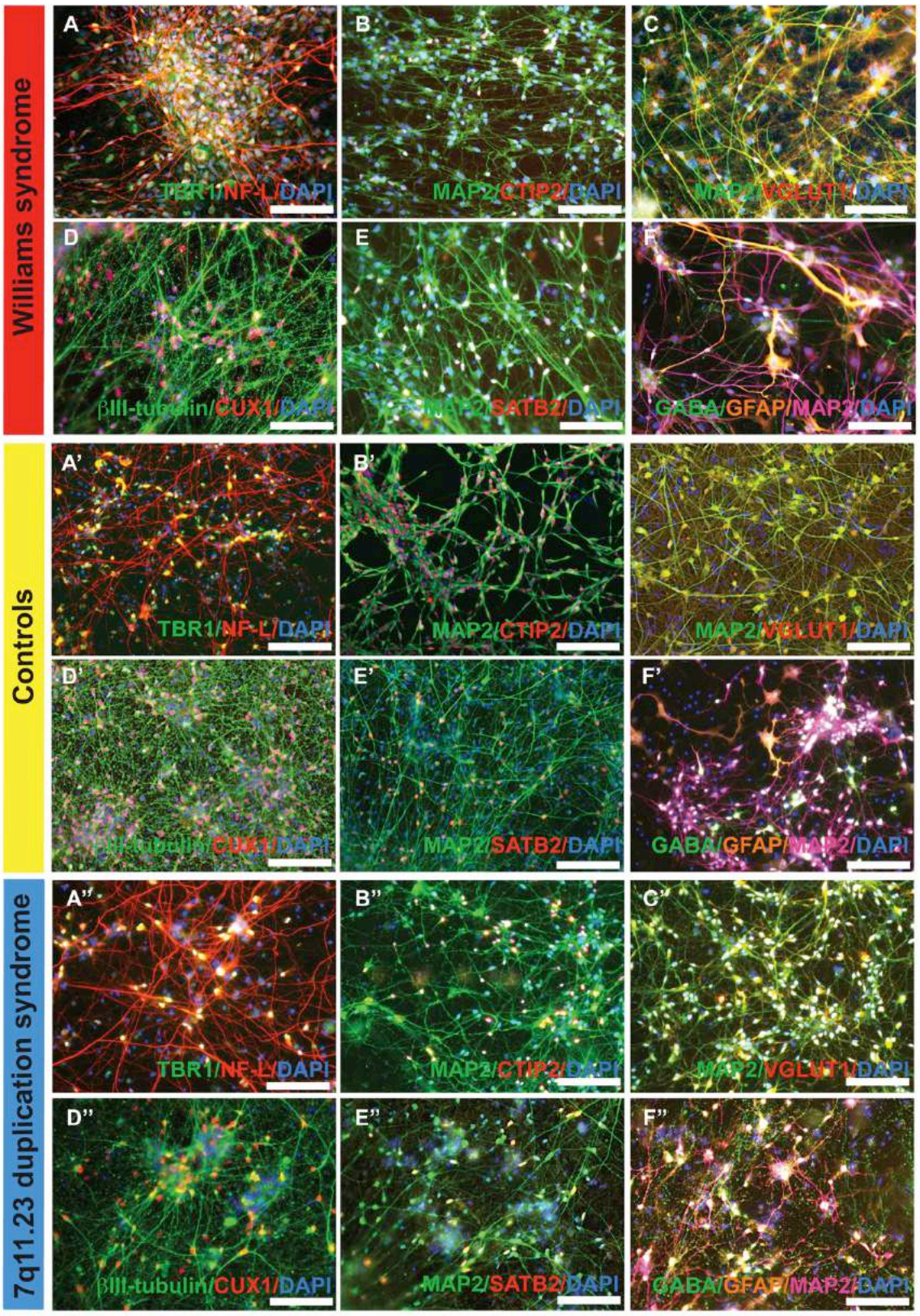


Figure 51: Immunocytochemical characterization of iPSCs-derived neocortical

neurons generated at the end of the in vitro differentiation paradigm. *In the top panel (A-F), middle panel (A'-F') and bottom panel (A''-F'') are reported representative images obtained from Williams syndrome, controls and 7q11.23 duplication syndrome iPSCs-derived neocortical neurons respectively. A-A'-A''. Representative images showing the presence in culture of layer 6 TBR1-positive pyramidal neurons. B-B'-B''. Representative pictures demonstrating the derivation of layer 5 CTIP2-positive pyramidal neurons. C-C'-C''. Representative images indicating the successful generation of excitatory pyramidal neurons as proven by the widespread positivity for VGLU1 in combination with MAP2. D-D'-D''. Representative pictures attesting the appearance of upper layer pyramidal neurons through CUX1 and MAP2 expression. E-E'-E''. Representative images verifying the presence of SATB2-positive callosal projection neurons. F-F'-F''. Representative pictures indicating that the successful generation of also inhibitory interneurons and minor proportions of glial cells as demonstrated by GABA and GFAP positivity.*

6.4 Morphological analysis of iPSCs-derived neocortical neurons from Williams syndrome patients and healthy subjects

iPSCs-derived neocortical neurons generated from four patients affected by Williams syndrome and three healthy control lines in combination with human post-mortem brains tissue have been used to validate the neuronal morphological abnormalities observed in Dnajc30 knockout mice.

Despite Williams syndrome is an hemizygous genetic condition, it was decided to investigate the homozygous KO mouse to provide the necessary resolution to understand Dnajc30 function. Validation of Dnajc30 decreased expression in heterozygous (Het) and knockout (KO) mice was confirmed by droplet digital PCR (ddPCR) (Figure 52 A). Macroscopic measurements revealed that Dnajc30 KO mice display ~20% decrease in body weight beginning at postnatal days (P) 9-10, persisting until at least 2 months old. Similarly, extracted brains from Dnajc30 KO mice were also ~20% smaller compared to wild type (WT) littermates. To investigate the effects of Dnajc30 on neuronal morphology, Golgi staining was performed on neocortical sections of Dnajc30 knockout mice and then Neurolucida was used to reconstruct the neuronal morphology and quantify several parameters of the pyramidal neurons' dendritic tree. Morphometric analyses on the principal apical dendrite showed that Dnajc30 KO mice have fewer spines (Figure 53 B), while the number of apical nodes, ends, and total length were also significantly reduced by roughly 50% % compared to wild types (WT) (Figures 53 C-E). Additionally, the number of basal dendrites (Figure 53 G), as well as the number of basal nodes, ends, and total dendritic length (Figures 53 H-J) were significantly reduced in KO mice, which also have smaller soma area (Figure 53 F). These findings suggest an essential role of Dnajc30 in promoting neocortical neurons dendritic arborization and morphological development. Furthermore, these results are also consistent with those found in mice hemizygous for all 26 genes comprised in 7q11.23 CNV, which exhibit craniofacial and behavioral defects, and like *Dnajc30* KO mice, have smaller brains and increased neuronal packing which reflects smaller neuron size.

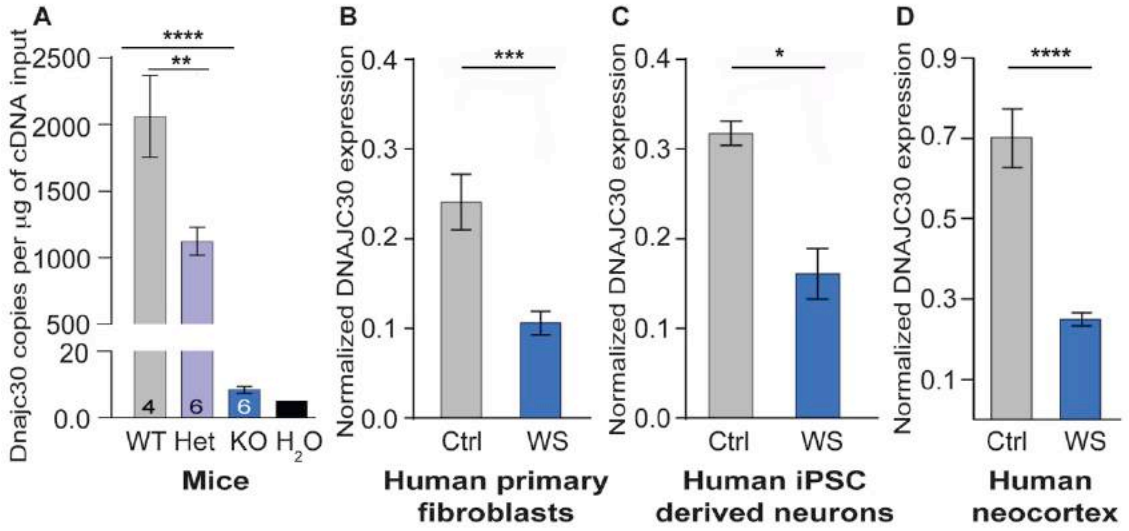


Figure 52: Dnajc30 reduced expression in heterozygous and knockout mice, Williams syndrome cells and brain tissues *A. ddPCR was used to confirm Dnajc30 reduced expression in Dnajc30 heterozygous and KO mice. B. Expression ratio (DNAJC30:TBP) in human fibroblasts (Ctrl, n=3; WS, n=9). C. Expression ratio (DNAJC30:TBP) in human iPSC-derived neurons (Ctrl, n=3; WS, n=4). D. Expression ratio (DNAJC30:TBP) in human neocortical samples (n=3 for each genotype).*

To extend these findings to humans, we performed the same analyses on healthy controls and Williams syndrome iPSCs-derived neocortical neurons generated at the end of the differentiation protocol and human brain tissues. Before examining the morphology of human cell cultures, we confirmed using digital droplet PCR (ddPCR) that Williams syndrome iPSCs-derived neurons had a decreased expression of DNAJC30 compared to controls of about 60% (Figure 52 C). To perform the morphometric analyses on iPSCs-derived neurons, cells were plated at low density and immunostained for the cytoskeletal marker MAP2 that was used for morphological quantifications along with CTIP2 (BCL11B), that was employed as a marker to distinguish layer 5 neocortical neurons (Figure 53 K). Also in this case Neurolucida was employed to reconstruct the neuronal morphology and quantify several parameters of the pyramidal neurons' dendritic tree. These analyses confirmed that the number of nodes and the total dendritic length were significantly reduced in Williams syndrome neurons as seen in Dnajc30 KO neurons (Figure 53 M-N), while the soma of Williams syndrome neurons was decreased but its reduction did not reached statistical significance (Figure 53 L). After having confirmed DNAJC30 decreased expression also in post-mortem Williams syndrome neocortical tissues through digital droplet PCR (Figure 52 D), a modified Golgi-Kopsch stain was used to compare pyramidal neurons morphologies *in situ* (Figure 53 O). Analyses of Williams syndrome and healthy controls neuronal morphology performed on layer 5 human primary motor cortex (M1C) confirmed that Williams syndrome pyramidal neurons were characterized by a statistically significant reduction in apical dendrite nodes, ends, and total length (Figure 53 P-R), and displayed a trending reduction between 5%-15% in soma area and basal dendritic architecture (data not shown).

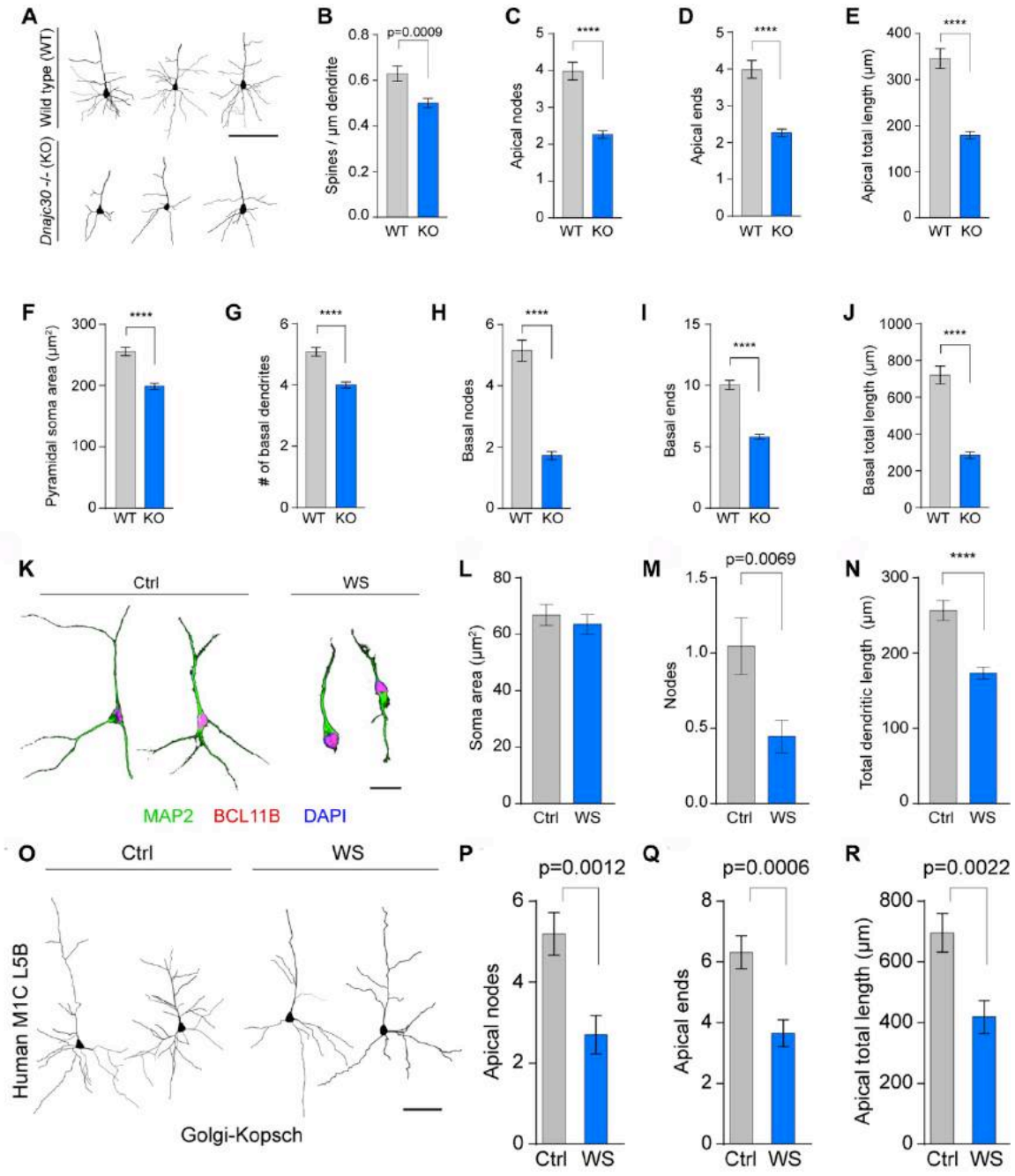


Figure 53: Morphological analysis of Dnajc30 knockout mice and Williams syndrome (WS) pyramidal neurons. *A. Representative images of Golgi-Cox stained WT and Dnajc30 KO pyramidal neurons illustrating the markedly reduced complexity of the dendritic arborization of Dnajc30 KO neocortical neurons. Scale bar: 100 mm. Neurons were analyzed using Neurolucida neuron tracing system. B-J Bar plots representing the results of the morphological analysis of spine density (B), apical and basal dendritic nodes (C, H), ends (D, I), and total dendritic length (E, J), as well as soma size (F) and number of basal dendrites (G), performed on WT and Dnajc30 KO neocortical neurons. All the morphometric parameter analyzed resulted significantly*

*reduced in Dnajc30 KO mice (WT: n = 64 neurons, 4 mice; Dnajc30: n = 93 neurons, 6 mice). K. Representative images of human iPSC-derived neurons immunofluorescently labeled for morphological analysis, scale bar: 40 μm. L-N. Quantification of soma area, nodes, and total dendritic length, respectively, of MAP2- and CTIP2-positive neurons. While the soma of WS cells was only moderately reduced compared to controls, the number of nodes and total dendritic length of WS neurons were significantly reduced (Ctrl: n = 45, 3 lines; WS: n = 45, 4 lines). O. Representative images from Golgi-stained human primary motor cortex (M1C) deep layer neurons in Ctrl (n=3) and WS (n=3), scale bar: 100 μm. P-R. Quantification of neuronal apical architecture from Golgi-stained human M1C deep layer neurons. All the morphometric parameter analyzed resulted significantly reduced in Williams syndrome patients compared to controls (Ctrl, n=19 neurons/3 specimens; WS, n=22 neurons/3 specimens). . Data are expressed as mean ± SEM and compared using unpaired t-test, * p < 0.05, ** p < 0.01, *** p < 0.005, **** p < 0.001.*

Overall, the comparison of the morphological abnormalities detected in Dnajc30 knockout mice, Williams syndrome iPSCs-derived neocortical neurons and human brain tissues confirms that the same morphological abnormalities observed in Dnajc30 knockout mice are also a distinctive cellular hallmark detected in patients affected by Williams syndrome, suggesting an essential role of Dnajc30 in mediating proper morphological maturation of neocortical neurons also in humans.

Discussion

Studies of human neurodevelopment are crucial to understand the regulatory mechanisms that give rise to the peculiar structure and functions of the human brain and to investigate the molecular basis of many neurological and psychiatric disorders that have a developmental origin. In this context, human pluripotent stem cells (PSCs) have attracted a lot of attention as their differentiation potential can be directed to generate with high efficiency the neural cell types of interest. Their competence to recapitulate in a dish the main *in vivo* neurodevelopmental events offers the possibility to reach an unprecedented spatial and temporal precision in the selection and analysis of specific neurodevelopmental stages, impossible to achieve employing other model systems such as human brain tissues and animal models. Among PSCs, especially induced pluripotent stem cells (iPSCs) have emerged as a powerful tool for disease modeling as they fully recapitulate the genetic background of diseased individuals allowing to bridge the gap between the genetic lesions and the clinical data. Furthermore, iPSCs can be easily experimentally manipulated through genetic engineering techniques to functionally validate the biological targets of interest. These unique advantages have fostered the development of many 2D and 3D human PSCs-based differentiation protocols to generate the neuronal populations of interest.

In this study, in order to generate *in vitro* human neocortical neurons, iPSCs have been subjected to a monolayer chemically-defined PSCs-based directed neocortical differentiation protocol. The neocortical differentiation protocol adopted integrates several previously published neocortical differentiation strategies [176, 375, 376] in an optimized synthesis to generate with high efficiency excitatory pyramidal neurons belonging to all neocortical layers and inhibitory interneurons at the end of the differentiation paradigm, accompanied by minor populations of glial cells.

Given the importance to rely on a solid and highly reproducible differentiation protocol that generates authentic neocortical neurons *in vitro* with high efficiency before applying it as a model system of human neurodevelopmental disorders, in the first phase of this study we have performed a comprehensive transcriptional, cellular and physiological characterization of the *in vitro* neurodevelopmental paradigm. Subsequently, we applied this extensively characterized differentiation paradigm to model *in vitro* two human neurodevelopmental disorders caused by symmetrical copy number variation (CNVs) of the Williams-Beuren syndrome chromosome region (WBSCR) located on the long arm (q) of chromosome 7 at position 11.23 (7q.11.23 locus).

Refinement and extensive characterization of a pluripotent stem cells based differentiation protocol for the generation of authentic human neocortical neurons *in vitro*

A comprehensive transcriptional, cellular and physiological characterization of the *in vitro* neocortical differentiation protocol adopted has been conducted on four iPSCs control lines derived from human fetal skin biopsies. Neocortical cells have been generated from two separate clones of each iPSCs line to better evaluate the reproducibility of the differentiation paradigm and to finely discriminate the degree of variability due to inter-line and intra-line differences.

Part II: Results

The transcriptional dynamics regulating *in vitro* neocortical differentiation have been investigated by performing RNA-sequencing (RNA-seq) at both population and single-cell level. Initially, we aimed to identify the contribution of different variables to the global transcriptional dynamics underlying *in vitro* neocortical differentiation. In order to answer this question we performed principal component analysis (PCA) on our RNA-sequencing dataset. Covariate analysis on the principal components revealed that the variable having the biggest effect on variability in our data is the differentiation stage. This finding was the first evidence confirming the correct progression of neocortical differentiation throughout the *in vitro* protocol and contributed to validate the quality of our data demonstrating that they are unrelated to sequencing-dependent variables. PCA also enabled to start addressing the question of protocol reproducibility and to discriminate the contributions to data variability due to inter-cell line and intra-cell line differences. All the iPSCs lines and clones examined followed the same neurodevelopmental trajectory as evident by their clustering mainly based on the differentiation stage. This finding constitutes the first evidence of the high reproducibility of our differentiation paradigm. Moreover, the absence of major differences in the neurodevelopmental dynamics of different cell lines indirectly suggested the homogeneity of the starting iPSCs lines. Only minor differences in neurogenesis progression were detected between the different iPSCs clones beginning from intermediate *in vitro* neurodevelopmental stages. This observation can be explained by the different differentiations rates at which *in vitro* neurons are maturing that becomes more evident at later differentiation time points. Considering also the contained data dispersion, this finding might be ascribed to biological variability even if it is not possible to definitely exclude specific iPSCs clone-related contributions.

A second set of questions we wanted to answer was to assess how well the transcriptional profiles of our iPSCs-derived neuronal cells reflected those of neurons obtained from human brain samples and to infer to which in vivo neurodevelopmental period were related the different *in vitro* differentiation stages. To address these points, we reperformed PCA including in the analysis transcriptional data obtained from both iPSCs-derived neuronal cells and human prenatal brain tissue specimens comprised in the reference brain's transcriptome atlas BrainSpan, generated in the laboratory. This analysis revealed the close similarity between *in vitro* and *in vivo* neocortical neurons and allowed to determine to which in vivo neurodevelopmental period were related the different *in vitro* differentiation stages. Neuroepithelial progenitor cells generated after 10-14 days of *in vitro* differentiation corresponded to neuroepithelial cells lining the central cavity of the neural tube at Carnegie stage 12 (4 pcw), iPSCs-derived neurons generated at intermediate stages of the differentiation protocol clustered closely with human early fetal dorsal telencephalon specimens (8 pcw) and neocortical neurons obtained at the end of the differentiation protocol were distinguished by transcriptional profiles similar to those characterizing human early mid fetal cortical tissues (13 pcw).

To better investigate the transcriptional changes occurring along *in vitro* neocortical development, we performed differential gene expression analysis across the four selected *in vitro* differentiation stages that identified a total number of 28094 differentially expressed genes (DEGs). As many DEGs were overlapping across pairwise comparisons between the different differentiation stages, we examined their distribution. At first we explored the distribution of DEGs comparing the three representative differentiation time points to the pluripotent state (day 0) and subsequently we also investigated DEGs distribution between different *in vitro* differentiation time points,

always comparing more differentiated *in vitro* stages to less differentiated ones. Gene ontology enrichment analysis and KEGG pathway analysis performed on all DEGs lists provided the first confirmation of the successful generation of the expected cell types at each *in vitro* time point analyzed and of the progressive acquisition of mature neocortical neurons identities with the progression of the differentiation paradigm. To further corroborate these findings we also examined the functional organization of the transcriptional changes underlying *in vitro* differentiation by performing weighted gene co-expression network analysis (WGCNA), that identified 46 modules on our dataset. Many of these modules contained genes that had a clear peak of expression at specific *in vitro* differentiation stages and followed a definite developmental trajectory. In agreement with the cell types expected to appear during the differentiation paradigm, modules characterized by a peak of expression of their genes in the pluripotent state and a progressively decreased expression at subsequent neocortical differentiation stages were distinguished by top-ranking GO categories related to cell cycle, cell proliferation and mitosis. This finding reflected the active cellular proliferation typical of the pluripotent state and its progressively decreased rates along the differentiation protocol when more mature postmitotic neuronal phenotypes start to appear. The generation of neuroepithelial cells after 10-14 days of *in vitro* differentiation was substantiated by the enrichment in many GO terms related to the acquisition of telencephalic identities, cell differentiation and cell maturation in modules characterized by a peaked expression at the end of the neural induction phase. The presence in culture of neocortical progenitor cells was confirmed by the identification of top-ranking GO categories related to neural stem cells maintenance, neural plate regionalization and neurogenesis combined with terms associated to cell cycle, mitosis and Notch signaling pathway detected in modules distinguished by a peaked expression in the day 10-14 and day 36-42 differentiation temporal windows. In these same modules, the presence of neocortical neurons was also suggested by the enrichment in multiple categories related to axonal growth, dendrite development, synaptogenesis and voltage-gated channels. Furthermore, in agreement with the successful generation of mature neocortical neurons at later stages of *in vitro* neurogenesis, modules distinguished by an increased expression along the differentiation paradigm or selectively at the end of the differentiation protocol displayed a marked enrichment in GO categories related to neocortical neurons development, development of complex neuronal morphologies and acquisition of physiologically mature neuronal phenotypes.

To further reveal the distinct neuronal populations obtained at different differentiation stages, we also investigated the transcriptional signatures of more than 1000 iPSCs-derived neuronal cells at single cell level. This analysis confirmed that differentiating cells become neuroepithelial progenitors (NEPs) and apical radial glial cells (vRGCs) at end of the neural induction phase, indicating the correct lineage-commitment and the onset of *in vitro* neurogenesis. The detection of later-born neocortical progenitors populations as intermediate progenitor cells (IPCs) and outer radial glial cells (oRGCs), intermingled with excitatory and inhibitory neocortical neurons starting from intermediate differentiation stages, confirmed the correct differentiation progression and successful generation of neocortical neurons after approximately 40 days of *in vitro* differentiation. Moreover, the presence in culture of additional excitatory and inhibitory neocortical neuronal subtypes combined with excitatory and inhibitory neurons and neocortical progenitors detected also at earlier time points, further proves the generation of a more complete set of mature neocortical neuronal populations belonging to both deep and upper layers by the end of differentiation paradigm.

These transcriptional results indicating the generation of specific neuronal cell types along the *in vitro* differentiation paradigm were validated by monitoring the expression of a selected panel of informative cell-fate and cell-stage specific markers through immunocytochemistry. The acquisition of neuroepithelial identities by the majority of cells in culture at the end of the neural induction phase was validated by their positivity for the transcription factors PAX6, SOX1 and OTX2 and the cytoskeletal marker Nestin. Moreover, the widespread positivity for the transcription factor FOXG1 confirmed their telencephalic commitment. Immuocytochemical analysis performed at intermediate stages of differentiation enabled also to assess the successful generation of deep layer neocortical pyramidal neurons intermingled with neocortical progenitor cells. The positivity for the transcription factor TBR2 and the cytoskeletal protein Nestin verified the presence in culture of neocortical progenitor cells while the generation of immature neurons was established by the co-expression of the transcription factor SOX1 and the microtubule element βIII-tubulin, which is a marker of both immature and postmitotic neurons. The presence in culture of iPSCs-derived pyramidal neurons was validated by the co-expression of VGLUT1, a protein involved in synaptic vesicles glutamate uptake, and the cytoskeletal postmitotic pan-neuronal marker MAP2. Furthermore, the appearance of postmitotic deep layer pyramidal neurons was confirmed by the co-expression of the cytoskeletal protein MAP2 and the layer-specific transcription factors TBR1 and CTIP2 employed as markers of layer 6 corticothalamic projection neurons and layer 5 subcortical projection neurons, respectively. At the end of the *in vitro* differentiation paradigm, the highly efficient generation of pyramidal neurons populating both deep and upper neocortical layers and their morphological maturity was assessed by monitoring the expression of the same pan-neuronal and layer-specific markers previously examined to validate the presence of these cell types at intermediate differentiation stages. In addition, the generation of upper layer pyramidal neurons was validated by the expression of the transcription factor CUX1 and the derivation of callosal projection neurons was confirmed by SATB2 positivity. Furthermore, we also wanted to confirm the presence in culture of GABAergic interneurons and minor populations of glial cells, as detected from single cell RNA-sequencing analysis. The generation of these cell types was corroborated by the positivity for the inhibitory neurotransmitter γ-aminobutyric acid (GABA) and the cytoskeletal intermediate filament protein glial fibrillary acidic protein (GFAP). Interestingly, inspection of the pan-neuronal cytoskeletal markers βIII-tubulin and MAP2 revealed that iPSCs-derived neocortical neurons had acquired full morphological maturity by the end of the differentiation paradigm, as proven by their highly interconnected and arborized dendritic trees and axons projecting to distant targets. The development of complex neuronal morphologies was also confirmed through stainings for the additional cytoskeletal marker neurofilament light (NF-L) which highlighted the formation of extended axon bundles. Overall, the appropriate temporal expression profiles and the widespread positivity for the analyzed markers confirmed our transcriptional findings and provided further evidence that the applied neocortical differentiation protocol is an efficient system for the generation of both deep and upper layer pyramidal neurons and inhibitory interneurons starting from human iPSCs.

To further assess the authenticity and the physiological maturity of our iPSCs-derived neocortical neurons obtained at the end of the *in vitro* differentiation protocol, we monitored their electrophysiological properties through whole-cell patch-clamp recordings performed both in current- and voltage-clamp mode. The ability to initiate at

least one action potential (AP) in response to the injection of progressively increasing steps of depolarizing currents displayed by the majority of iPSCs-derived neocortical neurons demonstrated that our neocortical cultures had reached functional maturity after 80 days of *in vitro* differentiation. This finding was further confirmed by the presence of voltage-dependent inward and outward ion currents in the majority of recorded cells. To better assess the level of functional development of our iPSCs-derived neocortical cells, we also classified neurons on the basis of their firing properties in response to suprathreshold depolarizing steps of current. This analysis revealed that nearly 60% of *in vitro*-derived mature neurons were competent to generate and sustain trains of action potentials organized into characteristic repetitive firing patterns. The comparison of our data with reference values obtained from human primary neocortical cultures derived from 7 to 9 postconceptional week fetal brain tissues [647] revealed that a fewer number of mature iPSCs-derived neurons was able to fire APs with respect to primary cultures (iPSCs-derived mature neurons = 82.6%; primary cultures = 96.3%). However, our iPSCs-derived neurons that were functionally mature (66%) seemed even more physiologically evolved as suggested by the higher percentage of cells able to sustain repetitive firing (iPSCs-derived mature neurons = 56.5%, primary cultures = 14.8%). Moreover, the additional comparison of our results to data obtained from cortical plate (CP) brain slices of 16 to 22 gestational weeks (gw) [648] and postnatal neocortical slices [649] suggested that our mature iPSCs-derived neocortical neurons are more physiologically developed than CP neurons up to 22 gw (single AP: iPSCs-derived neurons =26.1%, CP neurons = 14.9 %; repetitive firing: iPSCs-derived neurons = 56.5%, CP neurons = 14.3%) but do not acquired the full functional maturity of postnatal neurons (repetitive firing = 100%). Given the differences of culture environment between cell cultures and brain slices, the results obtained from these last comparisons should be interpreted just as an additional indication of the level of functional development reached by iPSCs-derived neocortical neurons at the end of the differentiation protocol. The further analysis of APs amplitude and duration revealed that iPSCs-derived neurons generated APs characterized by lower spikes amplitudes and significantly higher spike durations compared to gold-standard reference values obtained from human postnatal neocortical brain slices. Also this finding indicates that the majority of our iPSCs-derived neurons are functionally active but have not yet reached the full functional maturity distinguishing human neocortical postnatal neurons in agreement with the other electrophysiological analyses and the transcriptional data estimating a correlation of our *in vitro*-derived neurons with 13 pcw neurons derived from human brain tissues. Even if our iPSCs-derived neocortical neurons do not reach postnatal levels of physiological development, their functional maturity was also corroborated by recordings of their excitatory and inhibitory spontaneous activity. The majority of mature neurons (69%) was able to generate spontaneous excitatory postsynaptic currents (sEPSCs) and a substantial subgroup of these neurons (38%) presented also spontaneous inhibitory postsynaptic currents (sIPSCs) confirming the generation of a functional network of pyramidal neurons and inhibitory interneurons at the end of the differentiation paradigm. Overall, this articulated set of electrophysiological recordings allowed to prove that iPSCs-derived neocortical neurons differentiate appropriately to excitable cells able to fire APs, develop mature firing properties and sustain excitatory and inhibitory spontaneous activity by the end of *in vitro* neurogenesis confirming the transcriptional results.

Taken together the results obtained from this extensive panel of analyses enabled an in-depth characterization of the transcriptional, cellular and physiological signatures of the neuronal populations obtained at different differentiation stages, providing a comprehensive assessment of the neocortical differentiation paradigm adopted with a previously unachieved integration of different levels of analysis and resolution to our knowledge. Data obtained from a wide collection of samples and with different techniques were in agreement with each other corroborating the quality of our findings. Therefore, through this study we were able to thoroughly validate the reproducibility of the differentiation protocol and its efficiency and reliability in generating physiologically mature authentic neocortical pyramidal neurons belonging to both deep and upper neocortical layers and inhibitory interneurons by the end of the differentiation paradigm.

Human neocortical neurons derived from induced pluripotent stem cells as a model system to investigate the effects of 7q11.23 gene-dosage dependent dysregulation

After having established the quality of our differentiation strategy, we have applied it to model *in vitro* two human neurodevelopmental disorders caused by symmetrical copy number variations (CNVs) of the Williams-Beuren syndrome chromosome region (WBSCR). WBSCR is located on the long arm (q) of chromosome 7 at position 11.23 (7q.11.23 locus) and comprise 26-28 genes. CNVs of 7q11.23 locus are especially interesting as the two disorders due to the deletion (Williams syndrome, WS) and duplication (7q.11.23 duplication syndrome, 7q11DUP) of this region display behavioral and cognitive phenotypes characterized by both similar features and symmetrically opposite traits. Moreover, association of 7q11DUP to complex neurodevelopmental disorders, such as autism spectrum disorder (ASD) and schizophrenia, while WS is a well-characterized syndrome without clear overlap to any complex neurodevelopmental disorder, make the study of this locus particularly interesting to investigate the molecular mechanisms unique to each clinical condition, common to both syndromes and shared with other complex neurodevelopmental disorders. Furthermore, in order to identify the most significant molecular mechanisms that may serve as therapeutic targets, it is also highly relevant to determine single genotype-phenotype correlations and to investigate the molecular mechanisms behind each genes' contribution to the specific disease-related neuroanatomical and cellular abnormalities of neocortical neurons that have been observed in patients affected by these two disorders.

As the human spatiotemporal expression patterns of the 26-28 protein coding genes comprised in the 7q11.23 CNV have not been analyzed in great detail, we first performed system-level analyses utilizing human brain transcriptomic and protein-protein interaction datasets to investigate the molecular coordination and spatio-temporal expression specificity of genes comprised in this locus. Next, top functions for the identified modules of highly co-expressed genes were determined performing Gene Ontology enrichment analysis. Taken together, these analyses provided evidence for prenatal gene regulation, postnatal synaptic transmission and mitochondrial respiration as potential contributors to 7q11.23 phenotypes.

Building on this knowledge, two different lines of research were pursued. One was aimed to delineate the contribution of single genes to specific phenotypes of 7q11.23

syndromes. The other had as objective to investigate with enhanced resolution the spatiotemporal activity of 7q11.23 gene dosage imbalances, highlighting neuronal subtypes specific genes and gene networks having the most statistically significant relationship to these disorders, through single cell RNA-sequencing analysis.

To this aim, we have generated several iPSCs lines from a large cohort comprising four WS individuals, three 7q11DUP patients and two unaffected relatives of individuals presenting with typical WS, in addition to four control lines derived from unrelated individuals that were also used in the previous extensive protocol characterization study. The large size of our cohort and its composition, including in the control group also unaffected relatives of WS patients that share a similar genetic background with affected individuals, enables to identify disease-relevant molecular mechanisms with increased confidence.

Successful reprogramming of skin fibroblasts into iPSCs has been achieved by applying a non-integrating reprogramming method based on the transfection of episomal plasmid vectors slightly modified from Okita et al. [544]. The selected reprogramming approach allowed to generate iPSCs with high efficiency and to reduce the risk of residual transgene expression and insertional mutagenesis. Pluripotency validation of all 13 human iPSCs lines has been successfully accomplished. Briefly, the absence of residual transgene expression in our iPSCs clones was confirmed by analyzing their genomic DNA through PCR using a primer pair designed on a sequence common to all the nucleofected plasmids. Only transgene-free clones were further characterized to rule out the presence of chromosome abnormalities through karyotype analysis. To assess that iPSCs clones met the standard criteria for pluripotency definition, we checked through immunofluorescence the expression of a panel of classical pluripotency markers including the transcription factors OCT4, Nanog, SOX2 and the plasma membrane surface markers SSEA-4 and TRA-1-60. Moreover, human iPSCs differentiation potential was evaluated *in vivo* and for some lines also *in vitro* by performing teratoma formation assay and embryoid bodies formation assay respectively. Both assays confirmed that selected iPSCs clones were able to differentiate into cells belonging to all three germ layers. In addition to validate the pluripotency of our iPSCs clones, these results demonstrated that 7q11.23 CNVs do not prevent the successful reprogramming of patient-derived somatic cells into iPSCs, in agreement with other studies [470, 621].

One clone of each iPSCs line has been subjected to the previously extensively characterized neocortical differentiation protocol. The successful generation of human neocortical neurons at the end of the differentiation paradigm has been validated by the correct expression of the same panel of informative cell fate and cell-stage specific markers applied also in the protocol integrative analysis. These findings confirm the quality of our differentiations and indicate that 7q11.23 CNVs do not completely impair the generation of mature pyramidal neurons, inhibitory interneurons and minor populations of glial cells after protracted periods of *in vitro* differentiation.

To investigate with enhanced resolution if and to which extent specific neuronal subtypes are affected by 7q11.23 gene-dosage imbalances and to identify the molecular mechanisms having the most statistically significant relationship to these disorders, single-cell RNA-sequencing has been performed on samples of iPSCs-derived neocortical neurons obtained from all the cell lines included in the cohort under study. Analysis of single-cell RNA-sequencing data and functional validation of the identified

biological targets are still ongoing.

At the cellular level, we verified that 7q11.23 gene-dosage imbalances impair neuronal morphologies. WS and control human iPSCs-derived neocortical neurons plated at low density were immunostained for the transcription factor CTIP2 that was adopted as a marker of neocortical pyramidal neurons and the pan-neuronal cytoskeletal marker MAP2, which was employed to perform the morphological quantifications using Neurolucida neuron tracing software. Morphometric analysis revealed that the total dendritic length and number of branching nodes were significantly reduced in WS neurons while the soma of WS neurons was reduced but the decrease in its size did not reached statistical significance. Similar results indicating diminished neuronal morphologies and less complex neuronal architecture were confirmed also by performing the same morphometric analyses in the primary motor cortex of healthy individuals and Williams syndrome patients. Results obtained in humans were compared to findings obtained from a Dnajc30 knockout mice in the framework of a study aimed at investigating the specific contributions to the onset of disease-related neural phenotypes of the critical and functionally uncharacterized gene Dnajc30 comprised in 7q11.23 CNV. The morphology of Dnajc30 knockout mice and wild types neocortical neurons was investigated using Golgi-stained sections and also in this case Neurolucida software was used for neuron tracing and subsequent analyses. Morphometric analysis showed that apical and basal dendrites of Dnajc30 knockout mice neocortical neurons were characterized by a significantly reduced total length and a significantly decreased number of branching nodes and ends compared to wild types. These findings are consistent with neuronal morphological alterations observed in mice hemizygous for all 26 WBSCR genes [221] and suggest an essential role of Dnajc30 in promoting neocortical neurons dendritic arborization and morphological development. Overall, these findings confirm in humans many neuronal morphological abnormalities observed in Dnajc30 knockout mice, suggesting an essential role of DNAJC30 in mediating proper morphological maturation of neocortical neurons also in humans and a contribution DNAJC30 to specific aspects of the Williams syndrome phenotype.

PART III

Materials and Methods

Human Tissues

Human tissues procurement

Human fetal brain tissue samples were obtained from tissue collections of the Newcastle Brain Tissue Resource, the Department of Neurobiology at Yale University School of Medicine and NIH Neurobiobank. Fetal tissues were collected from patients that requested pregnancy terminations. All procedures were approved by the local research ethical committees, the Research Services division of Yale University and the National Institutes of Health. Tissues were handled in accordance with ethical guidelines and regulations for the research use of human brain tissue set forth by the WMA Declaration of Helsinki (http://www.wma.net/en/30publications/10policies/b3/index.html) and NIH (http:// bioethics.od.nih.gov/humantissue.html).

Selection criteria for brain specimens

To better ensure brain tissue samples quality, specific selection criteria were established and strictly followed when deciding whether to exclude or include a brain specimen.

- Brains with chromosomal or large-scale genomic abnormalities, detected by karyotype analysis and/or Illumina Human Omni-2.5, were excluded.
- Only brains free of obvious malformations or lesions were collected. Disqualifying characteristics included any obvious abnormality of the neural tube, forebrain and brainstem.
- Samples that tested positive for Hepatitis B, Hepatitis C or HIV were excluded.

Tissue processing and dissections

Depending on the condition and period of the procured specimens, different dissection methods were applied. Photos and/or video were used to document dissections using a digital camera and to ensure consistency between specimens. All dissections were performed by the same operator and the specific protocol applied to dissect brain regions and neocortical areas of interest depended upon the period of the specimen and its preservation method. During dissection, brains were placed onto a chilled plate on ice and dissected samples were collected in RLT buffer (Qiagen), immediately frozen in dry ice and stored at −80 °C for later RNA extraction.

Cell Culture

Human fibroblasts isolation and culture

Primary fibroblast cell lines HSB 286, HSB 292, HSB 311, HSB 314 were derived in Sestan laboratory staring from skin samples. WS 1574, WS 5437 primary fibroblasts lines were obtained from NIH Neurobiobank while the other primary fibroblasts lines from Williams syndrome individuals, healthy controls relatives of Williams syndrome patients and 7q11.23 duplication syndrome subjects included in the cohort under study were obtained from the Genomic and Genectic Disorders Biobank (GGDB, Network of

Telethon Genetic Biobanks, project no. GTB07001G). Concerning the primary fibroblast cell lines derived in the laboratory, all biopsies and subsequent sample processing steps for deriving primary fibroblast cell lines were conducted in accordance with the WMA Declaration of Helsinki and NIH guidelines. Briefly, skin or meninges biopsies were rinsed in sterile DPBS 1X (Life Technologies, 14190-144) containing 1% Penicillin/Streptomycin (Life Technologies, 15140-122) and manually dissected into smaller pieces (1-2 mm). Biopsies pieces where then transferred into 6-well tissue culture treated plates coated with 1:50 diluted growth factor reduced Matrigel (Matrigel Basement Membrane Matrix Growth Factor Reduced, Corning, 354230) and grown at 37 °C under hypoxic condition (5% CO_2 and 5% O_2) in low volumes of human fibroblasts medium (HF medium, volume < 800 uL for each well of 6-well plates) that was changed daily until tissue pieces were well adherent to the substrate. Then, media volumes were increased and changed every 2-3 days. All obtained and in-house derived primary human fibroblasts cell lines were cultured in human fibroblast (HF) medium composed of DMEM high glucose (Life Technologies, 11965-092) supplemented with 10% FBS (Life Technologies, 16140-071), 1% L-glutamine 200 mM (Life Technologies, 25030-081), 1% MEM Non-Essential Amino Acids solution (Life Technologies, 11140-050), 1% Sodium Pyruvate 100 mM (Life Technologies, 11360070), 1% Penicillin/Streptomycin (Life Technologies, 15140-122) for few passages before reprogramming. Once human fibroblasts cultures reached 90-100% confluence, cells were passaged using Trypsin-EDTA 0.25% (Life Technologies, 25200-056), pelleted at 200 g for 5 minutes and plated at 6-8*10^3 cells/cm^2.

Human induced pluripotent stem cells generation

Human iPSCs were generated by reprogramming human skin or meninges fibroblasts with episomal vectors. Episomal plasmids pCXLE-hOCT3.4-shp53-F (Addgene #27077), pCXLE-hSK (Addgene #27078), pCXLE-hUL (Addgene #27080), pCXLE-EGFP (Addgene #27082) were used for reprogramming. Human fibroblasts (2x10^6) harvested using Trypsin-EDTA 0.25% (Life Technolgoies**,** 25200-056) were nucleofected with 1.5 ug of each episomal vector using Amaxa Nucleofector II (Program U-23 and U-20, Lonza) with the Amaxa NHDF nucleofector kit (VPD-1001, Lonza). Nucleofected cells were resuspended in fibroblasts medium and seeded at 25000-33000 cells/cm^2 onto 100-mm tissue culture dishes coated with 1:50 diluted growth factor reduced Matrigel (Corning, 354230). Starting from one day after reprogramming, cells were grown under hypoxic conditions in Essential 6 medium (Life Technologies, A1516401) [628] supplemented with bFGF 100 ng/mL (R&D Systems, 4114-TC), hydrocortisone 100 nM (Sigma Aldrich, H0888) and sodium butyrate 0.5 mM (Stemgent, 04-0005). Hydrocortisone was added to promote fibroblasts proliferation while bFGF and sodium butyrate were administrated to improve the reprogramming efficiency. Cell health and proliferation rates were checked daily by visual inspection and medium was routinely changed every day. Reprogrammed cells usually started to appear 3–5 days after transduction and monitoring of GFP expression from day 3-4 enabled to ensure that transfection efficiency was equal or greater than 30%. Hydrocortisone was added until cell density exceeded >50% confluence, a condition that was generally reached around day 7-10 after nucleofection. When reprogrammed cultures become confluent around day 10-14, cells were detached using Trypsin-EDTA 0.25% (Life Technologies, 25200-056), replated at a density of 5000 cells/cm^2 on tissue culture treated dishes coated with 1:60 dilulted Matrigel (Corning, 354230) and maintained Essential 6 medium supplemented with bFGF 100 ng/mL. Sodium butyrate was removed from the medium 2-

4 days after replating. iPSCs colonies intermingled with fibroblasts and partially reprogrammed cells started to form around 18–25 days after transfection. Prior to manual picking, colonies were inspected for GFP transgene expression and only those GFP-negative were chosen for further expansion. Selected individual iPSCs colonies were manually isolated between day 20–30 onto dishes coated with 1: 50 diluted Matrigel (Corning, 354230) in Essential 8 (Life Technologies, A1517001) or Essential 8 Flex medium (Life Technologies, A2858501) supplemented with 10 μM rock inhibitor Y-27632 (Stemcell, 72308). After manual picking, single clones were grown under standard iPSCs culture conditions.

Induced pluripotent stem cells culture

iPSCs lines were routinely cultured at 37 °C under hypoxic conditions (5% CO_2 and 5% O_2) in either Essential 8 (Life Technologies, A1517001) or Essential 8 Flex medium (Life Technologies, A2858501) on tissue culture treated plates coated with 1:100 diluted growth factor reduced Matrigel (Corning, 354230). Cells were generally passaged every 4-5 days with a 0.5 mM EDTA solution when approximately 85% confluence was reached. Briefly, for passaging iPSCs were washed twice with DPBS 1X and incubated with 0.5 mM EDTA solution for 3-5 minutes at 37°C. The EDTA solution was then removed by aspiration and cells aggregates were collected with a small volume of corresponding media and directly replated. The suggested split ratio varies from 1:2 - 1:4 for early passages iPSCs and 1:3 - 1:12 for established cultures [628, 631].

Neocortical differentiation protocol

Human iPSCs directed differentiation into neocortical neurons was achieved by applying a monolayer feeder-free dual SMAD inhibition based [176] neocortical differentiation protocol modified from Shi *et al*. [375, 376]. Briefly, iPSC were dissociated to single cells through incubation with Accutase (Life Technologies, A11105-01) at 37 °C, seeded at 0.7-2 x 10^5 cells/cm^2 density on multiwall 6 plates coated with Matrigel diluted 1:50 (Corning, 354230) and cultured in Essential 8 Flex medium supplemented with 10 μM ROCK inhibitor Y-27632 (StemCell, 72308). ROCK inhibitor was withdrawn the next day, and the iPSCs were allowed to expand in Essential 8 FLEX medium until the cells were nearly confluent. Neural differentiation was started by changing the culture medium to neural induction medium composed of a 1:1 mixture of DMEM/F-12 (Life Technologies, 11330) and Neurobasal (Life Technologies, 21103-049) supplemented with N-2 (1:100, Life Technologies, 17502-048), B-27 (1:50, Life Technologies, 17504-044), 20 μg/ml Insulin (Insulin solution human, Sigma Aldrich, I9278), 1% L-glutamine 200 mM (Life Technologies, 25030-081), 1% MEM Non-Essential Amino Acids (Life Technologies, 11140-050), 0.1 mM 2-mercaptoethanol (Life Technologies, 21985), 100 nM LDN-193189 (Stemgent, 04-0074), 10 μM SB-431542 (TOCRIS, 1614) and 2 μM XAV 939 (Stemgent, 04-0046). Cells were cultured in neural induction medium for 11 days changing the medium daily. The efficiency of neural induction was monitored visually by assessing the appearance of cells with characteristic neuroepithelial morphology and the formation of neural rosettes. On day 12, neuroepithelial cells were harvested through incubation with Accutase (Life Technologies, A11105-01) at 37 °C for approximately 15-20 minutes and replated onto on 13 mm glass coverslips or 6-well plates previously coated with 10 μg/mL Poly-D-lysine (Sigma Aldrich, P6407) and 3 μg/mL Laminin (Natural Mouse Laminin, 23017-015) at 0.65-1.0 ×10^5 cells/cm^2 density in neural induction medium supplemented with 100 nM LDN-193189 (Stemgent, 04-0074) and 10

uM ROCK inhibitor (StemCell, 72308). After replating, the medium was progressively switched through partial media changes to neural differentiation medium which contained Neurobasal (Life Technologies, 21103-049) supplemented with N-2 (1:100, Life Technologies, 17502-048), B-27 (1:50, Life Technologies, 17504-044), 1% Penicillin/Streptomycin (Life Technologies, 15140-122), 1% GlutaMAX (Life Technologies, 35050-061) and 30 ng/mL BDNF human recombinant protein (Life Technologies, PHC7074). Cultures were maintained for up to day 80 of differentiation changing the medium partially every 2-3 days.

Pluripotency validation of iPSCs clones

Nucleic acid extraction

Genomic DNA was extracted from fibroblasts and iPSCs samples using DNeasy Blood and Tissue kit (Qiagen, 69506) according to manufacturer's instructions. The quality and concentration of DNA samples was assessed using a NanoDrop spectrophotometer (NanoDrop Technologies).

Integration check

To test for genomic integration and residual retention of episomal plasmids, PCR on iPSC clones (150 ng) grown at least until passage 4 was performed using GoTaq Green Master Mix (Promega, M712) and a primer pair designed to amplify a portion of the *EBNA-1* sequence derived from Epstein-Barr virus that is a region common to all four episomal vectors utilized (FW: TTTGCAAGCAGCAGATTACGCGCAG, REV: TTTGCTCACCCAGAAACGCTGGTGAAAG). Genomic DNA isolated from day 14-30 reprogramming samples (150 ng) and each episomal vector (1 ng) were used as positive controls, while genomic DNA isolated from fibroblasts (150 ng) was used as negative control. PCR products were separated on 1.5% agarose gels and visualised using ethidium bromide at ChemiDoc (Bio-Rad).

Karyotype analysis

Karyotype analysis was performed at the Molecular Cytogenetics Laboratory, Yale University School of Medicine (New Haven, CT). Chromosomes were classified according to the International System for Human Cytogenetic Nomenclature. At least 5 metaphases were analyzed per cell line.

Teratoma formation assay

For teratoma formation, 70-80% confluent iPS cell plates were harvested using 0.5 mM EDTA solution, collected in Essential 8 or Essential 8 Flex medium and centrifuged at 100 g for 5 minutes. iPSCs (1-2 × 10^6) resuspended in a 2:1 mixture of Essential 8 or Essential 8 Flex medium (Life Technologies, A2858501) and Matrigel (Corning, 354230) were subcutaneously injected into NOD-SCID mice (The Jackson Laboratory). Mice were housed normally and assessed regularly for tumor formation that started to appear 4-6 weeks after cell injection. Around 2 weeks after teratomas were first detected, tissues were fixed in 4% paraformaldehyde. Samples were then OCT embedded, sectioned, stained with hematoxylin and eosin and analyzed at the Mouse Research Pathology Service, Yale University School of Medicine (New Haven, CT).

Embryoid bodies formation assay

To perform the embryoid body formation assay, 70-80% confluent iPS cell plates were washed twice with DPBS 1X and incubated with pre-warmed Dispase solution (Life Technologies, Dispase II powder, 17105-041) for 15-20 minutes until cells began to lift up in small patches and float in the enzyme solution. Once iPSCs colonies were released into the enzyme solution they were harvested by adding 1:1 volume of Essential 6 medium (Life Technologies, A1516401) and transferred to 15 mL Faclon tubes to be centrifuged at room temperature twice at 100 × *g* for 2 minutes to completely remove Disapse residuals and to pellet the cells. The supernatant was carefully aspirated and the cell pellet was gently dislodged from the tube bottom by flicking it and resuspended in pre-warmed Essential 6 medium (Life Technologies, A1516401). iPSCs were then transferred into a 60-mm or a 100-mm non-tissue culture-treated dishes and let grown in suspension as embryoid bodies (EBs) under hypoxic conditions. EBs were cultured for 7 days changing their medium every other day by transferring the entire volume of the dish into a centrifuge tube and waiting until EBs had settled to the bottom of the tube, that took about 5-10 minutes, before removing the supernatant from the tube and transferring the EBs resuspended in fresh Essential 6 medium in new dishes.Then, EBs were mechanically dissociated, plated onto 0.1% gelatin-coated (Gelatin powder from porcine skin, Sigma, G1890-50DG) multiwells 12 and grown for 2 to 3 additional weeks in Essential 6 medium supplemented with 10% FBS (Life Technologies, 16140-071) in a 37°C, 5% CO2 incubator prior to fixation and subsequent immunocytochemical analysis.

Immunocytochemistry

Cell cultures were fixed with 4% paraformaldehyde (wt/vol) for 10 minutes at room temperature or with ice-cold methanol at -20 °C for 20 minutes. Cells were permeabilized for 10 minutes at room temperature with PBS 1X supplemented with 0.5% (vol/vol) Triton X-100 (Sigma, 93443), 100 mg/L $CaCl_2$ and 100 mg/L $MgCl_2$. When required, antigen unmasking was performed according to supplier instructions. Blocking has been achieved using a PBS 1X solution supplemented with 5% fetal bovine serum, 0.3% (vol/vol) Triton X-100, 100 mg/L $CaCl_2$ and 100 mg/L $MgCl_2$. Primary antibodies were incubated overnight at 4 °C diluted at the appropriate concentrations in PBS 1X supplemented with 3% fetal bovine serum, 0.2% (vol/vol) Triton X-100, 100 mg/L $CaCl_2$ and 100 mg/L $MgCl_2$. Specifications of the primary antibodies used to assess the pluripotent state (Table 1) and iPSCs-derived neuronal cells identity (Table 2). For primary antibody detection, species-specific Alexa-dye conjugated secondary antibodies (Life Technologies) in combination with DAPI (Sigma, #10236276001) were incubated for 1 hour at room temperature diluted according to supplier instructions (1:250) in PBS 1X supplemented with 3% fetal bovine serum, 0.2% (vol/vol) Triton X-100, 100 mg/L $CaCl_2$ and 100 mg/L $MgCl_2$. For long-term storage a PSB 1X solution supplemented with 100 mg/L $CaCl_2$, 100 mg/L $MgCl_2$ and 0.01% (wt/vol) Sodium Azide (Sigma Aldrich, S2002) was used. When staining cell cultures grow on a glass substrate, coverslips were mounted to a slide with droplets of Aqua Poly/Mount (Polysciences Inc, 18606). Images of iPSCs colonies and day 14 iPSCs-derived neocortical progenitors were acquired on Apotome.2 (Zeiss) combined with Observer.Z1 (Zeiss) inverted microscope at 10X and 20X magnification, respectively. Fluorescence images of day 40 and day 80 iPSCs-derived neocortical neurons were acquired on Apotome.2 microscope combined

with AxioImager M2 upright microscope (Zeiss) and confocal images at 20X and 40X were acquired at LSM 510 META (Zeiss).

Antibody	Company	Code	Species
Nanog	Abcam	ab21624	Rabbit IgG
Nanog	Stemgent	09-0020	Rabbit IgG
OCT3/4	Abcam	ab19857	Rabbit IgG
OCT3/4	Stemgent	09-0023	Rabbit IgG
SOX2	Millipore	ab5603	Rabbit IgG
SSEA4	BD	560218	Mouse IgG
SSEA4	Stemgent	09-0006	Mouse IgG3
TRA-1-60	BD	560122	Mouse IgM
TRA-1-60	Millipore	09-0010	Mouse IgM

Table 1. Primary antibodies used to validate the pluripotent state.

Antibody	Company	Code	Species
AFP	Santa Cruz	sc-130302	Mouse IgG
βIII-tubulin	Covance	MMS-43SP	Mouse IgG
CUX1	Santa Cruz	sc13024	Rabbit IgG
CTIP2	Abcam	ab18465	Rat
FOXA (HNF3β)	Santa Cruz	sc-6554	Goat
FOXG1	NCFAB	Stem Culture	Rabbit
GABA	Sigma	A2052	Rabbit
GFAP	Advanced ImmunoChemical	31223500	Guinea Pig
MAP2	Sigma	M4403	Mouse IgG
MAP2	Millipore	ab5622	Rabbit
Nestin	R&D	MAB1259	Mouse IgG1
Neurofilament light	Millipore	MAB1615	Mouse IgG
OTX2	Millipore	ab9566	Rabbit IgG
PAX6	BD	561462	Mouse IgG
Phospho-vimentin	Abcam	Ab22651	Mouse IgG
SATB2	GenWay	20-372-60065	Mouse IgG
SMA (ASM-1)	Millipore	CBL171	Mouse
SOX1	Cell Signaling	4194	Rabbit
TBR1	Santa Cruz	48816	Rabbit IgG
TBR2	Millipore	ab15894	Chicken
VGLUT1	Millipore	MAB5502	Mouse IgG
ZO-1	Invitrogen	61-7300	Rabbit IgG

Table 2. Primary antibodies used to characterize iPSCs-derived neuronal cell identity at different time points of the in vitro differentiation protocol.

Antibody	Company	Code	Species
DNAJC30	Novus	H00084277-B01P	Mouse
DNAJC30	Novus	48816	Rabbit
NRGN	R&D	AF7947	Sheep
GAD1	R&D	AF2086	Goat
GFAP	Sigma	G9269	Rabbit
GFAP	Encor	RPCA-GFAP	Rabbit
CNP	Encor	CPCA-CNP	Chicken

Table 3. Primary antibodies used to characterize the cellular expression of Dnajc30.

Electrophysiology

All recordings were performed at 21-23°C using the whole-cell patch-clamp technique in voltage- and current-clamp configurations. The extracellular solution was composed by 125 mM NaCl, 2.5 mM KCl, 26 mM NaHCO3, 15mM glucose, 1.3 mM MgCl2, 2.3 mM CaCl2, 1.25 mM NaH2PO4 (bubbled with 95% O2, 5% CO2). Cells were visualized using an Axioskop FS2 microscope equipped with a 5X scanning objective and a 40X water-immersion objective (Zeiss), and connected to a CCD camera. Pipettes were produced from borosilicate glass capillary tubes (Sutter Instruments) by means of a horizontal puller (P-2000, Sutter instruments) and filled with the following intracellular solution: 130 mM K-gluconate, 4 mM NaCl, 2 mM MgCl2, 1 mM EGTA, 5 mM creatine phosphate, 2mM Na2ATP, 0.3 mM Na3GTP, 10 mM HEPES (pH 7.3 with KOH). Membrane voltage was corrected off-line for a calculated liquid junction potential of -10 mV. Series resistance was always compensated by 70–80% and monitored throughout the experiment. Recordings were made with a MultiClamp 700B amplifier (Molecular Devices) and digitized with a Digidata 1322 computer interface (Molecular Devices). Data were acquired using the software Clampex 9.2 (Molecular Device), sampled at 20 kHz, and filtered at 10 kHz. Software Clampfit 10.2 (Molecular Devices) and OriginPro 8 (Microcal) were used for data analysis. Cell capacitance (C_m) was calculated by integrating the capacitive current evoked by a −10 mV pulse. Resting membrane potential (V_m) was measured without current injection. Neuronal input resistance (R_{in}) was calculated in the linear portion of the I–V relationship during depolarizing voltage responses near the resting potential. Activation of inward and outward currents mediated by voltage-dependent sodium and potassium channels was obtained by stimulating the cell with steps of voltage of increasing amplitude. Firing properties were assessed by recording the voltage responses to a series of supra-threshold current steps starting from a potential of -70 mV. The spontaneous excitatory (sEPSCs) and inhibitory (sIPSCs) post-synaptic currents were recorded at the holding potentials of -70 mV and 0 mV, respectively, and their nature was confirmed by bath application of specific synaptic blockers added to the extracellular solution. The $GABA_A$ receptor blocker bicuculline methiodide (10 μM) (bicuculline, Sigma-Aldrich) was used to abolish the sIPSCs, while the AMPA receptor agonist 2,3-dioxo-6-nitro-1,2,3,4-tetrahydrobenzo(f)quinoxaline-7-sulfonamide (10 μM) (NBQX, Sigma-Aldrich) was used to abolish the sEPSCs. Data are

expressed as mean ± s.e.m. Statistical significance was evaluated using two-tailed unpaired homoscedastic Student's t-test.

mRNA sequencing

RNA extraction from cell samples

Total RNA was collected from undifferentiated iPSCs (day 0) and iPSCs-derived neocortical neural progenitors and neurons specimens obtained at day 10-14, day 36-42 and day 80-83 of *in vitro* differentiation. Samples for RNA-sequencing were collected in RNAprotect cell reagent (Qiagen, #76526). Total RNA was extracted using the RNeasy plus mini kit (QIAGEN, #74136) according to manufacturer's instructions. Each sample was subjected to a DNase treatment (TURBO DNase, Ambion) as indicated in the product manual. Prior to perform cDNA library preparation, RNA quantity and quality was first checked using NanoDrop (Thermo Scientific) and then using High Sensitivity RNA ScreenTapes (Agilent, #5067-5578) at 2200 TapeStation (Agilent).

RNA extraction from brain tissue samples

A bead mill homogenizer (Bullet Blender, Next Advance) was used to lyse the pulverized tissue. Each pulverized tissue sample was transferred to a chilled safe-lock microcentrifuge tube (Eppendorf). A mass of chilled stainless steel beads (Next Advance, cat# SSB14B) equal to the mass of the tissue was added to the tube. Two volumes of lysis buffer were added to the tissue and beads. Samples were mixed in the Bullet Blender for 1 min at a speed of six. Samples were visually inspected to confirm desired homogenization and then incubated at 37 °C for 5 min. The lysis buffer was added up to 0.6 ml, and samples were mixed in the Bullet Blender for 1 min. Total RNA was extracted using RNeasy Plus Mini Kit (Qiagen, #74136). Each sample was subjected to a DNase treatment (TURBO DNase, Ambion) as indicated in the product manual. Prior to perform cDNA library preparation, RNA quantity and quality was first checked using NanoDrop (Thermo Scientific) and then using High Sensitivity RNA ScreenTapes (Agilent, #5067-5578) at 2200 TapeStation (Agilent).

RNA-sequencing Library preparation and RNA-sequencing

The TruSeq Stranded Total RNA LT Sample Prep Kit (Illumina) was used to prepare cDNA libraries according to manufacturer instructions. Briefly, 500 ng of total RNA was processed with Ribo-Zero Gold Set A-B to deplete samples of both cytoplasmic and mitochondrial rRNA. The remaining RNA was purified, fragmented, reverse transcribed into first strand cDNA and further processed to generate a second strand of blunt-ended cDNA. Blunt-ended cDNA fragments were adenylated, ligated to multiple indexing adapters and DNA fragments that have adapter molecules ligated to both ends were selectively amplified through PCR. Prior to samples submission for sequencing, the size and distribution of DNA library templates was assessed by loading 2 μL of each sample on High Sensitivity DNA 1000 ScreenTapes (Agilent, #5067-5584) read at 2200 TapeStation (Agilent). cDNA libraries were provided to the Yale Center for Genomic Analysis where paired-end RNA sequencing was performed on Illumina Hi Seq2000 to an average depth of 100M reads per sample.

mRNA sequencing Quality Assessment

Reads alignment, expression quantification, and quality assessment

The generated FASTQ files were input to Illumina's CASAVA pipeline to purify the low-quality and non-identified reads. Low quality reads defined as those that did not had at least 80% of bases with a quality score of 20 or more (conducted by fastq_quality_filter function) and reads left with > 10 bases after being trimmed with a quality score of < 20 (conducted by fastq_quality_trimmer function) were filtered out. FastQC was used to report the fundamental quality parameters as raw reads number, sequencing base quality score distribution, duplication level, adapter contamination, and GC content (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), which will sever as co-variables to perform association analyses to examine the unwanted batch effects.

Reads filtered by quality score were then mapped to the human reference genome (GRCh38/hg38) with default parameters using the Spliced Transcripts Alignment to a Reference (STAR) software [639]. The aligned reads stored as BAM format were converted to MRF format, and the expression level of genes and exons were quantified by count number and RPKM (reads per kilobase of exon model per million mapped reads) [660] using RSEQtools. Only genes with count number greater than 5 and RPKM greater than 1 in either condition were included in the further analyses. The number and the percentage of uniquely mapped reads were calculated for each sample. Samples with less than 20 million uniquely mapped reads were excluded from further analysis or re- sequenced.

Moreover, we developed some in-house R and Perl scripts and collaborated with some public available tool-kits to perform the in-depth quality assessment of our sequencing data sets, which were implemented throughout sample preparation and transcriptome data generation steps as described [148]. Samples that failed to pass the quality control measures were reprocessed or removed from the downstream analyses. For instances, the coverage by chromosome of the mapped reads was calculated using the depth function of the SAMtools version 0.1.19 module. Exon coverage was calculated by using the intersectBed function of the bedtools version 2.20.1 module. The density plots of RPKM values were generated by ggplot2 version 1.0.1 R package. RnaSeqMetrics function under Picard tools (v.1.102; http://picard.sourceforge.net) was used to compute the 5′–3′ coverage bias along gene body as well as the number of bases assigned to various classes of RNA. Subsequent data processing with the R package conditional quantile normalization (CQN) and ComBat function enabled to correct for gene length, gene GC content, sequencing site and batch effects respectively.

mRNA Sequencing Data Analysis

Hierarchical clustering and batch effect correction

All statistical analyses were performed on data previously corrected for batch effects using R function *ComBat* [661]. Hierarchical clustering performed using the function *hclust* in R with 1-correlation as distance, clearly separated samples according to their differentiation stage.

Principal component analysis

PCA was performed to visualize the relatedness of all the 35 cell samples separately and in combination with 6 brain tissue specimens using *prcom* package and the first two principal components were plotted using *rgl* package, both retrieved from R (http://www.r-project.org/). Our results showed that cell samples clustered by *in vitro* differentiation stage and clustered closely with human prenatal brain-derived neurons.

Identification of differentially expressed genes (DEGs)

DEGs genes were computed using R package *DESeq2* [662]. In this package, the count data were modeled with a negative binomial distribution to address the problem of overdispersion. *P* values were calculated by a likelihood ratio test comparing models and DEGs genes were called at a false discovery rate (FDR) of 0.01 [663].

Weighted gene co-expression network analysis (WGCNA)

Signed co-expression networks were built using the *WGCNA* package [664] in R. For all genes included, a pair-wise correlation matrix was computed, and an adjacency matrix was calculated by raising the correlation matrix to a power. For each pair of genes, a robust measure of network interconnectedness (topological overlap measure) was calculated based on the adjacency matrix. The topological overlap based dissimilarity was then used as input for average linkage hierarchical clustering. Modules were generated by hybrid dynamic tree-cutting. To detect rare expression patterns, we set the minimum module size to 10 genes, deepSplit to 5, and the minimum height for merging modules to 0.2. Each module was summarized by an eigengene, which is the first principal component of the scaled module expression. Thus, the module eigengene explained the maximum amount of variation of the module expression levels. The module membership (also known as module eigengene based connectivity kME) is defined as the correlation between gene expression values and the module eigengene. To obtain modules with clear clustering of expression patterns, we iteratively reassigned genes to the module with largest kME and genes with maximum kME < 0.5 were assigned to the M0.

Functional enrichment analysis

Functional enrichment analysis was performed using GeneCodis [665] and DAVID Bioinformatics Resources (http ://david.abcc.ncifcrf.gov/). All reported p-values adjusted were calculated using the Benjamini-Hochberg method. Data were plotted using Excel.

Single-Cell RNA-sequencing

Single-Cell capture and cDNA synthesis

Cells were captured with medium-sized (10-17μm) RNA-seq IFCs using the Fluidigm C1 system according to the manufacturer's instruction. Immediately before cell capturing, viability was accessed by trypan blue staining. In all experiment, the cell viability was close to 100%. A concentration of 1,000–3,000 cells/μL was used for cell loading. After capture, C1 chips were examined visually and the number of cells at each capture site was recorded manually. Cells captured by C1 were subsequently processed through lysis, reverse transcription and PCR amplification to generate single-cell cDNA libraries using the Smarter Ultra Low Input RNA kit for Fluidigm (Clontech). The cDNAs from all

capture sites were harvested the next morning in about 3 µL C1 harvest reagent and were transferred to a 96-well PCR plate in 10 µL C1 DNA dilution reagent.

Single-Cell RNA-sequencing library preparation

cDNA concentration was quantified with Quant-iT PicoGreen (Invitrogen, P7589). Samples were diluted to 200 ng/µL with C1 harvest reagent. The resulting cDNA samples were indexed using Nextera XT Library Prep Kit (Illumina). All sequencing libraries were assessed for quality by Agilent Bioanalyzer, using high sensitivity dsDNA assay.

Single-Cell RNA-sequencing

Libraries of single cells pooled from each C1 IFC were denatured using the Illumina protocol. The denatured libraries were diluted to 6 pM, followed by cluster generation on a single-end HiSeq flow cell using an Illumina cBOT, according to the manufacturer's instructions. The HiSeq flow cell was run for 100 cycles using a single-read recipe according to the manufacturer's instructions. We used Illumina CASAVA to purify the low -quality and non-identified reads and FastQC to report the fundamental quality parameters as raw reads number, sequencing base quality score distribution and GC content (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). To avoid adapter contamination and higher error rates in reads boundary, we trimmed 12 nucleotides in 5'-end and 13 nucleotides in 3'-end, leaving 75 nucleotide long reads for sequence alignment.

Reads alignment and gene expression quantification

STAR (v2.4.0) [639] was employed to uniquely align the filtered reads to human reference genome (GRCh38/hg38) with default parameters. The gtf formatted annotations for human (GENCODE, released version v21) were retrieved from GENCODE project. We used RSEQtools and SAMtools to calculate the gene expression [666, 667]. We used HTSeq (v0.6.1) to calculate the gene reads count values for each annotation entry. The gene RPKM values were calculated using in-house R scripts.

Data quality assessment

In addition to FastQC, we implemented a series of quality control measures. First, we counted uniquely mapping reads per cell and used only cells with at least 100,000 unique reads, which is also at least 10% of total reads, mapped to coding sequences. Next, we checked the reads distribution across different chromosomes. Cells with more than 10% unique reads mapped to mitochondrial chromosome were considered low quality and removed. Furthermore, we checked the reads aligned to exonic, intronic and intergenic regions. High quality cells should have at least a percentage of 30% on total reads mapped to exons. We checked gene coverage uniformity and chose those cells with fold change of average reads coverage within gene less than 2. The gene coverage uniformity is considered as an indicator for RNA degradation. Read counts were converted to RPKM and log2 transformed for downstream analysis.

Single-Cell RNA-sequencing data analysis

Identification of variable genes

To reduce computational burden, we sought to select variable genes for further analysis. Variable genes were defined as those that have higher coefficient of variation comparing to empirically predicted value. Briefly, means of log transformed expression level of each gene in all single cells were calculated, as well as the observed coefficient of variation of each gene. Then, a log linear regression model was used to generate a predicted coefficient of variation for each gene. Genes whose observed coefficient of variation was greater than estimation were considered as variable genes and were used for t-SNE.

Clustering analysis and t-SNE plots

We used t-distributed stochastic neighbor embedding [108] to present data in two dimensional coordinates. Graphs were generated using R package ggplot2. To identify different neuronal subtypes and cell states we used *pcaReduce* package in R [641] which adopts an agglomerative unsupervised clustering approach that integrates principal components analysis (PCA) and hierarchical clustering. Number of samples (nbt) was set to 1, and number of dimensions (q) was set to 30. Sampling based merging was applied.

Expression and protein network analysis of 7q11.23 CNV genes

Gene expression analysis was performed on the exon array [148] and publicly available BrainSpan RNA-seq datasets (www.BrainSpan.org) using the RNA-seq data for data presentation. The pairwise co-expression coefficient (Pearson correlation coefficient) among all 7q11.23 CNV genes was calculated and hierarchical clustering (Ward's method) was performed with the distance being defined as 1 - co-expression coefficient. Log2 transformed RPKM values were used for the co-expression calculation. The three clusters were identified by visual inspection of the dendrogram and heatmap.

Integration of the brain transcriptome and protein-protein interaction network was done using the non-redundant physical BioGRID protein-protein interaction network (3.4.132, Jan. 2016) as a backbone interactome. The interactome space was confined based on the availability and the abundance of the BrainSpan RNA-seq dataset. First all the nodes (proteins and corresponding genes) that are not profiled in the BrainSpan RNA-seq dataset were removed from the interactome. The analysis was further refined by removing genes with noise expression levels by grouping as expressed vs. non-expressed using a mixture model. This resulted in an interactome network of 12,367 nodes and 159,353 edges. Within the resulting interactome network, the pairwise co-expression of nodes significantly decayed over the shortest path distance between the nodes. This observation is consistent with several reports showing that the proximity of genes in the interactome network is indicative of the functional similarity of genes. Similar results were found when using expression level cutoff of RPKM>1 (data not shown). The data presented are without removing non-expressed genes to demonstrate the contribution of all genes comprised in the 7q11.23 locus. The significance of the observed mean pairwise co-expression of gene set (i.e. the group of genes belong to a cluster) was tested against an empirical null expectation. To draw a null expected mean co-expression for a set of genes in a cluster, we randomly sampled a set of matched

number of genes having the same number of protein interacting partners from the interactome network and averaged the absolute value of the pairwise co-expression coefficients. The significance p-value of the co-expression of the cluster was then determined by 1 - the number of times the mean absolute co-expression of null sets larger than that of a cluster gene set divided by the number of randomly sampled null sets. Here 1000 null sets were used.

The relative tissue-specific gene expression enrichment between groups of genes was examined using Genotype-Tissue Expression (GTEx) project RNA-Seq dataset (Consortium, 2013). We downloaded GTEx V6 Analysis RNA-Seq dataset available from http://www.gtexportal.org/home/datasets/GTEx_Analysis_v6_RNA-seq_RNA-SeQCv1.1.8_gene_rpkm.gct.gz. We summarized the expression values of genes per tissue group and performed the Wilcoxon rank sum test between two groups (i.e., C1 and C2 combined vs. C3 cluster) upon the quantile-normalization of the per-tissue mean expression values.

Identifying spatio-temporal context of 7q11.23 genes was done by grouping all the BrainSpan RNA-seq dataset into 36 spatio-temporal intervals: 4 spatial groups ((i) neocortex, (ii) hippocampus and amygdala, (iii) striatum and thalamus, and (iv) cerebellum) and 9 temporal windows (8 pcw – 40 years; see www.BrainSpan.org). The specificity of gene expression of all genes in the interactome network was evaluated using geneEnrichment profiler [668], which provides a quantitative measure of up-down regulation in a given spatio-temporal interval with respect to all the rest intervals. We standardized the specificity scores with mean=0 and standard deviation=1 in each spatio-temporal interval. Therefore, the resulting gene-expression specificity score of each gene provides a deviation (Z-score) from the normalized specificity score distribution in a given spatio-temporal interval.

Analysis of co-expressed genes was performed by retrieving normalized log2 signal intensity and RPKM values from BrainSpan datasets (exon array and RNA-seq, respectively), and then using a Pearson correlation, identifying the top 1000 correlation coefficients for each 7q11.23 gene. These gene lists were analyzed for Gene Ontology terms using the Functional Annotation Tool within DAVID v6.7 [669]. We reached similar observations from both exon array and RNA-seq platforms. All results were then filtered for FDR <0.01, and all subsequent Gene Ontology terms associated with genes from one cluster were submitted simultaneously to REVIGO [652], using medium similarity (0.7), and SimRel semantic similarity measures. The provided .xgmml files were imported into Cytoscape v3.3 for plotting [670].

Western blotting

Cells or tissues were lysed by sonication in PBS or TBS supplemented with 0.01% Tween-20, protease and phosphatase inhibitors. Following centrifugation at 20,000xg, supernatant protein concentrations were quantified and mixed with Laemmli loading buffer, boiled, electrophoresed on NuPage Bis-Tris gels and transfered to polyvinylidene fluoride (PVDF) membranes. Membranes were blocked and then incubated with primary and secondary antibodies, in 5% skim milk or 5% BSA, and subsequently developed using chemiluminescence. Bands were imaged using G:Box Chemi XRQ and quantified using GeneTools software.

Immunohistochemistry

Post-mortem brain tissue from adult humans, rhesus macaque, or mouse was used for immunohistochemistry (IHC), immuno-electron microscopy (immuno-EM), and immunofluorescence (IF). Standard procedures were followed for each application. For IHC, human sections were pre-treated with 0.3% H2O2 followed by incubation in blocking buffer (5% BSA, 1% normal donkey serum in PBS + 0.05% Tween-20) at room temperature and then incubation for 24-48 hours at 4°C in primary antibodies. Tissue was then incubated with biotin- labeled secondary antibodies, conjugated with Avidin-Biotin-Peroxidase Complex, and visualized with DAB-nickel following the manufacturer's recommended protocol. Developed sections were mounted on Superfrost Plus charged slides, dehydrated through an ethanol series, cleared with xylenes, and preserved with Permount and glass coverslips. Images were taken using Aperio ScanScope digital scanner. For immunofluorescence, a block of adult rhesus macaque STC was fixed and incubated in 30% sucrose. The block was then embedded in agarose, sectioned at 25μm, and perforated by sucrose-protected freeze/thaw treatment. Sections were blocked in blocking buffer using 0.05% *n*-Dodecyl β-D-maltoside, which was used for subsequent antibody dilutions. DAPI was used at 1ng/μL. Finished sections were mounted on slides and covered with Aqua-Poly/Mount. Fluorescent images were acquired on a Zeiss LSM 510 multi-photon microscope and modified using Zen software.

Droplet digital PCR

RNA was isolated from cells and tissues essentially as previously described [148]. Briefly, for frozen *Dnajc30* mouse cortices and WS neocortex, tissues were pulverized in a chilled mortar and pestle, and homogenized using chilled steel beads and a bead mill homogenizer. All lysates were then subjected to RNA purification using the RNeasy Plus Mini Kit. Concentration and RNA integrity number (RIN) values were obtained using R6K ScreenTape and a TapeStation 2200. One μg of total RNA was used to create cDNA with SuperScript III reverse transcriptase. One ng of cDNA was used for template in droplet digital PCR.

Neuronal morphometric analysis on mouse brains, human iPSCs-derived neocortical neurons and human brain tissues

Neuronal morphometric analysis was performed on mice neocortical sections obtained from WT (n=4) and Dnajc30 KO (n=6) mice aged 9-10 days stained with a FD Rapid GolgiStain kit (FD NeuroTechnologies Inc., PK401) by analyzing approximately 15 neurons per mouse. For the Golgi-Kopsch stain of human neocortical slices, a previously described protocol was followed (Riley, 1979). Briefly, blocks ~0.5-1.0cm3 of cortex were incubated in 3.5% potassium dichromate, 15% sucrose, and 0.6% formaldehyde for 7 days, followed by incubation in 0.75% silver nitrate for 3 days, and finally dehydrated overnight in 95% ethanol. Blocks were then sectioned on a vibratome at 200μm, cleared in Histo-Clear, and mounted on microscope slides with Permount and coverslips. Silver impregnation was generally poor in STC areas of all brains. The best cells that appeared fully impregnated (several branched apical and basal dendrites from easily identifiable soma) were in the deep layers of M1C, and for this reason only these cells were analyzed. The impregnation was not sufficient to reliably quantify spines/dendritic protrusions. The same morphological analysis was performed also on immunolabeled human iPSCs-derived neocortical neurons obtained from 3 healthy controls and 4 WS individuals after 80 days of *in vitro* differentiation, measuring also in this case about 15 neurons per cell line. Both mouse neocortical slices, human brain slices and day 80

iPSCs-derived neocortical neurons coverslips were viewed on a Zeiss Axio Imager.A1. Neuronal morphology analysis was performed using Neurolucida software (MBF Bioscience) that enabled to quantify the number of dendrites and nodes, dendrites lengths and soma area. For all these in person-based quantifications, researchers were blinded to mice and cell lines genotypes before their analyses.

References

1. Fischl, B. and A.M. Dale, *Measuring the thickness of the human cerebral cortex from magnetic resonance images.* Proc .Natl. Acad. Sci., 2000. **97**(20): p. 11050-11055.
2. Stephan, H., H. Frahm, and G. Baron, *New and revised data on volumes of brain structures in insectivores and primates.* Folia Primatol. , 1981. **35**(1): p. 1-29.
3. Rakic, P., *Radial unit hypothesis of neocortical expansion.* Novartis Found Symp., 2000. **228**(1528-2511): p. 30-42.
4. Rakic, P., *Evolution of the neocortex: a perspective from developmental biology.* Nat. Rev. Neurosci., 2009. **10**(10): p. 724-735.
5. Lewitus, E., I. Kelava, and W.B. Huttner, *Conical expansion of the outer subventricular zone and the role of neocortical folding in evolution and development.* Front. Hum. Neurosci., 2013. **7**(424).
6. Lewitus, E. and A.T. Kalinka, *Neocortical development as an evolutionary platform for intragenomic conflict.* Front. Neuroanat., 2013. **7**(2).
7. Zilles, K., et al., *The human pattern of gyrification in the cerebral cortex.* Anat. Embryol., 1988. **179**(2): p. 173-179.
8. Geschwind, D.H. and P. Rakic, *Cortical evolution: judge the brain by its cover.* Neuron, 2013. **80**(3): p. 633-647.
9. Elias, H. and D. Schwartz, *Surface areas of the cerebral cortex of mammals determined by stereological methods.* Science, 1969. **166**(3901): p. 111-113.
10. Zilles, K., N. Palomero-Gallagher, and K. Amunts, *Development of cortical folding during evolution and ontogeny.* Trends Neurosci., 2013. **36**(5): p. 275-284.
11. Kelava, I., E. Lewitus, and W.B. Huttner, *The secondary loss of gyrencephaly as an example of evolutionary phenotypical reversal.* Front. Neurosci., 2013. **7**(16).
12. Borrell, V. and I. Reillo, *Emerging roles of neural stem cells in cerebral cortex development and evolution.* Dev. Neurobiol., 2012. **72**(7): p. 955-971.
13. Yanez, I.B., et al., *Double bouquet cell in the human cerebral cortex and a comparison with other mammals.* J. Comp. Neurol., 2005. **486**(4): p. 344-360.

14. Azevedo, F.A., et al., *Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain.* J. Comp. Neurol., 2009. **513**(5): p. 532-541.
15. Silbereis, J.C., et al., *The Cellular and Molecular Landscapes of the Developing Human Central Nervous System.* Neuron, 2016. **89**(2): p. 248-268.
16. Herculano-Houzel, S., *The human brain in numbers: a linearly scaled-up primate brain.* Front. Hum. Neurosci., 2009. **3**: p. 31.
17. Tang, Y., et al., *Total regional and global number of synapses in the human brain neocortex.* Synapse, 2001. **41**(3): p. 258-273.
18. Gilbert, C.D., *Microcircuitry of the visual cortex.* Annu. Rev. Neurosci., 1983. **6**(0147-006X (Print)): p. 217-247.
19. Del Rio, J.A., et al., *Developmental history of the subplate and developing white matter in the murine neocortex. Neuronal organization and relationship with the main afferent systems at embryonic and perinatal stages.* Cereb. Cortex, 2000. **10**(8): p. 784-801.
20. Martin, K.A., *The Wellcome Prize lecture. From single cells to simple circuits in the cerebral cortex.* Q. J. Exp. Physiol., 1988. **73**(5): p. 637-702.
21. Szentagothai, J., *The Ferrier Lecture, 1977. The neuron network of the cerebral cortex: a functional interpretation.* Proc. R. Soc. Lond. B. Biol. Sci., 1978. **201**(1144): p. 219-248.
22. Kostovic, I., et al., *Laminar organization of the human fetal cerebrum revealed by histochemical markers and magnetic resonance imaging.* Cereb. Cortex, 2002. **12**(5): p. 536-544.
23. Barazany, D. and Y. Assaf, *Visualization of cortical lamination patterns with magnetic resonance imaging.* Cereb. Cortex, 2012. **22**(9): p. 2016-2023.
24. Mountcastle, V.B., *The columnar organization of the neocortex.* Brain, 1997. **120**(Pt 4): p. 701-722.
25. Rakic, P., *Confusing cortical columns.* Proc. Natl. Acad. Sci., 2008. **105**(34): p. 12099-12100.
26. Douglas, R.J. and K.A. Martin, *Neuronal circuits of the neocortex.* Annu. Rev. Neurosci., 2004. **27**: p. 419-451.
27. Mountcastle, V.B., P.W. Davies, and A.L. Berman, *Response properties of neurons of cat's somatic sensory cortex to peripheral stimuli.* J. Neurophysiol., 1957. **20**(4): p. 374-407.
28. Herculano-Houzel, S., et al., *The basic nonuniformity of the cerebral cortex.* Proc .Natl. Acad. Sci., 2008. **105**(34): p. 12593-12598.

29. Leone, D.P., et al., *The determination of projection neuron identity in the developing cerebral cortex.* Curr. Opin. Cell Biol., 2008. **18**(1): p. 28-35.
30. Molyneaux, B.J., et al., *Neuronal subtype specification in the cerebral cortex.* Nat. Rev. Neurosci., 2007. **8**(6): p. 427-437.
31. Parnavelas, J.G., *The origin and migration of cortical neurones: new vistas.* Trends Neurosci., 2000. **23**(3): p. 126-131.
32. DeFelipe, J., L. Alonso-Nanclares, and J.I. Arellano, *Microstructure of the neocortex: comparative aspects.* J. Neurocytol., 2002. **31**(3-5): p. 299-316.
33. Group, P.I.N., et al., *Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex.* Nat. Rev. Neurosci., 2008. **9**(7): p. 557-568.
34. Markram, H., et al., *Interneurons of the neocortical inhibitory system.* Nat. Rev. Neurosci., 2004. **5**(10): p. 793-807.
35. Monyer, H. and H. Markram, *Interneuron Diversity series: Molecular and genetic tools to study GABAergic interneuron diversity and function.* Trends Neurosci., 2004. **27**(2): p. 90-97.
36. Wonders, C.P. and S.A. Anderson, *The origin and specification of cortical interneurons.* Nat. Rev. Neurosci., 2006. **7**(9): p. 687-696.
37. Wang, X.J., et al., *Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory.* Proc .Natl. Acad. Sci. , 2004. **101**(5): p. 1368-1373.
38. Whittington, M.A. and R.D. Traub, *Interneuron diversity series: inhibitory interneurons and network oscillations in vitro.* Trends Neurosci., 2003. **26**(12): p. 676-682.
39. Hansen, D.V., et al., *Neurogenic radial glia in the outer subventricular zone of human neocortex.* Nature, 2010. **464**(7288): p. 554-561.
40. Hansen, D.V., et al., *Non-epithelial stem cells and cortical interneuron production in the human ganglionic eminences.* Nat. Neurosci., 2013. **16**(11): p. 1576-1587.
41. Ma, T., et al., *Subcortical origins of human and monkey neocortical interneurons.* Nat. Neurosci., 2013. **16**(11): p. 1588-1597.
42. Wichterle, H., et al., *In utero fate mapping reveals distinct migratory pathways and fates of neurons born in the mammalian basal forebrain.* Development, 2001. **128**(19): p. 3759-3771.
43. Letinic, K., R. Zoncu, and P. Rakic, *Origin of GABAergic neurons in the human neocortex.* Nature, 2002. **417**(6889): p. 645-649.
44. Metin, C., et al., *Cell and molecular mechanisms involved in the migration of cortical interneurons.* Eur. J. Neurosci., 2006. **23**(4): p. 894-900.

45. Ciceri, G., et al., *Lineage-specific laminar organization of cortical GABAergic interneurons.* Nat. Neurosci., 2013. **16**(9): p. 1199-1210.
46. Florio, M. and W.B. Huttner, *Neural progenitors, neurogenesis and the evolution of the neocortex.* Development, 2014. **141**(11): p. 2182-2194.
47. Wilsch-Brauninger, M., M. Florio, and W.B. Huttner, *Neocortex expansion in development and evolution - from cell biology to single genes.* Curr. Opin. Neurobiol., 2016. **39**: p. 122-132.
48. Rakic, P., *Radial versus tangential migration of neuronal clones in the developing cerebral cortex.* Proc .Natl. Acad. Sci., 1995. **92**(25): p. 11323-11327.
49. Rakic, P., *Developmental and evolutionary adaptations of cortical radial glia.* Cereb. Cortex, 2003. **13**(6): p. 541-549.
50. Kwan, K.Y., N. Sestan, and E.S. Anton, *Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex.* Development, 2012. **139**(9): p. 1535-46.
51. Elston, G.N., *Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function.* Cereb. Cortex, 2003. **13**(11): p. 1124-1138.
52. Jacobs, B., et al., *Regional dendritic and spine variation in human cerebral cortex: a quantitative golgi study.* Cereb Cortex, 2001. **11**(6): p. 558-571.
53. Spruston, N., *Pyramidal neurons: dendritic structure and synaptic integration.* Nat. Rev. Neurosci., 2008. **9**(3): p. 206-221.
54. DeFelipe, J. and I. Farinas, *The pyramidal neuron of the cerebral cortex: morphological and chemical characteristics of the synaptic inputs.* Prog. Neurobiol., 1992. **39**(6): p. 563-607.
55. Elston, G.N. and J. DeFelipe, *Spine distribution in cortical pyramidal cells: a common organizational principle across species.* Prog. Brain Res., 2002. **136**(0079-6123 (Print)): p. 109-133.
56. Sala, C. and M. Segal, *Dendritic spines: the locus of structural and functional plasticity.* Physiol. Rev., 2014. **94**(1): p. 141-188.
57. Kasai, H., et al., *Structural dynamics of dendritic spines in memory and cognition.* Trends Neurosci., 2010. **33**(3): p. 121-129.
58. Yang, G., F. Pan, and W.B. Gan, *Stably maintained dendritic spines are associated with lifelong memories.* Nature, 2009. **462**(7275): p. 920-924.
59. Holtmaat, A. and K. Svoboda, *Experience-dependent structural synaptic plasticity in the mammalian brain.* Nat. Rev. Neurosci., 2009. **10**(9): p. 647-658.

60. Kostovic, I., Rakic, P., *Cytology and time of origin of interstitial neurons in the white matter in infant and adult human and monkey telencephalon.* J. Neurocytol., 1980. **9**(2): p. 219-242.
61. Huttner, W.B. and M. Brand, *Asymmetric division and polarity of neuroepithelial cells.* Curr. Opin. Neurobiol., 1997. **7**(1): p. 29-39.
62. Taverna, E., M. Gotz, and W.B. Huttner, *The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex.* Annu. Rev. Cell Dev. Biol., 2014. **30**(1530-8995 (Electronic)): p. 465-502.
63. Weigmann, A., et al., *Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells.* Proc. Natl. Acad. Sci., 1997. **94**(23): p. 12425-12430.
64. Sykes, A.M. and W.B. Huttner, *Prominin-1 (CD133) and the Cell Biology of Neural Progenitors and Their Progeny.* Adv. Exp. Med. Biol., 2013. **777**(0065-2598 (Print)): p. 89-98.
65. Gotz, M. and W.B. Huttner, *The cell biology of neurogenesis.* Nat. Rev. Mol. Cell Biol., 2005. **6**(10): p. 777-788.
66. Kosodo, Y., et al., *Asymmetric distribution of the apical plasma membrane during neurogenic divisions of mammalian neuroepithelial cells.* EMBO J., 2004. **23**(11): p. 2314-2324.
67. Aaku-Saraste, E., A. Hellwig, and W.B. Huttner, *Loss of occludin and functional tight junctions, but not ZO-1, during neural tube closure--remodeling of the neuroepithelium prior to neurogenesis.* Dev. Biol., 1996. **180**(2): p. 664-679.
68. Zhadanov, A.B., et al., *Absence of the tight junctional protein AF-6 disrupts epithelial cell-cell junctions and cell polarity during mouse development.* Curr. Biol., 1999. **9**(16): p. 880-888.
69. Marthiens, V. and C. ffrench-Constant, *Adherens junction domains are split by asymmetric division of embryonic neural stem cells.* EMBO Rep., 2009. **10**(5): p. 515-520.
70. Chenn, A., et al., *Intrinsic polarity of mammalian neuroepithelial cells.* Mol. Cell Neurosci., 1998. **11**(4): p. 183-193.
71. Taverna, E. and W.B. Huttner, *Neural progenitor nuclei IN motion.* Neuron, 2010. **67**(6): p. 906-914.
72. Sauer, M. and B.E. Walker, *Radioautographic study of interkinetic nuclear migration in the neural tube.* Proc. Soc. Exp. Biol. Med., 1959. **101**(3): p. 557-560.
73. Spear, P.C. and C.A. Erickson, *Interkinetic nuclear migration: a mysterious process in search of a function.* Dev. Growth Differ., 2012. **54**(3): p. 306-316.

74. Rakic, P., *A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution.* Trends Neurosci., 1995. **18**(9): p. 383-388.
75. Bystron, I., C. Blakemore, and P. Rakic, *Development of the human cerebral cortex: Boulder Committee revisited.* Nat. Rev. Neurosci., 2008. **9**(2): p. 110-122.
76. Marin-Padilla, M., *The mammalian neocortex new pyramidal neuron: a new conception.* Front. Neuroanat., 2014. **7**(51).
77. Meyer, G., *Genetic control of neuronal migrations in human cortical development.* Adv. Anat. Embryol. Cell Biol., 2007. **189**(1-111).
78. Zecevic, N., *Synaptogenesis in layer I of the human cerebral cortex in the first half of gestation.* Cereb. Cortex, 1998. **8**(3): p. 245-252.
79. Super, H., E. Soriano, and H.B. Uylings, *The functions of the preplate in development and evolution of the neocortex and hippocampus.* Brain Res. Brain Res. Rev., 1998. **27**(1): p. 40-64.
80. Bystron, I., et al., *The first neurons of the human cerebral cortex.* Nat. Neurosci., 2006. **9**(7): p. 880-886.
81. Meyer, G., et al., *Embryonic and early fetal development of the human neocortex.* J. Neurosci., 2000. **20**(5): p. 1858-1868.
82. Rakic, S. and N. Zecevic, *Emerging complexity of layer I in human cerebral cortex.* Cereb. Cortex, 2003. **13**(10): p. 1072-1083.
83. Meyer, G., et al., *Expression of p73 and Reelin in the developing human cortex.* J. Neurosci., 2002. **22**(12): p. 4973-4986.
84. Tissir, F. and A.M. Goffinet, *Reelin and brain development.* Nat. Rev. Neurosci., 2003. **4**(6): p. 496-505.
85. Bystron, I., et al., *Tangential networks of precocious neurons and early axonal outgrowth in the embryonic human forebrain.* J. Neurosci., 2005. **25**(11): p. 2781-2792.
86. Huttner, W.B. and Y. Kosodo, *Symmetric versus asymmetric cell division during neurogenesis in the developing vertebrate central nervous system.* Curr. Opin. Cell Biol., 2005. **17**(6): p. 648-657.
87. Hartfuss, E., et al., *Characterization of CNS precursor subtypes and radial glia.* Dev. Biol., 2001. **229**(1): p. 15-30.
88. Kriegstein, A.R. and M. Gotz, *Radial glia diversity: a matter of cell fate.* Glia, 2003. **43**(1): p. 37-43.
89. Chenn, A. and S.K. McConnell, *Cleavage orientation and the asymmetric inheritance of Notch1 immunoreactivity in mammalian neurogenesis.* Cell, 1995. **82**(4): p. 631-641.
90. Gotz, M., *Glial cells generate neurons--master control within CNS regions: developmental perspectives on neural stem cells.* Neuroscientist, 2003. **9**(5): p. 379-397.

91. Halfter, W., et al., *A critical function of the pial basement membrane in cortical histogenesis.* J. Neurosci., 2002. **22**(14): p. 6029-6040.
92. Campbell, K. and M. Gotz, *Radial glia: multi-purpose cells for vertebrate brain development.* Trends Neurosci., 2002. **25**(5): p. 235-238.
93. Malatesta, P., E. Hartfuss, and M. Gotz, *Isolation of radial glial cells by fluorescent-activated cell sorting reveals a neuronal lineage.* Development, 2000. **127**(24): p. 5253-5263.
94. Rakic, P., *Mode of cell migration to the superficial layers of fetal monkey neocortex.* J. Comp. Neurol., 1972. **145**(1): p. 61-83.
95. Rakic, P., *Guidance of neurons migrating to the fetal monkey neocortex.* Brain Res., 1971. **33**(2): p. 471-476.
96. Noctor, S.C., et al., *Neurons derived from radial glial cells establish radial units in neocortex.* Neuron, 2001. **409**(6821): p. 714-720.
97. Miyata, T., et al., *Asymmetric inheritance of radial glial fibers by cortical neurons.* Neuron, 2001. **31**(5): p. 727-741.
98. Miyata, T., et al., *Asymmetric production of surface-dividing and non-surface-dividing cortical progenitor cells.* Development, 2004. **131**(13): p. 3133-3145.
99. Stancik, E.K., et al., *Heterogeneity in ventricular zone neural precursors contributes to neuronal fate diversity in the postnatal neocortex.* J. Neurosci., 2010. **30**(20): p. 7028-7036.
100. LaMonica, B.E., et al., *Mitotic spindle orientation predicts outer radial glial cell generation in human neocortex.* Nat. Commun., 2013. **4**(1665).
101. Betizeau, M., et al., *Precursor diversity and complexity of lineage relationships in the outer subventricular zone of the primate.* Neuron, 2013. **80**(2): p. 442-457.
102. Fietz, S.A. and W.B. Huttner, *Cortical progenitor expansion, self-renewal and neurogenesis-a polarized perspective.* Curr. Opin. Neurobiol., 2011. **21**(1): p. 23-35.
103. Fietz, S.A., et al., *OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling.* Nat. Neurosci., 2010. **13**(6): p. 690-699.
104. Pilz, G.A., et al., *Amplification of progenitors in the mammalian telencephalon includes a new radial glial cell type.* Nat. Commun., 2013. **4**(2125).
105. Anthony, T.E., et al., *Radial glia serve as neuronal progenitors in all regions of the central nervous system.* Neuron, 2004. **41**(6): p. 881-890.
106. Shitamukai, A., D. Konno, and F. Matsuzaki, *Oblique radial glial divisions in the developing mouse neocortex induce self-renewing*

progenitors outside the germinal zone that resemble primate outer subventricular zone progenitors. J. Neurosci., 2011. **31**(10): p. 3683-3695.
107. Noctor, S.C., et al., *Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases.* Nat. Neurosci., 2004. **7**(2): p. 136-144.
108. Dehay, C., H. Kennedy, and K.S. Kosik, *The outer subventricular zone and primate-specific cortical complexification.* Neuron, 2015. **85**(4): p. 683-694.
109. Lukaszewicz, A., et al., *G1 phase regulation, area-specific cell cycle control, and cytoarchitectonics in the primate cortex.* Neuron, 2005. **47**(3): p. 353-364.
110. Zecevic, N., Y. Chen, and R. Filipovic, *Contributions of cortical subventricular zone to the development of the human cerebral cortex.* J. Comp. Neurol., 2005. **491**(2): p. 109-122.
111. Lui, J.H., D.V. Hansen, and A.R. Kriegstein, *Development and evolution of the human neocortex.* Cell, 2011. **146**(1): p. 18-36.
112. Dehay, C. and H. Kennedy, *Cell-cycle control and cortical development.* Nat. Rev. Neurosci., 2007. **8**(6): p. 438-450.
113. Haubensak, W., et al., *Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: a major site of neurogenesis.* Proc. Natl. Acad. Sci. , 2004. **101**(9): p. 3196-3120.
114. Fietz, S.A., et al., *Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal.* Proc. Natl. Acad. Sci., 2012. **109**(29): p. 11836-11841.
115. Caviness, V.S., Jr., T. Takahashi, and R.S. Nowakowski, *Numbers, time and neocortical neuronogenesis: a general developmental and evolutionary model.* Trands Neurosci., 1995. **18**(9): p. 379-383.
116. Reillo, I., et al., *A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex.* Cereb. Cortex, 2011. **21**(7): p. 1674-1694.
117. Sidman, R.I. and P. Rakic, *Neuronal migration, with special reference to developing human brain: a review.* Brain Res., 1973. **62**(1): p. 1-35.
118. Rakic, P., *Neuronal migration and contact guidance in the primate telencephalon.* Postgrad. Med. J., 1978. **54**(Suppl. 1): p. 25-40.
119. Rakic, P., *Neurons in rhesus monkey visual cortex: systematic relation between time of origin and eventual disposition.* Science, 1974. **183**(4123): p. 425-427.

120. Workman, A.D., et al., *Modeling transformations of neurodevelopmental sequences across mammalian species.* J. Neurosci., 2013. **33**(17): p. 7368-7383.
121. Bayatti, N., et al., *A molecular neuroanatomical study of the developing human neocortex from 8 to 17 postconceptional weeks revealing the early differentiation of the subplate and subventricular zone.* Cereb. Cortex, 2008. **18**(7): p. 1536-1548.
122. Nadarajah, B. and J.G. Parnavelas, *Modes of neuronal migration in the developing cerebral cortex.* Nat. Rev. Neurosci., 2002. **3**(6): p. 423-432.
123. Nadarajah, B., et al., *Two modes of radial migration in early development of the cerebral cortex.* Nat. Neurosci., 2001. **4**(2): p. 143-150.
124. Marin-Padilla, M., *Cajal-Retzius cells and the development of the neocortex.* Trends Neurosci., 1998. **21**(2): p. 64-71.
125. Caviness, V.S., Jr., *Patterns of cell and fiber distribution in the neocortex of the reeler mutant mouse.* J. Comp. Neurol., 1976. **170**(4): p. 435-447.
126. Hoerder-Suabedissen, A. and Z. Molnar, *Development, evolution and pathology of neocortical subplate neurons.* Nat. Rev. Neurosci., 2015. **16**(3): p. 133-146.
127. Kornack, D.R. and P. Rakic, *Radial and horizontal deployment of clonally related cells in the primate neocortex: relationship to distinct mitotic lineages.* Neuron, 1995. **15**(2): p. 311-321.
128. Hoerder-Suabedissen, A. and Z. Molnar, *Molecular diversity of early-born subplate neurons.* Cereb. Cortex, 2013. **23**(6): p. 1473-1483.
129. Honig, L.S., K. Herrmann, and C.J. Shatz, *Developmental changes revealed by immunohistochemical markers in human cerebral cortex.* Cereb. Cortex, 1996. **6**(6): p. 794-806.
130. Kostovic, I. and P. Rakic, *Developmental history of the transient subplate zone in the visual and somatosensory cortex of the macaque monkey and human brain.* J. Comp. Neurol., 1990. **297**(3): p. 441-470.
131. Kanold, P.O. and H.J. Luhmann, *The subplate and early cortical circuits.* Annu. Rev. Neurosci., 2010. **33**(1545-4126 (Electronic)): p. 23-48.
132. Vasistha, N.A., et al., *Cortical and Clonal Contribution of Tbr2 Expressing Progenitors in the Developing Mouse Brain.* Cereb. Cortex, 2015. **25**(10): p. 3290-3302.
133. Anderson, S.A., et al., *Interneuron migration from basal forebrain to neocortex: dependence on Dlx genes.* Science, 1997. **278**(5337): p. 474-476.

134. Al-Jaberi, N., et al., *The early fetal development of human neocortical GABAergic interneurons.* Cereb. Cortex, 2015. **25**(3): p. 631-645.
135. De Carlos, J.A. and D.D. O'Leary, *Growth and targeting of subplate axons and establishment of major cortical pathways.* J. Neurosci., 1992. **12**(4): p. 1194-1211.
136. deAzevedo, L.C., C. Hedin-Pereira, and R. Lent, *Callosal neurons in the cingulate cortical plate and subplate of human fetuses.* J. Comp. Neurol., 1997. **386**(1): p. 60-70.
137. Goldman-Rakic, P.S., *Neuronal development and plasticity of association cortex in primates.* Neurosci. Res. Program Bull., 1982. **20**(4): p. 520-532.
138. Rakic, P., *Prenatal genesis of connections subserving ocular dominance in the rhesus monkey.* Nature, 1976. **261**(5560): p. 467-471.
139. Wise, S.P. and E.G. Jones, *Developmental studies of thalamocortical and commissural connections in the rat somatic sensory cortex.* J. Comp. Neurol., 1978. **178**(2): p. 187-208.
140. Allendoerfer, K.L. and C.J. Shatz, *The subplate, a transient neocortical structure: its role in the development of connections between thalamus and cortex.* Annu. Rev. Neurosci., 1994. **17**(0147-006X (Print)): p. 185-218.
141. Kostovic, I., *Prenatal development of nucleus basalis complex and related fiber systems in man: a histochemical study.* Neuroscience, 1986. **17**(4): p. 1047-1077.
142. Kostovic, I. and M. Judas, *The development of the subplate and thalamocortical connections in the human foetal brain.* Acta Paediatr., 2010. **99**(8): p. 1119-1127.
143. Marin, O. and J.L. Rubenstein, *Cell migration in the forebrain.* Annu. Rev. Neurosci., 2003. **26**(0147-006X (Print)): p. 441-483.
144. Angevine, J.B.J. and R.L. Sidman, *Autoradiographic study of cell migration during histogenesis of cerebral cortex in the mouse.* Nature, 1961. **192**: p. 766-768.
145. Shibata, M., F.O. Gulden, and N. Sestan, *From trans to cis: transcriptional regulatory networks in neocortical development.* Trends Genet., 2015. **31**(2): p. 77-87.
146. Samuels, B.A. and L.H. Tsai, *Nucleokinesis illuminated.* Nat. Neurosci., 2004. **7**(11): p. 1169-1170.
147. Kriegstein, A.R. and S.C. Noctor, *Patterns of neuronal migration in the embryonic cortex.* Trends Neurosci., 2004. **27**(7): p. 392-399.
148. Kang, H.J., et al., *Spatio-temporal transcriptome of the human brain.* Nature, 2011. **478**(7370): p. 483-9.

149. O'Leary, D.D., S. Chou, and S. Sahara, *Area patterning of the mammalian cortex.* Neuron, 2007. **56**(2): p. 252-269.
150. Batista-Brito, R. and G. Fishell, *The developmental integration of cortical interneurons into a functional network.* Curr. Top. Dev. Biol., 2009. **87**: p. 81-118.
151. Miller, J.A., et al., *Transcriptional landscape of the prenatal human brain.* Nature, 2014. **508**(7495): p. 199-206.
152. Kawauchi, T. and M. Hoshino, *Molecular pathways regulating cytoskeletal organization and morphological changes in migrating neurons.* Dev. Neurosci., 2008. **30**(1-3): p. 36-46.
153. Petanjek, Z., et al., *Extraordinary neoteny of synaptic spines in the human prefrontal cortex.* Proc .Natl. Acad. Sci., 2011. **108**(32): p. 13281-13286.
154. Tau, G.Z. and B.S. Peterson, *Normal development of brain circuits.* Neuropsychopharmacology, 2010. **35**(1): p. 147-168.
155. Larsen, C.C., et al., *Total number of cells in the human newborn telencephalic wall.* Neuroscience, 2006. **139**(3): p. 999-1003.
156. Spalding, K.L., et al., *Retrospective birth dating of cells in humans.* Cell, 2005. **122**(1): p. 133-143.
157. Sanai, N., et al., *Corridors of migrating neurons in the human brain and their decline during infancy.* Nature, 2011. **478**(7369): p. 382-386.
158. Howard, B.M., et al., *Radial glia cells in the developing human brain.* Neuroscientist, 2008. **14**(5): p. 459-473.
159. Jakovcevski, I., et al., *Oligodendrocyte development and the onset of myelination in the human fetal brain.* Front Neurosci., 2009. **3**(5).
160. Yeung, M.S., et al., *Dynamics of oligodendrocyte generation and myelination in the human brain.* Cell, 2014. **159**(4): p. 766-774.
161. Sigaard, R.K., M. Kjaer, and B. Pakkenberg, *Development of the Cell Population in the Brain White Matter of Young Children.* Cereb. Cortex, 2016. **26**(1): p. 89-95.
162. Petanjek, Z., et al., *Lifespan alterations of basal dendritic trees of pyramidal neurons in the human prefrontal cortex: a layer-specific pattern.* Cereb. Cortex, 2008. **18**(4): p. 915-929.
163. Schuurmans, C. and F. Guillemot, *Molecular mechanisms underlying cell fate specification in the developing telencephalon.* Curr. Opin. Neurobiol., 2002. **12**(1): p. 26-34.
164. Hebert, J.M. and G. Fishell, *The genetics of early telencephalon patterning: some assembly required.* Nat. Rev. Neurosci., 2008. **9**(9): p. 678-685.

165. Hemmati-Brivanlou, A. and D.A. Melton, *A truncated activin receptor inhibits mesoderm induction and formation of axial structures in Xenopus embryos.* Nature, 1992. **359**(6396): p. 609-614.
166. Hemmati-Brivanlou, A., O.G. Kelly, and D.A. Melton, *Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity.* Cell, 1994. **77**(2): p. 283-295.
167. Sasal, Y., et al., *Regulation of neural induction by the Chd and Bmp-4 antagonistic patterning signals in Xenopus.* Nature, 1995. **378**(6555): p. 419.
168. Piccolo, S., et al., *Dorsoventral patterning in Xenopus: inhibition of ventral signals by direct binding of chordin to BMP-4.* Cell, 1996. **86**(4): p. 589-598.
169. Lamb, T.M., et al., *Neural induction by the secreted polypeptide noggin.* Science, 1993. **262**(5134): p. 713-718.
170. Smith, W.C. and R.M. Harland, *Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in Xenopus embryos.* Cell, 1992. **70**(5): p. 829-840.
171. Munoz-Sanjuan, I. and A.H. Brivanlou, *Neural induction, the default model and embryonic stem cells.* Nat. Rev. Neurosci., 2002. **3**(4): p. 271-280.
172. Schmierer, B. and C.S. Hill, *TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility.* Nat. Rev. Mol. Cell Biol., 2007. **8**(12): p. 970-982.
173. Shi, Y. and J. Massague, *Mechanisms of TGF-beta signaling from cell membrane to the nucleus.* Cell, 2013. **113**(6): p. 685-700.
174. Feng, X.H. and R. Derynck, *Specificity and versatility in tgf-beta signaling through Smads.* Ann. Rev. Cell Dev. Biol., 2005. **21**(1081-0706 (Print)): p. 659-693.
175. Allendorph, G.P., W.W. Vale, and S. Choe, *Structure of the ternary signaling complex of a TGF-beta superfamily member.* Proc .Natl. Acad. Sci., 2006. **103**(20): p. 7643-7648.
176. Chambers, S.M., et al., *Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling.* Nat Biotechnol, 2009. **27**(3): p. 275-80.
177. Elkabetz, Y., et al., *Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage.* Genes Dev., 2008. **22**(2): p. 152-165.
178. Lee, H., et al., *Directed differentiation and transplantation of human embryonic stem cell-derived motoneurons.* Stem Cells, 2007. **25**(8): p. 1931-1939.

179. Mason, I., *Initiation to end point: the multiple roles of fibroblast growth factors in neural development.* Nat. Rev. Neurosci., 2007. **8**(8): p. 583-596.
180. Guillemot, F. and C. Zimmer, *From cradle to grave: the multiple roles of fibroblast growth factors in neural development.* Neuron, 2011. **71**(4): p. 574-588.
181. Bottcher, R.T. and C. Niehrs, *Fibroblast growth factor signaling during early vertebrate development.* Endocr. Rev., 2005. **26**(1): p. 63-77.
182. Londin, E.R., J. Niemiec, and H.I. Sirotkin, *Chordin, FGF signaling, and mesodermal factors cooperate in zebrafish neural induction.* Dev. Biol., 2005. **279**(1): p. 1-19.
183. Pera, E.M., et al., *Integration of IGF, FGF, and anti-BMP signals via Smad1 phosphorylation in neural induction.* Genes Dev., 2003. **17**(24): p. 3023-3028.
184. Albazerchi, A. and C.D. Stern, *A role for the hypoblast (AVE) in the initiation of neural induction, independent of its ability to position the primitive streak.* Dev. Biol., 2007. **301**(2): p. 489-503.
185. Marchal, L., et al., *BMP inhibition initiates neural induction via FGF signaling and Zic genes.* Proc .Natl. Acad. Sci., 2009. **106**(41): p. 17437-17442.
186. Storm, E.E., et al., *Dose-dependent functions of Fgf8 in regulating telencephalic patterning centers.* Development, 2006. **133**(9): p. 1831-1844.
187. Raballo, R., et al., *Basic fibroblast growth factor (Fgf2) is necessary for cell proliferation and neurogenesis in the developing cerebral cortex.* J. Neurosci., 2000. **20**(13): p. 5012-5023.
188. Sahara, S. and D.D. O'Leary, *Fgf10 regulates transition period of cortical stem cell differentiation to radial glia controlling generation of neurons and basal progenitors.* Neuron, 2009. **63**(1): p. 48-62.
189. Maric, D., et al., *Prospective cell sorting of embryonic rat neural stem cells and neuronal and glial progenitors reveals selective effects of basic fibroblast growth factor and epidermal growth factor on self-renewal and differentiation.* J. Neurosci., 2003. **23**(1): p. 240-251.
190. Palmer, T.D., et al., *Fibroblast growth factor-2 activates a latent neurogenic program in neural stem cells from diverse regions of the adult CNS.* J. Neurosci., 1999. **19**(19): p. 8487-8497.
191. Qian, X., et al., *FGF2 concentration regulates the generation of neurons and glia from multipotent cortical stem cells.* Neuron, 1997. **18**(1): p. 81-93.

192. Gunhaga, L., et al., *Specification of dorsal telencephalic character by sequential Wnt and FGF signaling.* Nat. Neurosci., 2003. **6**(7): p. 701-707.
193. Kuschel, S., U. Ruther, and T. Theil, *A disrupted balance between Bmp/Wnt and Fgf signaling underlies the ventralization of the Gli3 mutant telencephalon.* Dev. Biol., 2003. **260**(2): p. 484-495.
194. Viti, J., A. Gulacsi, and L. Lillien, *Wnt regulation of progenitor maturation in the cortex depends on Shh or fibroblast growth factor 2.* J. Neurosci., 2003. **23**(13): p. 5919-5927.
195. Sirko, S., et al., *Chondroitin sulfates are required for fibroblast growth factor-2-dependent proliferation and maintenance in neural stem cells and for epidermal growth factor-dependent migration of their progeny.* Stem Cells, 2010. **28**(4): p. 775-787.
196. Jen, Y.H., M. Musacchio, and A.D. Lander, *Glypican-1 controls brain size through regulation of fibroblast growth factor signaling in early neurogenesis.* Nerual. Dev., 2009. **4**(33).
197. Ghosh, A. and M.E. Greenberg, *Distinct roles for bFGF and NT-3 in the regulation of cortical neurogenesis.* Neuron, 1995. **15**(1): p. 89-103.
198. Temple, S. and X. Qian, *bFGF, neurotrophins, and the control or cortical neurogenesis.* Neuron, 1995. **15**(2): p. 249-252.
199. Vaccarino, F.M., et al., *Changes in cerebral cortex size are governed by fibroblast growth factor during embryogenesis.* Nat. Neurosci., 1999. **2**(9): p. 848.
200. Kang, W., et al., *The transition from radial glial to intermediate progenitor cell is inhibited by FGF signaling during corticogenesis.* J. Neurosci., 2009. **29**(46): p. 14571-14580.
201. McCabe, K.L. and M. Bronner-Fraser, *Molecular and tissue interactions governing induction of cranial ectodermal placodes.* Dev. Biol., 2009. **332**(2): p. 189-195.
202. Streit, A., *The preplacodal region: an ectodermal domain with multipotential progenitors that contribute to sense organs and cranial sensory ganglia.* Int. J. Dev. Biol., 2007. **51**(6-7): p. 447-461.
203. Liu, J.P., E. Laufer, and T.M. Jessell, *Assigning the positional identity of spinal motor neurons: rostrocaudal patterning of Hox-c expression by FGFs, Gdf11, and retinoids.* Neuron, 2001. **32**(6): p. 997-1012.
204. Crossley, P.H., et al., *Coordinate expression of Fgf8, Otx2, Bmp4, and Shh in the rostral prosencephalon during development of the telencephalic and optic vesicles.* Neuroscience, 2001. **108**(2): p. 183-206.

205. Cholfin, J.A. and J.L. Rubenstein, *Frontal cortex subdivision patterning is coordinately regulated by Fgf8, Fgf17, and Emx2.* J. Comp. Neurol., 2008. **509**(2): p. 144-155.
206. Garel, S., K.J. Huffman, and J.L. Rubenstein, *Molecular regionalization of the neocortex is disrupted in Fgf8 hypomorphic mutants.* Development, 2003. **130**(9): p. 1903-1914.
207. O'Leary, D.D. and S. Sahara, *Genetic regulation of arealization of the neocortex.* Curr. Opin. Biotechnol., 2008. **18**(1): p. 90-100.
208. Cholfin, J.A. and J.L. Rubenstein, *Patterning of frontal cortex subdivisions by Fgf17.* Proc .Natl. Acad. Sci., 2007. **104**(18): p. 7652-7657.
209. Shimogori, T., et al., *Embryonic signaling centers expressing BMP, WNT and FGF proteins interact to pattern the cerebral cortex.* Development, 2004. **131**(22): p. 5639-5647.
210. Paek, H., G. Gutin, and J.M. Hebert, *FGF signaling is strictly required to maintain early telencephalic precursor cell survival.* Development, 2009. **136**(4): p. 2457-2465.
211. Fukuchi-Shimogori, T. and E.A. Grove, *Neocortex patterning by the secreted signaling molecule FGF8.* Science, 2001. **294**(5544): p. 1071-1074.
212. Okada, T., et al., *FGF8 signaling patterns the telencephalic midline by regulating putative key factors of midline development.* Dev. Biol., 2008. **320**(1): p. 92-101.
213. Hoch, R.V., J.l. Rubenstein, and S. Pleasure, *Genes and signaling events that establish regional patterning of the mammalian forebrain.* Semin. Cell Dev. Biol. , 2009. **20**(4): p. 378-386.
214. Qian, X., et al., *Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells.* Neuron, 2000. **28**(1): p. 69-80.
215. Kessaris, N., et al., *Cooperation between sonic hedgehog and fibroblast growth factor/MAPK signalling pathways in neocortical precursors.* Development, 2004. **131**(6): p. 1289-1298.
216. Chandran, S., et al., *FGF-dependent generation of oligodendrocytes by a hedgehog-independent pathway.* Development, 2003. **130**(26): p. 6599-6609.
217. Hasegawa, H., et al., *Laminar patterning in the developing neocortex by temporally coordinated fibroblast growth factor signaling.* J. Neurosci., 2004. **24**(40): p. 8711-8719.
218. Smith, K.M., et al., *Midline radial glia translocation and corpus callosum formation require FGF signaling.* Nat. Neurosci., 2006. **9**(6): p. 787-797.

219. Itoh, N. and D.M. Ornitz, *Functional evolutionary history of the mouse Fgf gene family.* Dev. Dyn., 2008. **237**(1): p. 18-27.
220. Ornitz, D.M. and N. Itoh, *The Fibroblast Growth Factor signaling pathway.* Wiley Interdiscip. Rev. Dev. Biol., 2015. **4**(3): p. 215-266.
221. Beenken, A. and M. Mohammadi, *The FGF family: biology, pathophysiology and therapy.* Nat. Rev. Drug Discov., 2009. **8**(3): p. 235-253.
222. Wesche, J., et al., *FGF-1 and FGF-2 require the cytosolic chaperone Hsp90 for translocation into the cytosol and the cell nucleus.* J. Biol. Chem., 2006. **281**(16): p. 11405-11412.
223. Sorensen, V., et al., *Different abilities of the four FGFRs to mediate FGF-1 translocation are linked to differences in the receptor C-terminal tail.* J. Cell Sci., 2006. **119**(Pt 20): p. 4332-4341.
224. Goetz, R. and M. Mohammadi, *Exploring mechanisms of FGF signalling through the lens of structural biology.* Nat. Rev. Mol. Cell Biol., 2013. **14**(3): p. 166-180.
225. Eswarakumar, V.P., I. Lax, and J. Schlessinger, *Cellular signaling by fibroblast growth factor receptors.* Cytokine Growth Factor Rev., 2005. **16**(2): p. 139-149.
226. Dailey, L., et al., *Mechanisms underlying differential responses to FGF signaling.* Cytokine Growth Factor Rev., 2005. **16**(2): p. 233-247.
227. Kouhara, H., et al., *A lipid-anchored Grb2-binding protein that links FGF-receptor activation to the Ras/MAPK signaling pathway.* Cell, 1997. **89**(5): p. 693-702.
228. Hadari, Y.R., et al., *Binding of Shp2 tyrosine phosphatase to FRS2 is essential for fibroblast growth factor-induced PC12 cell differentiation.* Mol. Cell Biol., 1998. **18**(7): p. 3966-3973.
229. Lamothe, B., et al., *The docking protein Gab1 is an essential component of an indirect mechanism for fibroblast growth factor stimulation of the phosphatidylinositol 3-kinase/Akt antiapoptotic pathway.* Mol. Cell Biol., 2004. **24**(13): p. 5657-5666.
230. Manning, B.D. and L.C. Cantley, *AKT/PKB signaling: navigating downstream.* Cell, 2007. **129**(7): p. 1261-1274.
231. Firnberg, N. and A. Neubuser, *FGF signaling regulates expression of Tbx2, Erm, Pea3, and Pax3 in the early nasal region.* Dev. Biol., 2002. **247**(2): p. 237-250.
232. Yordy, J.S. and R.C. Muise-Helmericks, *Signal transduction and the Ets family of transcription factors.* Oncogene, 2000. **19**(55): p. 6503-6513.

233. Faedo, A., U. Borello, and J.L. Rubenstein, *Repression of Fgf signaling by sprouty1-2 regulates cortical patterning in two distinct regions and times.* J. Neurosci., 2010. **30**(11): p. 4015-4023.
234. Reich, N.C. and L. Liu, *Tracking STAT nuclear traffic.* Nat. Rev. Immunol., 2006. **6**(8): p. 602-612.
235. Xu, X., Y.l. Sun, and T. Hoey, *Cooperative DNA binding and sequence-selective recognition conferred by the STAT amino-terminal domain.* Science, 1996. **273**(5276): p. 794-797.
236. Carpenter, G. and Q. Ji, *Phospholipase C-gamma as a signal-transducing element.* Exp. Cell Res., 1999. **253**(1): p. 15-24.
237. Schlessinger, J., *Cell signaling by receptor tyrosine kinases.* Cell, 2000. **103**(2): p. 211-225.
238. Doherty, P. and F.S. Walsh, *CAM-FGF receptor interactions: a model for axonal growth.* Mol. Cell Neurosci., 1996. **8**(2-3): p. 99-111.
239. Browaeys-Poly, E., et al., *Grb14 inhibits FGF receptor signaling through the regulation of PLCgamma recruitment and activation.* FEBS Lett., 2010. **584**(21): p. 4383-4388.
240. Assimacopoulos, S., E.A. Grove, and C.W. Ragsdale, *Identification of a Pax6-dependent epidermal growth factor family signaling source at the lateral edge of the embryonic cerebral cortex.* J. Neurosci., 2003. **23**(16): p. 6399-6403.
241. Kornblum, H.I., et al., *Prenatal ontogeny of the epidermal growth factor receptor and its ligand, transforming growth factor alpha, in the rat brain.* J. Comp. Neurol., 1997. **380**(2): p. 243-261.
242. Eagleson, K.L., R.T. Ferri, and P. Levitt, *Complementary distribution of collagen type IV and the epidermal growth factor receptor in the rat embryonic telencephalon.* Cereb. Cortex, 1996. **6**(3): p. 540-549.
243. Scaltriti, M. and J. Baselga, *The epidermal growth factor receptor pathway: a model for targeted therapy.* Clin. Cancer Res., 2006. **12**(18): p. 5268-5272.
244. Conti, L. and E. Cattaneo, *Neural stem cell systems: physiological players or in vitro entities?* Nat. Rev. Neurosci., 2010. **11**(3): p. 176-187.
245. Conti, L., et al., *Niche-independent symmetrical self-renewal of a mammalian tissue stem cell.* PLoS Biol., 2005. **3**(9): p. e283.
246. Tropepe, V., et al., *Distinct neural stem cells proliferate in response to EGF and FGF in the developing mouse telencephalon.* Dev. Biol., 1999. **208**(1): p. 166-188.
247. Gage, F.H., *Mammalian neural stem cells.* Science, 2000. **287**(5457): p. 1433-1438.

248. Temple, S., *The development of neural stem cells.* Nature, 2001. **414**(6859): p. 112-117.
249. Ostenfeld, T. and C.N. Svendsen, *Requirement for neurogenesis to proceed through the division of neuronal progenitors following differentiation of epidermal growth factor and fibroblast growth factor-2-responsive human neural stem cells.* Stem Cells, 2004. **22**(5): p. 798-811.
250. Lillien, L. and H. Raphael, *BMP and FGF regulate the development of EGF-responsive neural progenitor cells.* Development, 2000. **127**(22): p. 4993-5005.
251. Ciccolini, F. and C.N. Svendsen, *Fibroblast growth factor 2 (FGF-2) promotes acquisition of epidermal growth factor (EGF) responsiveness in mouse striatal precursor cells: identification of neural precursors responding to both EGF and FGF-2.* J. Neurosci., 1998. **18**(19): p. 7869-7880.
252. Burrows, R.C., et al., *Response diversity and the timing of progenitor cell maturation are regulated by developmental changes in EGFR expression in the cortex.* Neuron, 1997. **19**(2): p. 251-267.
253. Morrow, T., M.R. Song, and A. Ghosh, *Sequential specification of neurons and glia by developmentally regulated extracellular factors.* Development, 2001. **128**(18): p. 3585-3594.
254. Caric, D., et al., *EGFRs mediate chemotactic migration in the developing telencephalon.* Development, 2001. **128**(21): p. 4203-4216.
255. Sun, Y., S.K. Goderie, and S. Temple, *Asymmetric distribution of EGFR receptor during mitosis generates diverse CNS progenitor cells.* Neuron, 2005. **45**(6): p. 873-886.
256. Nicholas, M.K., et al., *Epidermal growth factor receptor - mediated signal transduction in the development and therapy of gliomas.* Clin. Cancer Res., 2006. **12**(24): p. 7261-7270.
257. Okano, J., et al., *Akt/protein kinase B isoforms are differentially regulated by epidermal growth factor stimulation.* J. Biol. Chem., 2000. **275**(40): p. 30934-30942.
258. Maden, M., *Retinoic acid in the development, regeneration and maintenance of the nervous system.* Nat. Rev. Neurosci., 2007. **8**(10): p. 755-765.
259. Maden, M., *Retinoid signalling in the development of the central nervous system.* Nat. Rev. Neurosci., 2002. **3**(11): p. 843-853.
260. Melton, K.R., A. Iulianella, and P.A. Trainor, *Gene expression and regulation of hindbrain and spinal cord development.* Front Biosci., 2004. **9**: p. 117-138.

261. Wilson, L., et al., *Retinoic acid and the control of dorsoventral patterning in the avian spinal cord.* Dev. Biol., 2004. **269**(2): p. 433-446.
262. Maden, M., et al., *Vitamin A-deficient quail embryos have half a hindbrain and other neural defects.* Curr. Biol., 1996. **6**(4): p. 417-426.
263. Gale, E., M. Zile, and M. Maden, *Hindbrain respecification in the retinoid-deficient quail.* Mech. Dev., 1999. **89**(1-2): p. 43-54.
264. Glover, J.C., J.S. Renaud, and F.M. Rijli, *Retinoic acid and hindbrain patterning.* J. Neurosci., 2006. **66**(7): p. 705-725.
265. Novitch, B.G., et al., *A requirement for retinoic acid-mediated transcriptional activation in ventral neural patterning and motor neuron specification.* Neuron, 2003. **40**(1): p. 81-95.
266. Diez del Corral, R. and K.G. Storey, *Opposing FGF and retinoid pathways: a signalling switch that controls differentiation and patterning onset in the extending vertebrate body axis.* Bioessays, 2004. **26**(8): p. 857-869.
267. Wilson, L. and M. Maden, *The mechanisms of dorsoventral patterning in the vertebrate neural tube.* Dev. Biol., 2005. **282**(1): p. 1-13.
268. Wohl, C.A. and S. Weiss, *Retinoic acid enhances neuronal proliferation and astroglial differentiation in cultures of CNS stem cell-derived precursors.* J. Neurobiol., 1998. **37**(2): p. 281-290.
269. Giardino, L., C. Bettelli, and L. Calza, *In vivo regulation of precursor cells in the subventricular zone of adult rat brain by thyroid hormone and retinoids.* Neurosci. Lett., 2000. **295**(1-2): p. 17-20.
270. Wang, T.W., H. Zhang, and J.M. Parent, *Retinoic acid regulates postnatal neurogenesis in the murine subventricular zone-olfactory bulb pathway.* Development, 2005. **132**(12): p. 2721-2732.
271. Jang, Y.K., et al., *Retinoic acid-mediated induction of neurons and glial cells from human umbilical cord-derived hematopoietic stem cells.* J. Neruosci. Res., 2004. **75**(4): p. 573-584.
272. Okada, Y., et al., *Retinoic-acid-concentration-dependent acquisition of neural cell identity during in vitro differentiation of mouse embryonic stem cells.* Dev. Biol., 2004. **275**(1): p. 124-142.
273. Lee, S., et al., *Retinoid signaling and neurogenin2 function are coupled for the specification of spinal motor neurons through a chromatin modifier CBP.* (1097-4199 (Electronic)).
274. Ensini, M., et al., *The control of rostrocaudal pattern in the developing spinal cord: specification of motor neuron subtype identity is initiated by signals from paraxial mesoderm.* Develpment, 1998. **125**(6): p. 969-982.

275. Sockanathan, S. and T.M. Jessell, *Motor neuron-derived retinoid signaling specifies the subtype identity of spinal motor neurons.* Cell, 1998. **94**(4): p. 503-514.
276. Guidato, S., C. Barrett, and S. Guthrie, *Patterning of motor neurons by retinoic acid in the chick embryo hindbrain in vitro.* Mol. Cell Neurosci., 2003. **23**(1): p. 81-95.
277. Sockanathan, S., T. Perlmann, and T.M. Jessell, *Retinoid receptor signaling in postmitotic motor neurons regulates rostrocaudal positional identity and axonal projection pattern.* Neuron, 2003. **40**(1): p. 97-111.
278. Maden, M., *Role and distribution of retinoic acid during CNS development.* Int. Rev. Cytol., 2001. **209**: p. 1-77.
279. O'Byrne, S.M. and W.S. Blaner, *Retinol and retinyl esters: biochemistry and physiology.* J. Lipid Res., 2013. **54**(7): p. 1731-1743.
280. Kawaguchi, R., et al., *A membrane receptor for retinol binding protein mediates cellular uptake of vitamin A.* Science, 2007. **315**(5813): p. 820-825.
281. Sandell, L.L., et al., *RDH10 is essential for synthesis of embryonic retinoic acid and is required for limb, craniofacial, and organ development.* Genes Dev., 2007. **21**(9): p. 1113-1124.
282. Budhu, A.S. and N. Noy, *Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest.* Mol. Cell Biol., 2002. **22**(8): p. 2632-2641.
283. Bastien, J. and C. Rochette-Egly, *Nuclear retinoid receptors and the transcription of retinoid-target genes.* Gene, 2004. **328**: p. 1-16.
284. Balmer, J.E. and R. Blomhoff, *Gene expression regulation by retinoic acid.* J. Lipid Res., 2002. **43**(11): p. 1773-1808.
285. Takahashi, H. and F.C. Liu, *Genetic patterning of the mammalian telencephalon by morphogenetic molecules and transcription factors.* Birth Defects Res C. Embryo Today, 2006. **78**(3): p. 256-266.
286. Freese, J.L., D. Pino, and S.J. Pleasure, *Wnt signaling in development and disease.* Neurobiol. Dis., 2010. **38**(2): p. 148-153.
287. Gaulden, J. and J.F. Reiter, *Neur-ons and neur-offs: regulators of neural induction in vertebrate embryos and embryonic stem cells.* Hum. Mol. Genet., 2008. **17**(R1): p. R60-66.
288. Yoshikawa, Y., et al., *Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse.* Dev. Biol., 1997. **183**(2): p. 234-242.

289. Kelly, O.G., K.I. Pinson, and W.C. Skarnes, *The Wnt co-receptors Lrp5 and Lrp6 are essential for gastrulation in mice.* Development, 2004. **131**(12): p. 2803-2815.
290. Mukhopadhyay, M., et al., *Dickkopf1 is required for embryonic head induction and limb morphogenesis in the mouse.* Dev. Cell., 2001. **1**(3): p. 423-434.
291. Chenn, A. and C.A. Walsh, *Increased neuronal production, enlarged forebrains and cytoarchitectural distortions in beta-catenin overexpressing transgenic mice.* Cereb. Cortex, 2003. **13**(6): p. 599-606.
292. Zechner, D., et al., *beta-Catenin signals regulate cell growth and the balance between progenitor cell expansion and differentiation in the nervous system.* Dev. Biol., 2003. **258**(2): p. 406-418.
293. Backman, M., et al., *Effects of canonical Wnt signaling on dorso-ventral specification of the mouse telencephalon.* Dev. Biol., 2005. **279**(1): p. 155-168.
294. Lee, S.M., et al., *A local Wnt-3a signal is required for development of the mammalian hippocampus.* Development, 2000. **127**(3): p. 457-467.
295. Wilson, S.W. and J.L. Rubenstein, *Induction and dorsoventral patterning of the telencephalon.* Neuron, 2000. **28**(3): p. 641-651.
296. Salinas, P.C. and Y. Zou, *Wnt signaling in neural circuit assembly.* Ann. Rev. Neurosci. **31**: p. 339-358.
297. Niehrs, C., *On growth and form: a Cartesian coordinate system of Wnt and BMP signaling specifies bilaterian body axes.* Development, 2010. **137**(6): p. 845-857.
298. Wray, J. and C. Hartmann, *WNTing embryonic stem cells.* Trends Cell Biol., 2012. **22**(3): p. 159-168.
299. Clark, C.E., C.C. Nourse, and H.M. Cooper, *The tangled web of non-canonical Wnt signalling in neural migration.* Neurosignals, 2012. **20**(3): p. 202-220.
300. Goodrich, L.V., *The plane facts of PCP in the CNS.* Neuron, 2008. **60**(1): p. 9-16.
301. De, A., *Wnt/Ca2+ signaling pathway: a brief overview.* Acta Biochim. Biophys. Sin. **43**(10): p. 745-756.
302. Kikuchi, A., et al., *New insights into the mechanism of Wnt signaling pathway activation.* Int. Rev. Cell Mol. Biol., 2011. **291**: p. 21-71.
303. van Amerongen, R., A. Mikels, and R. Nusse, *Alternative wnt signaling is initiated by distinct receptors.* Sci. Signal., 2008. **1**(35): p. re9.

304. He, X., et al., *LDL receptor-related proteins 5 and 6 in Wnt/beta-catenin signaling: arrows point the way.* Development, 2004. **131**(8): p. 1663-1677.
305. Minami, Y., et al., *Ror-family receptor tyrosine kinases in noncanonical Wnt signaling: their implications in developmental morphogenesis and human diseases.* Dev. Dyn., 2010. **239**(1): p. 1-15.
306. Peradziryi, H., N.S. Tolwinski, and A. Borchers, *The many roles of PTK7: a versatile regulator of cell-cell communication.* Arch. Biochem. Biophys., 2012. **524**(1): p. 71-76.
307. Niehrs, C., *The complex world of WNT receptor signalling.* Nat. Rev. Mol. Cell Biol., 2012. **13**(12): p. 767-779.
308. Simons, M. and M. Mlodzik, *Planar cell polarity signaling: from fly development to human disease.* Annu. Rev. Genet., 2008. **42**: p. 517-540.
309. Wallingford, J.B. and R. Habas, *The developmental biology of Dishevelled: an enigmatic protein governing cell fate and cell polarity.* Development, 2005. **132**(20): p. 4421-4436.
310. Huang, H. and X. He, *Wnt/beta-catenin signaling: new (and old) players and new insights.* Curr. Opin. Cell Biol., 2008. **20**(2): p. 119-125.
311. Schwarz-Romond, T., C. Metcalfe, and M. Bienz, *Dynamic recruitment of axin by Dishevelled protein assemblies.* J. Cell Sci., 2007. **120**(Pt 14): p. 2402-2412.
312. Zeng, X., et al., *Initiation of Wnt signaling: control of Wnt coreceptor Lrp6 phosphorylation/activation via frizzled, dishevelled and axin functions.* Development, 2008. **135**(2): p. 367-375.
313. MacDonald, B.T. and X. He, *Frizzled and LRP5/6 receptors for Wnt/beta-catenin signaling. LID - 10.1101/cshperspect.a007880 [doi] LID - a007880 [pii].* Cold Spring Harb. Perspect. Biol., 2012. **4**(12): p. pii: a007880.
314. Davidson, G., et al., *Casein kinase 1 gamma couples Wnt receptor activation to cytoplasmic signal transduction.* Nature, 2005. **438**(7069): p. 867-872.
315. Logan, C.Y. and R. Nusse, *The Wnt signaling pathway in development and disease.* Ann. Rev. Cell Dev. Biol., 2004. **20**: p. 781-810.
316. Clevers, H., *Wnt/beta-catenin signaling in development and disease.* Cell, 2006. **127**(3): p. 469-480.
317. Ingham, P.W. and A.P. McMahon, *Hedgehog signaling in animal development: paradigms and principles.* Genes Dev., 2001. **15**(23): p. 3059-3087.

318. Ho, K.S. and M.P. Scott, *Sonic hedgehog in the nervous system: functions, modifications and mechanisms.* Curr. Opin. Neurobiol., 2002. **12**(1): p. 57-63.
319. Ericson, J., et al., *Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling.* Cell, 1997. **90**(1): p. 169-180.
320. Briscoe, J., et al., *A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube.* Cell, 2000. **101**(4): p. 435-445.
321. Patten, I. and M. Placzek, *The role of Sonic hedgehog in neural tube patterning.* Cell Mol. Life Sci., 2000. **57**(12): p. 1695-1708.
322. Briscoe, J. and J. Ericson, *Specification of neuronal fates in the ventral neural tube.* Curr. Opin. Neurobiol., 2001. **11**(1): p. 43-49.
323. Kohtz, J.D., et al., *Regionalization within the mammalian telencephalon is mediated by changes in responsiveness to Sonic Hedgehog.* Development, 1998. **125**(24): p. 5079-5089.
324. Lupo, G., H.W. A., and K.E. Lewis, *Mechanisms of ventral patterning in the vertebrate nervous system.* Nat. Rev. Neurosci., 2006. **7**(2): p. 103-114.
325. Shi, Y., et al., *Neural stem cell self-renewal.* Crit. Rev. Oncol. Hematol., 2008(1040-8428).
326. Reinchisi, G., et al., *Sonic Hedgehog modulates EGFR dependent proliferation of neural stem cells during late mouse embryogenesis through EGFR transactivation.* Front. Cell Neurosci., 2013. **7**(166).
327. Zhu, G., et al., *Sonic hedgehog and BMP2 exert opposing actions on proliferation and differentiation of embryonic neural progenitor cells.* Dev. Biol., 1999. **215**(1): p. 118-129.
328. Dahmane, N., et al., *The Sonic Hedgehog-Gli pathway regulates dorsal brain growth and tumorigenesis.* Development, 2001. **128**(24): p. 5201-5212.
329. Murone, M., A. Rosenthal, and F.J. de Sauvage, *Sonic hedgehog signaling by the patched-smoothened receptor complex.* Curr. Biol., 1999. **9**(2): p. 76-84.
330. Villavicencio, E.H., D.O. Walterhouse, and P.M. Iannaccone, *The sonic hedgehog-patched-gli pathway in human development and disease.* Am. J. Hum. Genet., 2000. **67**(5): p. 1047-1054.
331. Kogerman, P., et al., *Mammalian suppressor-of-fused modulates nuclear-cytoplasmic shuttling of Gli-1.* Nat. Cell Biol., 1999. **1**(5): p. 312-319.
332. Rimkus, T.K., et al., *Targeting the Sonic Hedgehog Signaling Pathway: Review of Smoothened and GLI Inhibitors. .* Cancers, 2016. **8**(2): p. pii: E22.

333. Meng, E., et al., *The Impact of Hedgehog Signaling Pathway on DNA Repair Mechanisms in Human Cancer.* Cancers, 2015. **7**(3): p. 1333-1348.
334. Fishell, G. and C. Hanashima, *Pyramidal neurons grow up and change their mind.* Neuron, 2008. **57**(3): p. 333-338.
335. Ross, M.E. and C.A. Walsh, *Human brain malformations and their lessons for neuronal migration.* Annu. Rev. Neurosci., 2001. **24**: p. 1041-1070.
336. Rash, B.G. and E.A. Grove, *Area and layer patterning in the developing cerebral cortex.* Curr. Opin. Neurobiol., 2006. **16**(1): p. 25-34.
337. Molyneaux, B.J., et al., *DeCoN: genome-wide analysis of in vivo transcriptional dynamics during pyramidal neuron fate selection in neocortex.* Neuron, 2015. **85**(2): p. 275-288.
338. Johnson, M.B., et al., *Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex.* Nat. Neurosci., 2015. **18**(5): p. 637-646.
339. Florio, M., et al., *Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion.* Science, 2015. **347**(6229): p. 1465-1470.
340. Zeisel, A., et al., *Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq.* Science, 2015. **347**(6226): p. 1138-1142.
341. Pollen, A.A., et al., *Molecular identity of human outer radial glia during cortical development.* Cell, 2015. **163**(1): p. 55-67.
342. Porteus, M.H., et al., *DLX-2, MASH-1, and MAP-2 expression and bromodeoxyuridine incorporation define molecularly distinct cell populations in the embryonic mouse forebrain.* J. Neurosci., 1994. **14**(11 Pt 1): p. 6370-6383.
343. Yun, K., S. Potter, and J.L. Rubenstein, *Gsh2 and Pax6 play complementary roles in dorsoventral patterning of the mammalian telencephalon.* Development, 2001. **128**(2): p. 193-205.
344. Molyneaux, B.J., et al., *Novel subtype-specific genes identify distinct subpopulations of callosal projection neurons.* J. Neurosci., 2009. **29**(39): p. 12343-12354.
345. Bulfone, A., et al., *Spatially restricted expression of Dlx-1, Dlx-2 (Tes-1), Gbx-2, and Wnt-3 in the embryonic day 12.5 mouse forebrain defines potential transverse and longitudinal segmental boundaries.* J. Neurosci., 1993. **13**(7): p. 3155-3172.
346. Hevner, R.F., et al., *Tbr1 regulates differentiation of the preplate and layer 6.* Neuron, 2001. **29**(2): p. 353-366.

347. Saito, T., et al., *Neocortical layer formation of human developing brains and lissencephalies: consideration of layer-specific marker expression.* Cereb. Cortex, 2011. **21**(3): p. 588-596.
348. Lodato, S., et al., *Gene co-regulation by Fezf2 selects neurotransmitter identity and connectivity of corticospinal neurons.* Nat. Neurosci., 2014. **17**(8): p. 1046-1054.
349. Kawaguchi, A., et al., *Single-cell gene profiling defines differential progenitor subclasses in mammalian neurogenesis.* Development, 2008. **135**(18): p. 3113-3124.
350. Arlotta, P., et al., *Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo.* Neuron, 2005. **45**(2): p. 207-221.
351. Greig, L.C., et al., *Molecular logic of neocortical projection neuron specification, development and diversity.* Nat. Rev. Neurosci., 2013. **14**(11): p. 755-769.
352. Visel, A., et al., *A high-resolution enhancer atlas of the developing telencephalon.* Cell, 2013. **152**(4): p. 895-908.
353. Han, W., et al., *TBR1 directly represses Fezf2 to control the laminar origin and development of the corticospinal tract.* PNAS, 2011. **108**(7): p. 3041-3046.
354. Chen, J.G., et al., *Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex.* PNAS, 2005. **102**(49): p. 17792-17797.
355. Kwan, K.Y., et al., *SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons.* PNAS, 2008. **105**(41): p. 16021-16026.
356. Britanova, O., et al., *Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex.* Neuron, 2008. **57**(3): p. 378-392.
357. Mitalipov, S. and D. Wolf, *Totipotency, pluripotency and nuclear reprogramming.* Adv. Biochem. Eng. Biotechnol., 2009. **114**: p. 185-199.
358. Clarke, D.L., et al., *Generalized potential of adult neural stem cells.* Science, 2000. **288**(5471): p. 1660-1663.
359. Reynolds, B.A. and S. Weiss, *Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system.* Science, 1992. **255**(5052): p. 1707-1710.
360. Evans, M.J. and M.H. Kaufman, *Establishment in culture of pluripotential cells from mouse embryos.* Nature, 1981. **292**(5819): p. 154-156.

361. Martin, G.R., *Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.* PNAS, 1981. **78**(12): p. 7634-7638.
362. *A Cell-ebration of Induced Pluripotency.* Cell, 2016. **164**(3): p. 331.
363. Thomson, J.A., et al., *Embryonic stem cell lines derived from human blastocysts.* Science, 1998. **282**(5391): p. 1145-1147.
364. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-676.
365. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors.* Cell, 2007. **131**(5): p. 861-872.
366. Takahashi, K., et al., *Induction of pluripotent stem cells from fibroblast cultures.* Nat. Protoc., 2007. **2**(12): p. 3081-3089.
367. Yu, J., et al., *Induced pluripotent stem cell lines derived from human somatic cells.* Science, 2007. **318**(5858): p. 1917-1920.
368. Doetschman, T.C., et al., *The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium.* J. Embryol. Exp. Morphol., 1985. **87**: p. 27-45.
369. Wobus, A.M., et al., *Characterization of a pluripotent stem cell line derived from a mouse embryo.* Exp. Cell Res., 1984. **152**(1): p. 212-219.
370. Itskovitz-Eldor, J., et al., *Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers.* Mol. Med., 2000. **6**(2): p. 88-95.
371. Dimos, J.T., et al., *Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons.* Science, 2008. **321**(5893): p. 1218-1221.
372. Si-Tayeb, K., et al., *Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells.* Hepatology, 2010. **51**(1): p. 297-305.
373. Lian, X., et al., *Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling.* PNAS, 2012. **109**(27): p. E1848-1857.
374. Lumelsky, N., et al., *Differentiation of embryonic stem cells to insulin-secreting structures similar to pancreatic islets.* Science, 2001. **292**(5520): p. 1389-1394.
375. Shi, Y., et al., *Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses.* Nat. Neurosci., 2012. **15**(3): p. 477-86, S1.

376. Shi, Y., P. Kirwan, and F.J. Livesey, *Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks.* Nat. Protoc., 2012. **7**(10): p. 1836-46.
377. Lancaster, M.A., et al., *Cerebral organoids model human brain development and microcephaly.* Nature, 2013. **501**(7467): p. 373-9.
378. Mariani, J., et al., *FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders.* Cell, 2015. **162**(2): p. 375-390.
379. Schwartz, S.D., et al., *Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies.* Lancet, 2015. **385**(9967): p. 509-516.
380. Gurdon, J.B., *The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles.* J. Embryol. Exp. Morphol., 1962. **10**: p. 622-640.
381. Wilmut, I., et al., *Viable offspring derived from fetal and adult mammalian cells.* Nature, 1997. **385**(6619): p. 810-813.
382. Hochedlinger, K. and R. Jaenisch, *Monoclonal mice generated by nuclear transfer from mature B and T donor cells.* Nature, 2002. **415**(6875): p. 1035-1038.
383. Eggan, K., et al., *Mice cloned from olfactory sensory neurons.* Nature, 2004. **428**(6978): p. 44-49.
384. Inoue, K., et al., *Generation of cloned mice by direct nuclear transfer from natural killer T cells.* Curr. Biol., 2005. **15**(12): p. 1114-1118.
385. Cowan, C.A., et al., *Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells.* Science, 2005. **309**(5739): p. 1369-1373.
386. Tada, M., et al., *Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells.* Curr. Biol., 2001. **11**(19): p. 1553-1558.
387. Schneuwly, S., R. Klemenz, and W.J. Gehring, *Redesigning the body plan of Drosophila by ectopic expression of the homoeotic gene Antennapedia.* Nature, 1987. **325**(6107): p. 816-818.
388. Davis, R.L., H. Weintraub, and A.B. Lassar, *Expression of a single transfected cDNA converts fibroblasts to myoblasts.* Cell. **51**(6): p. 987-1000.
389. Smith, A.G., et al., *Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides.* Nature, 1988. **336**(6200): p. 688-690.
390. Yamanaka, S., *Induced pluripotent stem cells: past, present, and future.* Cell Stem Cell, 2012. **10**(6): p. 678-684.

391. Niwa, H., J. Miyazaki, and A.G. Smith, *Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.* Nat. Genet., 2000. **24**(4): p. 372-376.
392. Nichols, J., et al., *Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4.* Cell, 1998. **95**(3): p. 379-391.
393. Avilion, A.A., et al., *Multipotent cell lineages in early mouse development depend on SOX2 function.* Genes Dev., 2003. **17**(1): p. 126-140.
394. Mitsui, K., et al., *The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells.* Cell, 2003. **113**(5): p. 631-642.
395. Chambers, I., et al., *Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells.* Cell, 2003. **113**(5): p. 643-655.
396. Silva, J., et al., *Nanog is the gateway to the pluripotent ground state.* Cell, 2009. **138**(4): p. 722-737.
397. Loh, Y.H., et al., *The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells.* Nat. Genet., 2006. **38**(4): p. 431-440.
398. Boyer, L.A., et al., *Core transcriptional regulatory circuitry in human embryonic stem cells.* Cell, 2005. **122**(6): p. 947-956.
399. Masui, S., et al., *Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells.* Nat. Cell Biol., 2007. **9**(6): p. 625-635.
400. Wang, J., et al., *A protein interaction network for pluripotency of embryonic stem cells.* Nature, 2006. **444**(7117): p. 364-368.
401. Rodda, D.J., et al., *Transcriptional regulation of nanog by OCT4 and SOX2.* J. Biol. Chem., 2005. **280**(26): p. 24731-24737.
402. Evans, P.M. and C. Liu, *Roles of Krupel-like factor 4 in normal homeostasis, cancer and stem cells.* Acta Biochim. Biophys. Sin., 2008. **40**(7): p. 554-564.
403. Li, Y., et al., *Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4.* Blood, 2005. **105**(2): p. 635-637.
404. Rowland, B.D., R. Bernards, and D.S. Peeper, *The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene.* Nat. Cell Biol., 2005. **7**(11): p. 1074-1082.
405. Lin, T., et al., *p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression.* Nat. Cell Biol., 2004. **7**(2): p. 165-171.

406. Meyer, N. and L.Z. Penn, *Reflecting on 25 years with MYC.* Nat. Rev. Cancer, 2008. **8**(12): p. 976-990.
407. Dang, C.V., *c-Myc target genes involved in cell growth, apoptosis, and metabolism.* Mol. Cell Biol., 1999. **19**(1): p. 1-11.
408. Hoffman, B. and D.A. Liebermann, *Apoptotic signaling by c-MYC.* Oncogene, 2008. **27**(50): p. 6462-6472.
409. Adhikary, S. and M. Eilers, *Transcriptional regulation and transformation by Myc proteins.* Nat. Rev. Mol. Cell Biol., 2005. **6**(8): p. 635-645.
410. Fernandez, P.C., et al., *Genomic targets of the human c-Myc protein.* Genes Dev., 2003. **17**(9): p. 1115-1129.
411. Vervoorts, J., et al., *Stimulation of c-MYC transcriptional activity and acetylation by recruitment of the cofactor CBP.* EMBO Rep., 2003. **4**(5): p. 484-490.
412. McMahon, S.B., et al., *The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins.* Cell, 1998. **94**(3): p. 363-374.
413. Dominguez-Sola, D., et al., *Non-transcriptional control of DNA replication by c-Myc.* Nature, 2007. **448**(7152): p. 445-451.
414. Shyh-Chang, N. and G.Q. Daley, *Lin28: primal regulator of growth and metabolism in stem cells.* Cell Stem Cell, 2013. **12**(4): p. 395-406.
415. Tsialikas, J. and J. Romer-Seibert, *LIN28: roles and regulation in development and beyond.* Development, 2015. **142**(14): p. 2397-2404.
416. Tang, F., et al., *Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis.* Cell Stem Cell, 2010. **6**(5): p. 468-478.
417. Melton, C., R.I. Judson, and R. Blelloch, *Opposing microRNA families regulate self-renewal in mouse embryonic stem cells.* Nature, 2010. **463**(7281): p. 621-626.
418. Suh, N., et al., *MicroRNA function is globally suppressed in mouse oocytes and early embryos.* Curr. Biol., 2010. **20**(3): p. 271-277.
419. Hanna, J., et al., *Direct cell reprogramming is a stochastic process amenable to acceleration.* Nature, 2009. **462**(7273): p. 595-601.
420. Okita, K., T. Ichisaka, and S. Yamanaka, *Generation of germline-competent induced pluripotent stem cells.* Nature, 2007. **448**(7151): p. 313-317.
421. Nienhuis, A.W., C.E. Dunbar, and B.P. Sorrentino, *Genotoxicity of retroviral integration in hematopoietic cells.* Mol. Ther., 2006. **13**(6): p. 1031-1049.

422. Soldner, F., et al., *Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors.* Cell, 2009. **136**(5): p. 964-977.
423. Woltjen, K., et al., *piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells.* Nature, 2009. **458**(7239): p. 766-770.
424. Kaji, K., et al., *Virus-free induction of pluripotency and subsequent excision of reprogramming factors.* Nautre, 2009. **458**(7239): p. 771-775.
425. Zhou, W. and C.R. Freed, *Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells.* Stem Cells. **27**(11): p. 2667-2674.
426. Okita, K., et al., *Generation of mouse induced pluripotent stem cells without viral vectors.* Science, 2008. **322**(5903): p. 949-953.
427. Fusaki, N., et al., *Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome.* Proc Jpn Acad Ser B Phys Biol Sci, 2009. **85**(8): p. 348-362.
428. Yu, J., et al., *Human induced pluripotent stem cells free of vector and transgene sequences.* Science, 2009. **324**(5928): p. 797-801.
429. Okita, K., et al., *A more efficient method to generate integration-free human iPS cells.* Nat Methods, 2011. **8**(5): p. 409-12.
430. Jia, F., et al., *A nonviral minicircle vector for deriving human iPS cells.* Nat. Methods. **7**(3): p. 197-199.
431. Warren, L., et al., *Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA.* Cell Stem cell, 2010. **7**(5): p. 618-630.
432. Kim, D., et al., *Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins.* Cell Stem cell, 2009. **4**(6): p. 472-476.
433. Zhou, H., et al., *Generation of induced pluripotent stem cells using recombinant proteins.* Cell Stem Cell, 2009. **4**(5): p. 381-384.
434. Hou, P., et al., *Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds.* Science, 2013. **341**(6146): p. 651-654.
435. Lin, T. and S. Wu, *Reprogramming with Small Molecules instead of Exogenous Transcription Factors.* Stem Cells Int., 2015: p. 794632.
436. Aasen, T., et al., *Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes.* Nat. Biotechnol., 2008. **26**(11): p. 1276-1284.
437. Eminli, S., et al., *Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression.* Stem Cells, 2008. **26**(10): p. 2467-2474.

438. Aoi, T., et al., *Generation of pluripotent stem cells from adult mouse liver and stomach cells.* Science, 2008. **321**(5889): p. 699-702.
439. Utikal, J., et al., *Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells.* J. Cell Sci., 2009. **122**(Pt 19): p. 3502-3510.
440. Stadtfeld, M., K. Brennand, and K. Hochedlinger, *Reprogramming of pancreatic beta cells into induced pluripotent stem cells.* Curr Biol., 2008. **18**(12): p. 890-894.
441. Hanna, J., et al., *Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency.* Cell, 2008. **133**(2): p. 250-264.
442. Li, W., et al., *Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors.* Cell Stem Cell, 2009. **4**(1): p. 16-19.
443. Liu, H., et al., *Generation of induced pluripotent stem cells from adult rhesus monkey fibroblasts.* Cell Stem Cell, 2008. **3**(6): p. 587-590.
444. Wu, Y., et al., *Generation of induced pluripotent stem cells from newborn marmoset skin fibroblasts.* Stem Cell Res., 2010. **4**(3): p. 180-188.
445. Ben-Nun, I.F., et al., *Induced pluripotent stem cells from highly endangered species.* Nat. Methods, 2011. **8**(10): p. 829-831.
446. Gallego Romero, I., et al., *A panel of induced pluripotent stem cells from chimpanzees: a resource for comparative functional genomics.* Elife, 2015. **4**(e07103).
447. Stadtfeld, M. and K. Hochedlinger, *Induced pluripotency: history, mechanisms, and applications.* Genes Dev., 2010. **24**(20): p. 2239-2263.
448. Chari, S. and S. Mao, *Timeline: iPSCs--The First Decade.* Cell, 2016. **164**(3): p. 580.
449. De Los Angeles, A., et al., *Hallmarks of pluripotency.* Nature, 2015. **525**(7570): p. 469-478.
450. Jeste, S.S. and D.H. Geschwind, *Developmental disorders.* Curr. Opin. Neurol. **28**(2): p. 89-90.
451. Hu, W.F., M.H. Chahrour, and C.A. Walsh, *The diverse genetic landscape of neurodevelopmental disorders.* Annu. Rev.Genomics Hum. Genet., 2014. **15**: p. 195-213.
452. Psych, E.C., et al., *The PsychENCODE project.* Nat Neurosci, 2015. **18**(12): p. 1707-12.
453. Sullivan, P.F., M.J. Daly, and M. O'Donovan, *Genetic architectures of psychiatric disorders: the emerging picture and its implications.* Nat. Rev. Genet., 2012. **13**(8): p. 537-551.

454. Lee, S.H., et al., *Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs.* Nat. Genet., 2013. **45**(9): p. 984-994.
455. Akbarian, S., et al., *The PsychENCODE project.* Nat. Neurosci., 2015. **18**(12): p. 1707-1712.
456. Bae, B.I., D. Jayaraman, and C.A. Walsh, *Genetic changes shaping the human brain.* Dev. Cell., 2015. **32**(4): p. 423-434.
457. DeFelipe, J., *Cortical interneurons: from Cajal to 2001.* Prog. Brain Res., 2002. **136**(0079-6123 (Print)): p. 215-238.
458. Jones, E.G., *The origins of cortical interneurons: mouse versus monkey and human.* Cereb. Cortex, 2009. **19**(9): p. 1953-1956.
459. Butti, C., et al., *Von Economo neurons: clinical and evolutionary perspectives.* Cortex, 2013. **49**(1): p. 312-326.
460. Rakic, P. and R.L. Sidman, *Telencephalic origin of pulvinar neurons in the fetal human brain.* Z Anat Entwicklungsgesch, 1969. **129**(1): p. 53-82.
461. Finlay, B.L. and R.B. Darlington, *Linked regularities in the development and evolution of mammalian brains.* Science, 1995. **268**(5217): p. 1578-1584.
462. Kelava, I. and M.A. Lancaster, *Stem Cell Models of Human Brain Development.* Cell Stem Cell, 2016. **18**(6): p. 736-48.
463. Humpel, C., *Organotypic brain slice cultures: A review.* Neuroscience, 2015. **305**: p. 86-98.
464. Darbinyan, A., et al., *Isolation and propagation of primary human and rodent embryonic neural progenitor cells and cortical neurons.* Methods Mol. Biol., 2013. **1078**: p. 45-54.
465. Xu, J., Y. Du, and H. Deng, *Direct lineage reprogramming: strategies, mechanisms, and applications.* Cell Stem Cell, 2015. **16**(2): p. 119-134.
466. Smith, J.R., et al., *Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm.* Dev. Biol., 2008. **313**(1): p. 107-117.
467. van den Ameele, J., et al., *Thinking out of the dish: what to learn about cortical development using pluripotent stem cells.* Trends Neurosci., 2014. **37**(4): p. 334-342.
468. Anderson, S. and P. Vanderhaeghen, *Cortical neurogenesis from pluripotent stem cells: complexity emerging from simplicity.* Curr. Opin. Neurobiol., 2014. **27**: p. 151-157.
469. Dolmetsch, R. and D.H. Geschwind, *The human brain in a dish: the promise of iPSC-derived neurons.* Cell, 2011. **145**(6): p. 831-834.

470. Adamo, A., et al., *7q11.23 dosage-dependent dysregulation in human pluripotent stem cells affects transcriptional programs in disease-relevant lineages.* Nat Genet, 2015. **47**(2): p. 132-41.
471. Marchetto, M.C., et al., *A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells.* Cell, 2010. **143**(4): p. 527-539.
472. Pollard, S.M., et al., *Adherent neural stem (NS) cells from fetal and adult forebrain.* Cereb. Cortex, 2006. **16**(Suppl. 1): p. i112-120.
473. Yin, X., et al., *Engineering Stem Cell Organoids.* Cell Stem Cell, 2016. **18**(1): p. 25-38.
474. Clevers, H., *Modeling Development and Disease with Organoids.* Cell, 2016. **165**(7): p. 1586-97.
475. Pasca, A.M., et al., *Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture.* Nat. Methods, 2015. **12**(7): p. 671-678.
476. Kadoshima, T., et al., *Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex.* Proc .Natl. Acad. Sci., 2013. **110**(50): p. 20284-20289.
477. Lancaster, M.A. and J.A. Knoblich, *Organogenesis in a dish: modeling development and disease using organoid technologies.* Science, 2014. **345**(6194): p. 1247125.
478. Qian, X., et al., *Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure.* Cell, 2016. **165**(5): p. 1238-1254.
479. Zein, N.N., et al., *Three-dimensional print of a liver for preoperative planning in living donor liver transplantation.* Liver Transpl., 2013. **19**(12): p. 1304-1310.
480. Seitz, H., et al., *Three-dimensional printing of porous ceramic scaffolds for bone tissue engineering.* J. Biomed. Mater Res. B Appl. Biomater., 2005. **74**(2): p. 782-788.
481. Hockaday, L.A., et al., *Rapid 3D printing of anatomically accurate and mechanically heterogeneous aortic valve hydrogel scaffolds.* Biofabrication, 2012. **4**(3): p. 035005.
482. Auger, F.A., L. Gibot, and D. Lacroix, *The pivotal role of vascularization in tissue engineering.* Annu. Rev. Biomed. Eng., 2013. **15**: p. 177-200.
483. Schimek, K., et al., *Integrating biological vasculature into a multi-organ-chip microsystem.* Lab Chip, 2013. **13**(18): p. 3588-3598.
484. van Duinen, V., et al., *Microfluidic 3D cell culture: from tools to tissue models.* Curr. Opin. Biotechnol., 2015. **35**: p. 118-126.

485. Rouwkema, J. and A. Khademhosseini, *Vascularization and Angiogenesis in Tissue Engineering: Beyond Creating Static Networks.* Trends Biotechnol., 2016. **34**(9): p. 733-745.
486. Carlson, A.L., et al., *Generation and transplantation of reprogrammed human neurons in the brain using 3D microtopographic scaffolds.* Nat. Commun., 2016. **7**(10862).
487. Langer, R. and J. Vacanti, *Advances in tissue engineering.* J. Pediatr. Surg., 2016. **51**(1): p. 8-12.
488. Esch, E.W., A. Bahinski, and D. Huh, *Organs-on-chips at the frontiers of drug discovery.* Nat. Rev. Drug Discov., 2015. **14**(4): p. 248-260.
489. Panchision, D.M., *Concise Review: Progress and Challenges in Using Human Stem Cells for Biological and Therapeutics Discovery: Neuropsychiatric Disorders.* Stem Cells, 2016. **34**(3): p. 523-36.
490. Inoue, H. and S. Yamanaka, *The use of induced pluripotent stem cells in drug development.* Clin. Pharmacol. Ther., 2011. **89**(5): p. 655-661.
491. Ferreira, L.M. and M.A. Mostajo-Radji, *How induced pluripotent stem cells are redefining personalized medicine.* Gene, 2013. **520**(1): p. 1-6.
492. Liang, N., et al., *Stem cell contributions to neurological disease modeling and personalized medicine. .* Prog. Neuropsychopharmacol. Biol. Psychiatry, 2017. **S0278-5846**(17): p. 30232-30234.
493. Boyle, C.A., et al., *Trends in the prevalence of developmental disabilities in US children, 1997-2008.* Pediatrics, 2011. **127**(6): p. 1034-1042.
494. Polderman, T.J., et al., *Meta-analysis of the heritability of human traits based on fifty years of twin studies.* Nat. Genet., 2015. **47**(7): p. 702-709.
495. Ronald, A. and R.A. Hoekstra, *Autism spectrum disorders and autistic traits: a decade of new twin studies.* Am. J. Hum. Genet. B. Neuropsychiatr. Genet., 2011. **156B**(3): p. 255-274.
496. Sandin, S., et al., *The familial risk of autism.* JAMA, 2014. **311**(17): p. 1770-1777.
497. Lichtenstein, P., et al., *The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood.* Am. J. Psychiatry, 2010. **167**(11): p. 1357-1363.
498. Kishino, T., M. Lalande, and J. Wagstaff, *UBE3A/E6-AP mutations cause Angelman syndrome.* Nat. Genet., 1997. **15**(1): p. 70-73.

499. Amir, R.E., et al., *Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2.* Nat. Genet., 1999. **23**(2): p. 185-188.
500. Verkerk, A.J., et al., *Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome.* Cell, 1991. **65**(5): p. 905-914.
501. Nelen, M.R., et al., *Localization of the gene for Cowden disease to chromosome 10q22-23.* Nat. Genet., 1996. **13**(1): p. 114-116.
502. Gratten, J., et al., *Large-scale genomics unveils the genetic architecture of psychiatric disorders.* Nat. Neurosci., 2014. **17**(6): p. 782-790.
503. Gandal, M.J., et al., *The road to precision psychiatry: translating genetics into disease mechanisms.* Nat. Neurosci., 2016. **19**(11): p. 1397-1407.
504. Geschwind, D.H. and J. Flint, *Genetics and genomics of psychiatric disease.* Science. **349**(6255): p. 1489-1494.
505. Hawi, Z., et al., *The molecular genetic architecture of attention deficit hyperactivity disorder.* Mol. Psychiatry, 2015. **20**(3): p. 289-297.
506. Sanders, S.J., et al., *Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci.* Neuron, 2015. **87**(6): p. 1215-1233.
507. Iossifov, I., et al., *The contribution of de novo coding mutations to autism spectrum disorder.* Nature, 2014. **515**(7526): p. 216-221.
508. Group, P.G.C.B.D.W., *Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4.* Nat. Genet., 2011. **43**(10): p. 977-983.
509. Tebbenkamp, A.T., et al., *The developmental transcriptome of the human brain: implications for neurodevelopmental disorders.* Curr. Opin. Neurol., 2014. **27**(2): p. 149-156.
510. Levy, D., et al., *Rare de novo and transmitted copy-number variation in autistic spectrum disorders.* Neuron, 2011. **70**(5): p. 886-897.
511. Pinto, D., et al., *Convergence of genes and cellular pathways dysregulated in autism spectrum disorders.* Am J Hum Genet, 2014. **94**(5): p. 677-694.
512. Iossifov, I., et al., *De novo gene disruptions in children on the autistic spectrum.* Neuron, 2012. **74**(2): p. 285-299.
513. Yu, T.W., et al., *Using whole-exome sequencing to identify inherited causes of autism.* Neuron, 2013. **77**(2): p. 259-273.
514. Kirov, G., et al., *De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia.* Mol Psychiatry, 2012. **17**(2): p. 142-153.

515. Manolio, T.A., et al., *Finding the missing heritability of complex diseases.* Nature, 2009. **461**(7265): p. 747-753.
516. Wray, N.R., M.E. Goddard, and P.M. Visscher, *Prediction of individual genetic risk of complex disease.* Curr. Opin. Genet. Dev., 2008. **18**(3): p. 257-263.
517. Sullivan, P.F., *The psychiatric GWAS consortium: big science comes to psychiatry.* Neuron, 2010. **68**(2): p. 182-186.
518. Pinto, D., et al., *Functional impact of global rare copy number variation in autism spectrum disorders.* Nature, 2010. **466**(7304): p. 368-372.
519. Hussman, J.P., et al., *A noise-reduction GWAS analysis implicates altered regulation of neurite outgrowth and guidance in autism.* Mol. Autism, 2011. **2**(1): p. 1.
520. Schork, N.J., et al., *Common vs. rare allele hypotheses for complex diseases.* Curr. Opin. Genet. Dev., 2009. **19**(3): p. 212-219.
521. Hirschhorn, J.N. and M.J. Daly, *Genome-wide association studies for common diseases and complex traits.* Nat. Rev. Genet., 2005. **6**(2): p. 95-108.
522. McCarthy, M.I., et al., *Genome-wide association studies for complex traits: consensus, uncertainty and challenges.* Nat. Rev. Genet., 2008. **9**(5): p. 356-369.
523. Voineagu, I., et al., *Transcriptomic analysis of autistic brain reveals convergent molecular pathology.* Nature, 2011. **474**(7351): p. 380-384.
524. Barak, B. and G. Feng, *Neurobiology of social behavior abnormalities in autism and Williams syndrome.* Nat. Neurosci., 2016. **19**(6): p. 647-655.
525. Abbas, E., et al., *The 7q11.23 Microduplication Syndrome: A Clinical Report with Review of Literature.* J. Pediatr. Genet., 2016. **5**(3): p. 129-140.
526. Parikshak, N.N., M.J. Gandal, and D.H. Geschwind, *Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders.* Nat. Rev. Genet., 2015. **16**(8): p. 441-458.
527. Parikshak, N.N., et al., *Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism.* Cell, 2013. **155**(5): p. 1008-1021.
528. Kou, Y., et al., *Network- and attribute-based classifiers can prioritize genes and pathways for autism spectrum disorders and intellectual disability.* Am. J. Med. Genet. C. Semin. Med. Genet., 2012. **160C**(2): p. 130-142.

529. Xu, B., et al., *De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia.* Nat. Genet., 2012. **44**(12): p. 1365-1369.
530. Glessner, J.T., et al., *Strong synaptic transmission impact by copy number variations in schizophrenia.* PNAS, 2010. **107**(23): p. 10584-10589.
531. Consortium, N.a.P.A.S.o.P.G., *Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways.* Nat. Neurosci., 2015. **18**(2): p. 199-209.
532. Elia, J., et al., *Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes.* Mol. Psychiatry, 2010. **15**(6): p. 637-646.
533. Cooper, G.M., et al., *A copy number variation morbidity map of developmental delay.* Nat. Genet., 2011. **43**(9): p. 838-846.
534. Girirajan, S., C.D. Campbell, and E.E. Eichler, *Human copy number variation and complex genetic disease.* Annu. Rev.Genet., 2011. **45**: p. 203-226.
535. Elsea, S.H. and S. Girirajan, *Smith-Magenis syndrome.* Eur. J. Hum. Genet., 2008. **16**(4): p. 412-421.
536. Pober, B.R., *Williams-Beuren syndrome.* N. Engl. J. Med., 2010. **362**(3): p. 239-252.
537. Driscoll, D.A., M.I. Budarf, and B.S. Emanuel, *A genetic etiology for DiGeorge syndrome: consistent deletions and microdeletions of 22q11.* Am. J. Hum. Genet., 1992. **50**(5): p. 924-933.
538. Malhotra, D. and J. Sebat, *CNVs: harbingers of a rare variant revolution in psychiatric genetics.* Cell, 2012. **148**(6): p. 1223-1241.
539. Koolen, D.A., et al., *A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism.* Nat. Genet., 2006. **38**(9): p. 999-1001.
540. Shaw-Smith, C., et al., *Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability.* Nat. Genet., 2006. **38**(9): p. 1032-1037.
541. Sharp, A.J., et al., *Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome.* Nat. Genet., 2006. **38**(9): p. 1038-1042.
542. de Vries, B.B., et al., *Diagnostic genome profiling in mental retardation.* Am. J. Hum. Genet., 2005. **77**(4): p. 606-616.
543. Williams, N.M., et al., *Genome-wide analysis of copy number variants in attention deficit hyperactivity disorder: the role of rare variants and duplications at 15q13.3.* Am. J. Psychiatry, 2012. **169**(2): p. 195-204.

544. Consortium, I.S., *Rare chromosomal deletions and duplications increase risk of schizophrenia.* Nature, 2008. **455**(7210): p. 237-241.
545. Turner, D.J., et al., *Germline rates of de novo meiotic deletions and duplications causing several genomic disorders.* Nat. Genet., 2008. **40**(1): p. 90-95.
546. Gu, W., F. Zhang, and J.R. Lupski, *Mechanisms for human genomic rearrangements.* Pathogenetics, 2008. **1**(1): p. 4.
547. Liburd, N., et al., *Novel mutations of MYO15A associated with profound deafness in consanguineous families and moderately severe hearing loss in a patient with Smith-Magenis syndrome.* Hum. Genet., 2001. **109**(5): p. 535-541.
548. McCarthy, S.E., et al., *Microduplications of 16p11.2 are associated with schizophrenia.* Nat. Genet., 2009. **41**(11): p. 1223-1227.
549. Weiss, L.A., et al., *Association between microdeletion and microduplication at 16p11.2 and autism.* N. Engl. J. Med., 2008. **358**(7): p. 667-675.
550. Bijlsma, E.K., et al., *Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals.* Eur. J. Hum. Genet., 2009. **52**(2-3): p. 77-87.
551. Mefford, H.C., et al., *Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes.* N. Engl. J. Med., 2008. **359**(16): p. 1685-1699.
552. Purcell, S.M., et al., *Common polygenic variation contributes to risk of schizophrenia and bipolar disorder.* Nature, 2009. **460**(7256): p. 748-752.
553. Meyer-Lindenberg, A., C.B. Mervis, and K.F. Berman, *Neural mechanisms in Williams syndrome: a unique window to genetic influences on cognition and behaviour.* Nat. Rev. Neurosci., 2006. **7**(5): p. 380-393.
554. Martens, M.A., S.J. Wilson, and D.C. Reutens, *Research Review: Williams syndrome: a critical review of the cognitive, behavioral, and neuroanatomical phenotype.* J. Child Psychol. Psychiatry, 2008. **49**(9): p. 576-608.
555. Udwin, O., M. Davies, and P. Howlin, *A longitudinal study of cognitive abilities and educational attainment in Williams syndrome.* Dev. Med. Child Neurol., 1996. **38**(11): p. 1020-1029.
556. Levy, Y., J. Smith, and H. Tager-Flusberg, *Word reading and reading-related skills in adolescents with Williams syndrome.* J. Child Psychol. Psychiatry, 2003. **44**(4): p. 576-578.
557. Mervis, C.B., et al., *The Williams syndrome cognitive profile.* Brain Cogn., 2000. **44**(3): p. 604-628.

558. Del Campo, M., et al., *Hemizygosity at the NCF1 gene in patients with Williams-Beuren syndrome decreases their risk of hypertension.* Am. J. Hum. Genet., 2006. **78**(4): p. 533-542.
559. Bayes, M., et al., *Mutational mechanisms of Williams-Beuren syndrome deletions.* Am J Hum Genet, 2003. **73**(1): p. 131-151.
560. Levinson, D.F., et al., *Copy number variants in schizophrenia: confirmation of five previous findings and new evidence for 3q29 microdeletions and VIPR2 duplications.* Am. J. Psychiatry, 2011. **168**(3): p. 302-316.
561. Marshall, C.R., et al., *Infantile spasms is associated with deletion of the MAGI2 gene on chromosome 7q11.23-q21.11.* Am. J. Hum. Genet., 2008. **83**(1): p. 106-111.
562. Heller, R., et al., *Partial deletion of the critical 1.5 Mb interval in Williams-Beuren syndrome.* J. Med. Genet., 2003. **40**(8): p. e99.
563. Morris, C.A., et al., *GTF2I hemizygosity implicated in mental retardation in Williams syndrome: genotype-phenotype analysis of five families with deletions in the Williams syndrome region.* Am. J. Hum. Genet. A, 2003. **123A**(1): p. 45-59.
564. Frangiskakis, J.M., et al., *LIM-kinase1 hemizygosity implicated in impaired visuospatial constructive cognition.* Cell, 1996. **86**(1): p. 59-69.
565. Stromme, P., P.G. Bjornstad, and K. Ramstad, *Prevalence estimation of Williams syndrome.* J. Clin. Neurol., 2002. **17**(4): p. 269-271.
566. Vos, T., et al., *Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010.* Lancet, 2012. **380**(9859): p. 2163-2196.
567. Collins, R.T., et al., *Long-term outcomes of patients with cardiovascular abnormalities and williams syndrome.* Am. J. Cardiol., 2010. **105**(6): p. 874-878.
568. Del Pasqua, A., et al., *New findings concerning cardiovascular manifestations emerging from long-term follow-up of 150 patients with the Williams-Beuren-Beuren syndrome.* Cardiol. Young, 2009. **19**(6): p. 563-567.
569. Kim, Y.M., et al., *Endocrine dysfunctions in children with Williams-Beuren syndrome.* Ann. Pediatr. Endocrinol. Metab., 2016. **21**(1): p. 15-20.
570. Palacios-Verdu, M.G., et al., *Metabolic abnormalities in Williams-Beuren syndrome.* J. Med. Genet., 2015. **52**(4): p. 248-255.

571. Greer, M.K., et al., *Cognitive, adaptive, and behavioral characteristics of Williams syndrome.* Am. J. Hum. Genet., 1997. **74**(5): p. 521-525.
572. Bhattacharjee, Y., *Friendly faces and unusual minds.* Science, 2005. **310**(5749): p. 802-804.
573. Klein-Tasman, B.P. and C.B. Mervis, *Distinctive personality characteristics of 8-, 9-, and 10-year-olds with Williams syndrome.* Dev. Neuropsychol., 2003. **23**(1-2): p. 269-290.
574. Bellugi, U., et al., *Towards the neural basis for hypersociability in a genetic syndrome.* Neuroreport, 1999. **10**(8): p. 1653-1657.
575. Mervis, C.B. and C.H. Pitts, *Children with Williams syndrome: Developmental trajectories for intellectual abilities, vocabulary abilities, and adaptive behavior.* Am. J. Hum. Genet. C Semin. Med. Genet., 2015. **169**(2): p. 158-171.
576. Howlin, P. and O. Udwin, *Outcome in adult life for people with Williams syndrome-- results from a survey of 239 families.* J. Intellect. Disabil. Res., 2006. **50**(Pt 2): p. 151-160.
577. Woodruff-Borden, J., et al., *Longitudinal course of anxiety in children and adolescents with Williams syndrome.* Am. J. Med. Genet. C. Semin. Med. Genet., 2010. **154C**(2): p. 277-290.
578. Pitts, C.H., et al., *Predictors of specific phobia in children with Williams syndrome.* J. Intellect. Disabil. Res., 2016. **60**(10): p. 1031-1342.
579. Leyfer, O.T., et al., *Prevalence of psychiatric disorders in 4 to 16-year-olds with Williams syndrome.* Am. J. Hum. Genet. B Neuropsychiatr. Genet., 2006. **141B**(6): p. 615-622.
580. Klein-Tasman, B.P., et al., *Overlap with the autism spectrum in young children with Williams syndrome.* J. Dev. Behav. Pediatr., 2009. **30**(4): p. 289-299.
581. Galaburda, A.M., et al., *Williams syndrome: neuronal size and neuronal-packing density in primary visual cortex.* Arch. Neurol., 2002. **59**(9): p. 1461-1467.
582. Jernigan, T.L. and U. Bellugi, *Anomalous brain morphology on magnetic resonance images in Williams syndrome and Down syndrome.* Arch. Neurol., 1990. **47**(5): p. 529-533.
583. Reiss, A.L., et al., *An experiment of nature: brain anatomy parallels cognition and behavior in Williams syndrome.* J. Neurosci., 2004. **24**(21): p. 5009-5015.
584. Galaburda, A.M. and U. Bellugi, *V. Multi-level analysis of cortical neuroanatomy in Williams syndrome.* J. Cogn. Neurosci., 2000. **12**(Suppl 1): p. 74-88.

585. Meyer-Lindenberg, A., et al., *Functional, structural, and metabolic abnormalities of the hippocampal formation in Williams syndrome.* J. Clin. Invest., 2005. **115**(7): p. 1888-1895.
586. Jones, W., et al., *Cerebellar abnormalities in infants and toddlers with Williams syndrome.* Dev. Med. Child Neurol., 2002. **44**(10): p. 688-694.
587. Reiss, A.L., et al., *IV. Neuroanatomy of Williams syndrome: a high-resolution MRI study.* J. Comp. Neurosci., 2000. **12**(Suppl. 1): p. 65-73.
588. Toga, A.W., P.M. Thompson, and E.R. Sowell, *Mapping brain maturation.* Trends Neurosci., 2006. **29**(3): p. 148-159.
589. Jernigan, T.L., et al., *Cerebral morphologic distinctions between Williams and Down syndromes.* Arch. Neurol., 1993. **50**(2): p. 186-191.
590. Tomaiuolo, F., et al., *Morphology and morphometry of the corpus callosum in Williams syndrome: a T1-weighted MRI study.* Neuroreport, 2002. **13**(17): p. 2281-2284.
591. Marenco, S., et al., *Genetic contributions to white matter architecture revealed by diffusion tensor imaging in Williams syndrome.* PNAS, 2007. **104**(38): p. 15117-15122.
592. Arlinghaus, L.R., et al., *Alterations in diffusion properties of white matter in Williams syndrome.* Magn. Reson. Imaging, 2011. **29**(9): p. 1165-1174.
593. Faria, A.V., et al., *Quantitative analysis of gray and white matter in Williams syndrome.* (1473-558X (Electronic)).
594. Hoeft, F., et al., *More is not always better: increased fractional anisotropy of superior longitudinal fasciculus associated with poor visuospatial abilities in Williams syndrome.* J. Neurosci., 2007. **27**(44): p. 11960-11965.
595. Chiang, M.C., et al., *3D pattern of brain abnormalities in Williams syndrome visualized using tensor-based morphometry.* Neuroimage, 2007. **36**(4): p. 1096-1109.
596. Jackowski, A.P. and R.T. Schultz, *Foreshortened dorsal extension of the central sulcus in Williams syndrome.* Cortex, 2005. **41**(3): p. 282-290.
597. Gaser, C., et al., *Increased local gyrification mapped in Williams syndrome.* Neuroimage, 2006. **33**(1): p. 46-54.
598. Thompson, P.M., et al., *Abnormal cortical complexity and thickness profiles mapped in Williams syndrome.* J. Neurosci., 2005. **25**(16): p. 4146-4158.

599. Schmitt, J.E., et al., *Increased gyrification in Williams syndrome: evidence using 3D MRI methods.* Dev. Med. Child Neurol., 2002. **44**(5): p. 292-295.
600. Kippenhan, J.S., et al., *Genetic contributions to human gyrification: sulcal morphometry in Williams syndrome.* J. Neurosci., 2005. **25**(34): p. 7840-7846.
601. Mobbs, D., et al., *Frontostriatal dysfunction during response inhibition in Williams syndrome.* Biol. Psychiatry, 2007. **62**(3): p. 256-261.
602. Meyer-Lindenberg, A., et al., *Neural basis of genetically determined visuospatial construction deficit in Williams syndrome.* Neuron, 2004. **43**(5): p. 623-631.
603. Thornton-Wells, T.A., et al., *Auditory attraction: activation of visual cortex by music and sound in Williams syndrome.* Am. J. Intellect Dev. Disabil., 2010. **115**(2): p. 172-189.
604. Meyer-Lindenberg, A., et al., *Neural correlates of genetically abnormal social cognition in Williams syndrome.* Nat. Neurosci., 2005. **8**(8): p. 991-993.
605. Munoz, K.E., et al., *Abnormalities in neural processing of emotional stimuli in Williams syndrome vary according to social vs. non-social content.* Neuroimage, 2010. **50**(1): p. 340-346.
606. N., V.d.A., et al., *Fourteen new cases contribute to the characterization of the 7q11.23 microduplication syndrome.* Eur. J. Med. Genet., 2009. **52**(2-3): p. 94-100.
607. Velleman, S.L. and C.B. Mervis, *Children with 7q11.23 Duplication Syndrome: Speech, Language, Cognitive, and Behavioral Characteristics and their Implications for Intervention.* Perspect. Lang. Learn Educ., 2011. **18**(3): p. 108-116.
608. Morris, C.A., et al., *7q11.23 Duplication syndrome: Physical characteristics and natural history.* Am. J. Hum. Genet. A., 2015. **167A**(12): p. 2916-2935.
609. Torniero, C., et al., *Cortical dysplasia of the left temporal lobe might explain severe expressive-language delay in patients with duplication of the Williams-Beuren locus.* Eur. J. Hum. Genet., 2007. **15**(1): p. 62-67.
610. Berg, J.S., et al., *Speech delay and autism spectrum behaviors are frequently associated with duplication of the 7q11.23 Williams-Beuren syndrome region.* Genet. Med., 2007. **9**(7): p. 427-441.
611. Mervis, C.B., et al., *Children with 7q11.23 duplication syndrome: psychological characteristics.* Am. J. Hum. Genet., 2015. **167**(7): p. 1436-1450.

612. Mervis, C.B., et al., *7q11.23 Duplication Syndrome BTI - GeneReviews(R).* GeneReviews (Internet), 2015.
613. Sanders, S.J., et al., *Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism.* Neuron, 2011. **70**(5): p. 863-885.
614. Prontera, P., et al., *Brief report: functional MRI of a patient with 7q11.23 duplication syndrome and autism spectrum disorder.* J. Autism Dev. Disord., 2014. **44**(10): p. 2608-2613.
615. Edelmann, L., et al., *An atypical deletion of the Williams-Beuren syndrome interval implicates genes associated with defective visuospatial processing and autism.* J. Med. Genet., 2007. **44**(2): p. 136-143.
616. Dai, L., et al., *Is it Williams syndrome? GTF2IRD1 implicated in visual-spatial construction and GTF2I in sociability revealed by high resolution arrays.* Am. J. Hum. Genet. A, 2009. **149A**(3): p. 302-314.
617. Segura-Puimedon, M., et al., *Heterozygous deletion of the Williams-Beuren syndrome critical interval in mice recapitulates most features of the human disorder.* Hum. Mol. Genet., 2014. **23**(24): p. 6481-6494.
618. Li, H.H., et al., *Induced chromosome deletions cause hypersociability and other features of Williams-Beuren syndrome in mice.* EMBO Mol. Med., 2009. **1**(1): p. 50-65.
619. Hoogenraad, C.C., et al., *Targeted mutation of Cyln2 in the Williams syndrome critical region links CLIP-115 haploinsufficiency to neurodevelopmental abnormalities in mice.* Nat. Genet., 2002. **32**(1): p. 116-127.
620. Lucena, J., et al., *Essential role of the N-terminal region of TFII-I in viability and behavior.* BCM Med. Genet., 2010: p. 11-61.
621. Chailangkarn, T., et al., *A human neurodevelopmental model for Williams syndrome.* Nature, 2016. **536**(7616): p. 338-43.
622. Li, D.Y., et al., *Elastin is an essential determinant of arterial morphogenesis.* Nature, 1998. **393**(6682): p. 276-280.
623. Capossela, S., et al., *Growth defects and impaired cognitive-behavioral abilities in mice with knockout for Eif4h, a gene located in the mouse homolog of the Williams-Beuren syndrome critical region.* Am. J. Pathol. **180**(3): p. 1121-1135.
624. Mervis, C.B., et al., *Duplication of GTF2I results in separation anxiety in mice and humans.* Am J Hum Genet, 2012. **90**(6): p. 1064-1070.
625. Okita, K., et al., *An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells.* Stem Cells, 2013. **31**(3): p. 458-466.

626. Hong, H., et al., *Suppression of induced pluripotent stem cell generation by the p53-p21 pathway.* Nature, 2009. **460**(7259): p. 1132-1135.
627. Xu, C., et al., *Feeder-free growth of undifferentiated human embryonic stem cells.* Nat. Biotechnol., 2001. **19**(10): p. 971-974.
628. Chen, G., et al., *Chemically defined conditions for human iPSC derivation and culture.* Nat Methods, 2011. **8**(5): p. 424-9.
629. Liu, W. and G. Chen, *Cryopreservation of human pluripotent stem cells in defined medium.* Curr Protoc Stem Cell Biol, 2014. **31**: p. 1C 17 1-13.
630. Valenzuela, D.M., et al., *Identification of mammalian noggin and its expression in the adult nervous system.* J. Neurosci., 1995. **15**(9): p. 6077-6084.
631. Beers, J., et al., *Passaging and colony expansion of human pluripotent stem cells by enzyme-free dissociation in chemically defined culture conditions.* Nat Protoc, 2012. **7**(11): p. 2029-40.
632. Cuny, G.D., et al., *Structure-activity relationship study of bone morphogenetic protein (BMP) signaling inhibitors.* Bioorg. Med. Chem. Lett., 2008. **18**(15): p. 4388-4392.
633. Vogt, J., R. Traynor, and G.P. Sapkota, *The specificities of small molecule inhibitors of the TGFss and BMP pathways.* Cell Signal., 2011. **23**(11): p. 1831-1842.
634. Inman, G.J., et al., *SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7.* Mol. Pharmacol., 2002. **62**(1): p. 65-74.
635. Laping, N.J., et al., *Inhibition of transforming growth factor (TGF)-beta1-induced extracellular matrix with a novel inhibitor of the TGF-beta type I receptor kinase activity: SB-431542.* Mol. Pharmacol., 2002. **62**(1): p. 58-64.
636. James, D., et al., *TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells.* Development, 2005. **132**(6): p. 1273-1282.
637. Merrill, B.J., *Wnt pathway regulation of embryonic stem cell self-renewal.* Cold Spring Harb. Perspect. Biol., 2012. **4**(9): p. a007971.
638. Huang, S.M., et al., *Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling.* Nature, 2009. **461**(7264): p. 614-620.
639. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
640. Zhang, B. and S. Horvath, *A general framework for weighted gene co-expression network analysis.* Stat. Appl. Genet. Mol. Biol., 2005. **4**(Article 17).

641. Zurauskiene, J. and C. Yau, *pcaReduce: hierarchical clustering of single cell transcriptional profiles.* BMC Bioinformatics, 2016. **17**(140).
642. Nat, R. and G. Dechant, *Milestones of directed differentiation of mouse and human embryonic stem cells into telencephalic neurons based on neural development in vivo.* Stem Cell Dev., 2011. **20**(6): p. 947-958.
643. Englund, C., et al., *Pax6, Tbr2, and Tbr1 are expressed sequentially by radial glia, intermediate progenitor cells, and postmitotic neurons in developing neocortex.* J. Neurosci., 2005. **25**(1): p. 247-251.
644. Kan, L., et al., *Dual function of Sox1 in telencephalic progenitor cells.* Dev. Biol., 2007. **310**(1): p. 85-98.
645. Bedogni, F., et al., *Tbr1 regulates regional and laminar identity of postmitotic neurons in developing neocortex.* Proc Natl Acad Sci U S A, 2010. **107**(29): p. 13129-13134.
646. Nieto, M., et al., *Expression of Cux-1 and Cux-2 in the subventricular zone and upper layers II-IV of the cerebral cortex.* J. Comp. Neurol., 2004. **479**(2): p. 168-180.
647. Onorati, M., et al., *Molecular and functional definition of the developing human striatum.* Nat. Neurosci., 2014. **17**(12): p. 1804-1815.
648. Moore, A.R., et al., *Electrical excitability of early neurons in the human cerebral cortex during the second trimester of gestation.* Cereb Cortex, 2009. **19**(8): p. 1795-1805.
649. Avoli, M., et al., *Electrophysiological and repetitive firing properties of neurons in the superficial/middle layers of the human neocortex maintained in vitro.* Exp Brain Res., 1994. **98**(1): p. 135-144.
650. Stark, C., et al., *BioGRID: a general repository for interaction datasets.* Nucleic Acid res., 2006. **34**: p. D535-539.
651. Lin, G.N., et al., *Spatiotemporal 16p11.2 protein network implicates cortical late mid-fetal brain development and KCTD13-Cul3-RhoA pathway in psychiatric diseases.* Neuron, 2015. **85**(4): p. 742-754.
652. Supek, F., et al., *REVIGO summarizes and visualizes long lists of gene ontology terms.* PLoS ONE, 2011. **6**(7): p. E21800.
653. Kwon, S.K., et al., *LKB1 Regulates Mitochondria-Dependent Presynaptic Calcium Clearance and Neurotransmitter Release Properties at Excitatory Synapses along Cortical Axons.* PLoS Biol., 2016. **14**(7): p. e1002516.
654. Kampinga, H.H. and E.A. Craig, *The HSP70 chaperone machinery: J proteins as drivers of functional specificity.* Nat. Rev. Mol. Cel. Biol., 2010. **11**(8): p. 579-592.

655. Van Essen, D.C., et al., *Symmetry of cortical folding abnormalities in Williams syndrome revealed by surface-based analyses.* J. Neurosci., 2006. **26**(20): p. 5470-5483.
656. Rivara, C.B., et al., *Stereologic characterization and spatial distribution patterns of Betz cells in the human primary motor cortex.* Anat. Rec. A Discov. Mol. Cell. Evol. Biol., 2003. **270**(2): p. 137-151.
657. Arion, D., et al., *Distinctive transcriptome alterations of prefrontal pyramidal neurons in schizophrenia and schizoaffective disorder.* Mol. Psychiatry, 2015. **20**(11): p. 1397-1405.
658. Calvo, S.E., K.R. Clauser, and V.K. Mootha, *MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins.* Nucleic Acid Res., 2016. **44**(D1): p. D1251-1257.
659. Huttlin, E.L., et al., *The BioPlex Network: A Systematic Exploration of the Human Interactome.* Cell, 2015. **162**(2): p. 425-440.
660. Mortazavi, A., et al., *Mapping and quantifying mammalian transcriptomes by RNA-Seq.* Nat. Methods, 2008. **5**(7): p. 621-628.
661. Johnson, W.E., C. Li, and A. Rabinovic, *Adjusting batch effects in microarray expression data using empirical Bayes methods.* Biostatistics, 2007. **8**(1): p. 118-127.
662. Anders, S. and W. Huber, *Differential expression analysis for sequence count data.* Genome Biol., 2010. **11**(10): p. R106.
663. Klipper-Aurbach, Y., et al., *Mathematical formulae for the prediction of the residual beta cell function during the first two years of disease in children and adolescents with insulin-dependent diabetes mellitus.* Med. Hypotheses, 1995. **45**(5): p. 486-490.
664. Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis.* BMC Bioinformatics, 2008. **9**(559).
665. Carmona-Saez, P., et al., *GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists.* EMBO Mol. Med., 2009. **1**(1): p. 50-65.
666. Habegger, L., et al., *RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries.* Bioinformatics, 2011. **27**(2): p. 281-283.
667. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-2079.
668. Benita, Y., et al., *Gene enrichment profiles reveal T-cell development, differentiation, and lineage-specific transcription factors including ZBTB25 as a novel NF-AT repressor.* Blood, 2010. **115**(26): p. 5376-5384.
669. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat. Protoc. **4**(1): p. 44-57.

670. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome Res., 2003. **13**(11): p. 2498-2504.