PhD degree in Molecular Medicine (curriculum in Molecular Oncology)
European School of Molecular Medicine (SEMM),
University of Milan and University of Naples “Federico II”
Settore disciplinare: MED/04

A unifying framework to study the genetic and environmental factors shaping human brain development

Nicolò Caporale
European Institute of Oncology
University of Milan
Matricola n. R11469

Supervisor: Prof. Giuseppe Testa
European Institute of Oncology, Milan
University of Milan, Milan

Anno accademico 2018-2019

Table of content

Index of figures

Index of tables

Abstract

The development of human brain is a fascinating and complex process that still needs to be uncovered at the molecular resolution. Even though animal studies have revealed a lot of its unfolding, the fine regulation of cellular differentiation trajectories that characterizes humans has become only recently open to experimental tractability, thanks to the development of organoids, human cellular models that are able to recapitulate the spatiotemporal architecture of the brain in a 3D fashion. Here we first benchmarked human brain organoids at the level of transcriptomic and structural architecture of cell composition along several stages of differentiation. Then we harnessed their properties to probe the longitudinal impact of GSK3 on human corticogenesis, a pivotal regulator of both proliferation and polarity, that we revealed having a direct impact on early neurogenesis with a selective role in the regulation of glutamatergic lineages and outer radial glia output. Moreover, we spearheaded the use of organoids for regulatory toxicology through the study of Endocrine disrupting chemicals (EDC), pervasive compounds that can interfere with human hormonal systems. Early life exposure to EDC is associated with human disorders, but the molecular events triggered remain unknown. We developed a novel approach, integrating epidemiological with experimental biology to study the mixtures of EDC that were associated with neurodevelopmental and metabolic adverse effects in the biggest pregnancy cohort profiled so far. Our experiments were carried out on two complementary models i) human fetal primary neural stem cells, and ii) 3-dimensional cortical brain organoids and we identified the genes specifically dysregulated by EDC mixture exposure, unravelling a significant enrichment for autism spectrum disorders causative genes, thereby proposing a convergent paradigm of neurodevelopmental disorders pathophysiology between genetic and environmental factors. Finally, while EDCs are everywhere, their impact on adverse health outcomes can vary substantially among individuals, suggesting that other genetic factors may play a pivotal role for the onset of the disorders. We took advantage of organoids multiplexing to recapitulate, at the same time, neurodevelopmental trajectories on multiple genetic backgrounds, and showed that chimeric organoids preserved the overall morphological organization and transcriptomic signatures of the ones generated from single lines. In conclusion our work shows the possibility to perform population level studies in vitro and use the deep resolution of molecular biology to dissect key aspects of human neurodevelopment.

Introduction

Theoretical framework

Experimental and computational advances in the field of developmental biology are revolutionizing our understanding of the basic mechanisms of human life (Atlasi & Stunnenberg, 2017; John & Rougeulle, 2018).

In order to properly understand the implications of such innovations however, it is important to contextualize the key concepts and theoretical background that is shaping this scientific evolution.

The study of human physiology and pathology is founded on the attempt to systematically observe, describe and classify different types of phenotypic traits, in order to understand their causative determinants during organismal development. The logic and rigor of the scientific methodology allows to build solid relationships between causative factors and observable outcomes. It is clear since centuries that the vast majority of the phenotypes are influenced by the interactions between genetic and environmental factors.

This dichotomy however is embodied in a continuous spectrum where different weights of influence exist. They oscillate between the two extremes of highly penetrant genetic mutation, such as the mutations in BRCA genes for breast and ovarian cancer (Petrucelli, Daly, & Pal, 2016) and high-risk environmental factors, like the exposure to asbestos for pleural mesothelioma (B. M. Robinson, 2012). However, the vast majority of the human diseases that biomedical research is trying to investigate are represented by various degrees of low penetrant genetic variants and environmental factors that are only contributing for a portion of the phenotype.

A similar debate actually dates back at the beginning of the 20th century, between the Mendelian vision of the discrete impact of single genetic variants versus the Galton approach for which continuous traits cannot be attributed to monogenic models, followed by the seminal resolution proposed by Fisher, who shows that the random sampling of alleles at each locus determines continuous, normally distributed phenotypes in the population, assuming many genes affecting a trait (Fisher, 1919). This led to the formulation of the infinitesimal model, for which one or several quantitative traits are described as the sum of lots of small genetic and a non-genetic components. This framework was then used to build better and more efficient statistical models for quantitative genetic analysis (Barton, Etheridge, & Véber, 2017).

In our historical period this is reflected by the rapid and massive increase of genome wide association studies (GWAS), in which hundreds of thousands of single-nucleotide polymorphisms (SNPs) are tested for association with a disease in hundreds or thousands of people. As a results lots of genomic loci have been associated with low risks to different disease traits (The Wellcome Trust Case Control Consortium ,2007; Manolio, 2010; M. H. Wang, Cordell, & Van Steen, 2019).

Complementary to GWAS, other analytical frameworks have been developed to try and explain how genetic variants can affect gene expression, thereby providing an intermediate step to associate a DNA sequence to a phenotypic manifestation. Expression quantitative trait loci (eQTLs) are loci that explain a part of the genetic variance of gene expression. They are typically measured in tens or hundreds of individuals. (Brynedal et al., 2017; Nica & Dermitzakis, 2013) and recently allowed the comprehensive profile of gene expression variation across human tissues, including publicly available genotype, gene expression, histological and clinical data for 449 human donors across 44 tissues , with the aim to mechanistically interpret the genetic basis of diseases (Gamazon et al., 2018).

Within the same kind of analytical models, also methylation QTL (meQTL) and chromatin accessibility QTL (caQTL) have been used to associate particular epigenetic states with the resulting effects on gene expression (Banovich et al., 2018).

Finally, the recent emergence of single cell multi-omics studies is also fostering an increasing interest in dissecting single cell eQTL, with the idea to precisely describe how genetic variants regulate gene expression in each specific cell type (Mandric et al., 2019; van der Wijst et al., 2018; Wills et al., 2013).

Those results are shading light on the molecular basis of complex traits, and elucidating a scenario characterized on the one hand by lots of common genetic variants (present in at least 1% of the population), among which the main drivers are situated in non-coding regulatory regions (Y. I. Li et al., 2016), that have a small effect size on the phenotype (Shi, Kichaev, & Pasaniuc, 2016), and on the other hand by rare variants (with <1% frequency in the population) with larger effect size, that have important roles, particularly for the traits related to cognitive and social abilities (De Rubeis et al., 2014).

All of these results, have recently brought to the proposal of the expanded omnigenic view of complex traits, for which gene regulatory networks are sufficiently interconnected such that all genes expressed in disease-relevant cells are liable to affect the functions of the ‘core’ disease-related genes (Boyle, Li, & Pritchard, 2017).

Interestingly, the omnigenic notion can be considered as a result of the profound debate between reductionism and systems biology, discussed deeply by (Brenner, 2010) that recent discoveries are pushing towards the systems logic, because of its better representation of the real dynamics that characterise biological events (Kevin Mitchell, 2019).

The complexity of the implications of the scientific discoveries on the molecular basis of life for our current and future biotechnological society has been comprehensively analyzed and discussed in (Nowotny & Testa, 2010), where the authors clearly illustrate the importance of contextualizing the building blocks of biology and the scientific discoveries to properly understand the different meanings and values they acquired when studied, used or manipulated in the lab, in the hospitals or at the institutional and political levels.

The scientific advances of our age, however, still entail a number of problems characterizing and limiting the impact and actionability of molecular biology to improve public health. The main issues are represented by insufficient diversity of participants, gender disparities in authorship and tightly knit social network of researchers and consortiums involved in these studies (Mills & Rahal, 2019). Moreover, the field still has fundamental unresolved questions, represented for example by the problem of how to prioritize actionable targets for large cohorts of patients, as well as the still highly debated criteria to select traits that can be analyzed with the statistical models previously described, particularly evident in the public debate that followed the recent publication of the largest study on the genetic basis of sexuality (Ganna et al., 2019).

Systems medicine, that can be defined as the interdisciplinary field that integrates the holistic vision of the human organism with quantitative mathematical models (*Systems Medicine*, 2016), attempts to include the vast spectrum of factors that influence human life in a unique framework, that resonate around a network of nodes and connections that together contribute to various effects. This concept is thus revolutionizing even our imagination and abstraction of biology, where informatic metaphors are better suited than the old mechanistic representations. Indeed, we can better study life science phenomena if we focus on the processing and the transmission of information at a system level. We should modify our abstraction of a static genome that determines the rules, towards a complex cellular environment that dynamically regulates gene expression for lineage differentiation in response to genetic and external stimuli, but

also feed information back to the genome in the context of evolutionary editing (Shapiro, 2011).

Looking back at the historical times when the human mind was represented for the first time, through abstractive efforts, with the logic of computational models, in the broader attempt, during the 20th century, to connect the structural knowledge of the physiology of the brain with the cognitive properties of the mind (only studied by psychology at the time), we find the transdisciplinary clash and match between medicine, sociology and mathematics that gave birth to the first steps of artificial intelligence development (Abraham, 2016) (W. S. (Warren S. McCulloch, 1970). The seminal paper by McCulloch and Pitt, respectively a physician and a mathematician, formulated the first mathematical model of neural networks, proposing to represent the relations between neurons by means of propositional logics and moving the focus towards the net of possible activities instead of the single nodes. In this way they built the basis of systems biology, network science and machine learning that nowadays are already very efficient for studying the biological basis of many human diseases or physiological processes (Eraslan, Avsec, Gagneur, & Theis, 2019; W. S. McCulloch & Pitts, 1943).

Moving one step beyond, from the research of the molecular mechanisms at the basis of human pathophysiology to the clinical world and the medical practice, we can observe that, even in this context, the approach to diagnosis and treatment of diseases is undergoing a profound change that is moving the critical decision making from the expertise of physicians to the computational power of machines that can use algorithms to stratify, predict and suggest to humans the best guide lines to follow in a number of different scenario. The main reason at the basis of this shift is the obvious observation that computers can process a number of information, both in terms of quantity of data for each patient and number of different individuals that can be cross-compared at the same time, that is not conceivable for human thinking.

As illustrated in a recent book by Eric Topol (Topol, 2019) even for a highly challenging process such as medical decision making, artificial intelligence is evolving rapidly and achieving degrees of efficiency that already overcome human expertise, as in the case of big tech companies developing products to monitor and record a huge amount of data, analyzed with advanced statistical models, to predict deviation from the health status before symptoms onset (Beam & Kohane, 2018; Dillon et al., 2015; Miotto, Li, Kidd, & Dudley, 2016; Turakhia et al., 2019). Moreover, with the exponential growth rate of the molecular characterization that a lot of patients and

healthy individuals are undergoing, it is evident that computational analysis will become increasingly important for medical practice.

There are several potential problems and challenges still associated to the path for reaching a smooth and automated medical decision making, especially due to the black-box nature of machine learning algorithms ("The Dark Secret at the Heart of AI - MIT Technology Review,"), biased learning of machines because of human data pattern such as racism and sexism ("Machines Learn a Biased View of Women | WIRED,"; "When It Comes to Gorillas, Google Photos Remains Blind | WIRED,"), as well as unresolved ethical issues related to the tremendous capability of artificial intelligence to threaten patient preference, safety, and privacy (Rigby, 2019). Nevertheless, this whole process represents a great opportunity for medical practitioners to increase the quality of their action, if they manage to take advantage of the efficiency of computational tools in diagnostic and therapeutic decisions, to dedicate more time to the unique edges of human relationships and communication with the patients, realizing the deep empathy that is needed for medical care, as underlined throw-out the history of medicine (Derksen, Bensing, & Lagro-Janssen, 2013; PEABODY, 1927; Topol, 2019).

Along the same research interests and goals, part of our work was dedicated to the study of an emerging innovation that could have a far-reaching impact on the way in which data is collected and circulated in the field of genomic research, namely blockchain technology applied to health and genomic research uses (Grishin, Obbad, & Church, 2019).

Over the last ten years, blockchain technology has acquired public awareness following the popularization of its most well-known application, the Bitcoin cryptocurrency (*Bitcoin: A Peer-to-Peer Electronic Cash System*, 2008). Blockchain is a database distributed across a peer-to-peer network that includes actors, or nodes, that voluntarily agree to take part in it. The database serves as a ledger, permanently and cryptographically registering all transactions between actors on the network. The network is decentralized, as the ledger is distributed between nodes without the need to rely on central servers of third-party authorities. Transaction rules are formed by common consensus and embedded into software that ensures rules are followed. Each actor's transactions are transparent and accessible to all other actors, although they cannot be retroactively modified (Vigna & Casey, 2016).

The method of following transactions that are permanently stamped into a historical register in the form of a shared yet secure ledger, can potentially be applied to many fields beyond money transactions.

Very recently, initiatives have developed, that apply blockchain technology to collect, circulate and use individuals' genomic and health information for a variety of goals, aiming to realize a personalized, more participative biomedicine (Grishin et al., 2019; Kain et al., 2019). As owners of their data, individuals are viewed as equal partners in the process of maintaining the genomic ledger and are incentivized by the blockchain transactional reward structure (Grishin et al., 2018; Mackey et al., 2019).

While many literature sources focus on blockchain as a form of distributed, decentralized ledger and its implications for money and forms of economic transaction, we focused on the notions of consensus and reward or incentive to explore the implications of applying blockchain for genomic and health information. To this purpose, we examined comparatively the business models of emerging platforms that represent different models of collecting and 'tokenizing' health and genomic data on blockchain (Table 1).

Table 1: Review of existing genomics blockchain platforms

Genomic Blockchain	Website	Whitepaper	Token	Blockchain platform
Nebula	Link	Link to pdf	Nebula	Ethereum, Blockstack
Zenome	Link	Link to pdf	ZNA	Ethereum
Ecrypgene	Link	Link to pdf	DNA	Ethereum
Genomes	Link	Link to pdf	GENE	Ethereum
LunaDna	Link	Not available	LUNA coin	Not available
Alphacon	Link	Link to pdf	ALP	Ethereum
Shivom	Link	Link to pdf	OMX	Ethereum
Dnatix	Link	Link to pdf	DNTX	Ethereum

We argue that understanding the novel concepts of distributed consensus is crucial for properly comprehending the fundamental innovations that blockchain can foster in

basic and translational research based on omics data. Within this context, omics applications of blockchain claim to decentralize the ecosystem to put users, developers and other stakeholders on the same level. This flattened structure can potentially allow for the dispersion of power. However, platforms compete, and are sometimes structured, to become points-of-entry that create a hierarchy between users or customers, and those who develop and control the specific blockchain application. Thus it is important that both single users of genetic tests, companies and institutions generating and analyzing the data, as well as regulatory agencies will develop a critical formation and consciousness for the correct management of genomic data in a decentralized environment.

To conclude this first introductory part, the high degree of human-machine hybridism that our society is facing, could foster the longstanding attempt to achieve a personalized, interdisciplinary and interactive way to practice medicine and increase the importance of the preventive approaches (Gameiro, Sinkunas, Liguori, & Auler-Júnior, 2018; Loomans-Kropp & Umar, 2019). This processes are indeed reflected by the importance of the concept of disease interception, in the structure of the international and visionary effort of shaping the next future of European biomedicine, in which we are directly contributing within the Lifetime initiative ("The LifeTime FET Flagship"). Since the advances of biomedical research are revolutionizing the scales and ways in which molecular markers can be analyzed and used for personalized medical care, our effort is to integrate the omics approaches, that are already interiorized in the research setting, also into the clinical practice, primary care and public health policies.

Human brain development

The brain is the most complex and fascinating organ of human species, especially to understand the higher cognitive functions that characterize the peculiarity of our evolutionary trajectories. In order to properly explore the fundamental mechanisms of brain functioning, our approach is to study its developmental unfolding, focusing in particular on the neocortex, the evolutionarily newest portion of the brain, composed by a high complexity of cell types and organized into a structure of six layers, that is particularly enlarged in primates and underpins higher mental functions for humans. Genetic and environmental factors interact throughout physiological brain development to fine tune the regulation of its complex differentiation during embryonic and post-natal life, needed to obtain the different cell types, structure and functions that we reach in the adult life. Basic research, until recently mostly performed on animal models, provided a lot of molecular insights into the spatial and temporal orchestration of neurodevelopment. Notwithstanding, the recurrent failure of translational attempts to bring the results obtained in the laboratory settings into successful strategies for human therapy in neuropsychiatric and neurological disorders underlines the need of models that dynamically project and recapitulate the unique properties of human brain development. One of the most important peculiarity of the human brain is the higher number of neurons that populate the cortex. Also, human cortical expansion is accompanied by increased folding of the pial surface, that generates a gyrencephalic rather than lissencephalic neocortex. The main reason that can explain these features however comes from processes that happen during the early stages of fetal life, reflecting the prolonged and increased proliferation of neural progenitor cells (NPCs). After neural tube closure, in the area identified as ventricular zone (VZ), neural stem cells form the primary germinal layer that starts to proliferate and establish a pool of neuronal progenitors, called apical radial glia (aRG). aRG cells are elongated cells, with protrusion towards both the ventricular (apical) and pial (basal) surfaces of the neocortex and begin the asymmetric cell division that generate the progenitors that migrate superficially into the subventricular zone (SVZ). These basal progenitors that populate the SVZ in human, however, undergo a rapid (between gestational week 11 and 17) expansion that constitute the outer SVZ, where outer radial glia (oRG) cells proliferate and generate the vast majority of the cells that will differentiate to become cortical neurons, such as the ones that will populate the middle and upper layers of the cortex. One of the reasons that differentiate the proliferative capacity of aRG vs oRG is that while the former can only divide next to the ventricle,

the latter can divide throughout the SVZ. oRG cells are morphologically similar to aRG cells, with the pial extensions that constitute a fundamental characteristic for girenchephalic species in order to guide neuronal migration. Even though they express specific markers, such as HOPX and ARHGAP11B, it is only the combination of their gene expression signature, characteristic morphology and localization that allows to unequivocally identify this cell type. oRG have also a specific mode of cell proliferation, called mitotic somal translocation because of the jump-like behaviour of the daughter cell after division. Moreover, with proliferation, each oRG can produce hundreds of neurons that migrate to different cortical layers and also astrocytes, depending on the timing of differentiaion. After the progenitor expansion, indeed, cortical plate neurons are generated following an inside-out order that sees the migration of deep layers neurons followed by upper layers ones, afterwards. Moreover, within the layers of the cortex there is the additional integration of cell lineages with different developmental origin, like excitatory and inhibitory neurons that come from the dorsal pallium and the ganglionic eminence of the ventral subpallium, respectively. Finally, different developmental trajectories also intersect among each others for the non - neuronal cell componenet of the brain cortex, with the differentiation and integration of the astrocytes, mainly deriving from the dorsal telenchefalon, the oligedendrocytes from the ganglionic eminence and microglial cells that instead populate the central nervous system after a differetiation from the mesodermal lineage.

In addition to the radial domain of the laminar structure, the human cortex is also organised in a tangential fashion, with different cortical areas (orbital (OFC), dorsolateral (DFC), ventrolateral (VFC), medial (MFC), and primary motor (M1C) cortices of the frontal lobe; the primary somatosensory (S1C) and posterior inferior (IPC) cortices of the parietal lobe; the primary auditory (A1C), posterior superior (STC), and anterior inferior (ITC) cortices of the temporal lobe; and the primary visual (V1C) cortex of the occipital lobe) reflecting the different computational functions they have to perform. Those functional differences are however determined by different trajectories that start already during the fetal period of neuronal maturation, as illustrated by the hourglass pattern of transcriptomic regulation between cortical areas, across different time of development, with the greatest interareal differences present in the prenatal period, followed by a decrease in transcriptional divergence during infancy and childhood and an increase from adolescence onward. Several studies showed that different morphogens, during early phases of neurodevelopment,

can pattern the specific cortical areas differentiation, in particular the gradients along dorso-ventral and anterio-posterior axes of extracellular signals, such as Sonic hedgehog (Shh), Fibroblast Growth Factors (FGFs), Retinoic Acid (RA), Bone Morphogenetic Proteins (BMPs), the WNT family of secreted proteins, and epidermal growth factors (EGFs). Those evidences points to the protomap and radial unit hypotheses for which neocortical areas share similar dimensions and organization among different species and are independent from the connections with extracortical neurons. On the other hand, it was shown that the wiring of thalamocortical afferents (TCA) is highly plastic and both cortical and subcortical connections can interact with the early neurons in the subplate for re-specifying their phenotype before they assume their final position in the cortex, following the protocortex hypothesis. Even though different mammals show similar fundamental mechanisms of areal specification, subtle differences could illuminate key aspects of cortical evolution, such as the expression patterns of the transcription factors SP8 and COUP-TFI that differ between human and rodents, and could explain the increased size and complexity of the association areas of the ventro-temporal cortex including the ventral stream of cognitive visual processing. (Alfano & Studer, 2013; Bystron, Blakemore, & Rakic, 2008; Clowry et al., 2018; M. Florio & Huttner, 2014; Marta Florio, Borrell, & Huttner, 2017; Kyrousi & Cappello, 2019; Lui, Hansen, & Kriegstein, 2011; Miller, Bhaduri, Sestan, & Kriegstein, 2019; Molnár et al., 2019; Namba & Huttner, 2017; Noctor, Martinez-Cerdeño, & Kriegstein, 2007; Pletikos et al., 2014; Pollen et al., 2015; Taverna, Götz, & Huttner, 2014; Wilsch-Bräuninger, Florio, & Huttner, 2016)

Figure 1: A schematic illustration of cortical development

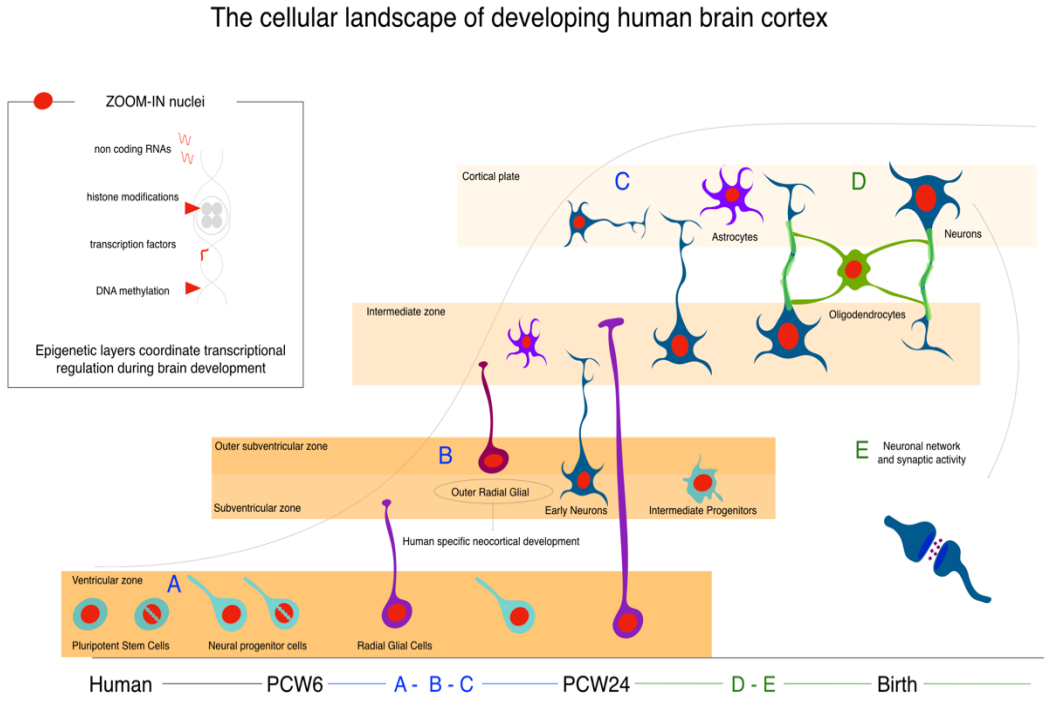
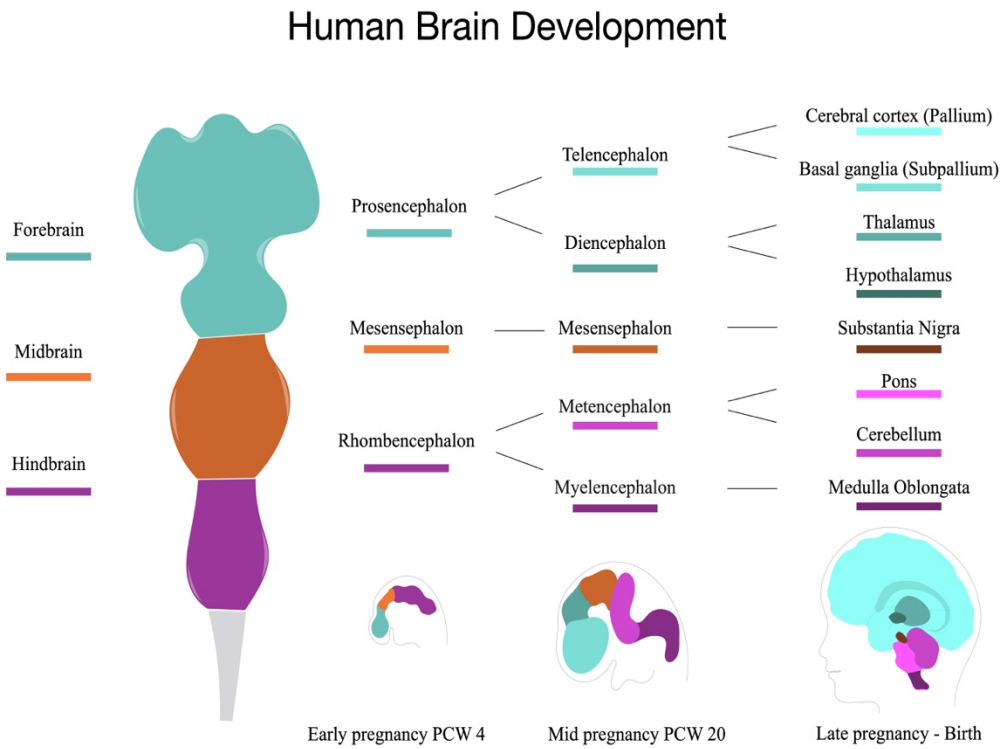


Figure 2: A schematic illustration of fetal brain development



Given the premise on the molecular description of the main neurodevelopmental processes, the vast apparatus of developmental epigenetics, also needs to be briefly reviewed and dissected in its multiplicity of meanings, to properly understand how the dynamic process of brain development can be studied and manipulated. As we recently elaborated in the book chapter (Caporale & Testa, 2019), the study of the regulation of gene expression, and the epigenetic states of the early embryo, constitutes the current challenge and promise to molecularly describe, quantitatively measure and intervene on the principles that govern developmental biology. The entire epigenetic landscape, including the broad set of molecular layers, ranging from histone marks and chromatin states, to DNA modifications and non-coding RNA, is indeed the key player for controlling developmental trajectories. As a matter of fact, a mutation in any of the known chromatin remodeller and/or transcriptional regulator is the ultimate cause of several disorders of the central nervous system as well as cancer, depending on the developmental stage in which the mutation takes place (Gabriele, Lopez Tobon, D'Agostino, & Testa, 2018). On the one hand epigenetics can serve developmental biology research as a tool in reprogramming, to regenerate tissues in the personalized approach that is allowed by the use of induced pluripotent stem cells (iPSC). On the other hand, epigenetic is a way to measure organism's developmental time, a tool to analyze and predict functional states *in vitro*. The latter sense, however, requires itself two lineages of epigenetic meanings: if epigenetic is meant just in the shallow "epi" sense, then all that matters is to verify which of the epigenetic layers are most relevant for verifying cell identities and are most amenable to tractability *in vitro*, an approach that has been recently used to demonstrate the ability of experimental systems to recapitulate in vivo dynamics (Bouschet et al., 2017; Luo et al., 2016). If instead epigenetic is meant in the sense of developmental and heritable states, it may be possible to check, even in the early stages of an experimental model differentiating *in vitro*, if the epigenetic patterns are properly in place so that can be predictive of future developmental milestones. Seminal works in line with the predictive principle were recently carried out and show that it is possible to use epigenetic information, in particular DNA methylation profiles, to predict complex phenotypic traits and diseases (McCartney et al., 2018; Shah et al., 2015). As it was in-depth argued before, the very same term epigenetic has the blurred contours required for representing the flexible and versatile concepts that entangle "the whole complex of developmental processes" that connect genotype and phenotype (Meloni & Testa, 2014).

Neurodevelopmental disorders

Neurodevelopmental Disorders (NDDs) encompass a broad group of diseases characterized by alterations in the development of the central nervous system (CNS) resulting in varying degrees of cognitive symptoms. Several conditions are grouped under NDD diagnosis, including intellectual disability (ID), learning disorders, communication disorders, motor disorders, attention deficit hyperactivity disorder (ADHD) , autism spectrum disorder (ASD) as well as schizophrenia among the others ("DSM-5," 2016; Thapar, Cooper, & Rutter, 2017).

The worldwide prevalence of NDDs is estimated to be around 4% according to the systematic analysis of the Global Burden of Disease Study 2016 (Global Research on Developmental Disabilities Collaborators et al., 2018). This value doesn't take into account visual and hearing loss that were the most prevalent developmental disabilities recorded worldwide, and pools together intellectual disabilities, ASD, epilepsy and ADHD. On the other hand, a systematic review analyzing intellectual disability alone, gives estimates around 1% (McKenzie, Milton, Smith, & Ouellette-Kuntz, 2016). NDDs prevalence is different across socio economic classes, with a disproportionally higher rate in low-middle income countries (Bitta, Kariuki, Abubakar, & Newton, 2018). Moreover, compared to the analysis performed by the same initiative in 1990, there was only a marginal decline in developmental disabilities prevalence, in contrast with the strong decrease of mortality among children younger than 5 years, reflecting the absence of any systematic global initiative to curtail this problem. Thus NDDs entail an enormous global burden in term of personal suffering, health care expenditures and lost productivity (Willsey et al., 2018).

Even if NDDs are very heterogenous, the boundaries between them are not clearly defined, neither at the level of clinical manifestations, nor for the underlying multi factorial etiology and pathophysiology. With GWAS performed on large cohorts, indeed, increasing evidences are showing common genetic variants that are shared between different NDDs and also other neuropsychiatric diseases, such as bipolar disorders and major depressive disorders (The Brainstorm Consortium et al., 2018).

As recently systematized in a comprehensive review by our lab, mutations in chromatin regulators and transcription factors figure prominently among the most common genetic causes of NDDs, with 152 listed in the Simons Foundation Autism Risk Initiative (SFARI) database. This molecular convergence into epigenetic regulation constitutes a first step to elucidate the broader phenotypical links between multiple NDDs, underlining the importance of identifying shared gene regulatory

networks to better shape the classification of brain disorders and develop new therapeutic potentials (Gabriele et al., 2018).

Moreover our lab provided key contributions to the characterization of such conditions, through i) the first systematic characterization of the transcriptional dysregulation of a paradigmatic Copy Number Variation (CNV) causing Williams-Beuren syndrome and 7q11.23 duplication syndrome, using the reprogramming of induced pluripotent stem cells, and finding that the deletion and the duplication of the chromosomal region 7q11.23, already at the early stage of pluripotency, alters the expression of genes that are involved in the health domains connected to the specific clinical impairment of the syndromes (Adamo et al., 2015), ii) the recent discovery of a new neurodevelopmental disorder caused by the haploinsufficiency of YY1 (Gabriele et al., 2017; Nabais Sá, Gabriele, Testa, & de Vries, 2019), and iii) the identification of the downstream targets of the dystonia-causing histone methyltransferase KMT2B that is selectively required for neuronal reprogramming (Barbagiovanni et al., 2018).

Fully penetrant NDD causing mutations in genes that regulate transcription are particularly relevant as they operate at a level of regulatory control that almost invariably affects cell fate and hence most likely translates into developmental phenotypes. Amidst the rising awareness of the impact of the three-dimensional genome organization on development and disease dynamics, as shown in the case of cell type specific genome organization in the context of schizophrenia (Rajarajan et al., 2018), our lab recently proposed the notion of the epigenetic manifold underlying human cognition (Vitriolo, Gabriele, & Testa, 2019)**:** a new paradigm for studying NDD that aims at systematically anchoring cognitive functions to the dynamic hierarchy of the 3D regulatory genome unfolding through neural development at the level of individual cells in their mutual set of proximal and distal connections. These intellectual trajectories led us to develop one of the key conceptual innovations for our future research directions, namely the use of NDDs to leverage the convergences and divergences across their phenotypes to illuminate the evolutionary trajectories of the modern human brain.

The promise and challenge of the next decade, indeed, is represented by the advance of precision psychiatric medicine in moving beyond the discovery of genes that are associated to neuropsychiatric traits, towards the elucidation of the mechanistic relations that connect molecular mechanisms to clinical outcomes (Iakoucheva, Muotri, & Sebat, 2019).

Genetics architecture of Autism Spectrum Disorder

The genetic architecture of NDDs is characterized by heterogenous causes among different patients and a high complexity of causes within individuals (Mitchell, 2015). Autism spectrum disorder (ASD) is a broad set of NDDs characterized by i) young age of insurgence, ii) impairment in communication and social abilities, iii) restricted interests and repetitive behaviors, and iv) symptoms that affect patients' function in various areas of their life ("DSM-5," 2016). The degree of symptoms' severity has wide variance among the affected individuals and the genetic background has a strong influence on clinical onset (Yoo, 2015).

The prevalence of ASD is estimated to be between 1 and 2%, a value that has increased in the last decades (Baio et al., 2018; "Data & Statistics on Autism Spectrum Disorder | CDC," n.d.).

Two paradigms have been proposed to address the genetic contribution to ASD: the common variants and rare variants models. The first derives from the observation that lots of common variants have small contribution to ASD onset, with the hypothesis that clinically diagnosed ASD individuals can be considered as the extreme portion of a spectrum of genetically influenced behaviors that extend over the entire population, that carry a peculiar combination of polygenic mutations. Common polygenic variants, distributed all over the genome, have been estimated to account for at least 20% of ASD liability and act additively as risk factors (Anney et al., 2012; Gaugler et al., 2014; Klei et al., 2012). On the other hand, the second model reflects the presence of rare, highly penetrant variants. Those variants can be single point mutation or CNV that arise de novo in the germinal cells of the patients and, being affected by negative selection in evolution because of their negative impact on reproductive fitness, are characterized by low frequency in the general population (Saffen, 2015).

Importantly, in contrast to other neuropsychiatric disorders, that usually manifest later in life, such as schizophrenia, bipolar disorder, and major depression, where GWAS have mainly identified a polygenic scenario with multiple alleles of very small effect (Ruderfer et al., 2018; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Wray et al., 2018), progress in ASD genetic research has advanced with the discovery of a lot of rare, de novo, coding, heterozygous, germline mutations that contribute to about 20%–30% of clinical cases of ASD (De Rubeis et al., 2014; Iossifov et al., 2014; Sanders et al., 2015) commonly observed in global developmental delay, in addition to autism itself (Deciphering Developmental Disorders Study,

2015), and that offer transformative potential to illuminate ASD pathophysiology and critically important for new therapeutic discoveries (Sestan & State, 2018).

Moreover the systematic work carried out by the Simons Foundation Autism Research Initiative (SFARI) already characterized and scored 1089 genes for their strength of association to ASD, including 450 (15.5%) common variants and 2462 (84.5%) rare variants (for which missense variants represent the most common type of mutation) ("Statistics - SFARI Gene").

Finally, a recent attempt to extend the transmission disequilibrium approach to encompass polygenic risk scores, showed that common polygenic predictors encompassing ASD, schizophrenia and years of educational attainment, are unambiguously associated to ASD risk and that they also contribute to the phenotype in the cases of patients carrying de novo deleterious mutations (Weiner et al., 2017).

The variability and complexity embodied in the genetic basis of ASD inspired the effort of recent neuroscience to look for convergent molecular pathways that can group a number of different genetic mutations, into categories that share developmental features and can explain the similarity of phenotypes between heterogenous conditions. The main biological pathways that have shown convergency in the context of ASD are represented by the epigenetic regulation of gene expression, post-transcriptional alternative splicing, regulation of translation, and synaptic activity, particularly evident for example in the context of mammalian target of rapamycin (mTOR), mitogen activated protein kinase (MAPK) and Wingless (WNT) signaling pathways, as well as higher order neuronal differentiation processes, such as migration, regional patterning, axonal and dendritic harboring, needed to establish the correct brain circuits (Gandal et al., 2018; Geschwind, 2011; Iakoucheva et al., 2019; Quesnel-Vallières, Weatheritt, Cordes, & Blencowe, 2019; Ruzzo et al., 2019). This effort can be summarized by the comprehensive Psychiatry Cell Map Initiative (PCMI), whose goal is to uncover new molecular, cellular, and circuit level understanding of neuropsychiatric disorders to reveal new targets for future therapies and bridge the gap between gene discovery and translational biology (Willsey et al., 2018).

Environment factors and endocrine disruptors

In addition to the complexity of the genetic architecture underlying ASD, this spectrum of disorders is characterized by the concomitant presence of several specific environmental factors with an influence on its pathogenesis.

Environmental factors that have an impact on ASD onset include maternal nutrition, hormonal equilibrium and stress status, as well as drug use and exposure to environmental chemicals, including air pollutants, pesticides, plastics derivatives and metals (Homberg, Kyzar, Scattoni, et al., 2016; Pelch, Bolden, & Kwiatkowski, 2019). Prenatal exposure to alcohol is one example of non-genetic risk factor, that seems to act through the dysregulation of the SHH pathway, and has an estimated prevalence that ranges from 6 to 9 per 1000 children (May et al., 2014). Also, many studies observed a higher risk for ASD or ADHD symptoms in subjects prenatally exposed to tobacco smoke (Tran & Miyake, 2017).

Maternal nutrition is very important to ensure the correct fetal supply of nutrients, in particular for fat-soluble vitamins (A, D, E), tryptophan and nutrients related to single carbon metabolism (choline, vitamins B2, B6, B12 and folate) that can contribute to the methylation of metastable epialleles in the progeny, that is persistent also in differentiated tissues (Dominguez-Salas et al., 2014). The maternal stress status, for which the neuroendocrine axis is very sensitive, has also been associated to alteration of epigenetic programming that can impair correct neuronal development (Maccari, Krugers, Morley-Fletcher, Szyf, & Brunton, 2014). Moreover, a recent study in mice described how maternal care deprivation can alter the function of progeny's brains, by modulating the activity of the retrotransposon LINE-1 that is a key source of somatic mosaicism in the brain, as well as altering the methylation of binding sites of the transcription factor YY1 and the expression of the DNA Methyltransferase DNMT3A (Bedrosian, Quayle, Novaresi, & Gage, 2018).

As for drug use, epidemiological studies showed that prenatal exposure to SSRI (Boukhris, Sheehy, Mottron, & Bérard, 2016) and valproic acid (Christensen et al., 2013) increase the risk of ASD in the offspring, and even prenatal exposure to paracetamol has an impact on language development (C.-G. Bornehag et al., 2018).

The early phases of human life, in particular prenatal life, has huge importance for the correct development and functioning of the human brain (Joseph M. Braun, 2017). Environmental chemical exposure is a key factor alone for ASD risk, because of its pervasive nature and its ability to influence the early phases of fetal development (Zoeller et al., 2012). This is particularly relevant for the endocrine disruptive

chemicals (EDC), defined by the WHO as an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse effects in an intact organism, or its progeny, or (sub)population ("WHO | Global assessment of the state-of-the-science of endocrine disruptors," 2013). The definition of EDC is a non-trivial process, that entails complex mechanisms that have to be taken into account and has huge implication for the society, given the immense number of stakeholders connected to this area, as it is shown by the recent effort of European Commission in mediating between industry, society, institutions and scientific results to define EDC and the policies for their usage ("Process to set scientific criteria to identify endocrine disruptors | Public Health, European Commission"). As an example, it is still not clear if heavy metals such as lead, cadmium, mercury, arsenic, manganese, and zinc, that are known neurotoxicants, should be classified as EDCs. Heavy metals are very diffuse as environmental contaminants and the exposure is indeed associated to NDDs (Tran & Miyake, 2017).

The exposure of fetal brains to EDC, in any case, is of particular concern because, through the dysregulation of the delicate hormonal balance, epigenetic responses can mediate effects on neuronal progenitors that have long term adverse health impact (Baccarelli & Bollati, 2009; Schug, Blawas, Gray, Heindel, & Lawler, 2015; Tran & Miyake, 2017). EDC adverse effects on neurodevelopment, indeed, can act at different scales and times, such as progenitor proliferation and migration, as well as neuronal maturation or the synthesis, transportation and release of neurotransmitters (Schug et al., 2015), and their association to ASD has been hypothesized to be mediated by the alteration of estrogen and NRF1 signaling pathways, in experimental studies, as well as epigenetic modifications that yet need to be systematically characterized (Moosa, Shu, Sarachana, & Hu, 2018). The relative costs for health care systems and society are as a consequence immense, given estimates over $1.9 billion in the United States and $265 million in the European Union (Lee, 2018).

In the following paragraph the compounds that are better known in terms of their molecular effect on neurodevelopment are reviewed.

Phthalates derive from a multitude of consumer products, including personal care products, medications, and plastics. They are ingested, inhaled and absorbed from the derma, and they can also cross the placenta. (Carl-Gustaf Bornehag et al., 2005; Joe M Braun et al., 2014; A. R. Singh, Lawrence, & Autianx, 1975). Because of their chemical properties, they may interfere with the action or metabolism of androgens, thyroid hormones, and glucocorticoids. Prospective cohort studies have associated

phthalates with ADHD and ASD behaviors, reduced mental and psychomotor development, emotional problems, and reduced IQ (Engel et al., 2010; Whyatt et al., 2012).

Bisphenol A (BPA) is used in several plastics that are present in day life consumer products. It is known to act through the binding of estrogen receptors α and β as a weak agonist (Milligan, Balasubramanian, & Kalita, 1998), but it can also interfere with androgen and thyroid signaling pathways (Romano et al., 2015). Different studies suggest that BPA exposure is associated with behavioral and cognitive problems in children, but there are inconsistencies with regard to the period of life with the greatest vulnerability and sex-specific effects (Joseph M. Braun, 2017).

Perfluoroalkyl Substances (PFAS) are used in water-resistant materials, industrial surfactant and food containers, thus, like phthalates and bisphenols are spread everywhere in the environment ((EFSA), 2008). They are very resistant so they bioaccumulate in human for several years. They are active on several endocrine axes, in particular interacting with the peroxisome proliferator activated receptor (PPAR), glucocorticoid and thyroid pathways (Boas, Feldt-Rasmussen, & Main, 2012; Vanden Heuvel, Thompson, Frame, & Gillies, 2006; Ye, Guo, & Ge, 2014). Their exposure has been linked to ASD and ADHD risk by epidemiological studies (Liew et al., 2015; Ode et al., 2014).

Polychlorinated biphenyl (PCB) were used as coolants, plasticizers, and flame retardants, among other uses, until their production was banned. Because of their resistance to degradation, however, human continue to be exposed to them. Their mechanisms of action are not well known but exposure has been associated to problems such as inattention, impulsiveness, and other ADHD–related behaviors (Marczylo, Jacobs, & Gant, 2016; Schug et al., 2015).

Finally, neurotoxicological studies on EDC have brought to the attention of scientists and regulators three important issues that are usually underestimated and, in general, not properly addressed in other contexts: the importance of low doses, long term effects and mixtures. It was observed, indeed, that even EDC doses similar to the environmental levels can have significant effects in experimental models, showing peculiar dose-response patterns that are not monotonic (Vandenberg et al., 2012), even if this phenomenon seems to be not so common (Beausoleil et al., 2016). Moreover, EDC can cause effects that only manifest much later in life than the period of exposure (Barouki, Gluckman, Grandjean, Hanson, & Heindel, 2012), most likely because of epigenetic mechanisms of information transmission (Skinner, Manikkam, & Guerrero-

Bosagna, 2010). As far as EDC mixtures are concerned, there are still huge gaps of knowledge regarding their mechanisms of action, especially if considering molecules that interfere with different hormonal pathways, for which additive predictive models are not suited and additional experimental evidences and testing systems are required (Kortenkamp, 2014). To conclude EDC are challenging the whole field of toxicology that now needs to find innovative tools and models to study and regulate the potential danger associated to chemical exposure and thus embodies both the challenge and the opportunity to translate new scientific discoveries into policies (Barouki, 2017).

Interplay genetics vs environment

The results previously presented highlight the complex interplay between genetic and environmental factors in the context of ASD pathophysiology. This is for example demonstrated by the strong contribution of the genetic background for the response to environmental stressors, such as the differential susceptibility to estrogen exposure in mice (Spearow, Doemeny, Sera, Leffler, & Barkley, 1999), or by the evidence of single nucleotide polymorphisms that in mice confer resistance to the adverse effects induced by the exposure to the endocrine disruptor Di(2-ethylhexyl)phthalate (DEHP). In this case the effect is mediated by the increased expression of the estrogen receptor as well as alteration of specific promoter methylation, that are regulated in opposite fashion by the genetic variants and the chemical exposure (Stenz, Rahban, Prados, Nef, & Paoloni-Giacobino, 2019). Moreover twins studies have shown, through the analysis of structural brain measures, in a cohort including ASD patients, that while cerebral and cerebellar gray matter (GM) and white matter (WM) volume, surface area, and cortical thickness, were primarily influenced by genetic factors, mean curvature appeared to be primarily influenced by environmental factors (Hegarty et al., 2019).

However, the current knowledge and scientific data about gene per environment (GxE) interactions are still limited and additional experiments, with proper design need to be further performed to illuminate the most relevant interactions at the population level (Esposito, Azhari, & Borelli, 2018). This is especially challenging for the GxE of low penetrant common variants, since a lot of research effort focused until now only in dissecting the impact of specific environmental factors on the NDD highly penetrant mutation (Dick et al., 2015). The increasing resolution of environmental data records ("The Human Exposome Project,"), coupled with the higher number and depth of genomic profiles, will likely help to perform more efficient statistical GxE analysis on larger cohorts and push the field closer to the stratification of developmental disorders based on both genetic and environmental marker (Esposito et al., 2018; Loo & Martens, 2007; Torkamani, Andersen, Steinhubl, & Topol, 2017).

To conclude, an efficient strategy to address the challenges of translational neuropsychiatry must include i) a better understanding of the molecular mechanisms of action of the drugs currently used for NDDs (Homberg, Kyzar, Stewart, et al., 2016); ii) the identification of common targets across different disorders (Stewart & Kalueff, 2015); iii) the tracing of the temporal dimension of NDDs (Xu et al., 2014); and iv) the focus on the role of epigenetics to bridge the gap between genetic and

chemical factors (De Rubeis et al., 2014). Of positive note, the molecular elucidation of the environmental influence on neurodevelopment, together with the observation that ASD patients, even if stressed during common activities, often have strong abilities in specific tasks, suggests that rather than fitting ASD patients with the 'standard' lifestyle shaped for the majority of the human population, a personalized environmental therapy should be beneficial for these patients (Homberg, Kyzar, Scattoni, et al., 2016).

Brain Organoids

The possibility of deriving human pluripotent stem cells, either from embryos (embryonic stem cells, ESCs) or from somatic cells (induced pluripotent stem cells, iPSCs) has transformed our ability to model and study the early stages of human brain development, allowing experimental access for the first time. The scientific advancements of the last decades with 3D cultures, are enabling, for the first time, an accurate recapitulation of the physiological paths that go from the pluripotent state to the terminally differentiated specific cell types, reproducing the structural and functional properties of in vivo organogenesis, thus opening the doors to the molecular dissection of genetic and environmental influence on NDDs.

The first evidences showing the potency of PSCs to form embryoid bodies and then differentiate into the neural lineage come from the exposure to patterning factors, such as retinoic acid (RA), as illustrated in (Bain, Kitchens, Yao, Huettner, & Gottlieb, 1995; Kawasaki et al., 2000; M. Li, Pevny, Lovell-Badge, & Smith, 1998).

Afterwards, a major improvement in the capacity to funnel differentiation towards specific neuronal types was achieved through the modulation of the SMAD pathway, whose inhibition pushes the cells efficiently towards ectodermal fate (Chambers et al., 2009; Elkabetz et al., 2008; Watanabe et al., 2005).

While this kind of patterning tries to recapitulate the gradual stages of neurodevelopment, parallel research efforts have been carried out to push the differentiation of neurons in a fast and reproducible way. This led to the discovery of the Ngn2 protocol, in which the transcription factor over-expression is able to induce the maturation of functional neurons from PSCs in less than one month. (Qi et al., 2017; Yingsha Zhang et al., 2013), as well as protocols to robustly differentiate non neuronal cell-types of the CNS, such as astrocytes, oligodendrocytes and microglia (Abud et al., 2017; Ehrlich et al., 2017; Ye Zhang et al., 2016).

On the other hand, the interest for recapitulating *in vitro* the complexity of human brain development, given the importance of the early and intermediate stages as well as the interactions between different cell types to understand brain diseases, ushered the development of the transforming model of organoids, 3D structures grown from stem cells and consisting of organ-specific cell types that self-organizes. Brain organoids can be derived from both pluripotent stem cells, ESCs and iPSCs. Starting from the pluripotent state, when cells grow on a matrix in a monolayer, most of the organoids protocols are characterized by the generation of embryoid bodies as a first step. There are several approaches that aim at recapitulating different regions and/or complexity

of the CNS development, each of which is better suited for addressing different biological questions, but there are common aspects to all brain organoid protocols, which include i) the formation of self-organizing structures, in particular rosettes which resemble the embryonic stratification of the epithelium of the neural tube with apico-basal polarity and subdivided into proliferative and differentiating zones, ii) the generation of different subpopulations of progenitors usually absent in the 2D counterparts and iii) a considerable degree of basic compartimentalization at the extracellular level that includes the production of their own rudimentary extracellular matrix (ECM) (Di Lullo & Kriegstein, 2017; Giandomenico & Lancaster, 2017; Kelava & Lancaster, 2016; Kyrousi & Cappello, 2019; Quadrato, Brown, & Arlotta, 2016).

Different brain organoids can be on the other hand very different in their specificity to recapitulate cerebral structures, such as whole brain organoids (Lancaster et al., 2017, 2013; Quadrato et al., 2017), forebrain organoids (Qian et al., 2016; Velasco et al., 2019), midbrain organoids (Jo et al., 2016), hippocampal organoids (Sakaguchi et al., 2015), hypothalamic organoids (Qian et al., 2016), thalamic organoids (Xiang et al., 2019), pituitary organoids (Ozone et al., 2016) and specific brain regions such as the cortical organoids or spheroids (Paşca et al., 2015).

Moreover, within the recapitulation of the cerebral cortex, there is the possibility to differentiate the most dorsal part of the forebrain (the pallium), that give rise to glutamatergic neurons or the ventral one (the subpallium) that, resembling the ganglionic eminence, give origin to GABAergic interneurons (Bagley, Reumann, Bian, Lévi-Strauss, & Knoblich, 2017; Birey et al., 2017; Paşca et al., 2015; Sloan et al., 2017). In this way, 3D differentiation of pluripotent stem cells into cortical brain organoids can recapitulate the physiological emergence of the six distinctive neuronal layers of the cortex, along with astrocytes (Sloan et al., 2017), oligodendrocytes (Madhavan et al., 2018; Marton et al., 2019) and the outer radial glia subset of progenitors that underlies the specific expansion of the human neocortex (Bershteyn et al., 2017; Pollen et al., 2019, 2015). Moreover recent studies show their ability to reproducibly reproduce the cell diversity of human brain, as well as producing patterns of electrical activities whose oscillations resemble the prenatal human electroencephalography (Trujillo et al., 2019; Velasco et al., 2019).

Finally, their recapitulation of neural differentiation has been corroborated with functional studies assessing connectivity. Transplantation of human brain organoids into the mouse brain showed integration of the grafted tissue with axonal projections

to different host brain regions (Mansour et al., 2018). Strikingly, co-culture of brain organoids with sections of the spinal column dissected from embryonic mice, in which the associated peripheral nerves and paraspinal muscles were intact, demonstrated the ability of cortical organoids of forming dense axon tracts that innervated the mouse spinal cord resulting in muscle contraction (Giandomenico et al., 2019).

Because of their capacity to recapitulate human neurodevelopment, one of the most exciting and useful application of organoids is the opportunity to use them to model human diseases. This experimental platform, indeed, opens up the possibility to study the molecular mechanisms of several diseases, both for genetic and environmental causes. It allows to isolate one of the two factors (genes or environment) and selectively interrogate its impact on the disease. Neurodevelopmental and neurodegenerative diseases, in particular, are a promising and expanding area research for organoids-based disease modelling. For studying NDDs, indeed, human cellular systems are particularly relevant because the animal models that have been used for a long time, are often insufficient and inappropriate to examine the pathogenesis of such complex brain disorders that can involve cellular types, in particular neuronal progenitors, that are human specific.

A number of different brain disorders were indeed already studied and manipulated taking advantage of the use of brain organoids, such as microcephaly (Lancaster et al., 2013), for which researchers reprogrammed iPSC from a patient carrying an heterozygous truncating mutations in *CDK5RAP2*, differentiated brain organoids and observed a premature neuronal maturation phenotype. In the case of lissencephaly (Bershteyn et al., 2017), iPSC and organoid differentiation showed an impairment of neuronal migration and mitosis for a specific class of neuronal progenitors, namely the outer radial glia. A similar approach was used also for studying an environmental cause of neurodevelopmental delay, the Zika virus exposure (Garcez et al., 2016; Qian et al., 2016), where cerebral organoids exposed to the virus were affected by increased cell death and reduced proliferation. Idiopathic autism spectrum disorders (ASD) was also studied by means of iPSC derived brain organoids (Marchetto et al., 2017; Mariani et al., 2015). The transcriptomic analysis revealed an imbalanced overproduction of GABAergic interneurons that is hypothesized to be dependent on the dysregulation of the gene *FOXG1*. As for genetically caused ASD, patients affected by Timothy syndrome were studied with the use of cortical glutamatergic and GABAergic cortical organoids and the results of the analysis showed an impairment of the migration of the inhibitory neurons (Birey et al., 2017). When researchers applied the same approach

for exploring the transcriptional dysregulation of the brain development of patients with schizophrenia, they identified an altered proliferation of neuronal progenitors that was hypothesized to be linked to the alteration of the expression of the gene *FGFR1* (Stachowiak et al., 2017). Finally Alzheimer's Disease (Abud et al., 2017; Raja et al., 2016) organoids were suitable to test the effects of different drugs relative to the secretion of the peptides amyloid beta. All those results are elucidating, through the use of the same experimental platform, the molecular mechanisms that are dysregulated in a number of severe brain disorders, with the potential to improve the current therapeutic effectiveness.

Advances in bioengineering, microfluidic systems and biomaterials are transforming the potential to automatize the process of organoids generation and maintenance, thereby significantly increasing the throughput. This implies that drug screening, and population level genomic and epigenomic studies will eventually be addressable through the use of organoids. Moreover, the versatility of fluidic chips permit to finely control the fluxes that feed the organoids as well as the matrices around which organoids develop. This approach indeed permits to reach a high precision in the control of spatiotemporal dynamics over the microenvironment that govern the morphogenetic signals for stem cells niches (Nikolce Gjorevski et al., 2016; Nikolche Gjorevski, Ranga, & Lutolf, 2014).

Recent advances in biomaterials and nanotechnology have enabled extensive progress toward such systems often named organ-on-chip for their characteristic top-down approach in constructing the models, in contrast to relying on biological self-organization and assembly of organoids. Different chips are designed for different modelling aims, consisting of a complex array of signaling mechanisms from the niche support cells, to the extracellular matrix, often mimicked by hydrogels, for which mechanical properties can be tuned by changing parameters such as pore size, crosslinking density, and topology (S. Kim, Lee, Chung, & Jeon, 2013), as well as physical and chemical conditions such as oxygen and pH levels (Yin et al., 2016). Another active area of research is exploring the development of the vascular system in relation to the organ of interest. In a recent study, perfusable network of blood vessels *in vitro* were created within microfluidic channels through the spatially controlled co-culture of endothelial cells and stromal fibroblasts, pericytes or cancer cells. Secretion of proangiogenic growth factors and extracellular matrix proteins by the stromal cells supported angiogenesis of the endothelial cells, producing perfusable networks that show strong barrier function and stability. The *in vitro* derived microvasculature

allows the delivery of nutrients and chemical compounds into the luminal space of the endothelium (S. Kim et al., 2013).

Interestingly also, researchers are characterizing how the shape and the geometry of the matrices used as scaffolds can pattern stem cells for differentiation through the modulation of the Notch and Hippo signaling pathway (Totaro et al., 2017; Yui et al., 2018). Those are promising directions to get closer to recapitulate the human physiological development with an *in vitro* experimental system.

Single cell omics

The capacity of multicellular organisms to generate a panel of diversified structures and functions, using the information encoded in a single genome is one of the most interesting questions in developmental and evolutionary biology. This is achieved by compartimentalization at multiple scales: from molecules, to tissues and organs, through which the cells, during the differentiation processes are guided to express genes in specific ways to achieve a particular function.

Historically, in order to build a taxonomy of cell types using a Linnean framework (Linné, 2013), the classifications of different cells within a tissue or across different tissues has been based on microscopy and, more recently, on fluorescence-activated cell sorting (FACS). Even though the resolution of those observation can achieve the level of single cells, the number of molecules or features that can be recorded at the same time is very limited. These taxonomies can be useful especially in the case of somatic static tissue, however they have lots of limits when trying to classify complex developing organ or whole organisms, because of the lack of a precise hierarchy in the continuum that characterize embryonic development and because often different molecular mechanisms at the basis of cell differentiation are tight enough that they cannot be physically separated from their function.

The recent development of single-cell omics represents a disruptive technique since we now have the opportunity to quantify simultaneously a high number of molecular states for each cell that constitute a complex tissue or organ. It is now possible to genome-wide profile RNA, DNA, histone modifications, chromatin accessibility, DNA methylation, nuclear lamina interactions, chromosomal contacts, and the protein signatures of single cells. Among those, single cell RNA seq is at the forefront because of the higher throughput and accuracy, and its application to the clinics is now closer. (Chappell, Russell, & Voet, 2018; Tanay & Regev, 2017)

These advances opened up the possibility to overcome the inherent limitations associated to the averaged read-out of bulk sequencing for a mixture of heterogenous cell populations, revealing at unprecedented details the developmental landscape of several systems and represents the current frontier to precisely dissect the cellular and molecular architecture of organ development (Clark, Lee, Smallwood, Kelsey, & Reik, 2016; Kelsey, Stegle, & Reik, 2017). Indeed, large consortia, like the Human Cell Atlas Project, are trying to build reference maps of the molecular states of all the cell types in healthy human tissues to study physiological states, developmental trajectories, regulatory circuitry and interactions of cells, and also provide a framework

for understanding cellular dysregulation in human disease (Ecker et al., 2017; *Lifetime*; Regev et al., 2017; Taylor et al., 2019).

One of the most interesting potentials of single cell omics data analysis is the reconstruction and analysis of developmental trajectories. Single cell transcriptomes can be projected, after dimensionality reduction, in a common analytical space, through different algorithms that can define a distance between the single cells coming from a complex tissue, on the basis of the similarity of their transcriptomes, thus defining a new temporal concept defined pseudotime (Haghverdi, Büttner, Wolf, Buettner, & Theis, 2016; Tritschler et al., 2019). This is recently acquiring even a dynamic spin, with the introduction of the concept of RNA velocity, which models the derivative of mRNA abundance through an estimate of spliced vs unspliced transcripts, and thus infer the direction of transcriptional regulation between cells in the pseudotime trajectory (Gorin, Svensson, & Pachter, 2019; La Manno et al., 2018; "scVelo").

Similarly to other areas of data analysis, however, there are a lot of challenges, particularly related to the attempt of integrating different experiments and correcting batch effects, due to the variability that can arise from different sources, from the library preparation and sequencing, to the different platforms that can be used to perform single cell experiments. In the case of single cell datasets, this issue is even more difficult to solve, especially because of the sparsity of the data due to the dropout events (Tracy, Yuan, & Dries, 2019), those genes that can have a zero value for technical rather than biological reasons. In the recent period, indeed, a lot of efforts in the scientific community are converging to try and develop new methods for data integration and correction (Stuart & Satija, 2019)

In addition to the high resolution of the high dimensional molecular landscape of many single cells, current advancements are improving the accuracy and feasibility of performing spatial transcriptomics experiments, in which it is possible to couple the transcriptome of single cells with the information of its spatial position within a tissue (Burgess, 2019; Rodriques et al., 2019; Weinstein, Regev, & Zhang, 2019).

Moreover, an emerging strategy to reduce the costs and technical variability of single cell experiments, thereby increasing the number of biological replicates that can be profiled in each library and sequencing run, is the multiplexing of samples. This is a general idea that can be implemented in different ways. One possibility is to take advantage of the genetic variation between individuals to deconvolute in silico single cells that comes from different genetic backgrounds on the basis of SNPs, as in the

case of (Heaton et al., 2019; Huang, McCarthy, & Stegle, 2019; Kang et al., 2018a). The other option is to use barcodes, that can be coupled to each single cell with the use of specific antibodies targeting membrane proteins or lipids, as in the case of (McGinnis et al., 2019; Mimitou et al., 2019; Stoeckius et al., 2017, 2018) that give the opportunity to not only multiplex together many samples at the same time, but also to couple the transcriptomic readout with protein expression data, or to perform single cell CRISPR screening with multi modal read out.

Finally, the possibility to integrate organoidogenesis with single cell analysis can be leveraged to elaborate a unique high dimensional analytical space, with a vast amount of molecular features, that is providing the scientific community with new insights into the definition of developmental trajectories and is already revolutionizing the criteria to define the identity of a cell. These processes are thus re-shaping our understanding of developmental biology, prompting also towards a more dynamic and functional definition of developmental regions (Martinez Arias & Steventon, 2018).

We can thus conclude that the burgeoning field of single cell omics and organoids, that is directed towards an holistic "omni-omics" dissection of cells, organs and organisms, provides the scientific community with a new conceptual resource and an operational toolkit that bring back the Waddingtonian idiom of developmental paths into the epistemic feature that has become tractable and quantifiable through robust experimental iterations.

Evolutionary perspective

As anticipated previously, we recently elaborated in the lab the innovative idea to study evolutionary trajectories within the network of NDDs, integrating their molecular and clinical features. In particular, for one of the domain that have arguably undergone the most salient changes between anatomically modern humans (AMH) and archaic humans (AH), the brain, our interest is focused on globularization. The reason is that while the theory of domestication observed a reduction of the skull, and thus of the brain, in all domesticated species, this is clearly not the case for H. sapiens, in whom differences in brain shape have been in place of a reduction of brain volume. This specific phenotype, termed globularization, refers to the well-known skull shape difference between AMH, who have a globular neurocranium, and extinct species of the genus Homo, who had a more elongated neurocranium.

It has been argued that this shape change is the outcome of differences in brain growth trajectory early in life (Gunz, Neubauer, Maureille, & Hublin, 2010; Hublin, Neubauer, & Gunz, 2015), which impact skull growth to yield the distinctive, globular shape associated with AMH. Importantly, this distinct brain growth trajectory is uneven and therefore reconfigures the modern brain in specific areas, most clearly affecting the parietal lobe (especially the precuneus), the cerebellum, and the frontal pole, but also impacting the temporal lobe and the olfactory bulbs.

A mechanistic dissection of this evolutionary process has been hampered till now by several key limitations.

First, the thus far understandably prevalent focus on fixed AMH alleles (i.e. variations present in all AMH and excluded from all AH) did not take into account either the extent of AMH genetic variation or the complex interbreeding among different hominin lineages. As recently demonstrated by our lab, widening the spectrum to include nearly fixed changes (i.e. present in at least 90% in the present-day human population while absent from all 5 high quality ancestral genomes) has proven of key value not in expanding the catalog of relevant AMH v. AH changes to new gene networks (Zanella et al., 2019).

Second, the functional dissection of AMH v. AH alleles has thus far progressed on a one-by-one basis, leaving the reconstruction of the multi-allele network logic underlying our modern features wholly unscrutinized.

Third, the few single-gene mechanistic studies have been anyway carried out in model organisms, with obvious caveats for inferring their impact on quintessential AMH features and in the specific regulatory context of modern human tissues.

Cerebral Organoids (CO) thus come into play as a key tool to comprehensively dissect globularization. CO can be used to investigate outstanding AMH features genome wide, and across a spatiotemporal axis of brain morphogenesis with functional readout. It is possible to apply a combination of CRISPR-based perturbations (Boettcher et al., 2018; Datlinger et al., 2017; Dixit et al., 2016; Xie, Duan, Li, Zhou, & Hon, 2017a), scRNAseq, and advanced microscopy, coupled to electrophysiological profiles of neuronal function and network activity, in a dedicated panel of brain organoid models, selected to identify and probe the impact of the given *AMH-circuit(s)* on its predicted phenotype.

The brain shape change in the evolutionary context of globularization, can be studied and dissected into the interplay between the protomap and the protocortex hypothesis about brain cortex arealization, isolating patterning factors and thalamocortical connections during multiple stages of differentiation. The early phases of patterning can me modelled with the chimeric organoids by (Cederquist et al., 2019), in particular by generating gradients of FGF8 (the one among the FGF family with the most clear-cut effects in areal size and positioning) and SP8 expression for the frontal regions as well as NR2F1 and EMX2 for the caudal ones. On the other hand, the combination of cortical organoids with the thalamic ones, using the assembloids protocol described in (Xiang et al., 2019), allows to study how thalamocortical axons can modulate neuronal fate and connectivity potential after the early patterning.

Processing in parallel wild type lines and iPSC edited for the AMH specific circuits, evolutionary relevant molecular alterations can be described, alongside their interference with cortical arealisation, thus shading lights into the dynamics of globularization events that characterized AMH.

Aim of the thesis

The societal and public health impact of recent biomedical discoveries is massive, and it is giving rise to a complex discussion on how to deal with the promises and pitfalls that these processes entails, as clearly illustrated in the recent debate on the New England Journal of Medicine and the Lancet about "Precision public health" (Chowkwanyun, Bayer, & Galea, 2018, 2019; Health, 2019; Horton, 2018).
This PhD thesis would like to be a small step into the translational and interdisciplinary integration of systems medicine approaches. For this work I started from my clinical and statistical background and moved to explore the mechanistic gaze of molecular biology and bioinformatics, coupled with epidemiology, in order to study some of the processes by which genetic and environmental factors can shape the development of human brain. The hope is that our results will be helpful for the development of better public health policies as well as novel understanding of the basic mechanisms that can be translated into medical diagnosis, therapy and preventive care in the context of neuropsychiatric disorders.

Specific aims:

- benchmark the use of cortical brain organoids to recapitulate the salient features of human brain development, by means of systematic characterization of bulk and single cell transcriptomes, as well as imaging analysis of structures and architecture of cellular composition, of multiple stages of organoids differentiation;
- dissect the molecular mechanisms and developmental trajectories that characterize peculiar aspects of human brain corticogenesis, through the specific chemical manipulation of the GSK3 pathway, given its key role in regulating fundamental processes of early brain development, such as neuronal progenitor proliferation and polarity;
- define the mixtures of Endocrine Disruptive Chemicals of potential danger for human health and unravel their mechanisms of action, by integrating epidemiology with experimental testing on human in vitro models, to improve EDC risk assessment and policy making and develope better tools for regulatory neurotoxicology;

- implement and validate innovative experimental and computational systems to disentangle the effects of genetic variants and environmental factors on neurodevelopmental dynamics, by multiplexing at different levels organoidogenesis of different individual iPSC lines, towards the idea of realizing experimental epidemiology *in vitro*.

Results

Human cortical brain organoids recapitulate fetal neurodevelopment

To model the development of the human brain cortex we imported and improved the protocol to differentiate cortical organoids (CO) through the patterning of the dorsal telencephalon (Birey et al., 2017; Paşca et al., 2015; Sloan et al., 2017). This protocol recapitulates the neuroectodermal induction with the first 5 days of DUAL-SMAD inhibition, followed by the expansion of the pool of the neuronal progenitors of the dorsal telenchephalon during the 18 days of EGF and FGF exposure, ushering in the organization of radial glia cells in ventricular zone-like structures. Afterwards there is the gradual differentiation of glutamatergic neurons with the BDNF and NT3 supplement and the long-term maintenance in Neurobasal that also allows the maturation of deep and upper layers neurons as well as astrocytes.

Immunofluorescence of cortical brain organoids across differentiation

The complex set of cell populations in the CO organize themselves spontaneously around ventricular-like structures with the progenitor cells in the inner most regions and the mature cells in the more cortical regions of the organoids.

Immunostainings revealed indeed the sequential presence of an organized architecture of cells expressing markers of apical progenitors (PAX6 and NESTIN), intermediate progenitors (TBR2), outer radial glia (HOPX), neurons of different maturation stages (TUJ1, MAP2, TBR1, CTIP2) and astrocytes (GFAP) between day 12 and 180, that are represented in Figure 3 along with the scheme of differentiation stages of the protocol by (Paşca et al., 2015). We systematically performed immunofluorescence characterization of the early (Day20-25), intermediate (Day50), and advanced (>Day100) stages of organoids differentiation for at least 3 different organoids derived from at least 3 different iPSC lines.

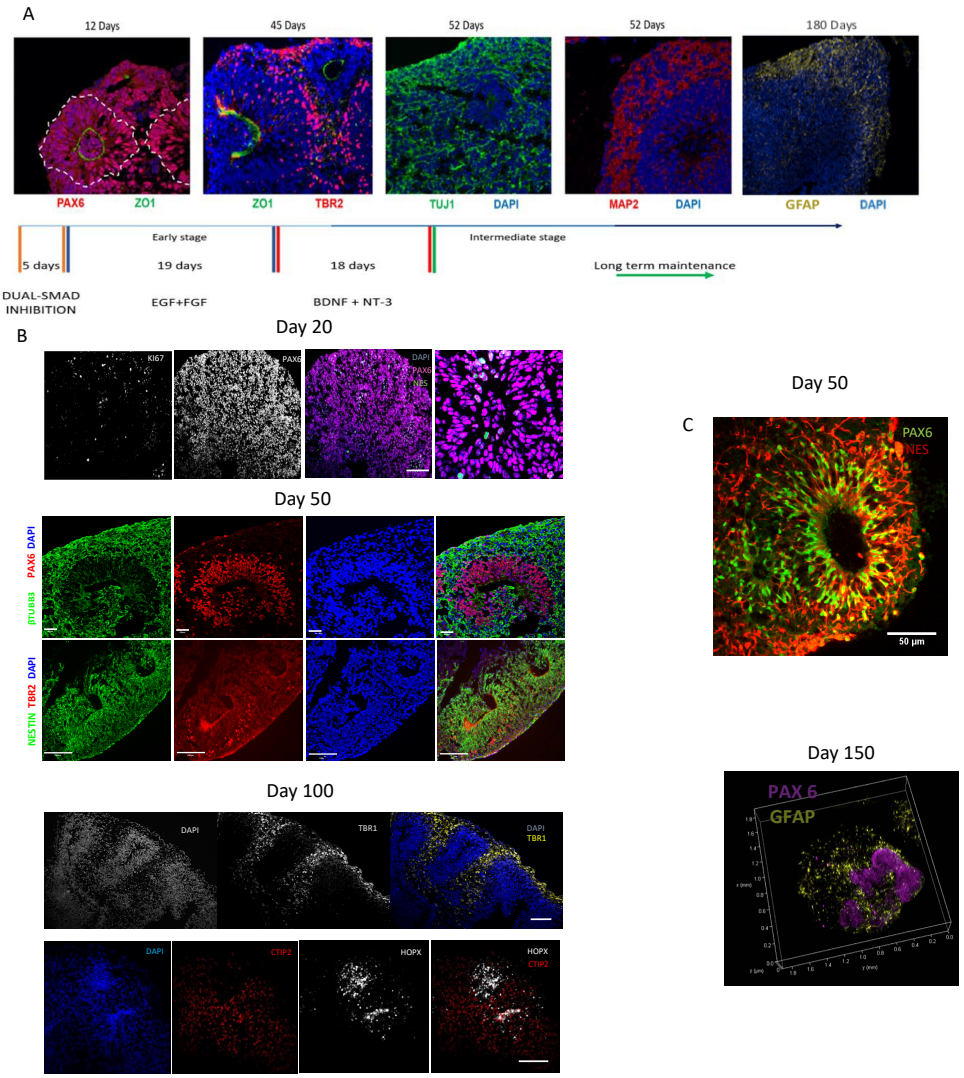


Figure 3. Immunofluorescence of cortical brain organoids across differentiation

A. Representative images of organoids at various stages of differentiation following the protocol by (Paşca et al., 2015). Ventricular-like structures with typical radial organization are observed. Apical progenitors markers PAX6 and NESTIN are expressed. TBR2-positive intermediate progenitors are found outside the layer of apical progenitors. The post-mitotic neuronal marker TUJ1 and MAP2 as well as the astrocytic marker GFAP are observed in more advanced stages.
B. Representative confocal and widefield images of organoids at 20, 50 and 100 days of differentiation. Apical progenitors markers PAX6 and NESTIN are expressed. TBR2-positive intermediate progenitors are found outside the layer of apical progenitors. The post-mitotic neuronal marker TUJ1, and the lower layer neurons markers CTIP2 and TBR1 are observed as well as HOPX that is expressed by outer radial glia cells. Scale bar 20 um.
C. Representative 3D images of organoids clarified with the iDISCO protocol

Transcriptomic characterization of cortical brain organoids across differentiation

We then characterized the CO model through RNAseq by analyzing its longitudinal dynamics in 4 healthy donor lines and two differentiation rounds, from iPSC up to 200-day-old organoids. We also cross compared this dataset with our in-house cohort of primary fetal central nervous system tissues and 2D fetal progenitors cultures. Moreover we compered our data to publicly available transcriptomic profiles of human fetal brain from the BrainSpan Atlas. We first performed principal component analysis (PCA) of CO transcriptomes along with primary fetal central nervous system tissues and 2D cultures of fetal cortex. PCA is a dimensionality reduction technique that allows to project and visualize the overall contribution of the transcriptme of each sample in a low dimensional space, where the principal components are the axis that capture the main sources of variability across the samples. This uncovered a clear temporal dynamic of CO towards recapitulation of the primary tissues and culminating with the close clustering of day 200 organoids with fetal cortex and in sharp contrast to the separate positioning of 2D cultures (Figure 4A). As a next step, stage-wise differential expression analysis of the CO unveiled a first phase of rapid evolution (till day 100), followed by a second phase of more gradual transcriptional modulation (Figure 4B). Gene modulation trajectories through development, defined by Weighted Gene Co-expression Network Analysis (WGCNA), confirmed the progressive downregulation of cell proliferation markers with the concomitant up-regulation of neuronal maturity signatures characteristic of the dorsal pallium (also confirmed by immunostaining, Figure 4C). Importantly, benchmarking this model against fetal corticogenesis datasets from post-conceptional week 8 through 37 (BrainSpan) highlighted a high and stage-specific degree of concordance that also recapitulated the velocity of fetal cortex gene expression programs (Figure 4D). Finally our CO cohort, including 4 individual iPSC lines in 2 rounds of differentiation, resulted very reproducible across replicates with overall similar maturation profiles during the entire time of differentiation (Figure 4A and 4D), thus reliably and reproducibly mimicking the human fetal cortical maturation.

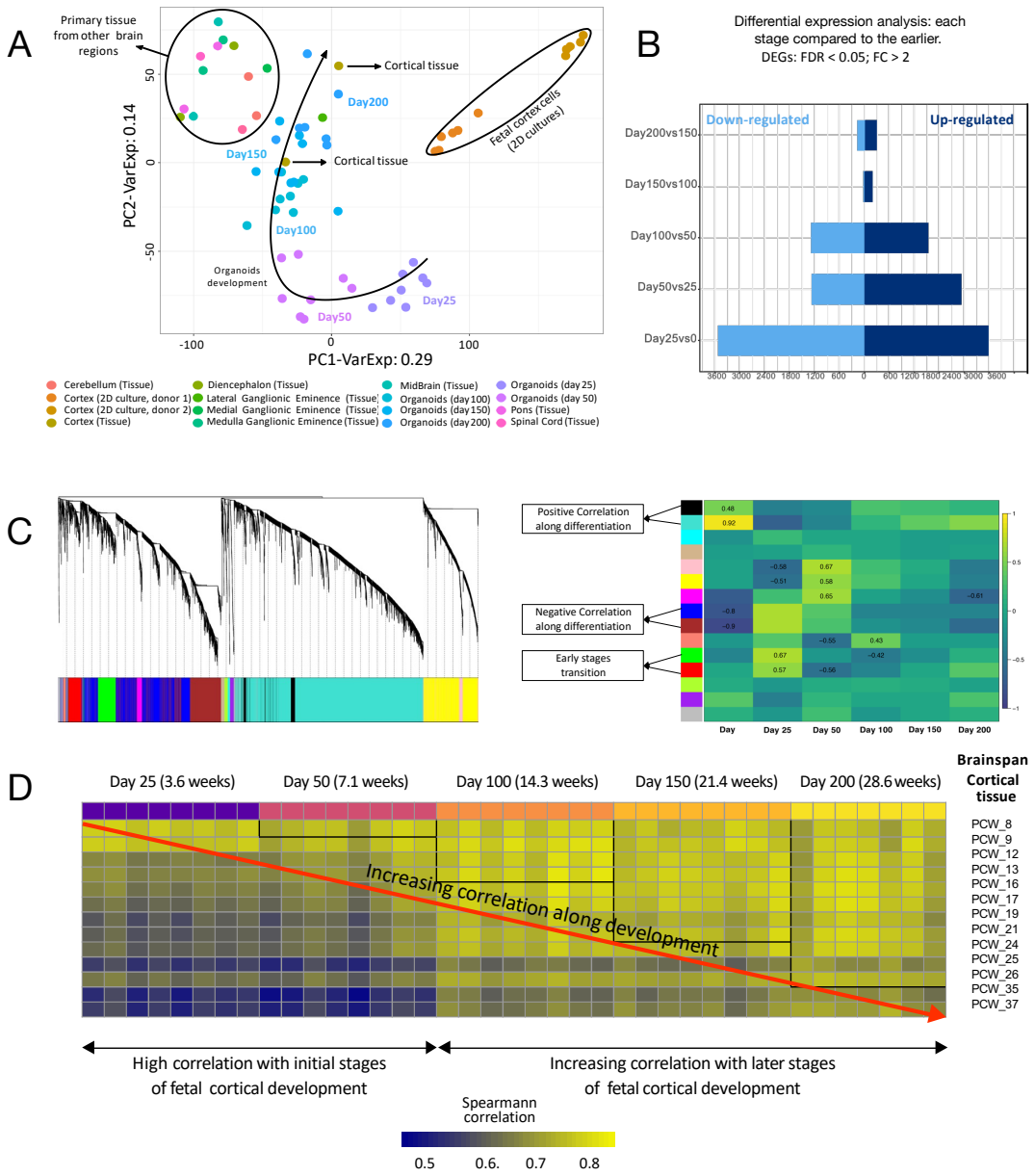


Figure 4 Transcriptomic characterization of cortical brain organoids across differentiation

A. Principal component analysis including transcriptomes from CO at day 25, 50, 100, 150 and 200 of differentiation, 2D cultured cortical tissue from PCW13 and 19 embryos and freshly isolated tissues from different brain regions at 13 and 19 PCW.
B. Number of DEGs observed performing differential expression analysis of CBOs at each stage against the previous.
C. Dendrogram of the WGCNA analysis of CO at all stages of differentiation highlighting the 14 modules identified. Heatmap showing the Spearman correlation analysis of the identified modules with developmental stage as a continuous variable or stage by stage.
D. Heatmap showing Spearman correlation of CO bulk transcriptomes at the indicated stages vs human fetal cortex transcriptomes from the BrainSpan dataset at the indicated postconceptional weeks (PCW).

Human cortical brain organoids can be used to dissect specific aspect of human corticogenesis

Having benchmarked our protocol to differentiate multiple cell lines into several stages of CO, we harnessed this model to investigate the role of the GSK3 pathway in the molecular landscape of human brain development. (López-Tobón, Villa, Cheroni, Trattaro, Caporale, et al., 2019)

The regulation of the proliferation and polarity of neural progenitors is crucial for the development of the brain cortex and even if animal studies have implicated glycogen synthase kinase 3 (GSK3) as a pivotal regulator of both proliferation and polarity, its functional relevance for the unique features of human corticogenesis remains to be elucidated. We thus chronically exposed human CO to a specific GSK3b inhibitor to probe its longitudinal impact through multiple developmental stages, and different layers of molecular and phenotypic profiles (figure 5).

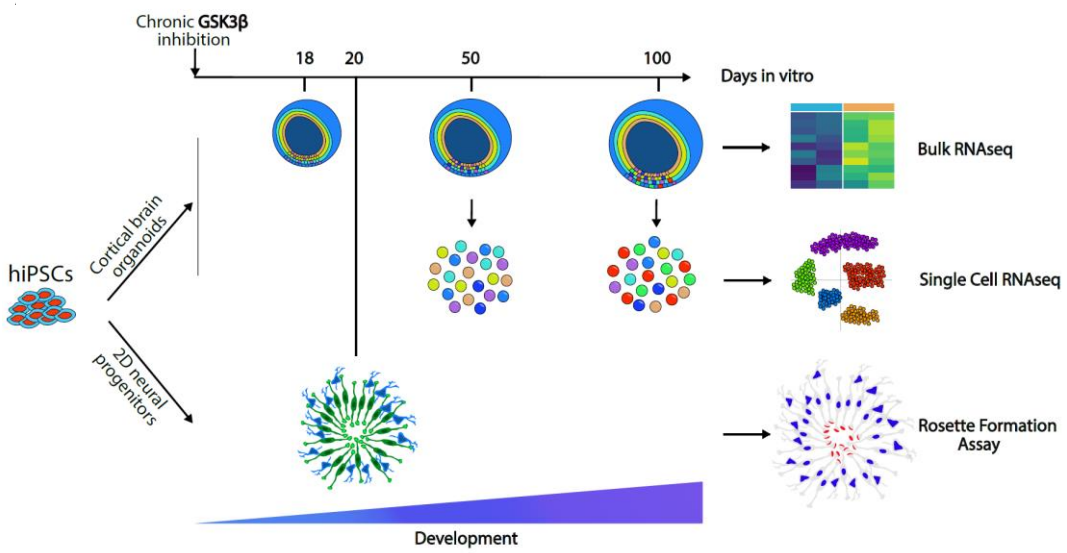


Figure 5 Experimental design

hPSCs differentiated following two parallel protocols, in 3D (up) cortical organoids or 2D (down) dual-smad inhibition. In both cases, parallel rounds were either exposed or not to GSK3 inhibitor CHI99021 (1 uM) starting from day 0 until indicated sample collection timepoint.

GSK3 is essential for cortical organoid morphogenesis

Chronic GSK3b inhibition resulted in an increase in organoid size concomitant with a virtually complete loss of the ventricular like structures (Figure 6), with a strong effect at day 50. This is consistent with the known role of the GSK3b receptor tyrosin kinase in regulating progenitor proliferation, in particular through its interactions with the Wnt and sonic hedgehog signaling pathways (W.-Y. Kim et al., 2009; McCubrey et al., 2016). Moreover GSK3 is involved in neurodevelopment through the phosphorylation of a broad set of substrates, including transcription factors essential for brain development, such as CREB (Grimes & Jope, 2001) neurogenin2 (Ma et al., 2008), β-catenin (Aberle, Bauer, Stappert, Kispert, & Kemler, 1997) and multiple microtubule-associated proteins (Izumi, Fumoto, Izumi, & Kikuchi, 2008). Indeed, studies in animal models have provided a wealth of evidence linking GSK3 activity to the regulation of early and late neurogenesis and to the maintenance of the overall polarity of the radial glia scaffold (Yokota et al., 2010). Unexpectedly, the difference we observed in organoid size and radial organization was accompanied no significant difference at day 50 in the proportion of PAX6+ cells. Moreover, staining of early neuron markers TBR1 and DCX revealed a profound disarray in tissue architecture (Figure 6) with only a slight reduction in TBR1+ cells, indicating that GSK3 activity is critical for the correct morphogenesis of the developing cortex while having a selective impact on the differentiation of specific cell subtypes.

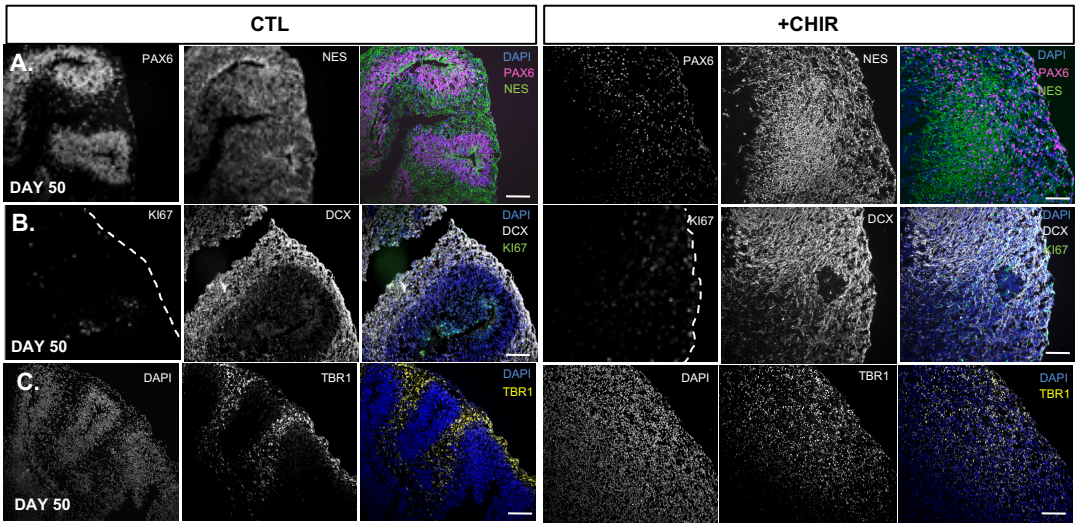


Figure 6 Morphogenetic alterations caused by chronic GSK3 inhibition

Representative images from day 50 organoids immunostained with: **A.** anti-PAX6 (red), anti-Nestin (green), DAPI (blue), **B.** anti-DCX (white), anti-KI67 (green), DAPI (blue), **C** anti-TBR1 (yellow), DAPI (blue), scale bar = 50 μm

Cortical organoids recapitulate the main features of mid-fetal human corticogenesis at single cell resolution

Given the impact of GSK3 in organoids structure, we harnessed single cell transcriptomics combined with distance-based analytical tools to break-down the effects of the inhibitor exposure in terms of population frequencies and developmental trajectories found during this time frame. We carried out droplet-based single-cell mRNA sequencing to profile over 30000 cells (N =33293) in 11 biological samples from unexposed and exposed CO at day 50 and day 100 of differentiation, attaining 15 unsupervised cell clusters defined using Louvain modularity algorithm (Šubelj & Bajec, 2014). By projecting the expression levels of canonical population markers over uniform manifold approximation and projection (UMAP) (Figure 7 A,B), combined with the overlap of cluster-specific genetic signatures from single-cell RNA-seq datasets of fetal human brain samples (Pollen et al., 2015) (Figure 7E) and considering their position over diffusion map (Figure 7C), we grouped the Louvain clusters into five main population identities, including proliferating radial glia (RG), Intermediate progenitors (IPC) (including a discernible subset of outer radial glia, oRG), early neurons (EN), Neurons (N) and Choroid (C). We then investigated the underlying developmental trajectories using diffusion map (Figure 7C-D). This confirmed the presence of a hierarchical progression from RGs to postmitotic neurons linked by intermediate populations (Figure 7C). The application of pseudotime, with an origin anchored in the early progenitors, reproduced the organization of populations from RG to N (Figure 7D). Together, these results confirm that patterned cortical organoids recapitulate cardinal features of human corticogenesis including the presence of early and intermediate progenitors, the emergence of oRGs and their hierarchical positioning in neurogenic trajectories.

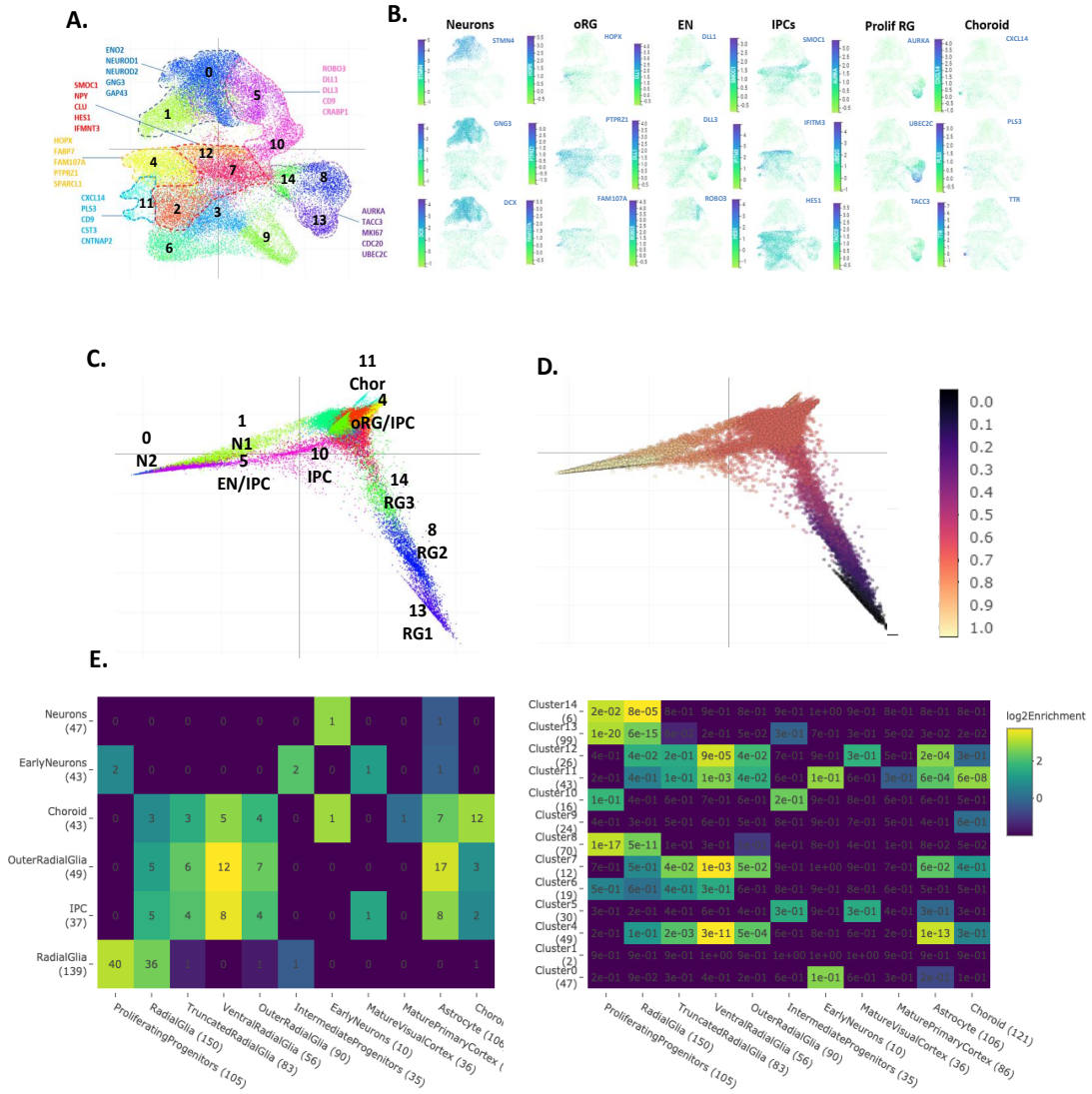


Figure 7 Cortical organoids recapitulate the main features of cortical development at the single cell transcriptomic level

A. Louvain clusters in UMAP plot colored by cluster identity; lines depict population areas defined by contrast with markers obtained from human fetal brain dataset (Radial glia: 8, 13, 14; Intermediate progenitor cells: 2, 7, 12; outer radial glia: 4; early neurons: 5, 10; Neurons: 0, 1; Choroid: 11).
B. UMAP plots. For each sub- panel, cells (represented as dots) are colored according to the expression levels of representative cell type markers (STMN4, GNG3, DCX: neurons; HOPX, PTPRZ1, FAM107A: outer radial glia; DLL1, DLL3, ROBO3: early neurons; SMOC1, IFITM3, HES1: intermediate progenitors; AURKA, UBEC2C, TACC3: proliferating progenitors; CXCL14, PLS3, TTR: choroid).
C-D. Diffusion map representing the developmental trajectory of the system. Cells (dots) are colored according to cluster identity (C) and to pseudotime trajectory (D), from origin in black to terminal state in light yellow according to wishbone algorithm.
E. Heatmap of the overlaps between marker genes characterizing the internal organoids clusters and external gene signatures of the relevant single cell clusters from human fetal brains (Nowakowski et al., 2017). Colors are based on the log2Enrichment computed for each pair of comparisons. * indicates an overlap with p-value < 0.05.

GSK3 inhibition differentially affects specific domains of corticogenesis

A comparative analysis of subpopulations revealed a selective impact of GSK3 inhibition on the relative proportion of specific cell subtypes (Figure 8 A-H).

A salient effect was the almost complete loss of a subpopulation characterized by the expression of genes expressed in choroid cells (CXCL14, TPD52L1, PCP4, EMX2) (Figure 8 A,B), as well as a noticeable decrease in frequency (FQ) of NEUROD2-expressing cells at day 100 in neuronal clusters (0,1,5) without affecting NEUROD6 or NEUROG2 expressing cells (Figure 8 C,D), underscoring the selectivity of GSK3 activity on these neurogenic pathways.

Noteworthy, analysis of the frequency and distribution of cells expressing a canonical oRG gene signature (HOPX, TNC, FAM107A, PTPRZ1) (Pollen et al., 2015) (Figure 8 E,F), showed a marked reduction in both frequency and expression levels upon CHIR-treatment at day 100, which was further confirmed by a reduction of HOPX+ cells in the tissue (Figure 8 I).

Given the cardinal role of oRG on neuronal production and cortical expansion, we checked the distribution of expression of telencephalic markers including forebrain determinant (FOXG1), early neurons (DCX), lower layer (BCL11B, RBFOX3) and upper layer (SATB2, POU3F2) markers. While upper layer markers were either not detected or present at very low frequencies, consistent with their surge at later stages of organoid corticogenesis, the expression of lower layer and general neuronal identity markers was unchanged by CHIR (Figure 8 H), whereas FOXG1 was significantly reduced at day 50 (Figure 8 G,H). Interestingly, despite the mild changes in frequency and expression levels of BCL11B at day 50 and day 100 (Figure 8 H), staining for its protein product (CTIP2) + cells showed a drastic reduction in d100 CHIR-treated organoids (Figure 8 I), pointing to regulatory mechanisms at the post-transcriptional level reported to affect CTIP2 activity and stability.

Finally, the application of pseudotime, independently to all conditions and stages, revealed two developmental trajectories, one that goes from RG to Neurons through a subset of IPCs and one that becomes apparent by day 100 and ends in an oRG identity (Figure 8 K). Strikingly, GSK3 inhibition resulted in a complete loss of the oRG-generating trajectory at day 100 (Figure 8 K). In agreement, HOPX distribution of expression on pseudotime peaked at day 100 in control organoids, while remained stalled in day 100 CHIR-treated organoids (Figure 8 K). Together, these results indicate that GSK3 inhibition results in a severe reduction of HOPX expressing cells with the attending impact on oRG- dependent lineages.

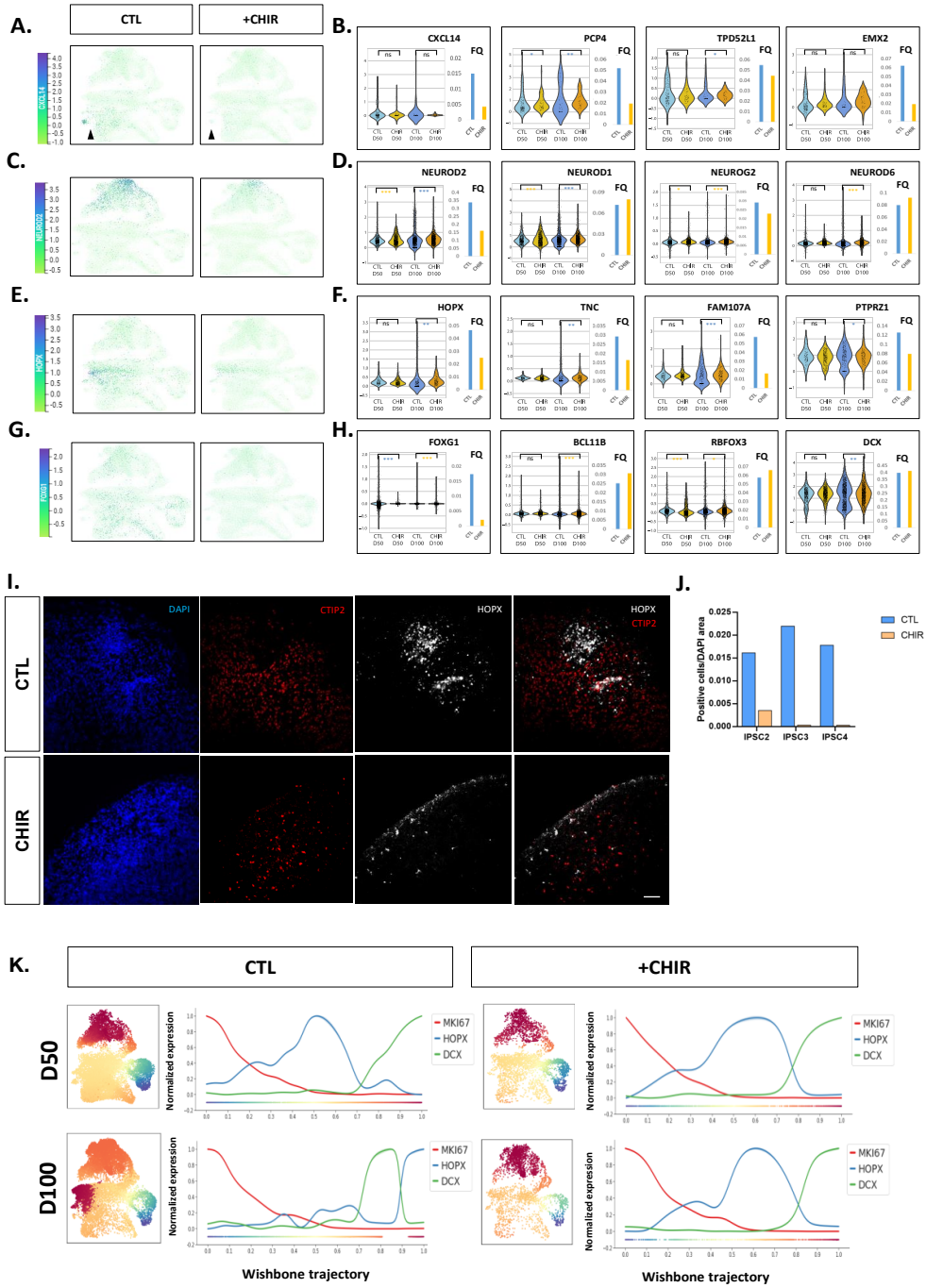


Figure 8 Effects of GSK3 inhibition at a single cell level in cortical organoids.

Visualization of normalized expression levels for genes identifying specific populations either in UMAP divided by CHIR-treated and control samples (A, C, E, G). Sub-sampled cluster-specific expression levels (violin plots), Fisher test, significance p-value *** = < 0.001, ** = < 0.01, * = < 0.05. FQ =normalized cell frequencies of cells expressing detectable levels/total population (bar plots) (B, D, F, H). The following genes are examined: (A-B) CXCL14 in UMAP and CXCL14, PCP4, TDP52L1, EMX2 in violin plots on cluster 11 for choroid population; (C-D) NEUROD2 in UMAP and NEUROD2, NEUROD1, NEUROG2, NEUROD6 on clusters 0, 1, 5 for neuronal markers; (E-F) HOPX in UMAP and HOPX, TNC, FAM107A, PTPRZ1 on cluster 4 for oRG markers; (G-H) FOXG1 in UMAP and FOXG1 (total) and BCL11B, RBFOX3 and DCX on clusters 0, 1, 5 for dorsal telencephalon markers. I. Representative widefield fluorescence images from day 100 organoids immunostained with anti-CTIP2 (red), anti-HOPX (grey), DAPI (blue), scale bar = 20 um. K. Over-imposition of pseudotime analyses performed separately for each experimental condition as color-scale on UMAP calculated on the complete system. Blue (origin); dark red (terminal state) according to wishbone trajectories. Visualization of the expression levels of representative genes along the condition-specific pseudotime.

From cohort to molecules: adverse impact of endocrine disruptors on neurodevelopment

In addition to the detailed elucidation of key aspects of human corticogenesis, the work previously presented demonstrates the possibility to use the CO model to investigate the effects of chemical exposure on brain development, a potential that we then used in a different context to study the impact of environmental compounds on neurodevelopment.

Human populations are exposed to a large number of chemicals with known or suspected endocrine disrupting properties (EDCs) (System & Others, 2014). This represents a major unmet challenge because while exposure to single EDCs has repeatedly been associated with major diseases and impaired development (Bergman Heindel, J.J, Jobling, S., Kidd, K.A., Zoeller, R.T., 2013), real life entails simultaneous exposure to multiple EDCs in mixtures, with additive effects at lower doses than experimental effect thresholds for single compounds (Kortenkamp, 2014). In addition, experimental evidence with mixtures is often limited to combinations within the same chemical class or to observational measurements on more complex mixtures, thus lacking causative weight to link actual population-based exposure with adverse health outcomes in humans. Here we pursued a systematic integration of epidemiological and experimental evidence to elucidate the molecular pathways affected by EDC mixtures that are causally related to adverse outcomes in humans.

An integrated epidemiological-experimental design assessing the impact of EDC mixtures on human health and development

To assess health outcomes of real-life EDC exposures we harnessed: i) the power of a population-based mother-child pregnancy cohort to measure prenatal EDC exposures, combined with novel biostatistical tools to infer associations between specific EDC mixtures and two child health domains: neurodevelopment and metabolism/growth (Figure 9); ii) complementary assays in human systems, to establish causality and deconvolute gene regulatory networks and *in vitro* cellular responses dysregulated by these EDC mixtures in concentrations corresponding to human exposure (Figure 9).

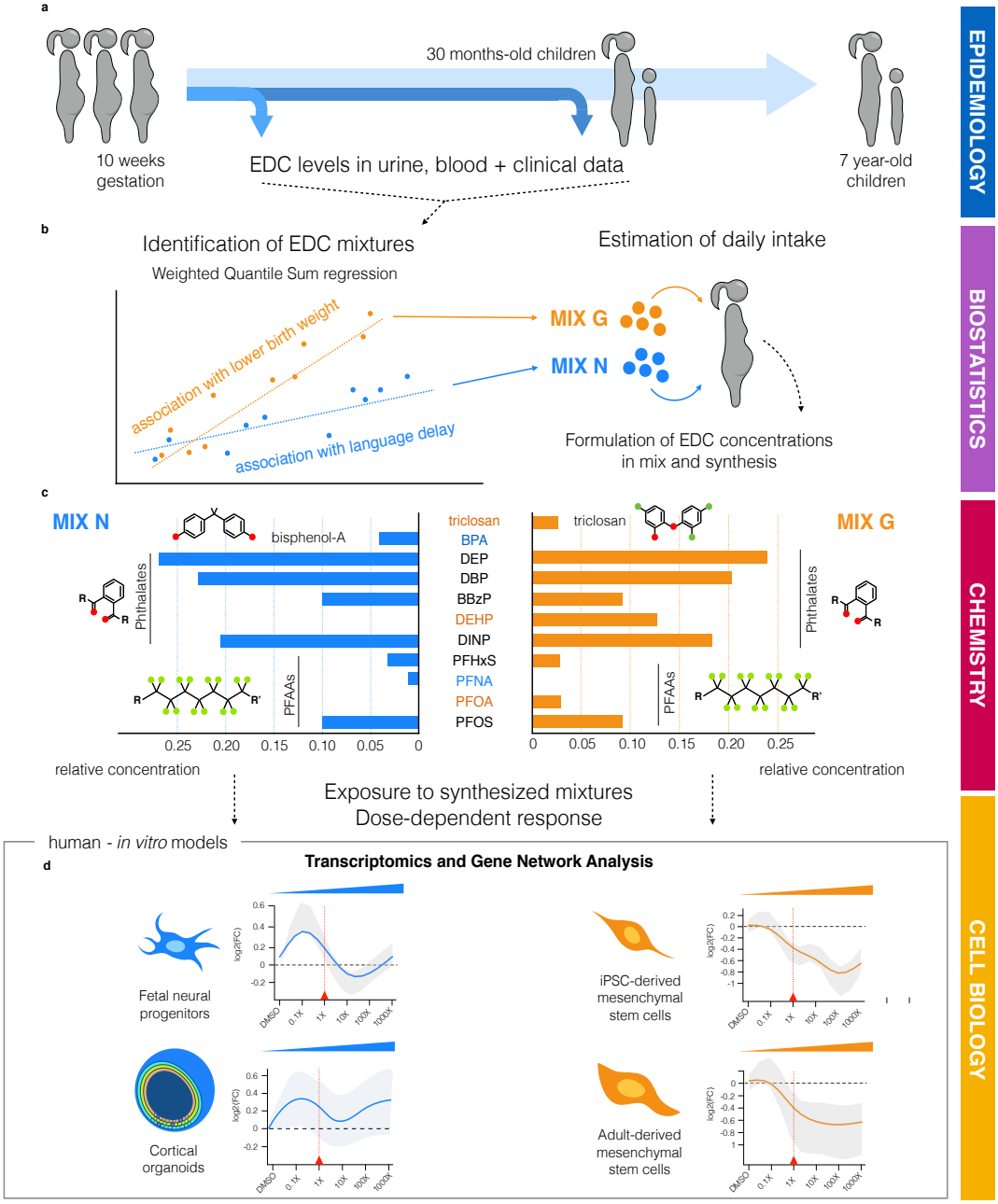


Figure 9 Overview of the study

A. Identification of two EDC mixtures that are associated with adverse health outcomes in two health domains, neurodevelopment and growth. In the SELMA pregnancy study, EDCs were measured in urine or serum of women around pregnancy week 10. B. Associations between these exposures and language delay of the children at age 2.5 years or birth weight were established using weighted quantile sum regression. This resulted in the identification of compounds that contributed to the association with the adverse health outcome (language delay or reduced birth weight) in the mixture. C. Based on their ratios found in the SELMA women's serum, the chemicals were mixed to compose MIX N (based on the association with language delay, blue coloured) and MIX G (based on the association with birth weight, orange coloured) for subsequent use in the experimental systems in concentrations corresponding to 0.1X, 1X, 10X, 100X, and 1000X serum concentrations in the SELMA mothers.
D. Identification of gene regulatory networks and cellular responses dysregulated by MIX N and MIX G, along with their dose-response relationships. Transcriptome analyses were carried out in Human Foetal Primary Neural Stem Cells and cortical organoids as well as in iPSC-derived or adult mesenchymal stem cells (MSCs) upon 48 h treatment with 0.1-1000X MIX N and MIX G, respectively. Significant transcriptional changes were detected already at 1X concentrations.

Definition and establishment of EDC-mixtures impacting human neurodevelopment and metabolism

Humans are exposed to several classes of EDCs including phthalates, phenols and perfluorinated alkyl acids (PFAAs) (System & Others, 2014). We focused on prenatal exposures to mixtures of 15 parent compounds (comprising 20 analytes/metabolites), measured in the Swedish Environmental Longitudinal, Mother and child, Asthma and allergy (SELMA) study. SELMA is a population-based pregnancy cohort that recruited more than 2,300 women in the first trimester from prenatal clinics in Värmland county, Sweden, from November 2007 to March 2010. Detailed recruitment and sample collection procedures have been described previously (C. G. Bornehag & Moniruzzaman - Paediatric and …, 2012). The chemicals compounds was selected because of their known or suspected endocrine disrupting properties. However, the selection of chemicals can be seen as a paradigm for how to evaluate mixture exposures from a health and developmental point of view rather than including all potential health relevant EDCs (Bergman Heindel, J.J, Jobling, S., Kidd, K.A., Zoeller, R.T., 2013).

We identified mixtures of EDCs in prenatal urine and serum in SELMA that were associated with two major child health outcomes: neurodevelopment, measured as language delay at 30 months of age and metabolism/growth, measured as birth weight. Delays in language development in early childhood as assessed by validated tests have been shown to be predictive of later academic achievement and the need for special education (Miniscalco, Nygren, Hagberg, Kadesjö, & Gillberg, 2006). Therefore, language delay, assessed in early childhood, is an important indicator of later neurodevelopmental impairment. Birth weight is an early marker for in utero growth and a low birth weight is associated with metabolic syndrome including obesity and glucose intolerance (Kopec, Shekhawat, & Mhanna, 2017) as well as other health issues during childhood.

The current analyses of SELMA cohort includes 1874 pregnant women (Table 2), assessed for their urinary or serum EDC levels around the 10^{th} week of gestation (Table 3). Specifically, we profiled urine levels of 10 metabolites of 5 phthalates, bisphenol A (BPA) and triclosan (TCS), as well as serum levels of 8 perfluorinated alkyl acids (PFAAs).

Specific mixtures (in terms of both composition and dose) were associated with the two health outcomes in a three-step procedure (Figure 9). Firstly, we identified the prenatal exposure to EDCs that was associated with lower birth weight or language

delay in children by using weighted quantile sum (WQS) regression (Carrico, Gennings, Wheeler, & Factor-Litvak, 2015). EDCs included in these exposures are hereafter referred to as sEDCs ('selected EDCs based on WQS regression weights). Secondly, we estimated the equivalent daily intake (DI) of sEDCs measured in the urine (i.e., phthalates and alkyl phenols), and estimated serum concentrations from the DI for these urinary measurement-based compounds. Thirdly and finally, we used the geometric means, on a molar basis, for either the measured or estimated serum levels of all compounds and established mixing proportions to prepare, for experimental validation, the two mixtures associated to language delay (MIX N) and lower birth weight (MIX G) (Figure 9). Mixtures were tested across concentrations (0.1X, 1X, 10X, 100X, 1000X) corresponding to human exposure (Table 3), where 1X denotes the geometric mean of exposure levels in SELMA pregnant women.

Table 2 Description of the study population including 1,874 pregnant women and their children in the SELMA study

Pregnant women		**Children**	
Age at enrollment (year); mean (SD)	31.0 (4.8)	**Gestational age at birth (week); mean (SD)**	39.3 (1.8)
Weight at enrollment (kg); mean (SD)	69.6 (13.6)	**Birth length (cm); mean (SD)**	51.2 (2.6)
Gestational age for biosampling (week); mean (SD)	9.9 (2.1)	**Sex (girls)**	47.2%
Smoking during pregnancy (yes)	4.7%	**APGAR score (<10)**	12.7%
Parity (null parity)	44.8%	**Birth weight (g); mean (SD)**	3.619 (584)
Education (university or higher)	61.4%	**Language delay (<50 words)**	10.0%

Table 3 Distribution of phthalate and phenol metabolites in urine and perfluorinated compounds (PFAS) in serum analysed in 1st trimester of 1,874 pregnant women in the SELMA study

Compound	**Metabolite**	**Phthalate and phenol metabolites in urine (ng/mL)**		
		Median	**95%**	**GM (95% CI)**
DEP	MEP	62.6	507.7	68.7 (65.3-72.3)
DBP	MBP	71.9	233.1	69.0 (66.5-71.5)
BBzP	MBzP	16.8	99.4	16.6 (15.8-17.4)
DEHP	MEHP	3.8	15.6	3.8 (3.6-3.9)
	MEHHP	16.6	66.6	16.3 (15.7-17.0)
	MEOHP	11.2	45.0	11.1 (10.7-11.6)
	MECPP	15.7	62.7	15.8 (15.2-16.4)
DiNP	MHiNP	5.9	54.6	6.2 (5.9-6.6)
	MOiNP	2.7	19.2	2.9 (2.8-3.0)
	MCiOP	8.7	74.9	9.8 (9.3-10.2)
BPA		1.5	6.2	1.5 (1.4-1.6)
Triclosan		0.8	351.4	1.3 (1.2-1.5)
Compound		**Perfluorinated compounds (PFAS) in serum (ng/mL)**		
		Median	**95%**	**GM (95% CI)**
PFOA		1.60	3.96	1.60 (1.56-1.64)
PFOS		5.35	12.29	5.30 (5.18-5.43)
PFNA		0.53	1.29	0.54 (0.53-0.55)
PFDA		0.25	0.59	0.26 (0.25-0.27)
PFUnDA		0.23	0.54	0.21 (0.21-0.22)
PFDoDA		0.03	0.08	0.03 (0.03-0.03)
PFHxS		1.23	3.71	1.32 (1.29-1.36)
PFHpA		0.02	0.09	0.02 (0.02-0.02)

MIXN disrupts human neurodevelopmental pathways

To define the molecular impact of MIX N, we first employed primary neural stem cells sourced from the cortex of human foetuses at post-conception week (PCW) 11 and 19 (henceforth Human Foetal Primary Neural Stem Cells (HFPNSC)). Given the potentially non-linear and non-monotonic dose-response patterns associated with EDC mixtures (Beausoleil et al., 2016), the experimental design included 5 doses of MIX N, ranging from 0.1X to 1000X and a global assessment of impact on gene expression. To this end, RNA-seq was performed after 48h MIX N exposure and patterns of EDC dose-dependent transcriptional responses determined using an analysis that considers MIX N dilutions (including the DMSO control) as distinct categories. This unbiased approach, which does not assume any particular response pattern (e.g., linearity or monotony), allowed us to define a core list of differentially-expressed genes (DEGs), which follow similar dose-responses in the two fetal cortical lines (figure 10 A). This core set includes the genes EPHB2 and CLSTN2, which are of particular relevance for the original phenotype scored in the SELMA cohort (ie. language delay), insofar as they are downregulated by acute MIXN exposure (Figure 10 B) and belong to the Simons Foundation Autism Research Initiative (SFARI) list of genes causing ASD through a loss of function mutation.

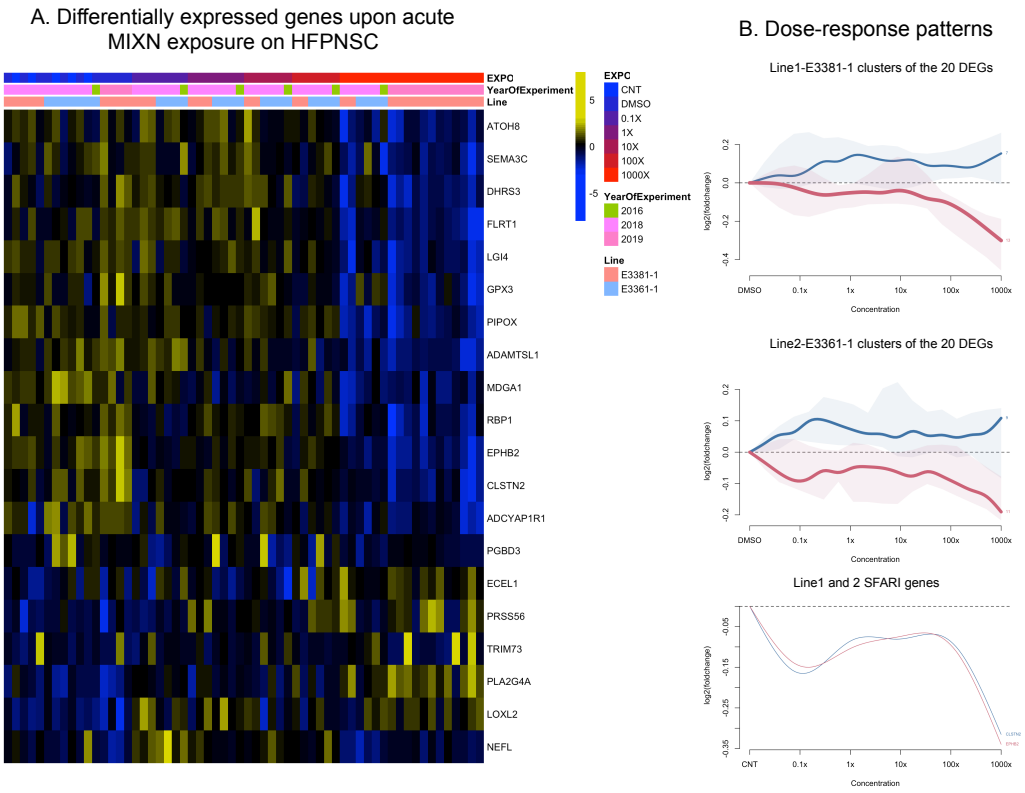


Figure 10 Transcriptional effect of acute MIXN exposure on HFPNSC

A. Heatmap of the differentially expressed genes (DEGs) identified through categorical analysis
B. DEGs were clustered into major dose-response patterns, mean smoothed fold change were plotted across the different MIX N dilutions

MIXG disrupts human metabolic pathways

Following the same logic applied to MIX N, we evaluated the molecular impact of MIX G on two growth/metabolism-relevant human models, bone marrow-derived mesenchymal stem cells (adult MSCs) and iPSC-derived mesenchymal stem cells (iPSC-MSC) which display comparable molecular hallmarks, allowing validation across experimental systems. Both MSC models were exposed to increasing concentrations of MIX G for 48h and profiled by RNA-Seq. Differential expression analysis showed dose dependent increases or decreases of gene expression upon MIX G exposure starting already at the 1X concentration and consistent across lines (Figure 11 A). Moreover we found that 6 of those DEGs are also associated to low birth weight in GWAS (Figure 11 B).

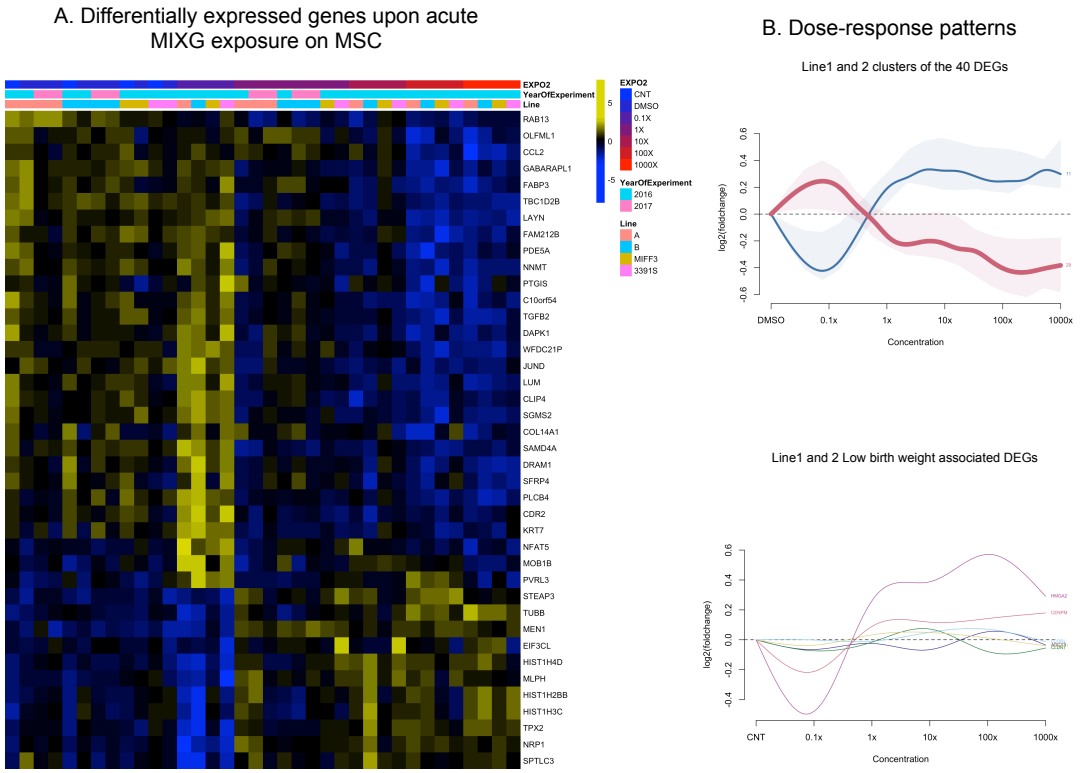


Figure 11 Transcriptional effect of acute MIXG exposure on Mesenchymal Stem Cells

A. Heatmap of the differentially expressed genes (DEGs) identified through categorical analysis
B. DEGs were clustered into major dose-response patterns, mean smoothed fold change were plotted across the different MIX N dilutions

Dissecting the impact of MIX N versus MIX G on human development

Despite having similar chemical compositions, MIX N and MIX G had been linked to different health outcomes on the basis of epidemiological evidence. We thus sought the molecular basis of this distinction by cross-exposure of representative models to the alternative mixture. HFPNSC were exposed to MIX G, using the same five dilutions as for MIX N. Strikingly the mixtures showed marked differences in the affected genes, in particular with respect to ASD and ID associated targets that were only significantly enriched in MIX N DEGs (figure 12 A), consistent with the association of MIX N exposure to early verbal skills.

Finally, to evaluate the specificity of MIX N and MIX G on cellular responses, we compared their effects on lipid droplet accumulation in adult MSCs. Whereas adult MSCs exposed to MIX G showed significant increase in lipid accumulation already at 1X (figure 12 B), MIX N showed significant increase only at 100X and only in the male line (figure 12 B). Together, these results provide experimental evidence of the mixture-to-phenotype dissection that had been originally only inferred at the population level, establishing the power of such integrated approaches for defining the molecular traces of EDC exposure across the population, organismal and cellular scales.

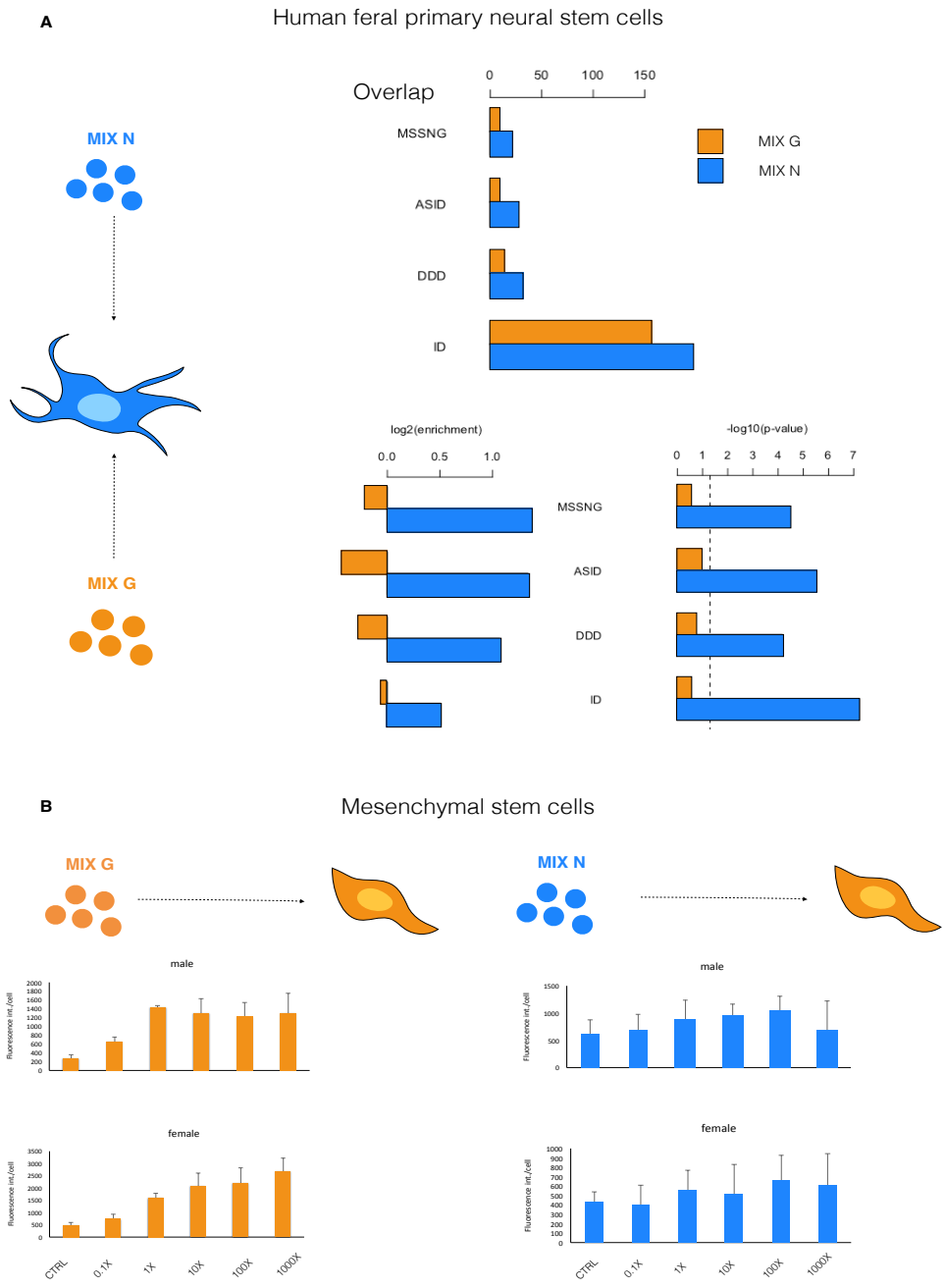


Figure 12 Dissection of MIXN vs MIXG

A. Overlaps, enrichment and significance between neurodevelopmental disorder relevant genes and, respectively, MIX N- and MIX G-associated DEGs in HFPNSC.
B. For adult MSCs, lipid droplet accumulation was quantified using Bodipy 493/503 staining upon treatment with the indicated concentrations of MIXN and MIXG for 3 weeks. Values are normalised to nuclei count and representative of three independent experiments for each of the 2 donors, shown as mean ± SD from 3 to 6 replicates of a single experiment.

Chronic expo on Human Fetal Primary Neural Stem Cells

While most proximally representing the target tissue of real-life EDC exposure (ie-developing fetal brain tissue), HFPNSC are inherently limited, as a non self-renewing experimental system, by the inherent features, technical and ethical alike, of their procurement. Hence, the observation that even an acute exposure to MIXN was able to trigger convergent dysregulation in a core set of phenotypically-relevant genes in two genetically independent fetal samples, sourced at different developmental stages, gave us the confidence to investigate the molecular impact of MIXN exposure in a setting that more closely approached human in vivo exposure. Specifically, we selected the two most relevant doses, 1X and 1000X and exposed cortical HFPNSC over two weeks to mimic the continuous exposure that occurs in vivo, profiling transcriptomes in 4 replicates of the cortical line already profiled in the acute setting.

Differential expression analysis revealed a major impact of MIXN exposure, mainly concordant between the two doses, specific for the mix, as compared to DMSO, and confirm the significant dysregulation of 3 of the genes, namely DHRS3, LGI4 and GPX3, previously identified in the acute setting. As for MIXN specific effects, the analysis revealed 151 DEGs (FDR < 0.05 and logFC > 0.5) significantly modulated by at least one of the 2 concentrations of MIXN exposure that, as shown both in the heatmaps and the scatterplot follow the same directions of dysregulation (Figure 13 A,B). Gene ontology categories enriched upon 1X MIXN exposure, are related to chromatin regulation and cell proliferation among the upregulated genes, while metabolic processes and response to hormonal signaling are enriched among the downregulated genes, in particular GPX3 (glutathione peroxidase), FXYD (related to the regulation of ion transport channels) are relevant down-regulated genes while DSCAML1 (adhesion molecules involved in neuronal differentiation and also associated to Down syndrome) is significantly up-regulated. Moreover, gene set enrichment analysis (GSEA) revealed an enrichment among the upregulated genes for neurodegenerative diseases categories and oxidative phosphorylation (Figure 13 C,D). Furthermore, given our epidemiological evidence associating MIXN exposure to language delay, and the central role of chromatin dysfunction in autism and intellectual disabilities we tested whether MIXN-induced DEGs were enriched for genes associated with these conditions. We found significant enrichment for genes associated with Intellectual Disabilities, Developmental Disorders and Autism Spectrum Disorders (ASD) (Figure 13 E).

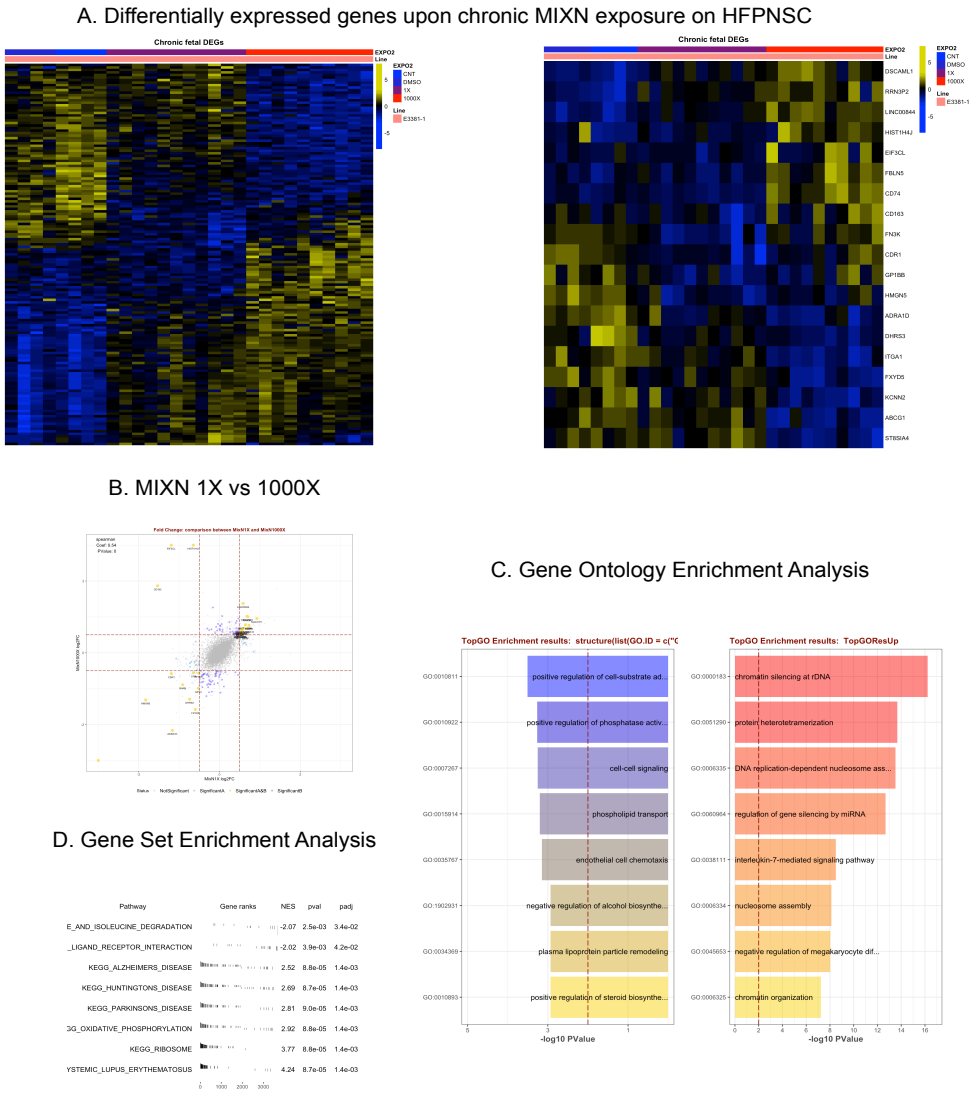


Figure 13 Transcriptional effect of chronic MIXN exposure on HFPNSC

A. Heatmap of the differentially expressed genes (DEGs) identified through categorical analysis, (FDR < 0.05 and logFC > 0.5) in the left panel, (FDR < 0.05 and logFC > 1) in the right panel
B. Scatterplots representing the fold-change induced by 1X vs 1000X exposure relative to DMSO. DEGs shared between the 2 conditions are reported in yellow, while DEGs specific of 1 of the 2 in blue.
C. Results of gene ontology enrichment analysis performed on the DEGs, divided for up- regulated and down-regulated genes. Bar plots depict the p-values for the top-10 Cellular Component GO terms.
D. Gene sets significantly associated with MIXN exposure by GSEA.
E. Overlaps, enrichment and significance between DEGs and genes associated to intellectual disability, developmental disorders and autism spectrum disorders in published databases. Gene expression across exposure conditions for the DEGs that overlap with the list of SFARI genes.

Human cortical brain organoids for developmental neurotoxicology

While primary neural stem cells directly sourced from fetal telencephali represent the arguably most proximal model of human neurodevelopment, their availability is limited and hence they are ill suited for large-scale and iterative studies required to advance regulatory toxicology. We thus explored the transcriptomic impact of MIX N using a neurodevelopmental model based on self-renewing sources of human induced pluripotent stem cells (iPSC): 3D cortical brain organoids (henceforth CO) that recapitulate human in vivo corticogenesis, as previously shown by us (López-Tobón et al., 2019). We thus followed the same experimental design used for fetal samples, having this time the opportunity to mimic even better the real life scenario, exposing in 4 replicates, cortical brain organoids from day 0 to day 50 in a chronic exposure setting that simulates the physiological conditions of early brain development during the first 10 weeks of pregnancy, when the women of the SELMA cohort were profiled. The results of the transcriptome analysis confirmed the major impact on gene expression already at 1X concentration, mainly concordant between the 2 doses of MIXN and a specific effect of MIXN exposure as compared to all the other conditions (Figure 14 A,B).

As for MIXN effects, we found 662 DEGs (FDR < 0.05 logFC>=0.5, fig 5 C). Interestingly we found that MIXN significantly repress SLC5A7, a sodium and chloride transporter involved in synaptic formation, while it up-regulates genes related to neuronal fate NEUROG1, FEZF2 and GAD2 (which is consistent also in fetal). There is an enrichment for biological processes related to detection of chemical stimulus and cell-matrix adhesion among the downregulated genes, while negative regulation of neuron differentiation and brain development are enriched among the up-regulated genes (Figure 14 C,D).

Furthermore, following the same analysis we did for the fetal data. we found significant enrichment for genes associated with Intellectual Disabilities, Developmental Disorders and Autism Spectrum Disorders (ASD) (figure 14 E). In particular, plotting the expression of the most relevant genes for ASD, among MIXN DEGS, from the SFARI database we observed that the mutations causing ASD for a loss of function, are also down-regulated in our datasets (Figure 14 E).

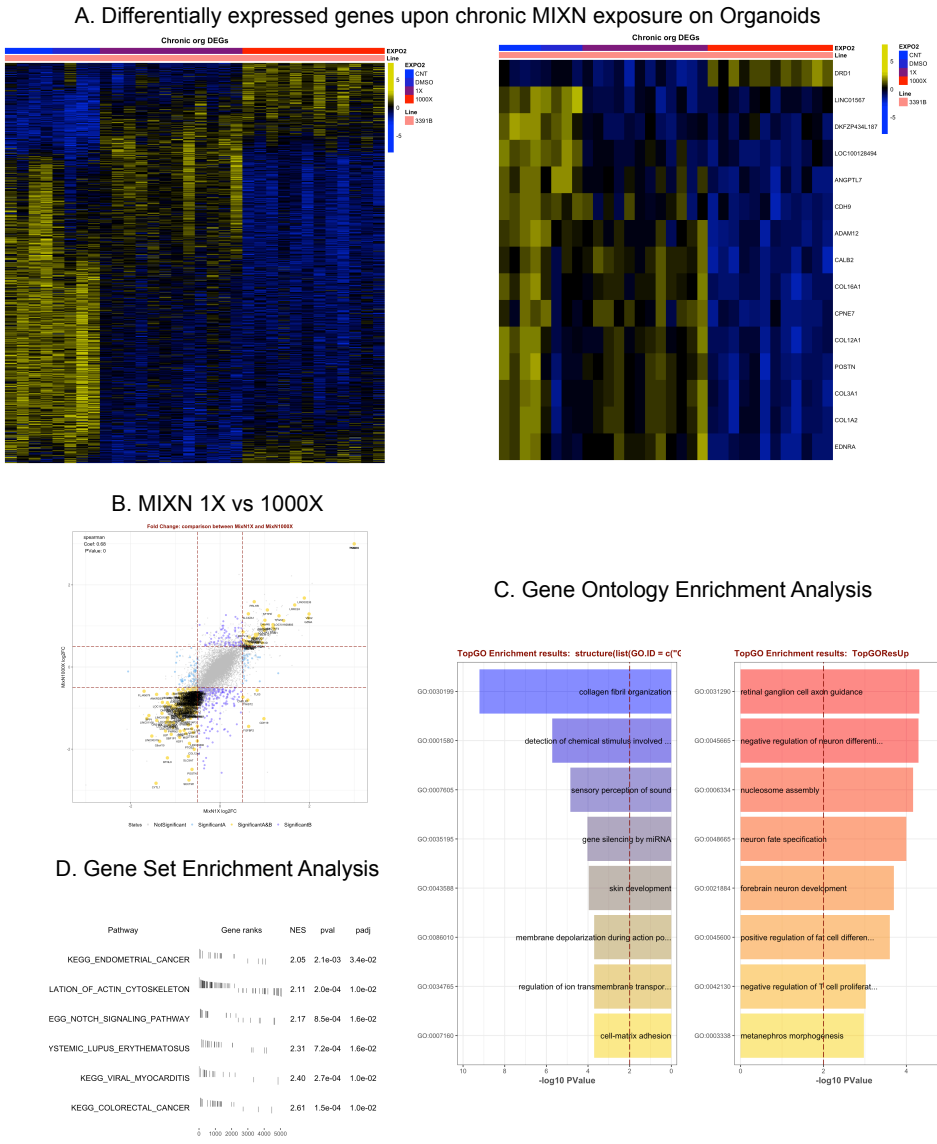


Figure 14 Transcriptional effect of chronic MIXN exposure on cortical organoids

A. Heatmap of the differentially expressed genes (DEGs) identified through categorical analysis, (FDR < 0.05 and logFC > 0.5) in the left panel, (FDR < 0.001 and logFC > 1) in the right panel
B. Scatterplots representing the fold-change induced by 1X vs 1000X exposure relative to DMSO. DEGs shared between the 2 conditions are reported in yellow, while DEGs specific of 1 of the 2 in blue.
C. Results of gene ontology enrichment analysis performed on the DEGs, divided for up- regulated and down-regulated genes. Bar plots depict the p-values for the top-10 Cellular Component GO terms.
D. Gene sets significantly associated with MIXN exposure by GSEA.
E. Overlaps, enrichment and significance between DEGs and genes associated to intellectual disability, developmental disorders and autism spectrum disorders in published databases. Gene expression across exposure conditions for the DEGs that overlap with the list of score1-2 SFARI genes.

Specificity of Mixture effects and comparison with single compound exposure

Conscious of the complexity of dealing with the molecular events triggered by endocrine disruption and in addition to that, by mixtures of EDCs, our experimental design included a panel of additional exposures, performed on the same human in vitro systems, used as controls to ensure the specificity of the results observed.

To further investigate the transcriptional effects specific of MIXN, in addition to 0.1% DMSO as a negative control, 10μM T3 exposure was used as a modulator of thyroid hormone responsive genes, given its essential roles in both brain development (Williams, 2008) and metabolism (Mullur, Liu, & Brent, 2014) and previous epidemiological (Morgenstern et al., 2017) and experimental (Fini et al., 2017) evidence implicating the main chemical classes present in both mixtures as TH disruptors. We thus exposed the human cellular models to a concentration that is significantly higher than the physiological dose of T3 in the developing human brain (Calvo, Roda, Obregón, & de Escobar, 1998; Kester et al., 2004). In addition, given the complexity of MIXN and MIX G composition, and the fact that BPA and TCS are two of the few compounds that differ between MixN and MixG, these chemicals were also used for treatment as a single compounds, at the same concentration in which they are present in 1X MIXN or 1X MIX G. Furthermore, we tested in the same experiment different batches of MIXN, synthesised in different time points and stored in different labs for different periods, and performed a batch correction so to avoid technical variability. Figure 15 A,B shows that, both for HFPNSC and CO the genes identified are specifically dysregulated by MIXN exposure and that BPA alone cannot explain the transcriptional effect of MIXN. Figure 15 C shows that triclosan alone cannot explain the transcriptional effect of MixG.

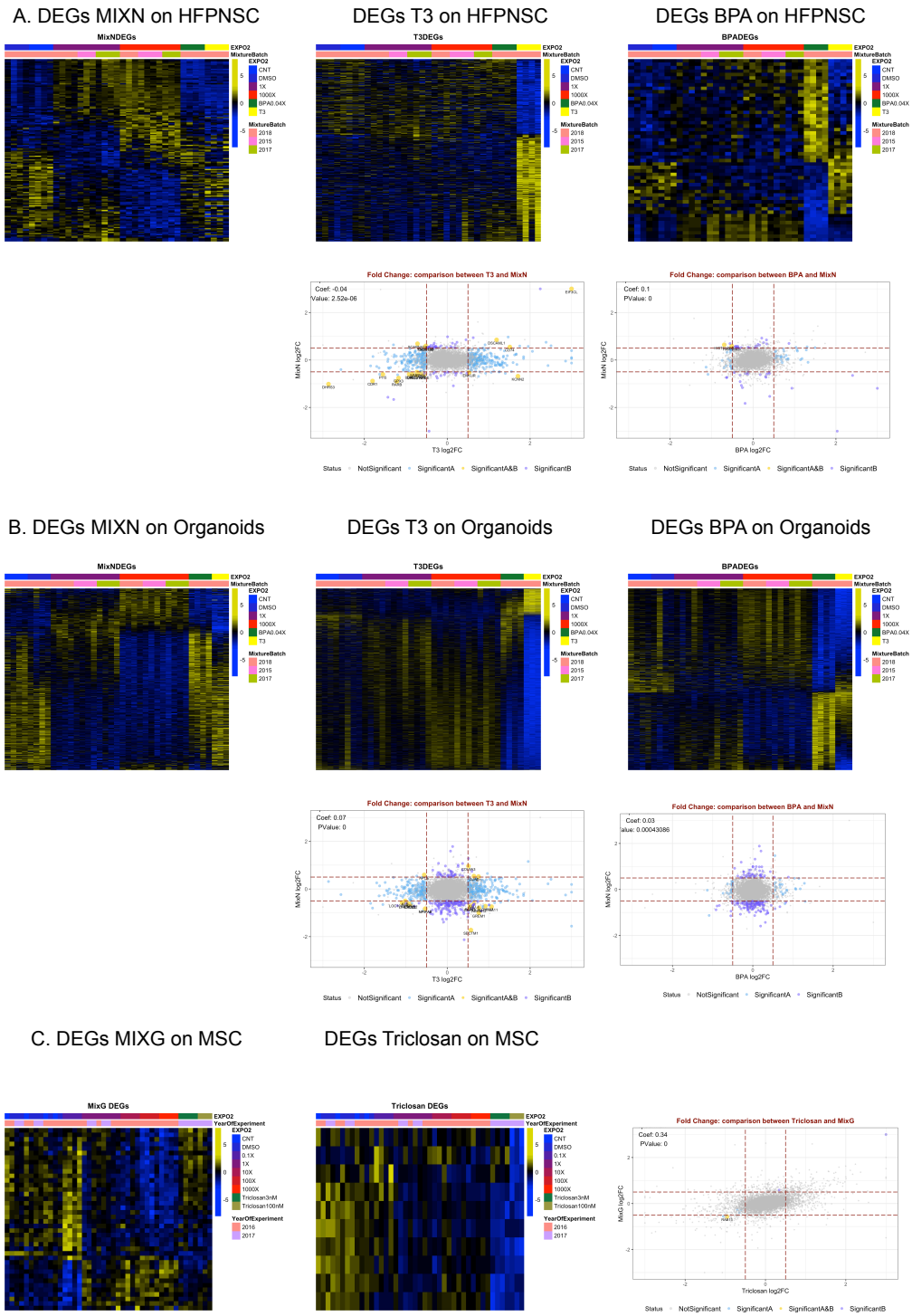


Figure 15 Dissection of mixture vs single compounds effects

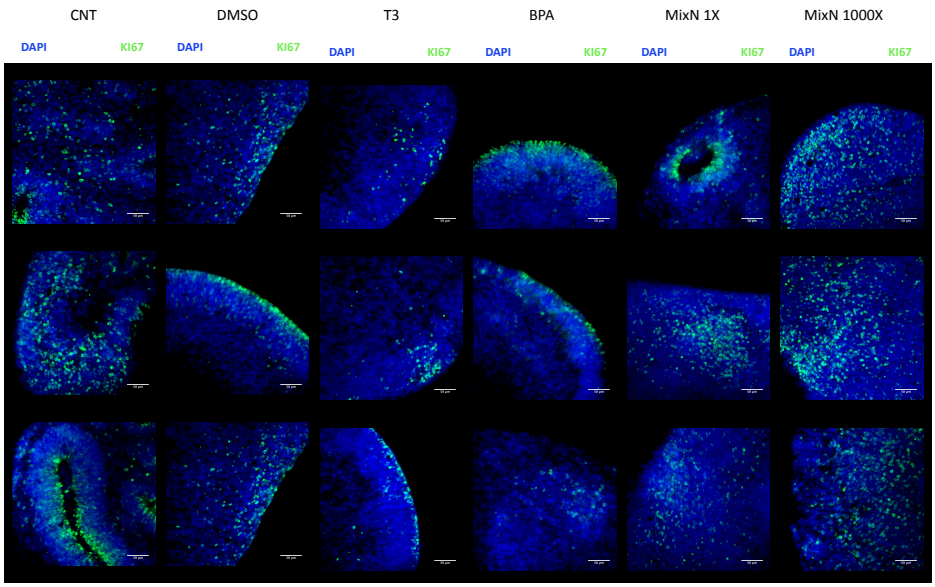
Heatmap of the differentially expressed genes (DEGs) identified through categorical analysis upon exposure to 1X MIXN, T3, BPA or Triclosan. Scatterplots representing the fold-change induced by 2 exposure conditions compared between each other, relative to DMSO. DEGs shared between the 2 conditions are reported in yellow, while DEGs specific of 1 of the 2 in blue. Results for the HFPNSC dataset in A, cortical organoids in B, mesenchymal stem cells in C.

Cellular phenotypes induced by EDC mixtures

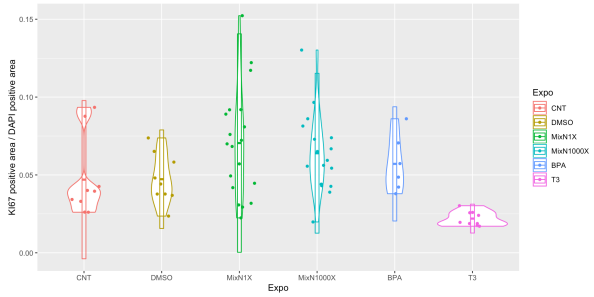
In addition to the molecular dissection of MIX N effects, the same organoids that were profiled at the transcriptomic level after the chronic 50 days of exposure, were also used to perform immunofluorescence analysis in order to score the cellular phenotypes and understand if MIX N had an impact on the developmental processes of neuronal differentiation. In particular we stained all the conditions for KI67 that mark proliferating progenitors and DCX, a protein involved in neuronal migration and represents a marker of early neurons. The results show an increased number of KI67 positive cells coupled with a decrease of DCX expressing cells in the MIX N exposed organoids, thus suggesting an effect of EDC that is favoring neural progenitors proliferation while hindering neuronal differentiation (Figure 16,17 A, B)

To further dissect this aspect, we interrogated in a supervised approach the transcriptomics data for a larger subset of genes related to proliferation (MKI67, CCNB1, CCNB2, CDC20, CDC20B, CDCA8, and HMGB2) and neural differentiation (DCX, SATB2, NEUROG1, SYP, MAP2, RBFOX3, and L1CAM), confirming the results observed at the protein level, also at the level of gene expression (Figure 16,17 C).

A. Immunofluorescence of organoids: cell proliferation



B. Immunofluorescence quantification (3 different organoids per condition)



C. Cell proliferation genes in the transcriptomic dataset

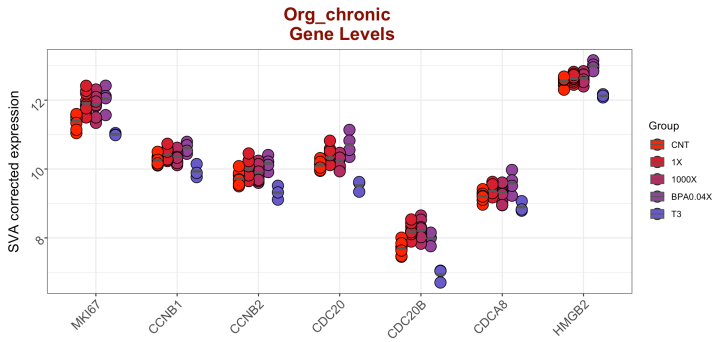
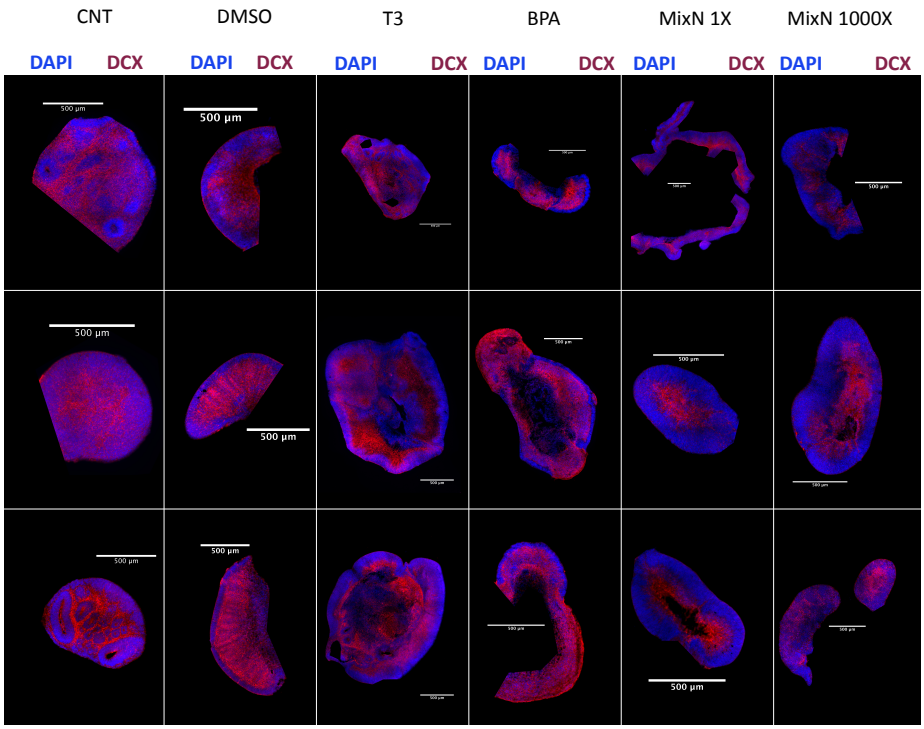


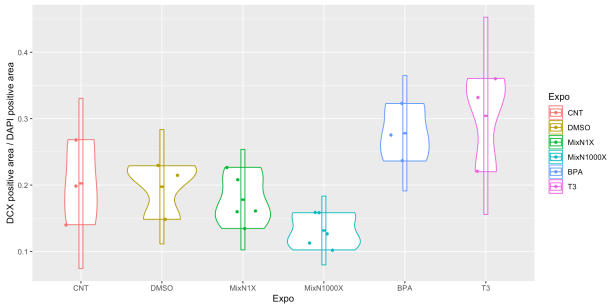
Figure 16 Cell proliferation effects of chronic MIXN exposure on cortical organoids

A. Representative captions from immunostaining performed for anti-KI67(green)/DAPI(blue), widefield fluorescence images acquisition 40X.
B. Violin plot of the ratio between the positive areas for KI67, normalised on the DAPI positive areas across exposure conditions. Areas were measured on several fields per organoid for at least 3 different organoids per condition.
C. Gene expression across exposure conditions for a selected set of genes regulating cell proliferation.
.

A. Immunofluorescence of organoids: neuronal differentiation



B. Immunofluorescence quantification (3 different organoids per condition)



C. Neuronal maturation genes in the transcriptomic dataset

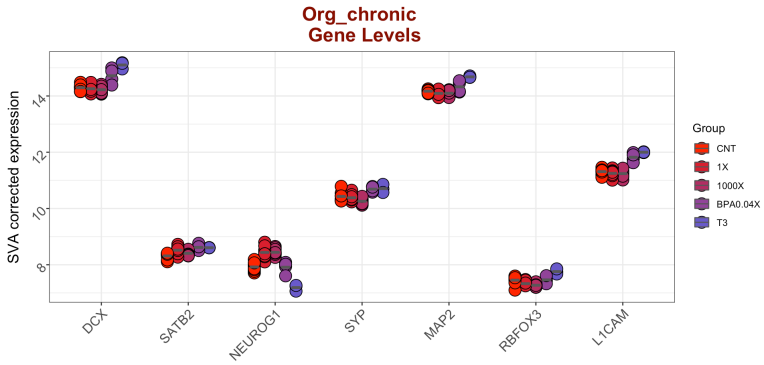


Figure 17 Neuronal maturation effects of chronic MIXN exposure on cortical organoids

A. Representative captions from immunostaining performed for anti-DCX(red)/DAPI(blue), widefield fluorescence images acquisition 20X.
B. Violin plot of the ratio between the positive areas for DCX, normalised on the DAPI positive areas across exposure conditions. Areas were measured on several fields per organoid for at least 3 different organoids per condition.
C. Gene expression across exposure conditions for a selected set of genes regulating neuronal maturation.

Transcription factors identified as master regulators of the EDC mixture-dependent transcriptional effect

As a next step, we reasoned that the best way to identify the key regulators responsible of the transcriptomic phenotype and, as a consequence of the cellular and clinical adverse effects, was to perform a master regulator analysis starting from the dysregulation observed in the chronic exposure setting. Our analytical approach for this purpose takes advantage of 2 key novelties represented by the opportunity to leverage the recently released data from the Psychencode consortium (D. Wang et al., 2018) to have better information about the regulatory activity of transcription factors, specifically in the neurodevelopmental context, and by the use of a curated regulon (see methods and online resources). Using this approach we identified key transcription factors, whose dysregulation relayed the transcriptional impact of MIX N. In the fetal dataset we identified transcription factors known to be fundamental for regulating neuronal development and implicated in neurodevelopmental disorders, such as FOXG1 (Vegas et al., 2018) and CUX1 (Platzer et al., 2018), as well as in organoids where we identified ARID2 (Shang et al., 2015) and NFKB1 (Yonggang Zhang & Hu, 2012) at the basis of the transcriptional dysregulation (Figure 18).

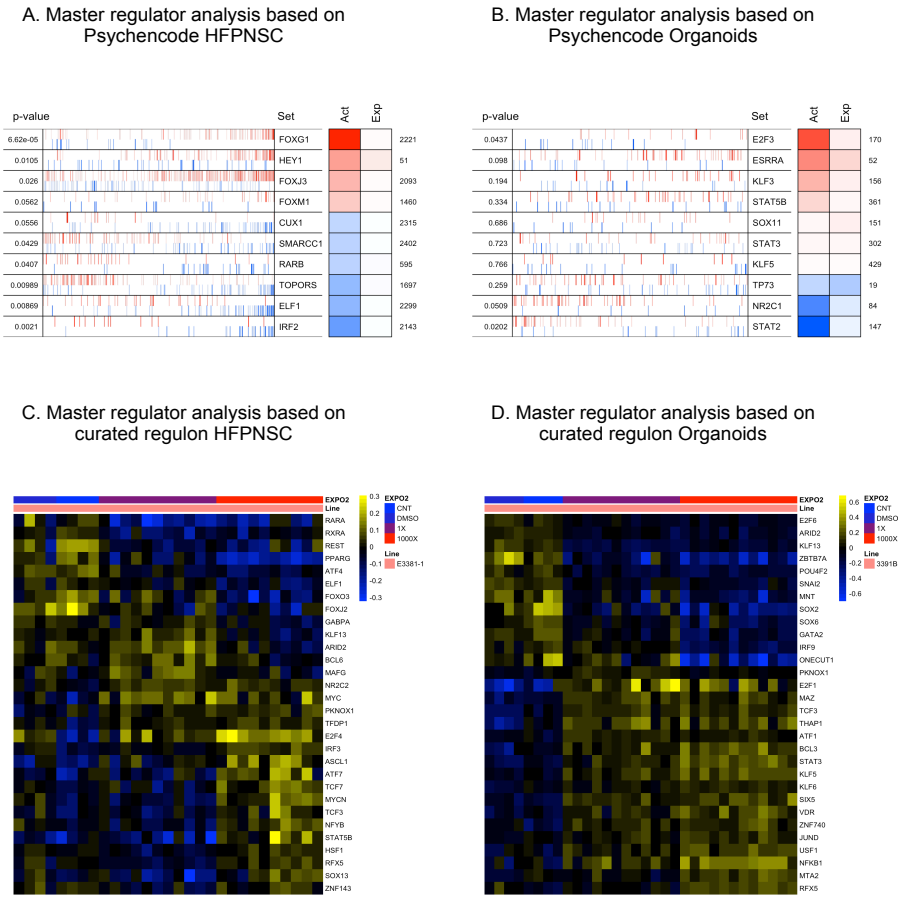


Figure 18 Master regulator analysis on genes dysregulated by MIXN chronic exposure

VIPER plots (Alvarez et al., 2016) showing the projection of the negative (repressed, shown in blue color) and positive (activated, shown in red color) targets for each TF, as inferred by ARACNE and correlation analysis when reverse engineering the regulatory network (vertical lines resembling a bar-code), on the gene expression signature (GES) (x-axis), where the genes in the GES were rank-sorted from the one most down-regulated to the one most upregulated in the 'test' vs 'reference' conditions. The two-columns heatmap displayed on the right side of the figure shows the inferred differential activity (first column) and differential expression (second column), with the rank of the displayed genes in the GES (shown all the way to the right). Results for the HFPNSC in A and for the organoids in B.
C. Heatmap of differentially activity for the transcription factors identified upstream of HFPNSC DEGs
D. Heatmap of differentially activity for the transcription factors identified upstream of organoids DEGs

Characterization of the hormonal pathways related to the genes affected by the EDC mixtures

Finally, in order to systematically characterise how the different endocrine pathways are related to MIXN effects, we analysed the hormonal pathway information associated to the DEGs and TFs previously identified, from the repository most commonly used to perform gene set enrichment analysis (Subramanian et al., 2005), as well as manually curated information for hormonal related genes (Chatonnet, Flamant, & Morte, 2015; B. K. Singh, Sinha, & Yen, 2018). The results show a strong intersections between the DEGs and transcription factors identified in our experiments and genes involved in all the main endocrine pathways, thus confirming the complex mechanisms of action of mixtures of endocrine disruptors. Among the 802 DEGs identified in HPFNSC and CO under MIX N exposure, 264 were involved in six endocrine pathways (estrogen, androgen, thyroid, PPARg, progesterone, corticoid). The most significantly regulated signalling pathways associated with the most DEGs identified were thyroid (143) > estrogen (76) > progesterone (50) > PPARg (36) > androgen (23) > corticoid (12). Genomatix Pathway System (GEPS) program was used to identify critical hubs of disruption by common environmental chemicals through hormonal pathways. This software determines interactions through both publicly available and in-house hand-curated data (Cartharius et al., 2005; Frisch, Klocke, Haltmeier, & Frech, 2009). The network was then visualized with Cytoscape (Shannon et al., 2003) (figure 19).

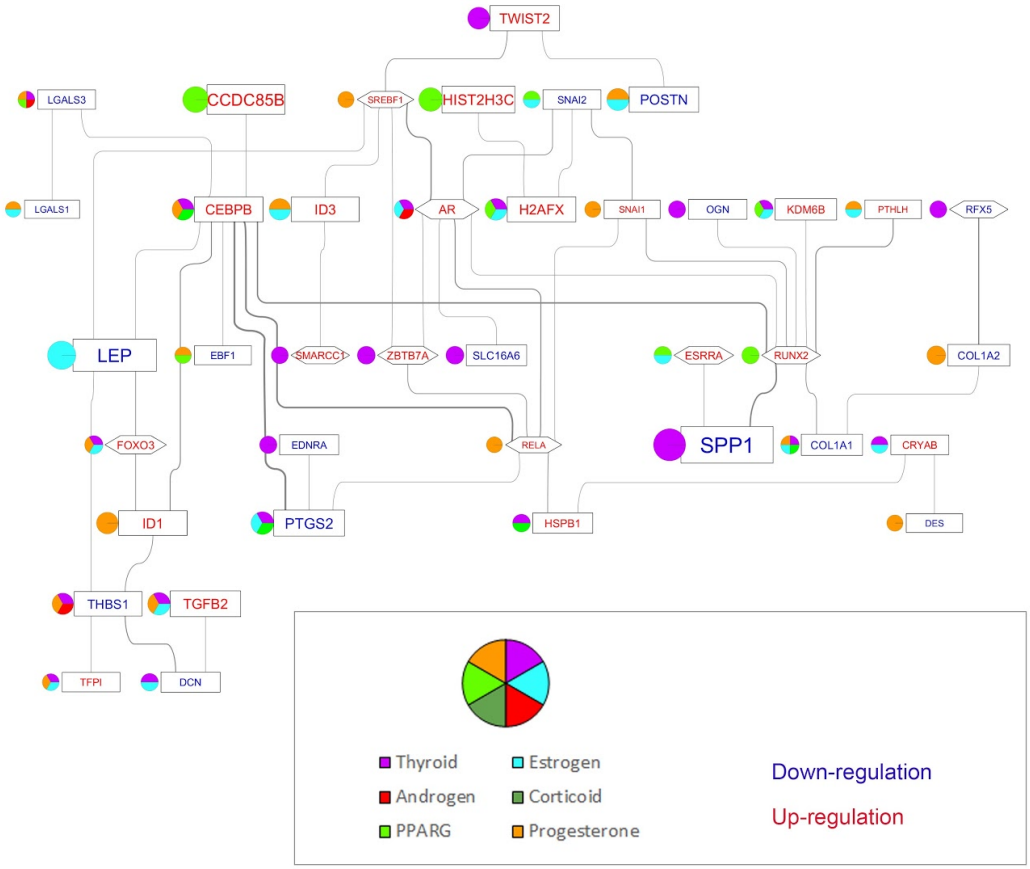


Figure 19 Gene network interactions of hormonal pathways related to the genes dysregulated by MIXN chronic exposure

Hormonal pathways gene interactions were generated by the Genomatix GEPS program that connects differentially expressed genes and transcription factors identified upon chronic exposure of MIXN on HPFNSC and CO.

Dual strategy for multiplexed organoidogenesis

We showed that cortical brain organoids are instrumental for regulatory toxicology, in particular to properly study the effect of chemical exposure during relevant developmental windows. They can also help to increase the throughput of chemical screening being a system that starts from self-renewable cells. Moreover, the opportunity to differentiate organoids from iPSC reprogrammed from different individuals, allows to test the effect of environmental factors on different genetic backgrounds. With the vision of applying chemical exposures on organoids cohorts of hundreds of iPSC lines, and thus dissect the differential impact of genetic variants on environmental factors, we set up and optimised 2 multiplexing strategies, in parallel with our previous experimental designs, that aim to study, at single cell transcriptomic level, the effects of inter-individual variability in response to EDCs, taking advantage of the reduction in technical variability, costs and workload needed to differentiate organoids from several cell lines. In the first approach, namely the downstream, we multiplexed samples just before library preparation, reducing the variability across library and sequencing runs while also decreasing the costs. In the second approach, the upstream, different pluripotent stem cell lines were multiplexed during the generation of the organoids, obtaining a chimeric model that could overcome also the variability related to culturing conditions (Figure 20).

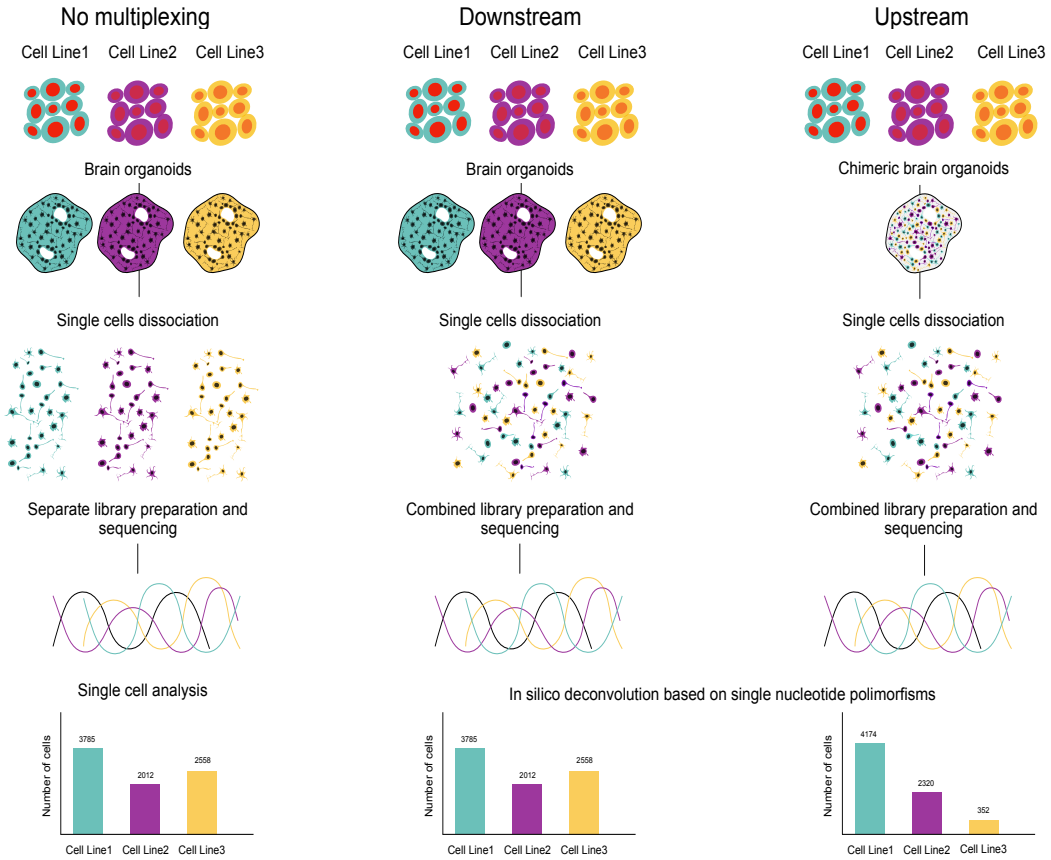


Figure 20 Schematic representation of multiplexing organoidogenesis with parallel strategies

Fluorescent tagged iPSC lines and benchmarking of the chimeric organoid model

First, in order to unravel the spatial distribution of different cell lines inside chimeric brain organoids within the upstream multiplexing strategy, we generated 2 different iPSCs lines that stably express either GFP or mCherry in their cytoplasm to follow their distribution in the organoids. Since the expression of the reporters was not homogeneous among cells, we performed fluorescence-assisted cell sorting (FACS) to select only cells with a high expression of the fluorescent reporters. Also, we performed cell proliferation assay on the different iPSC lines used, in order to verify that the starting proliferation rate at the pluripotent state was homogeneous. As a next step we generated chimeric organoids by mixing an equal number of cells during the first step of embryoid body generation. We performed multiple rounds of chimeric organoids generation and differentiation, testing the efficiency of the system when using 3 or 4 different iPSC lines, also trying different combinations of lines, as shown in Table 4. After the first period of 25 days of differentiation, we performed short tandem repeat (STR) profiling as a first proof that all the lines were retained in the organoids.

Table 4 Cell lines used and experiments performed for multiplexing

Cell line ID	Cell line name	specie	origin	sex
hiPSC 1	KOLF	human	Healthy control (fibroblast)	Male
hiPSC 2	3391B	human	Healthy control (fibroblast)	Female
hiPSC 3	MIFF1	human	Healthy control (fibroblast)	Male
hiPSC 4	809.1.5	human	Healthy control (fibroblast)	Female
HiPSC 5	A15461	human	Healthy control (fibroblast)	Male

Cell lines/ combinations	Chimeric org. 1	Chimeric org. 2	Chimeric org. taggged	Downstream
hiPSC 1	X	X		X
hiPSC 2	X	X	X (+GFP)	X
hiPSC 3	X	X	X	X
hiPSC 4	X		X (+mCherry)	
hiPSC 5		X		X

Model/ Experiments	Cell Proliferation	STR analysis	IF	Live imaging	FACS	scRNAseq
Chimeric org. 1		X	X			X
Chimeric org. 2		X	X			X
Chimeric org. “tagged”		X	X	X	X	
Pure cell lines	X	X	X		X (tagged cell	Downstream

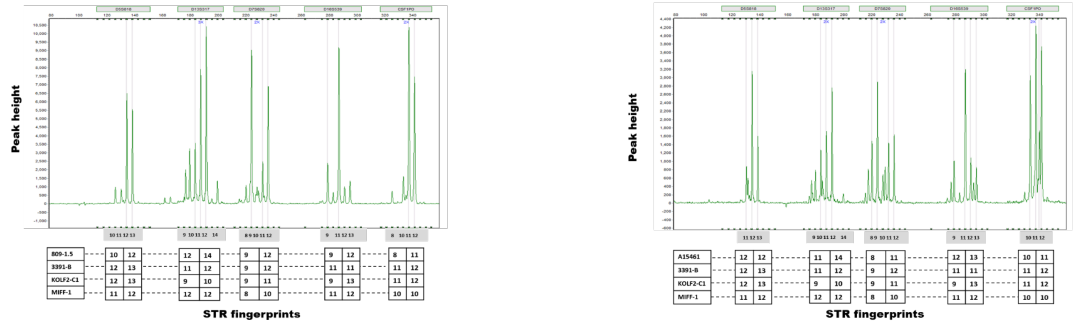


Figure 21 STR analysis

Two different combination of 25 days old chimeric brain organoids shows the presence of all the four mixed genotypes in both models. STR fingerprints of each pure cell line were previously determined both in donors' fibroblasts and in their corresponding

As a next step we performed immunofluorescence stainings to characterise the expected markers and morphological organization of the chimeric organoids. After 50 days of differentiation chimeric brain organoids fully recapitulated the structure of pure lines-derived organoids displaying ventricle-like structures organized in a radial pattern and mainly composed by PAX6 positive apical progenitors. They also showed the expected expression and distribution of the neuronal marker Tuj1 (Figure 22 A). Immunofluorescence of 100 days old organoids also showed the presence of cortical lower layers, CTIP2^{+} and Reelin^{+} neurons (Figure 22 B). Finally, we were able to image the chimeric organoids generated with the fluorescence tagged lines (Figure 22 C) and observe the absence of a rigid regionalization of the different lines in the organoids.

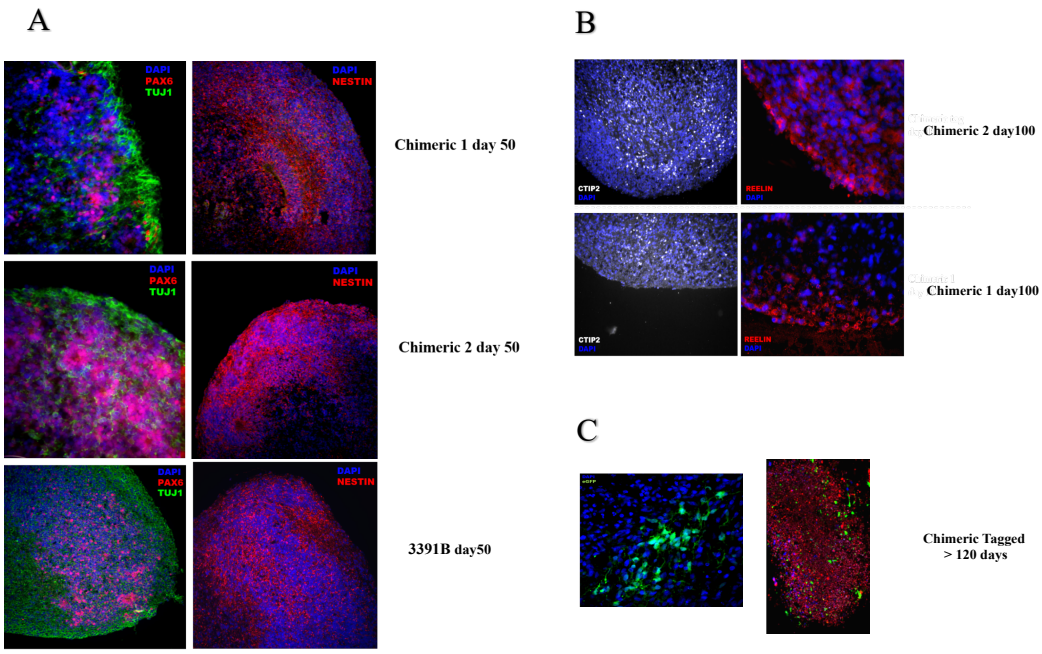


Figure 22 Immunofluorescence of chimeric brain organoids

A. Selected panels for Tuj1, Nestin and Pax6 stainings in 50 days organoids confirm the similarity between chimeric model structural organization and those of pure lines-derived organoids.
B. Selected panels for CTIP2 and Reelin confirms chimeric brain organoids differentiation proceeds giving rise to specialized and spatially organised cell populations.
C. Confocal image of 150 days old chimeric tagged organoid sections. It's possible to observe the three distinct cell populations that compose the organoid, namely GFP^{+} cells, $mCherry^{+}$ cells and double-negative cells which can be identified by the presence of $Hoecst^{+}$ nuclei that do not co-localize with fluorescent reporters.

Single cell transcriptomics deconvolution

As indicated in the experimental design, we performed downstream and upstream multiplexing of different cell lines before single cell library preparation, using the 10X Genomics platform and RNA sequencing. Afterwards we had the opportunity to demultiplex the data and associate the identity of each cell line, thanks to their SNP profiles (Figure 23).

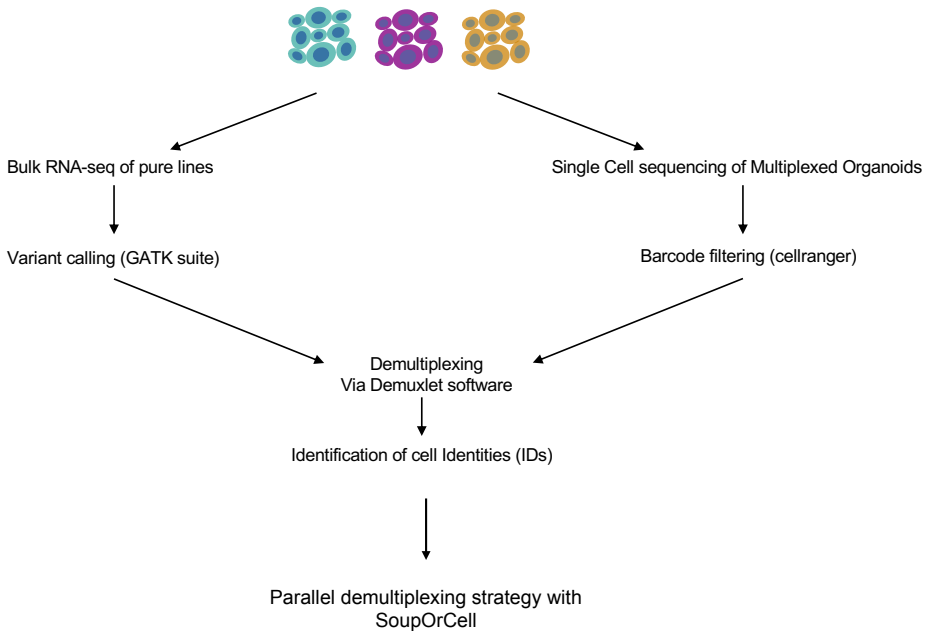


Figure 23 Scheme of the demultiplexing pipeline

Briefly we called the genetic variants of each line from the sequencing of the bulk transcriptomes that we already had, and then we associated the identity of every single cell, following 2 different approaches, namely Demuxlet (Kang et al., 2018a) and SoupOrCell (Heaton et al., 2019) that proved to be highly concordant for our data. Following this pipeline we observed that the differences between the lines in terms of cell numbers per multiplexed samples, are higher in the upstream approach. In both cases we found the presence of all the starting iPSC lines, however, while for the downstream approach the range of variability in cell numbers is comparable with the experiments performed without multiplexing, for the upstream approach we found in 2 out of 3 experiments that 1 of the 4 lines was heavily under represented and it was not possible to attribute the result to a specific line (Figure 24).

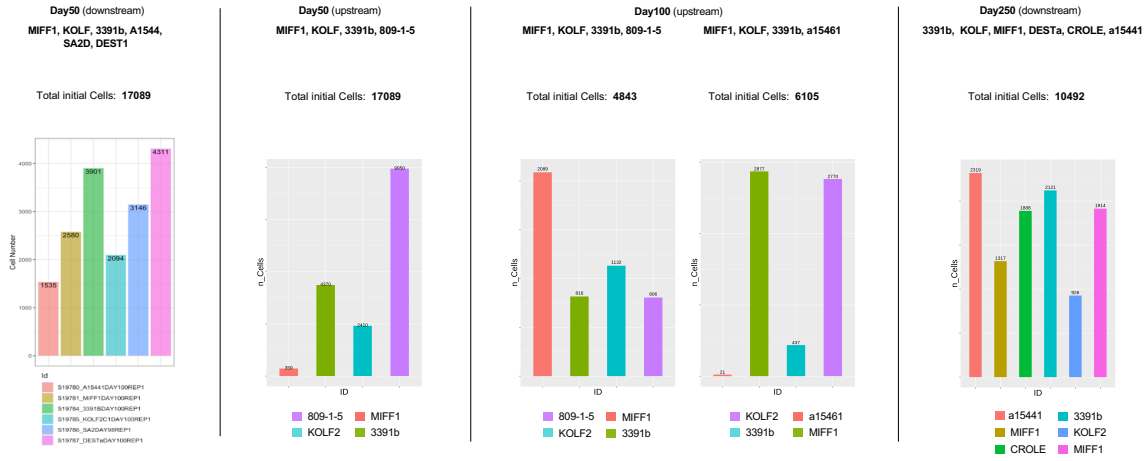


Figure 24 Deconvolution of cell line identities in upstream and downstream multiplexed samples

Barplot representing the number of cells assigned to each individual cell line after association of the SNP signatures, by Demuxlet, between the single cell dataset and the variants selected from the bulk transcriptomes, for each round of multiplexing experiments

After the deconvolution of cells' genetic identity, we analysed the data following the same pipelines applied to the other single cell datasets, performing the filtering, normalization, dimensionality reduction and clustering. Remarkably both for the downstream and upstream samples, the cells from the different lines show an homogeneous distribution across the manifold, with the principal component of variability driven by neuronal differentiation genes, thus confirming the big advantage of reducing the batch effect that is usually observed when plotting different samples in the same space without any batch correction. We were also able to assign the identity of the different cell clusters to the expected populations and observed that all different cell lines were represented in each of the clusters.

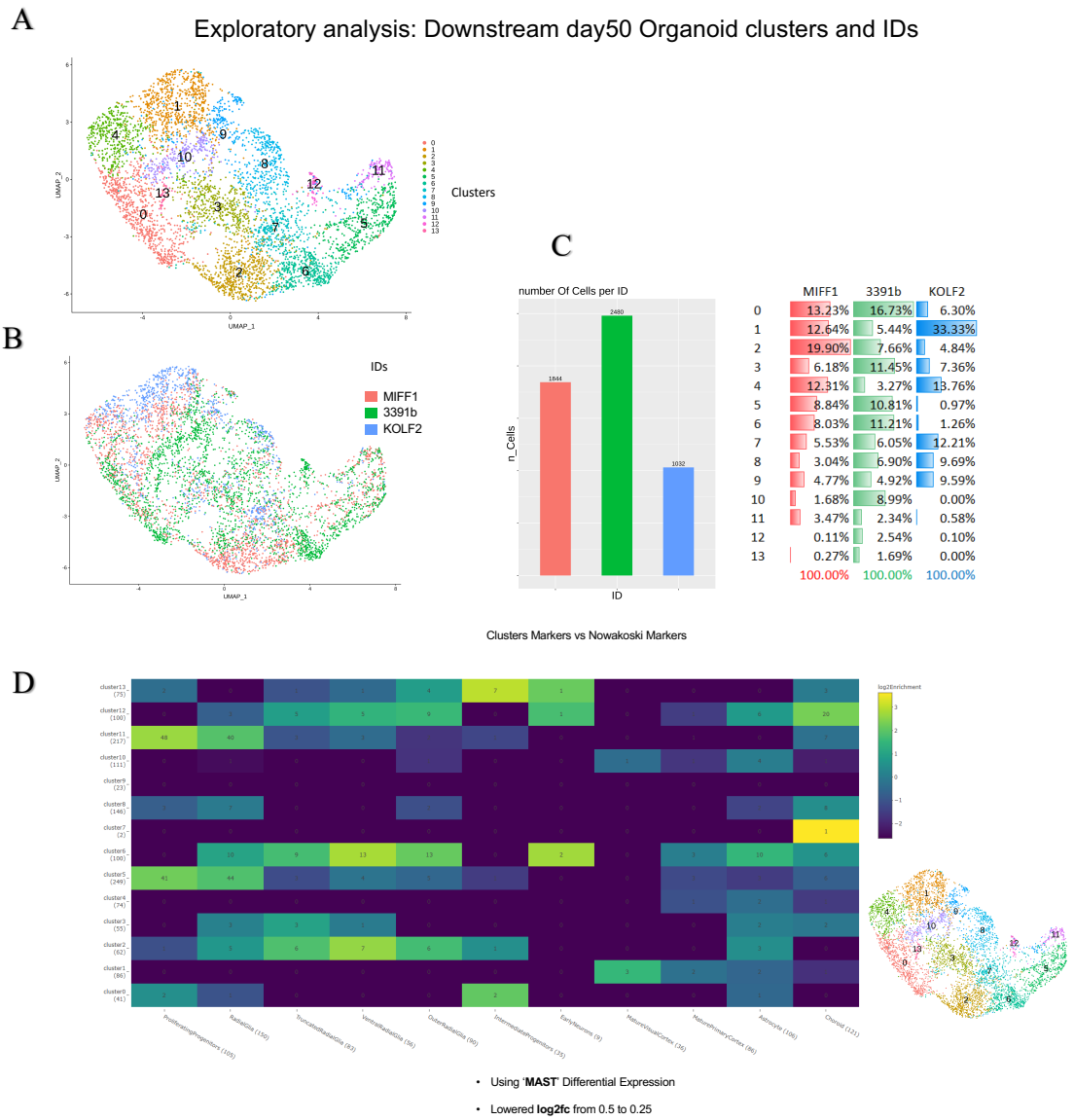


Figure 25 Cell populations decomposition in multiplexed brain organoids

A.Shared nearest neighbor (SNN) modularity optimization based clusters (computed in Seurat) and plotted using UMAP.
B. Cell identity assigned using Demuxlet and plotted using UMAP
C. Barplot of number of cells assigned to each identity and distribution of the cell indemnities in each of the clusters previously identified
D.Heatmap of the overlaps between marker genes characterizing the internal organoids clusters and external gene signatures of the relevant single cell clusters from human fetal brains (Nowakowski et al., 2017). Colors are based on the log2Enrichment computed for each pair of comparisons.

Discussion and future directions

The current vision for improving regulatory decision making relies on the transforming potential of high throughput and high content data to elucidate and quantify the molecular, cellular and organismal responses to chemicals (Califf et al., 2016). In the context of chemical regulation most authorities, including the Organisation for Economic Co-operation and Development (OECD), recommend integrated approaches to testing and assessment (IATA) that incorporate results from multiple methodologies. Emphasis is placed on molecular initiating events (MIE) that lead to physiologically measurable adverse outcome pathways (AOP). Here we first identified the adverse outcomes (language delay or low birth weight) in humans, then proceeded to determine the prenatal chemical mixtures associated with these outcomes in children and, finally, established the causative molecular and cellular impacts using in vitro human models. By making EDC mixtures experimentally tractable as the 'real life'-relevant unit of exposure, these complementary methodologies allowed us to uncover the gene networks specifically altered by neurodevelopment- or growth-targeting EDC mixtures and define endocrine axis of vulnerability. Furthermore, by establishing the value of human reprogrammed models based on self-renewing sources, we both expand their reach to regulatory toxicology and enrich the latter experimental human insight. Together, this approach allowed us to define dysregulated gene networks, identified and validated through complementary methods, that can be exploited to link, mechanistically, MIEs to AOPs. As such it represents a methodology and approach broadly applicable within the new regulatory frameworks globally.

We then moved forward to develop new experimental platforms within the vision of incorporating the contribution of human genetic variability to the assessment of chemical exposure by multiplexing organoidogenesis.

Our approach, in the context of chemical exposure was limited to iPSC lines derived from healthy individuals, but the chimeric model clearly represents an ideal platform to study and test the cell type specific effects of genetic diseases if a mutant line is grown in the same organoid with a wild type line.

Moreover, taking advantage of the high number of cell replicates within each cell population, given by single cell transcriptomics, it is possible to associate a genetic variant to the effect on gene expression in a specific cell population using the eQTL framework (Kang et al., 2018a; Shabalin, 2012; van der Wijst et al., 2018).

In the case of our data, since the number of individuals is still far from the minimum required for significant de novo eQTL calling, we are instead focusing on a supervised

approach in which we are starting from SNPs that are already associated to gene expression alteration on relevant cell types for neurodevelopment, or to neurodevelopmental disorders ("Autism-SNPedia"; "dbGaP PsychENCODE Consortium"; "GENEX-FB2"; "GTEx Portal - eQTL dashboard"; "LIBD eQTL Browser") to study more in-depth their association to neurodevelopmental dynamics. In this way we can take advantage of the longitudinal profiling of our organoids cohort and of the high number of differentiation replicates that we can compare across the different multiplexing strategies. The innovative idea we are proposing is thus to look for the associations between genetic variants and the developmental trajectories that, for each individual iPSC lines we can capture throughout organoids differentiation thanks to single cell transcriptomics deconvolution. As it is shown in the recent preprint by (Cuomo et al., 2019), pseudotime alignment of the single cells, improve the power to detect cell type specific eQTL. In our design we are analyzing how previously known SNPs correlate with the shape and velocity of developmental trajectory, to define, as a proof-of-concept, the idea of eDTL (expressed developmental trait loci) that could be generalized to the study of different organ development and applied to study gene regulatory mechanisms when dealing with the appropriate number of iPSC lines.

There are a number of challenges however to take into account in multiplexing experimental design, represented for example by the impact of cell autonomous vs interactions in the chimeric model (especially if the platform is used to combine a disease with a wild type line), or by the differences in the proliferative rates of each cell line that can lead to an imbalance of the composition of chimeric organoids. We are implementing, indeed, experimental screenings to test the number of lines that can be simultaneously grown in a single chimeric organoids, given the importance of this same approach for all the different pooled CRISPR based perturbations that we are envisioning for different projects of the lab (Boettcher et al., 2018; Datlinger et al., 2017; Mimitou et al., 2019; Xie, Duan, Li, Zhou, & Hon, 2017b).

To conclude, our multiplexing experiments represent the attempt to make a concrete example of experimental epidemiology *in vitro*, with the cohort of organoids that can be used to study the specific impact of genetic variants on neurodevelopment and that we defined as "Europe in a dish".

Acknowledgments

The work in our lab, and particularly my PhD project was carried out as a highly collaborative and interdisciplinary effort, involving the interactions of different skills, expertise as well as practical and intellectual contributions that are briefly summarized in the following section:

Giuseppe Testa is my mentor and supervisor. He conceived the projects, my hybrid wet and computational training, and our constant, in-depth and productive discussions are shaping all the aspects of our research and future perspectives.

Together with Alejandro Lopez Tobon and Sebastiano Trattaro ("the organoiders team") we imported and set up the use of organoids as an experimental system for human disease modelling since 4 years ago. This included the optimization of the culture conditions and of all the downstream assays that we can now perform. We have developed during this time different lines of research, so while we were sharing all the experimental work of organoids differentiation and profiling in the initial stages, now we continue to collective organize the experiments that are followed in a more independent fashion.

Alejandro has been my laboratory tutor in the first phase of my PhD and he taught me everything about stem cell culture, molecular biology and imaging techniques. Pierre-Luc Germain has been my computational tutor in the first phase of my PhD and, together with Cristina Cheroni and Emanuele Villa in the later stages, they taught me everything about computational biology and bioinformatics.

I followed all the work, both experimental and computational on endocrine disruptors, with the main contributions of Pierre-Luc and Cristina for the bioinformatics analysis. The work was carried out within the highly interdisciplinary and international environment of EDCMixRisk (https://edcmixrisk.ki.se/), where it was possible to integrate epidemiology and biostatistics with experimental biology, risk assessment and policy making, the last aspects spearheaded by Nadav Even Chorev in our lab. The success of this approach evolved in the recent award of the European project ENDpoiNT (https://endpoints.eu/) in which we are directly involved and that gives us the opportunity to follow and push forward these lines of research.

I contributed together with Sebastiano Trattaro and Cristina Cheroni, the main authors and others, to the work on the systematic analysis of the cohort of wild type cortical organoids relative to different external organoids protocols (not shown in this thesis) and the human fetal data.

I contributed together with Alejandro, Cristina and Emanuele, the main authors and others, to the work on the chronic inhibition of the GSK3 pathway during organoids differentiation.

I co-supervised Marco Tullio Rigoli for his master thesis, during which he carried out most of the experimental work on the multiplexing of organoids.

Together with Davide Castaldi, under the supervision of Cristina and Emanuele we are analyzing the single cell transcriptomic datasets of the multiplexing experiments.

Giuseppe D'Agostino, while sharing productive scientific discussions, taught me how to make illustrations for the figures and contributed to the graphics of the EDCMixRisk project.

Erika Tenderini and Flavia Troglio assisted the experimental work, particularly in the phases of long term organoids culture and library preparation for transcriptomics.

Together with Giuseppe we conceived the work on blockchain for omics data that we are elaborating with Nadav Even Chorev and Luca Marelli.

All the member of the lab have contributed importantly in different parts of our research projects with scientific discussions, constructive critiques and practical suggestions.

Materials and methods

Cell lines culture and quality control

Human induced pluripotent stem cells have been previously validated in our laboratory (Adamo et al., 2015). They were cultured in TeSR/E8^{TM} medium (Stemcell technology) supplemented with penicillin (100 U/mL) and streptomycin (100 µg/mL), with daily media change at 37 °C, 5 % CO2 and 3 % O2 in standard incubators. Before hiPSCs plating, cell culture dishes were coated with matrigel as follows: matrigel stock solution (Corning) is diluted 1:40 in DMEM/F12 medium (Euroclone) and the resulting matrigel solution was used to coat dishes at least 10 minutes at 37 °C. hiPSCs were normally splitted 1:6 to 1:10 using ReLeSR^{TM} (Stemcell technology) when confluency reached approximately 60-70%. When single-cell dissociation was needed, accutase solution (Sigma-Aldrich) was used. While ReleSR^{TM} contains only EDTA salts and detaches cells in clumps, accutase solution contains a proteolytic enzyme that dissociate colonies at single cell level. When using accutase, ROCK inhibitor 5uM (Sigma Aldrich) was used to enhance cell survival. To cryopreserve hiPSC lines, cells were dissociated when 60% confluent with accutase and resuspended in TeSR/E8 medium, 10 % DMSO and ROCK inhibitor 5µM. Cell lines were kept mycoplasma-free by routinely testing for contaminations.

They were routinely assessed with STR profiling to confirm the correct identity of the lines using the kit GenePrint^{R} 10 system (Promega). Cell lines were profiled with Array-CGH for verifying the absence of chromosomal rearrangements, using the Agilent kit SurePrint G3 Human CGH Microarray 8x60K.

Human foetal primary neural stem cells (HFPNSC) were provided by Dr. Steve Pollard's laboratory. They were derived from the cortex of post-conception week 11 and 19, male embryo and from the ganglionic eminence of post-conception week 8, male embryo. They were cultured in our cell culture in the media Sigma D8437 DMEM/HAMS-F12 (which comes with Hepes and Glutamine) complemented with:

For Wash Media = 500ml bottle Sigma D8437

+ 5mls Pen-Strep (Gibco 15140-122)

+1ml BSA Soln 7.5% (Gibco 15260-037)

For Complete Media = 500ml bottle Sigma D8437

+7.25ml Glucose (Sigma G8644)

+5ml MEM NEAA 100x (Gibco 11140-035)

+ 5mls Pen-Strep (Gibco 15140-122)

+800uL BSA Soln 7.5% (Gibco 15260-037)

+1mL bMercETOH 50mM (Gibco 31350-010 (20ml))

+5ml B27 Supplement 50x (LifeTech/Gibco 17504-044)

+2.5ml N2 Supplement 100x (LifeTech/Gibco 17502-048)

All reagents brought to 37degree before adding; N2 supplement swirled gently to avoid precipitation. Before adding to the cells, the complete media is supplemented with mouseEGF to final concentration 10ng/ml (peprotech) and humanFGF to final concentration 10ng/ml (peprotech)/Laminin to final concentration 1ug/ml (Sigma L2020-1MG/ML). For detaching the cells Accutase Solution from Sigma(A6964 100ml) is used. The standard method of remove media, PBS rinse, accutase – cells detach – add wash media and collect cells, spin to pellet cells(then resuspend in complete media + EGF + FGF + Laminin. Depending on cell line a rough guide is usually a 1 in 4/ 1 in 5 split for Human cells and approx. 1 in 8/1 in 10 for mouse cells. For freezing the cells Wash Media + 10%DMSO is used.

Adult human bone marrow derived mesenchymal stem cells (hMSCs) from 2 donors were a kind gift of Dr. Katarina Leblanc (Center of Hematology and Regenerative Medicine, Department of Medicine, Karolinska Institutet, Sweden) and were cultured in Dulbecco's Modified Eagle's medium (DMEM, Gibco ® by Life technologies) supplemented with 10% heat-inactivated fetal bovine serum (FBS) (Gibco® by Thermo Fisher Scientific) (see also SI). Human indcuded pluripotent stem cell-derived mesenchymal stem cells (hiPSC-MSCs) were obtained and cultured from neural crest stem cells derived from iPSCs in our lab according to the protocol described by (Menendez et al., 2013) and previously used in (Adamo et al., 2015). Adult MSCs were routinely cultured at 37°C, 5% CO2 in growth media (GM): Dulbecco's Modified Eagle's medium (DMEM, Gibco ® by Life technologies) supplemented with 10% heat-inactivated FBS (Gibco® by Thermo Fisher Scientific), 1% penicillin/streptomycin (Gibco® by Thermo Fisher Scientific) and 2% glutamine (Gibco® by Thermo Fisher Scientific). iPSC-MSCs were cultured under the same conditions but in DMEM/F12 (Gibco® by Thermo Fisher Scientific) medium supplemented with 10% heatinactivated FBS, 1% penicillin/streptomycin and 1% glutamine.

Cortical Brain Organoids

For the differentiation of cortical brain organoids we used the protocol described by (Paşca et al., 2015), with minor modifications to improve its efficiency.

For the work on GSK3b inhibition, hPSCs were plated onto cell cycle-arrested mouse embryonic feeders (MEFs)(Millipore) for one passage, colonies grown for at least 48h and then enzymatically detached by incubation with 0.7 mg/ml dispase (Invitrogen: 17105-041) for approx. 30 min. Suspended colonies were subsequently transferred into ultra-low- attachment 100 mm plastic plates (Corning) in FGF2-free knockout serum medium. For the first 24 h (day 0), the medium was supplemented with the ROCK inhibitor Y-27632 (EMD Chemicals). For neural induction, dorsomorphin (Merck, 5 μM) and SB-431542 (Tocris, 10 μM) were added to the medium until day 5. From day 6 onward, organoids were moved to neural medium (NM) containing Neurobasal (Invitrogen 10888), B-27 serum substitute without vitamin A (Invitrogen 12587), GlutaMax 1:100 (Fisher 35050071), 100 U/ml penicillin and streptomycin (Invitrogen) and 50 mM b-Mercaptoethanol (Gibco 31350010). The NM was supplemented with 20 ng/ml FGF2 (Thermo) and 20 ng/ml EGF (Tocris) for 19 days with daily medium change in the first 10 days, and every other day for the subsequent 9 days. On day 12, floating organoids were moved to orbital shaker (VWR Standard Orbital Shaker, Model 1000) and kept on constant shaking at 50 rpm to promote nutrient and oxygen exchange. To induce neurogenesis, FGF2 and EGF were replaced with 20 ng/ml BDNF (Peprotech) and 20 ng/ml NT3 (Peprotech) starting at day 25, while from day 43 onwards only NM without growth factors was used for medium changes every other day.

For the other projects, since we optimized the protocol to avoid the use of MEF, the following procedures were followed: when the hiPSC line reached 80% confluency in a 10 cm dish, colonies were dissociated with Accutase and centrifuged to remove the enzymatic suspension. After resuspension in TeSR/E8 medium supplemented with 5uM ROCK inhibitor cells were counted with a TC20 automatic cell counter (Biorad) and resuspeded to get a final concentration of 2 x 10^5 cells/mL.

100 uL/well of cell suspension were seeded into PrimeSurface^R 96 well plates (SystemBio) and then the plates spinned at 850 rpm for 3 minutes to promote the formation of embryoid bodies. The utilization of ultra-low attachment plates is fundamental to avoid cell adhesion. The day of the EB generation is referred to as day -2. On day -1, medium is not changed, leaving EBs undisturbed. On day 0 media change is performed adding differentiating medium one, composed of 80%

DMEM/F12 medium (1:1), 20% Knockout serum (Gibco), 1 mM Non-essential amino acids (Sigma), 0.1 mM cell culture grade 2-mercaptoethanol solution (Gibco), GlutaMax (Gibco, 1:100), penicillin 100 U/mL, streptomycin (100 μg/mL), 7 uM Dorsomorphin (Sigma) and 10 uM TGFβ inhibitor SB431542 (MedChem express). Dorsomorphin and TGFβ inhibitor are used to perform DUAL-SMAD inhibition, promoting the induction of neuroectoderm.

From day 0 to day 4, media change is performed every day with the very same medium while on day 5 differentiation medium two is added until day 24 with daily media change for the first 7 days.

Differentiation medium two is composed of neurobasal medium (Gibco) supplemented with B-27 supplement w/o vitamin A (Gibco, 1:50), GlutaMax (1:100), penicillin 100 U/mL, streptomycin (100 μg/mL), 20 ng/mL FGF2 (Thermo) and 20 ng/mL EGF (Thermo).

On day 12 organoids are moved to 10cm ultra low attachment dishes and grown on shakers to enhance oxygen and nutrient supply. From day 12 onwards media change is performed every other day. On day 24, FGF2 and EGF are replaced with 20 ng/mL brain-derived neurotrophic factor (BDNF, Thermo) and 20 ng/mL neurotrophin-3 (NT3, Thermo) to promote differentiation of neural progenitors. From day 42 onwards, grow factors are removed.

The volume of medium needed for each plate vary according to the size and number of organoids per plate as it must be sufficient to ensure they are completely covered. As a general rule, media volume increases while proceeding with differentiation, reaching a maximum of approximately 35 mL/plate.

For EDC chemical exposure of HFPNSC, were seeded in 6-well plates. When confluency was reached 0.1% DMSO or the other compounds diluted in DMSO, MIX N, MIX G in 5 different concentrations (from 0.1X to 1000X), 0.04 nM BPA, 10 uM T3 were added to the culture media and used to culture cells for 48 hours or for 15 days (in the chronic setting), adding the chemical every time medial change was performed (every other day). After the exposure cells were collected for RNA extraction.

For EDC chemical exposure of organoids, 0.1% DMSO, or the other compounds diluted in DMSO, MIXN in the 2 different concentrations used (1X and 1000X), 0.04 nM BPA and 10 uM T3 were added to the culture media and used to culture cells from day 0 to day 50. After the exposure, organoids were collected for both RNA extraction and fixed for immunofluorescence analysis.

For the EDC chemical exposure on MSC, cells were seeded in 6 well plates and expanded until they reached 70-80% confluence. Growth media was replaced by treatment media consisting of DMEM supplemented with 10% charcoal stripped FBS, 1% penicillin/streptomycin and 2% glutamine two days before experiment start, and cells were kept in this media throughout all experiments. Treatment was performed for 48h before the cells were lysed. For adipogenic induction and lipid accumulation assessment, cells were seeded in 96 well plates and treatments were performed in treatment media for 14-21 days where medium was changed twice a week.

RNA sequencing

Total RNA was isolated with the RNeasy Micro Kit (Qiagen, Hilden,Germany) according to the manufacturer's instructions. RNA was quantified with Nanodrop and then the integrity was evaluated with Agilent 2100 Bioanalyzer (only if the quality ratios were not optimal after Nanodrop analysis). TruSeq Stranded Total RNA LT Sample Prep Kit, Illumina was used to run the library for each sample. Sequencing is performed with the Illumina HiSeq 2000 platform, sequencing on average 10 millions 50bp paired-end reads per sample.

RNA-seq data analysis (EDC Mix project)

All the data and the code have been organized in a repository that will be open upon publication of the paper. Gitlab access can be granted upon specific request.

RNA-seq quantification was performed directly from the reads using Salmon 6.1, using the hg38 Refseq annotation. Only genes with at least 20 reads in each of at least 2 concentrations of the mixture using the same cell lines were included for further analysis; small (<200nt) genes, ribosomal RNA genes, and fusion genes were excluded. Batch correction was applied to get rid of technical variability due to experiments performed in different rounds using surrogate variable analysis (Parker et al., 2014). Differential expression analysis was performed on the estimated counts after TMM normalization with edgeR (M. D. Robinson, McCarthy, & Smyth, 2010) using a likelihood ratio test on the coefficients of a negative binomial model including the genetic background and the mix concentration. The concentration was treated as a categorical variable (i.e., converted to factor), and tested for any non-zero coefficient. Genes identified through this method were then kmeans-clustered on the basis of their smoothed fold change upon each concentration (using the NbClust R package to determine the consensus number of clusters). The mean smoothed fold change pattern for the main cluster(s) were plotted for dose-response patterns.

Enrichment analysis

Gene Ontology (GO) enrichment analyses were performed with the goseq R package, and topGO (version 2.30.1) including correction for eventual RNA-seq transcript-length bias and excluding genes without annotation. Terms with at least 10 but no more than 1,000 associated genes were considered, and Fisher’s exact test was used. The tested genes (excluding small and lowly-expressed genes) were used as a background. Parent terms with significantly enriched children terms were filtered out to improve the specificity of the enrichments. Unless stated otherwise, other enrichment tests were performed using the hypergeometric test.

Master regulator analysis

The psychencode network (D. Wang et al., 2018) was reconstructed using human fetal and adult brain datasets, combining general information about motifs, accessibility, 3D contact (HiC), eQTLs and correlations with target gene expression. It includes not only regulatory interactions proximal to the TSS, but also distal enhancers. It is particularly valuable because it is specific to brain tissues. Of course, it can detect only interactions for which there is variation among the population and during development. The downside is that none of the interactions are validated. The curated regulon is instead limited to TF bindings close to the TSS, and only to a small subset of TFs. The interactions are also not at all tissue-specific. The advantage is that there's a lot of experimental binding evidence in there (e.g. ChIPseq), as well as curated interactions, and that the interactions are weighted by amount of evidence. Master regulator analysis and visualization was performed with the VIPER package (Alvarez et al., 2016).

RNA-seq data analysis (general)

Bulk transcriptome analysis. *Differential gene expression.* Gene expression quantification at the gene level was performed by Salmon (version 0.8.2) (Patro et al., 2017), using hg38 RefSeq annotation. To estimate differential expression, the matrix of gene counts was analyzed by edgeR (version 3.20.9) (Robinson et al., 2009). For each time point, genes with an expression level of at least 2 cpm (count per million) in at least 3 samples were selected for the analysis. Small genes, ribosomal genes and fusion genes were excluded. After TMM normalization, differential expression analysis comparing treated to untreated samples was performed using a likelihood ratio test on the coefficients of a negative binomial model. Significantly modulated genes were selected setting an absolute value of log2 fold change (Log2FC) higher than 1 and a false discovery rate (FDR) lower than 5%. Log2 cpm values, were used for heatmap representation of gene expression profiles (visualized as z-scores). Heatmaps were produced with pheatmap R package (version 1.0.10, Raivo Kolde (2018). pheatmap: Pretty Heatmaps.). Analyses were performed in R version 3.4.4. Functional annotation of biological functions was performed by Gene ontology analysis and Gene set enrichment analysis (GSEA) using as set source H1 collection from the Molecular Signature Database (Liberzon et al., 2015). *Gene Ontology Enrichment Analysis.* Gene ontology enrichment analysis for the Cellular Component domain of the ontology was performed on the 898 DEGs identified at day 50, split in up-regulated and down-regulated genes. The pool of tested genes (as selected for differential expression

analysis) was used as background. The analysis was performed by topGO (version 2.30.1) (Adrian Alexa and Jorg Rahnenfuhrer 2016). topGO: Enrichment Analysis for Gene Ontology), relying on Fisher test and Weight01 method to take into account ontology hierarchy; minimum node size was set at 15. After imposing an enrichment cut-off of 2, a 0.01 p-value cut off was applied to select significantly enriched GO terms. Barplot in supplementary figure 4 shows the top-10 categories ranked for p-value.

Gene set enrichment analysis. GSEA was applied to each developmental stage with GSAA software, version 1.2 (Xiong et al., 2014). Raw reads for the same genes tested for differential expression were analyzed by GSAASeqSP (permutation type 'gene set').

Single cell library preparation and sequencing

Briefly, a small volume (6 - 8 µl) of single-cell suspension at a density of 1000 cells/µl was mixed with RT-PCR master mix and immediately loaded together with Single-Cell 3′ gel beads and partitioning oil into a single-cell 3′ Chip. The gel beads were coated with unique primers bearing 10× cell barcodes, unique molecular identifiers (UMI) and poly(dT) sequences. The chip was then loaded onto a Chromium instrument (10× Genomics) for single-cell GEM generation and barcoding. RNA transcripts from single cells were reverse-transcribed within droplets to generate barcoded full-length cDNA using Clontech SMART technology. After emulsion disruption, cDNA molecules from one sample were pooled and preamplified. Finally, amplified cDNAs were fragmented, and adapter and sample indices were incorporated into finished libraries which were compatible with Illumina sequencing. The final libraries were quantified by Qubit system (Thermo) and calibrated with an in- house control sequencing library. The size profiles of the pre-amplified cDNA and sequencing libraries were examined by Agilent Bioanalyzer 2100 using a High Sensitivity DNA chip (Agilent). Two indexed libraries were equimolarly pooled and sequenced on Illumina NOVAseq 6000 platform using the v2 Kit (Illumina, San Diego, CA) with a customized paired-end, dual indexing format according to the recommendation by 10× Genomics. Using proper cluster density, a coverage around 250 M reads per sample (2000–5000 cells) were obtained corresponding to at least 50,000 reads/cell.

Organoids were collected at day 50, 100 or 250. 3-5 organoids per condition were dissociated by incubation with a solution of 0,5 mg/ml trypsin + 0,22 mg/ml EDTA (Euroclone) with 10 ul of DNaseI 1000 U/ml (Zymo Research) for 30 – 45 min according to organoid size. Digested suspensions were passed once through 0.4 uM

Flowmi™ cell strainers, resuspended in PBS and counted using a TC20 Automated Cell Counter (Biorad). Droplet-based single-cell partitioning and single-cell RNA-Seq libraries were generated using the Chromium Single-Cell 3′ Reagent v2 Kit (10× Genomics, Pleasanton, CA) following manufacturer's instructions (Zheng et al., 2017). Briefly, a small volume (6 – 8 µl) of single-cell suspension at a density of 1000 cells/µl was mixed with RT-PCR master mix and immediately loaded together with Single-Cell 3′ gel beads and partitioning oil into a single-cell 3′ Chip. The gel beads were coated with unique primers bearing 10× cell barcodes, unique molecular identifiers (UMI) and poly(dT) sequences. The chip was then loaded onto a Chromium instrument (10× Genomics) for single-cell GEM generation and barcoding. RNA transcripts from single cells were reverse-transcribed within droplets to generate barcoded full-length cDNA using Clontech SMART technology. After emulsion disruption, cDNA molecules from one sample were pooled and pre-amplified. Finally, amplified cDNAs were fragmented, and adapter and sample indices were incorporated into finished libraries which were compatible with Illumina sequencing. The final libraries were quantified by real -time quantitative PCR and calibrated with an in-house control sequencing library. The size profiles of the pre-amplified cDNA and sequencing libraries were examined by Agilent Bioanalyzer 2100 using a High Sensitivity DNA chip (Agilent). Two indexed libraries were equimolarly pooled and sequenced on Illumina NOVAseq 6000 platform using the v2 Kit (Illumina, San Diego, CA) with a customized paired end, dual indexing (26/8/0/98-bp) format according to the recommendation by 10× Genomics. Using proper cluster density, a coverage around 250 M reads per sample (2000–5000 cells) were obtained corresponding to at least 50,000 reads/cell.

Single cell transcriptome analysis

Before downstream analyses, data deriving from the 11 samples was integrated by Seurat v3.0-alpha analytical framework (Stuart et al., 2018). After normalization, anchors for data integration were identified considering 3000 anchor points (genes) and 40 dimensions. For data reduction, UMAP was applied with 50 nearest neighbors (nn); cluster initial positions were set considering PAGA node position (Scanpy v1.3.1) (Wolf et al., 2018). On the integrated dataset, clusters were identified by applying Louvain with Multilevel Refinement from Seurat with resolution parameter at 0.7. This resulted in the identification of 15 clusters. For cluster annotation, we applied the FindMarker Seurat function, using MAST as test and filtering for up-regulated genes with adjusted P value < 0.05. The obtained lists were compared in an

overlap analysis with gene lists derived from two published single cell datasets : from the WGCNA analysis of single cell clusters of human fetal brains (paper by Nowakowski et al, DOI: 10.1126/science.aap8809), data were downloaded from the web portal at https://cells.ucsc.edu/?ds=cortex-dev and upregulated genes (vg_diff.float>1.5) were selected for each relevant cluster, from the differential expression analysis of single nuclei clusters (Table S6 of the paper by Amiri et al, DOI: 10.1126/science.aat6720), upregulated genes were selected for each relevant cluster fold change > 0 and adjusted P value < 0.05. P values and enrichment coefficients were computed relative to a universe of 3000 genes used for all the other single cell analysis. To test for unbalance in the number of highly expressing cells for representative genes per cluster or cluster set, a threshold of expression at the 90th percentile was fixed, and the number of cells above or below the threshold in treated and untreated cells was compared by Fisher test. The identification of the threshold and sub-sequent analysis was performed on the sub-sampled dataset, separately for Day 50 and Day 100.

Cell cycle analysis were performed using Scanpy function score_genes_cell_cycle, relying on the genes from (Kowalczyk et al., 2015). Diffusion map algorithm for dimensionality reduction was performed with Scanpy with 50 nn. Pseudotime analysis for lineage branching reconstruction was applied using wishbone algorithm (Setty et al., 2016). The analysis was performed on the complete dataset, as well as separately for each of the four biological conditions in order to infer stage or treatment-selective trajectories; the origin was identified with the same method applied on complete dataset. Trajectories were reproduced defining using at least 3 different markers. Partition-based graph abstraction (PAGA) algorithm was applied on the complete dataset, as well as separately for each of the four biological conditions and plotted with layout Reingold Tilford. The position of the nodes identified on the complete dataset was exploited in the graph for each biological condition.

Neurodevelopmental Disorder relevant genes

I collected in the following repository a pipeline to analyse a list of genes relative to public repositories that are relevant in the context of NDD.

https://crockol.github.io/NDDrelevantGenes/

Immunohistochemistry for neuronal systems.

Cortical organoids were fixed with PFA 4% overnight, then transferred to sucrose 30% between 24 / 48h until precipitated and mounted in Killik cryopreservation media for later sectioning, or processed for paraffinization. Sections were washed with PBS and incubated 30 min in sodium citrate buffer pH 6.0 for antigen retrieval. Section were then blocked in 10% normal goat serum (NGS), 0.5% Triton X100 diluted in PBS for one hour at room temperature. The sections were then incubated overnight at 4 °C with primary antibodies diluted in PBS solution containing 2% NGS and 0.1% Triton X-100. PBS was used to wash the primary antibodies and the cryosections were incubated with secondary antibodies containing 2% NGS and 0.1% Triton X-100 for 1 h.

Primary antibodies were prepared in PBS + 5% normal donkey serum (Jackson Immuno Resarch) overnight at 4°C. The following primary antibodies and dilutions were used: anti-PAX6, 1:200 (Biolegend); anti-KI67, 1:200 (Abcam 15580); anti-NESTIN, 1:500 (Millipore); anti– DCX, 1:1000 (BD Biosciences 611706); anti-TBR1, 1:200 (Abcam 31940). After primary incubation, sections were washed three times with PBS and the incubated with appropriate secondary antibodies conjugated to Alexa fluorophores 488 or 594 (Molecular Probes, Invitrogen) diluted 1:500 in blocking solution and incubated for 2 h at RT. Before mounting, sections were incubated with Hoechst 33258 (5 μg/mL; Molecular Probes, Invitrogen) or DAPI 1:5000 (Merck), as indicated on each caption.

Quantification of nuclear markers was done by using the automatic cluster counter ITCN plugin from FIJI (v.1.49 NIH-USA) for at least 3 organoids from 3 independent lines. Images were RGB converted and cluster counts were done over pre-defined concentric grids from VLS. The relative number of positive cells was calculated as a percentage of total DAPI+ cells. Images were acquired with a Leica DMI 6000B microscope (10x, 20x and 40x objectives) and analyzed with LAS-AF imaging software and then processed using Image J (v1.49 NIH, USA) to adjust contrast for optimal RGB rendering. Semi-quantitative measurements were made in imageJ, the background noise was removed with a sliding paraboloid filter of dimension 500 px, then a gaussian filter was applied. A positive cell was considered the local maxima with a minimum threshold of signal to noise. To evaluate the density of cells we consider organoid slice stained with DAPI, from the images of the whole organoid we remove, with a sliding paraboloid filter, the background and then apply a common

threshold on all the conditions. The area of the resulting mask is considered as area occupied by nuclei and used as normalization.

Lipid droplet accumulation assay

Cells were seeded in black-walled 96 well plates with μCLEAR bottom (Greiner Bio One) and exposed to DMSO or the indicated concentrations of MIX G or N as described above six replicate wells. Staining was performed using Bodipy 493/503 and Hoechst 33342 as described in SI and references therein. Images were acquired immediately using the Image Xpress Micro High-Content Analysis System (Molecular Devices, Sunnyvale California USA). Images were taken in FITC and DAPI channel at 10x magnification, at 16 sites per well. Images were further analyzed with the MetaXpress High-Content Image Acquisition and Analysis software (Molecular Devices, Sunnyvale California USA). Using the Transfluor HT analysis module, lipid droplets were quantified by measuring the integrated granule intensity and this value was normalized to nuclei count. The average signal for treatments are presented as ratios compared to the average signal of the DMSO control on the same plate.

Epidemiological exposure assessment

Using data from the Swedish Environmental Longitudinal Mother and child Asthma and allergy (SELMA) pregnancy cohort[8] (described in SI), mixtures of prenatal EDC exposures of relevance for health outcomes in children were identified. Exposure was measured in urine and serum taken in week 3-27 of pregnancy (median week 10, and 96% of the samples were taken before week 13). First morning void urine samples were analyzed for 10 phthalate metabolites (Mono-ethyl phthalate (MEP), metabolite of DEP; Mono-n-butyl phthalate (MnBP), metabolite of DBP; Monobenzyl phthalate (MBzP), metabolite of BBzP; Mono-(2-ethylhexyl) phthalate (MEHP), Mono-(2-ethyl-5-hydroxylhexyl) phthalate (MEHHP), Mono-(2-ethyl-5-oxohexyl) phthalate (MEOHP), Mono-(2-ethyl-5-carboxypentyl) phthalate (MECPP), metabolites of DEHP; Mono-hydroxy-iso-nonyl phthalate (MHiNP), Mono-, oxo-iso-nonyl phthalate (MOiNP), Mono-carboxy-iso-octyl phthalate (MCiOP), metabolites of DiNP); and alkyl phenols including Bisphenol A (BPA) and Triclosan (TCS)). Serum samples were analyzed for 8 perfluorinated alkyl acids (perfluoroheptanoic acid (PFHpA), perfluorooctanoic acid (PFOA), perfluorononanoic acid (PFNA), perfluorodecanoic acid (PFDA), perfluoroundecanoic acid (PFUnDA), perfluorododecanoic acid (PFDoDA), perfluorohexane sulfonate (PFHxS) and perfluorooctane sulfonate (PFOS)) as
described in SI and publications therein.

Health examinations

For a measure of metabolism and growth in the children, we used birth weight data from the Swedish national birth register. For a measure of neurodevelopment we used data from a routinely made language screening of the children when they were 30 months old. Language development was assessed by nurse's evaluation and parental questionnaire, including the number of words the child used (<25, 25-50 and >50). A main study outcome was parental report of use of fewer than 50 words, termed language delay (LD) corresponding to a prevalence of 10%.

Biostatistical analyses

Weighted quantile sum (WQS) regression[9], adjusted for covariates, was used to establish associations between mixture exposures and lower birth weight or language delay in children (see SI). In short, WQS regression is a strategy for estimating empirical weights for a weighted sum of concentrations most associated with the health outcome. The results are a beta coefficient associated with the weighted sum (estimate, SE and p value) and the empirical weights (which are the average weights,

constrained to sum to 1, from an ensemble step - here, 100 bootstrap samples)). The components most associated with the health outcomes have non-negligible weights, and were treated as sEDCs when the estimated weight exceeds the equi-weight case. . Next, we estimated the equivalent daily intake (DI) of sEDCs measured in the urine (i.e., phthalates and alkyl phenols), and estimated serum concentrations from the DI for these urinary measurement-based compounds (see SI). Finally, we used the geometric means, on a molar basis, for either the measured (PFAAs) or estimated serum levels (phthalates and alkyl phenols) and established mixing proportions to prepare one mixture associated with low birth weight (MIX G) and one associated with language delay (MIX N). These two mixtures were used in the experimental studies.

Similar Mixture Approach (SMACH)

We conducted a 'similar mixture approach' (SMACH) using the SELMA pregnancy cohort linking human exposures with estimates of adverse outcomes in zebrafish and xenopus. In short, we modeled the total distance each zebrafish swam using 4 connected 10-minute flexible curves fit using truncated power cubic splines (Figure A). Characteristics of these curves were evaluated across zebrafish as a function of exposure concentrations. Herein, the average maximum distance across the 4 cycles was modeled to reflect the change in swimming patterns due to exposure to Mixture N. We used a piecewise nonlinear model to approximate the concentration-response relationship (Figure B). For the XETA, Results from three individual experiments were combined with the T3 condition as control. The read-out was a normalized level of fluorescence – i.e., where the control mean is 100%. A piecewise model was used to approximate the concentration-response relationship (Figure B).

A benchmark dose (BMD) from each assay was estimated and its lower one-sided 95% confidence limit (BMDL) calculated for use in the 'similar mixture risk index' (SMRI). The benchmark response (BMR) was based on either established guideline values (i.e., XETA) or 1 STD decrease from the control mean (i.e., zebrafish average maximum distance). The SMACH was conducted by first testing whether the concentrations in SELMA pregnant women were sufficiently similar to the domain-specific reference mixture (Marshall et al, 2015). For those determined to be sufficiently similar, the SMRI was calculated – i.e., the sum of concentrations relative to the estimated BMDL from experimental studies. The percentages of women with SMRI exceeding 1 were tabulated to provide a level of concern. Finally, adjusted models were used to determine if children associated with the highest and lowest decile of SMRI were significantly different in language delay for MIX N.

Composition of the mixtures

The chemicals needed for the mixtures were obtained from commercial custom synthesis laboratories or vendors BPA, Dimethylsulfoxide (DMSO), MBzP, PFHxS, PFNA, PFOA and PFOS were obtained from Sigma-Aldrich
Inc. (St. Louis, MO, USA). Triclosan was purchased from Dr. Ehrenstorfer (Augsburg), MEP and MiNP were obtained from Toronto Research Chemicals (North York, ON, Canada). MBP and MEHP were purchased from TCI, Tokyo Chemical Industry Co., Ltd (Japan). For MIX N, 1M solutions in DMSO were prepared using, MEP, MBP, MBzP, MiNP, BPA, PFHxS, PFNA, PFOS. For MIX G, 1M solutions in DMSO were prepared using, MEP, MBP, MBzP, MEHP, MINP, Triclosan, PFHxS, PFOA, and PFOS.

Generation of iPSC lines tagged with mCherry or GFP

To permanently tag 2 different pluripotent cell lines the following commercial lentiviral 3rd generation vectors were used:

Plasmid CD550A-1 (system biosciences) was used to tag the hiPSC line 2 with a cytoplasmic eGFP tag under a GSK promoter.

Plasmid pSicoR-EF1α-mCherry-puro (Trono Lab) was used to tag the hiPSC line 4 with a cytoplasmic mCherry tag under an EF1α promoter.

Lentiviral vectors preparation

3rd generation packaging plasmids were used to generate lentiviral vectors. Plasmid-bearing bacterial strains were grown overnight in Luria-Bertani broth supplemented with 1 ug/mL ampicillin. The bacterial culture has been kept in a standard incubator at 37° C, 5% CO2 under shaking conditions. Plasmids were extracted and purified with NucleoBond Maxi kit (Macherey-Nagel) following manufacturer specifications and quantified using a nanodrop spectrophotometer.

To verify plasmid purity 1ug of each plasmid has been digested with restriction enzyme EcoIII (New England BioLabs) following the protocol recommended from the enzyme's manufacturer. The digested products were run for 1 hour in a 1,25% agarose gel in presence of 2 ng/ml SYBR safe dye (ThermoFischer) and then visualized under a UV lamp.

The preparation of lentiviral vectors bearing the transgenic constructs has been performed according to the protocol published by Dull and colleagues (Dull et al., 1998).

Lentiviral infection and antibiotic selection

The 2 previously mentioned hiPSCs lines were dissociated with acutase and plated in 6 wells at a concentration of $2x10^4$ cells/well in mTeSR medium supplemented with ROCK inhibitor 5uM (Clinisciences).

The day after, media was changed to remove ROCK inhibitor and 10, 7, 4, 1 or 0,1 uL of lentiviral suspension were added to 5 wells of the 6-wells plate where the specific hiPSC line was growing, to assess the best-working viral concentration. The untreated well served as a control for selection. Cells were then grown for 48 hours performing daily media changes to allow the integration and expression of the viral transgene. From the third day after infection cells were treated for 5 days with puromycin (1ug/mL) to select infected cells. Puromycin titration was not necessary as the working concentration was already assessed and, after 24 hours, all the cells of the negative control were dead. When almost confluent, wells were inspected under a fluorescent

microscope to visually select the cell population infected with the lowest viral concentration able to give fluorescent cells. The selected cell population was detached with ReLeSR and expanded for a minimum of 3 passages before further utilization.
For both lines 1 uL of the mCherry or GFP lentiviral suspension was enough to obtain a homogeneously detectable signal of the fluorescent transgene.

Multiplexing strategy

In our experimental design we adopted two distinct multiplexing strategies, namely upstream and downstream, that differ for the moment in which the different cell lines have been mixed.

In the downstream approach we generated brain organoids by taking each of the four hiPSC lines separately to obtain 3D structure containing only one specific genotype. These organoids have been grown in separate dishes according to their cell line of origin. Upon reaching the timepoint, organoids derived from different cell lines were dissociated separately, the cells counted and $2{,}5x10^5$ cells was taken from each genotypically diverse cell suspension and mixed together in the same vial. Due to the massive presence of cell debris that interfered with cell counting procedures, we considered total count a representative counting of our samples. Volume was subsequently adjusted to reach a concentration of $1x10^6$ cells/mL.

In the upstream approach instead, we generated brain organoids by mixing equal amounts of PSCs derived from each cell line to obtain the so-called chimeric brain organoids. Briefly, after dissociation of hiPSc colonies, cells were counted and mixed in equal proportions, namely 25% from each line) to reach a global amount of cells of $2{,}4x10^6$ which in our case represented the number of cells needed to seed an entire 96MW of embryoid bodies. The resulting cell suspension was subsequently diluted to reach the working concentration and organoids were generated as explained in the following chapter. A comprehensive summary of the models used for the different experiments can be found in table 3, at the end of "materials and methods" section

Chimeric brain organoids generation

Pure lines-derived and chimeric brain organoids were generated using a slightly modified version of the protocol developed by Pasca and collegues (Paşca et al., 2015), which allows to obtain dorsal telencephalon cortical organoids. The cell lines used in this project are shown in Table 4. To generate organoids hiPSC were expanded in 10 cm plates following the general rules previously described. When the most abundant hiPSC line reached 80% confluency, colonies were dissociated with accutase and centrifuged to remove the enzymatic suspension. After resuspension in TeSR/E8

medium supplemented with 5uM ROCK inhibitor cells were counted with a TC20 automatic cell counter (Biorad) and equal amount of each cell line were mixed at the final concentration of 2 x 10^5 cells/mL.

100 uL/well of cell suspension were seeded into PrimeSurface^R 96 well plates (SystemBio) and then the plates spinned at 850 rpm for 3 minutes to promote the formation of embryoid bodies. The utilization of ultra-low attachment plates is fundamental to avoid cell adhesion. The day of the EB generation is referred to as day -2. On day -1, medium is not changed, leaving EBs undisturbed. On day 0 media change is performed adding differentiating medium one, composed of 80% DMEM/F12 medium (1:1), 20% Knockout serum (Gibco), 1 mM Non-essential amino acids (Sigma), 0.1 mM cell culture grade 2-mercaptoethanol solution (Gibco), GlutaMax (Gibco, 1:100), penicillin 100 U/mL, streptomycin (100 µg/mL), 7 uM Dorsomorphin (Sigma) and 10 uM TGFβ inhibitor SB431542 (MedChem express). Dorsomorphin and TGFβ inhibitor are used to perform DUAL-SMAD inhibition, promoting the induction of neuroectoderm.

From day 0 to day 4, media change is performed every day with the very same medium while on day 5 differentiation medium two is added until day 24 with daily media change for the first 7 days.

Differentiation medium two is composed of neurobasal medium (Gibco) supplemented with B-27 supplement w/o vitamin A (Gibco, 1:50), GlutaMax (1:100), penicillin 100 U/mL, streptomycin (100 µg/mL), 20 ng/mL FGF2 (Thermo) and 20 ng/mL EGF (Thermo).

On day 12 organoids are moved to 10cm ultra low attachment dishes and grown on shakers to enhance oxygen and nutrient supply. From day 12 onwards media change is performed every other day. On day 24, FGF2 and EGF are replaced with 20 ng/mL brain-derived neurotrophic factor (BDNF, Thermo) and 20 ng/mL neurotrophin-3 (NT3, Thermo) to promote differentiation of neural progenitors. From day 42 onwards, grow factors are removed.

The volume of medium needed for each plate vary according to the size and number of organoids per plate as it must be sufficient to ensure they are completely covered. As a general rule, media volume increases while proceeding with differentiation, reaching a maximum of approximately 35 mL/plate.

Timepoints chosen for analysis were day 50 (+/- 3 days), day 100 (+/- 2 days) and day 150 (+/- 5 days)

Proliferation assay

The 6 lines of PSCs were simultaneously dissociated at cell level with accutase, counted and resuspended in TeSR/E8 medium supplemented with 5uM ROCK inhibitor.

Meanwhile, 4 different 24 wells (one per each timepoint) were coated with matrigel solution 1:40 in DMEM/F12 and the PSC lines was seeded in four replicates per line on each plate at a concentration of 10 000 cells/well. Negative control was obtained by coating 4 wells with matrigel and adding the same amount of medium without cells. Analytical timepoints were set at 3, 27, 51 and 75 hours after plating. Media change was performed daily at the same time starting from 24 hours after seeding. ROCK inhibitor was removed since the first media change.

To evaluate proliferation CellTiter 96^{R} aqueous one solution (Promega) was used. This ready-to-use solution contains a novel type of tetrazolium salt, MTS, that is oxidized by mitochondrial succinate DH to produce a formazan salt. The enzyme is active only in live cells and thus formazan production is directly dependent upon the total number of live cells. MTS-derived formazan has an absorbance peak around 490 nm which allows its measurement through conventional spectrophotometers. Briefly, at each timepoint 20 uL of MTS solution per well were added to 100 uL of fresh TeSR/E8 media, according to manufacturer specifications. The plate was incubated for 60 minutes at 37 °C, 3% O_2, 5% CO_2 and right after absorbance was measured at λ=490nm using a GlowMax spectrophotometer.

STR analysis

Chimeric brain organoids were collected after 26 days of differentiation taking one single organoid for each replicate to verify the presence in each organoid of all the original genotypes mixed. For genomic DNA extraction the Dneasy® Blood & Tissue Kit (Qiagen) was used according to manufacturer instructions. DNA was eluted in 50 µL of buffer AE. Nanodrop spectrophotometer was used to quantify genomic DNA. 1 µL of deionized water was used to initialize the machine. The program for dsDNA was selected and then 1 µL of elution buffer was used as blank. Samples were quantified by the instrument measuring the

absorbance at 260 nm. 10ng of DNA from each sample were then processed with the GenePrint^{R} 10 system (Promega) to amplify 8 different STR loci according to quality ATCC guidelines. Detection of STR amplicons was performed using the Applied Biosystem^{R} 3500 genetic analyzer using POP-4 polymer as recommended in the GenePrint^{R} 10 system protocol.

Organoids live imaging

Brain organoids were collected after 200 days of differentiation, rinsed with sterile PBS and embedded in 3% low melting agarose for vibratome cutting. Slices were cut 100 um using a vibratome and transferred in 24 wells with glass bottom in Neurobasal medium. The day after media was removed and slices were incubated for 30 minutes in PBS solution containing 5 ug/mL Hoechst 33258 (SantaCruz) to stain nuclei. After two 5 minutes rinses with sterile PBS to remove unbound dye, media was added again to allow cell survival and live imaging was performed straightaway using a confocal microscope

Organoids paraffinization

Organoids were harvested on day 50, 100 and 150, fixed overnight at 4 °C in paraformaldehyde 4%/PBS solution (SantaCruz). After rinsing with PBS, organoids were embedded in 2% low melting agarose dissolved in PBS to facilitate inclusion. After agarose solidification, blocks were put in 70% ethanol and kept at 4 °C before sending to the facility for paraffin embedding, sectioning and routine haematoxylin/eosin staining.

Organoids immunofluorescence

Deparaffinization and rehydration was achieved by consecutive passages of 5 minutes each in the following solutions: 2 x histolemon, 100% ethanol, 95% ethanol, 80% ethanol and water. Sections were then incubated for 45 min at 95 °C with 10mM Sodium citrate buffer (Normapur) + Tween 20 0,05% (Sigma) for simultaneous antigen retrieval and permeabilization; then left to cool for at least 2 hours at RT. For GFP staining, antigen retrieval was not performed since it disrupts the epitope/s recognized by all the 3 available commercial anti-GFP antibodies. After 30 minutes incubation in blocking solution (5% normal donkey serum in PBS), Primary antibodies were prepared in PBS + 5% normal donkey serum overnight at 4 °C. The day after, secondary antibodies were diluted in PBS and applied to the sections for 1 hour. DAPI was added for 5 minutes at room temperature. After each incubation 3 x 5 minutes washing steps with TBS buffer were performed. After a final rinse in deionized water, slides were dried and mounted using Mowiol mounting media.

The following primary antibodies were used: Pax6 (Rabbit, 1:200, Biolegend), Tuj1 (Mouse, 1:1000, Biolegend), Nestin (Mouse, 1:500, Millipore), GFP (Rabbit, 1:200, Evrogen), mCherry (Mouse, 1:200, Thermo), Reelin (Mouse, 1:400, Millipore), CTIP2 (Rat, 1:400, Abcam). Secondary antibodies were taken from Campus technical kitchen and all diluted 1:300 in PBS. Images were acquired using a standard

fluorescence microscope (Leica) at 20X, 40X and 63X magnification and then processed with FIJI software.

Total RNA extraction and bulk sequencing, single cell transcriptomics and analysis as detailed before

Demultiplexing

Alignment of RNA-seq reads on reference Genome (GRCh38-1.2.0 was used) was performed using STAR 2.6.1 aligner. During this phase, A first Index was generated starting from reference fasta sequence and then the first alignment cycle (per sample) has been performed in order to detect possible exon junctions splitting the reads alignment. At this point a second index has been generated using the gathered informations in the previous alignment. Finally, the actual alignment of row reads was carried out Generating an unsorted BAM file for every sample multiplexed.

For the Variant Calling we must specify we used GATK suite starting from bulk transcriptomic data obtained from each pure line, which to date has not a validated pipeline for joint genotyping of sample cohorts that was our final purpose. Therefore, we proceeded modifying the original DNA-seq joint variant call pipeline by introducing data pre-processing of RNA-seq method. Then, we applied a raw filtering on discovered variants based on the number of SNP detected within specific bp windows (at least 3 SNP in 35 BP windows). This approach led us to the production of gVCF files containing variants with a quality score distribution different and less accurate than the one used in the implementation of Demuxlet which was obtained from standard whole genome sequencing (Kang et al., 2018b).

In the end, genetic identity deconvolution was performed on the filtered dataset using Demuxlet algorithm. Demuxlet was ran with standards and GT field of the VCF file used as reference value for the Identity Call.

Bibliography

(EFSA), E. F. S. A. (2008). Perfluorooctane sulfonate (PFOS), perfluorooctanoic acid (PFOA) and their salts Scientific Opinion of the Panel on Contaminants in the Food chain. *EFSA Journal*, *6*(7). https://doi.org/10.2903/j.efsa.2008.653

Aberle, H., Bauer, A., Stappert, J., Kispert, A., & Kemler, R. (1997). β-catenin is a target for the ubiquitin–proteasome pathway. *The EMBO Journal*, *16*(13), 3797–3804. https://doi.org/10.1093/emboj/16.13.3797

Abraham, T. H. (2016). *Rebel genius : Warren S. McCulloch's transdisciplinary life in science*. Retrieved from https://mitpress.mit.edu/books/rebel-genius

Abud, E. M., Ramirez, R. N., Martinez, E. S., Healy, L. M., Nguyen, C. H. H., Newman, S. A., … Blurton-Jones, M. (2017). iPSC-Derived Human Microglia-like Cells to Study Neurological Diseases. *Neuron*. https://doi.org/10.1016/j.neuron.2017.03.042

Adamo, A., Atashpaz, S., Germain, P.-L., Zanella, M., D'Agostino, G., Albertin, V., … Testa, G. (2015). 7q11.23 dosage-dependent dysregulation in human pluripotent stem cells affects transcriptional programs in disease-relevant lineages. *Nat. Genet.*, *47*(2), 132–141. https://doi.org/10.1038/ng.3169

Alfano, C., & Studer, M. (2013). Neocortical arealization: Evolution, mechanisms, and open questions. *Developmental Neurobiology*, *73*(6), 411–447. https://doi.org/10.1002/dneu.22067

Alvarez, M. J., Shen, Y., Giorgi, F. M., Lachmann, A., Ding, B. B., Ye, B. H., & Califano, A. (2016). Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nature Genetics*, *48*(8), 838–847. https://doi.org/10.1038/ng.3593

Anney, R., Klei, L., Pinto, D., Almeida, J., Bacchelli, E., Baird, G., … Devlin, B. (2012). Individual common variants exert weak effects on the risk for autism spectrum disorders. *Human Molecular Genetics*, *21*(21), 4781–4792. https://doi.org/10.1093/hmg/dds301

Atlasi, Y., & Stunnenberg, H. G. (2017). The interplay of epigenetic marks during stem cell differentiation and development. *Nat. Rev. Genet.*, *18*(11), 643–658. https://doi.org/10.1038/nrg.2017.57

Autism-SNPedia. (n.d.). Retrieved from https://www.snpedia.com/index.php/Autism

Baccarelli, A., & Bollati, V. (2009). Epigenetics and environmental chemicals. *Current Opinion in Pediatrics*, *21*(2), 243–251. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19663042

Bagley, J. A., Reumann, D., Bian, S., Lévi-Strauss, J., & Knoblich, J. A. (2017). Fused cerebral organoids model interactions between brain regions. *Nature Methods*, *14*(7), 743–751. https://doi.org/10.1038/nmeth.4304

Bain, G., Kitchens, D., Yao, M., Huettner, J. E., & Gottlieb, D. I. (1995). Embryonic stem cells express neuronal properties in vitro. *Developmental Biology*. https://doi.org/10.1006/dbio.1995.1085

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., … Dowling, N. F. (2018). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR. Surveillance Summaries*, *67*(6), 1–23. https://doi.org/10.15585/mmwr.ss6706a1

Banovich, N. E., Li, Y. I., Raj, A., Ward, M. C., Greenside, P., Calderon, D., … Gilad, Y. (2018). Impact of regulatory variation across human iPSCs and differentiated cells. *Genome Research*, *28*(1), 122–131. https://doi.org/10.1101/gr.224436.117

Barbagiovanni, G., Germain, P.-L., Zech, M., Atashpaz, S., Lo Riso, P., D'Antonio-Chronowska, A., … Testa, G. (2018). KMT2B Is Selectively Required for Neuronal Transdifferentiation, and Its Loss Exposes Dystonia Candidate Genes. *Cell Rep.*, *25*(4), 988–1001. https://doi.org/10.1016/j.celrep.2018.09.067

Barouki, R. (2017). Endocrine disruptors: Revisiting concepts and dogma in toxicology. *Comptes Rendus Biologies*, *340*(9–10), 410–413. https://doi.org/10.1016/J.CRVI.2017.07.005

Barouki, R., Gluckman, P. D., Grandjean, P., Hanson, M., & Heindel, J. J. (2012). Developmental origins of non-communicable disease: Implications for research and public health. *Environmental Health*, *11*(1), 42. https://doi.org/10.1186/1476-069X-11-42

Barton, N. H., Etheridge, A. M., & Véber, A. (2017). The infinitesimal model: Definition, derivation, and implications. *Theoretical Population Biology*, *118*, 50–73. https://doi.org/10.1016/J.TPB.2017.06.001

Beam, A. L., & Kohane, I. S. (2018). Big Data and Machine Learning in Health Care. *JAMA*, *319*(13), 1317. https://doi.org/10.1001/jama.2017.18391

Beausoleil, C., Beronius, A., Bodin, L., Bokkers, B. G. H., Boon, P. E., Burger, M., … Zilliacus, J. (2016). Review of non-monotonic dose-responses of substances for human risk assessment. *EFSA Supporting Publications*, *13*(5). https://doi.org/10.2903/sp.efsa.2016.EN-1027

Bedrosian, T. A., Quayle, C., Novaresi, N., & Gage, F. H. (2018). Early life experience drives structural variation of neural genomes in mice. *Science (New York, N.Y.)*, *359*(6382), 1395–1399. https://doi.org/10.1126/science.aah3378

Bergman Heindel, J.J, Jobling, S., Kidd, K.A., Zoeller, R.T., A. (2013). State of the Science of Endocrine Disrupting Chemicals. *World Health Organization*, 296. Retrieved from http://www.who.int/iris/bitstream/10665/78101/1/9789241505031_eng.pdf

Bershteyn, M., Nowakowski, T. J., Pollen, A. A., Di Lullo, E., Nene, A., Wynshaw-Boris, A., & Kriegstein, A. R. (2017). Human iPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia. *Cell Stem Cell*, *20*(4), 435-449.e4. https://doi.org/10.1016/j.stem.2016.12.007

Birey, F., Andersen, J., Makinson, C. D., Islam, S., Wei, W., Huber, N., … Paşca, S. P. (2017). Assembly of functionally integrated human forebrain spheroids.

Nature, *545*(7652), 54–59. https://doi.org/10.1038/nature22330

Bitcoin: A Peer-to-Peer Electronic Cash System. (n.d.). Retrieved from https://bitcoin.org/bitcoin.pdf

Bitta, M., Kariuki, S. M., Abubakar, A., & Newton, C. R. J. . (2018). Burden of neurodevelopmental disorders in low and middle-income countries: A systematic review and meta-analysis. *Wellcome Open Research*, *2*, 121. https://doi.org/10.12688/wellcomeopenres.13540.3

Boas, M., Feldt-Rasmussen, U., & Main, K. M. (2012). Thyroid effects of endocrine disrupting chemicals. *Mol. Cell. Endocrinol.*, *355*(2), 240–248. https://doi.org/10.1016/j.mce.2011.09.005

Boettcher, M., Tian, R., Blau, J. A., Markegard, E., Wagner, R. T., Wu, D., … McManus, M. T. (2018). Dual gene activation and knockout screen reveals directional dependencies in genetic networks. *Nature Biotechnology*, *36*(2), 170. https://doi.org/10.1038/NBT.4062

Bornehag, C.-G., Reichenberg, A., Hallerback, M. U., Wikstrom, S., Koch, H. M., Jonsson, B. A., & Swan, S. H. (2018). Prenatal exposure to acetaminophen and children's language development at 30 months. *European Psychiatry*, *51*, 98–103. https://doi.org/10.1016/j.eurpsy.2017.10.007

Bornehag, C. G., & Moniruzzaman - Paediatric and …, S. (2012). The SELMA study: a birth cohort study in Sweden following more than 2000 mother–child pairs. *Wiley Online Library*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1365-3016.2012.01314.x/full

Bornehag, Carl-Gustaf, Lundgren, B., Weschler, C. J., Sigsgaard, T., Hagerhed-Engman, L., & Sundell, J. (2005). Phthalates in Indoor Dust and Their Association with Building Characteristics. *Environmental Health Perspectives*, *113*(10), 1399–1404. https://doi.org/10.1289/ehp.7809

Boukhris, T., Sheehy, O., Mottron, L., & Bérard, A. (2016). Antidepressant Use During Pregnancy and the Risk of Autism Spectrum Disorder in Children. *JAMA Pediatrics*, *170*(2), 117. https://doi.org/10.1001/jamapediatrics.2015.3356

Bouschet, T., Dubois, E., Reynès, C., Kota, S. K., Rialle, S., Maupetit-Méhouas, S., … Varrault, A. (2017). In Vitro Corticogenesis from Embryonic Stem Cells Recapitulates the In Vivo Epigenetic Control of Imprinted Gene Expression. *Cereb. Cortex*, *27*(3), 2418–2433. https://doi.org/10.1093/cercor/bhw102

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*, *169*(7), 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038

Braun, Joe M, Just, A. C., Williams, P. L., Smith, K. W., Calafat, A. M., & Hauser, R. (2014). Personal care product use and urinary phthalate metabolite and paraben concentrations during pregnancy among women from a fertility clinic. *Journal of Exposure Science & Environmental Epidemiology*, *24*(5), 459–466. https://doi.org/10.1038/jes.2013.69

Braun, Joseph M. (2017). Early Life Exposure to Endocrine Disrupting Chemicals and Childhood Obesity and Neurodevelopment. *Nature Reviews. Endocrinology*, *13*(3), 161. https://doi.org/10.1038/NRENDO.2016.186

Brenner, S. (2010). Sequences and consequences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1537), 207–212. https://doi.org/10.1098/rstb.2009.0221

Brynedal, B., Choi, J., Raj, T., Bjornson, R., Stranger, B. E., Neale, B. M., … Cotsapas, C. (2017). Large-Scale trans-eQTLs Affect Hundreds of Transcripts and Mediate Patterns of Transcriptional Co-regulation. *American Journal of Human Genetics*, *100*(4), 581–591. https://doi.org/10.1016/j.ajhg.2017.02.004

Burgess, D. J. (2019). Spatial transcriptomics coming of age. *Nature Reviews Genetics*, *20*(6), 317–317. https://doi.org/10.1038/s41576-019-0129-z

Bystron, I., Blakemore, C., & Rakic, P. (2008). Development of the human cerebral cortex: Boulder Committee revisited. *Nature Reviews Neuroscience*, Vol. 9, pp. 110–122. https://doi.org/10.1038/nrn2252

Califf, R. M., Robb, M. A., Bindman, A. B., Briggs, J. P., Collins, F. S., Conway, P. H., … Sherman, R. E. (2016). Transforming Evidence Generation to Support Health and Health Care Decisions. *N. Engl. J. Med.*, *375*(24), 2395–2400. https://doi.org/10.1056/NEJMsb1610128

Calvo, R. M., Roda, J. M., Obregón, M. J., & de Escobar, G. M. (1998). Thyroid hormones in human tumoral and normal nervous tissues. *Brain Research*, *801*(1–2), 150–157. https://doi.org/10.1016/S0006-8993(98)00576-9

Caporale, N., & Testa, G. (2019). At the Intersection of Epigenetics and Regeneration: An Analysis of the Experimental Outlook of Organoid Technology. In *Epigenetics and Regeneration*. https://doi.org/10.1016/b978-0-12-814879-2.00017-0

Carrico, C., Gennings, C., Wheeler, D. C., & Factor-Litvak, P. (2015). Characterization of Weighted Quantile Sum Regression for Highly Correlated Data in a Risk Analysis Setting. *JABES*, *20*(1), 100–120. https://doi.org/10.1007/s13253-014-0180-3

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., … Werner, T. (2005). MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*, *21*(13), 2933–2942. https://doi.org/10.1093/bioinformatics/bti473

Cederquist, G. Y., Asciolla, J. J., Tchieu, J., Walsh, R. M., Cornacchia, D., Resh, M. D., & Studer, L. (2019). Specification of positional identity in forebrain organoids. *Nature Biotechnology*, *37*(4), 436–444. https://doi.org/10.1038/s41587-019-0085-3

Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M., & Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature Biotechnology*. https://doi.org/10.1038/nbt.1529

Chappell, L., Russell, A. J. C., & Voet, T. (2018). Single-Cell (Multi)omics Technologies. *Annual Review of Genomics and Human Genetics*, *19*(1), 15–41. https://doi.org/10.1146/annurev-genom-091416-035324

Chatonnet, F., Flamant, F., & Morte, B. (2015). A temporary compendium of thyroid hormone target genes in brain. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1849*(2), 122–129. https://doi.org/10.1016/J.BBAGRM.2014.05.023

Chowkwanyun, M., Bayer, R., & Galea, S. (2018). "Precision" Public Health — Between Novelty and Hype. *New England Journal of Medicine*, *379*(15), 1398–1400. https://doi.org/10.1056/NEJMp1806634

Chowkwanyun, M., Bayer, R., & Galea, S. (2019). Precision public health: pitfalls and promises. *The Lancet*, *393*(10183), 1801. https://doi.org/10.1016/S0140-6736(18)33187-8

Christensen, J., Grønborg, T. K., Sørensen, M. J., Schendel, D., Parner, E. T., Pedersen, L. H., & Vestergaard, M. (2013). Prenatal Valproate Exposure and Risk of Autism Spectrum Disorders and Childhood Autism. *JAMA*, *309*(16), 1696. https://doi.org/10.1001/jama.2013.2270

Clark, S. J., Lee, H. J., Smallwood, S. A., Kelsey, G., & Reik, W. (2016). Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. *Genome Biol.*, *17*, 72. https://doi.org/10.1186/s13059-016-0944-x

Clowry, G. J., Alzu'bi, A., Harkin, L. F., Sarma, S., Kerwin, J., & Lindsay, S. J. (2018). Charting the protomap of the human telencephalon. *Seminars in Cell & Developmental Biology*, *76*, 3–14. https://doi.org/10.1016/j.semcdb.2017.08.033

Consortium, T. B., Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., … Neale, B. M. (2018). Analysis of shared heritability in common disorders of the brain. *Science*, *360*(6395), eaap8757. https://doi.org/10.1126/SCIENCE.AAP8757

Cuomo, A. S., Seaton, D. D., McCarthy, D. J., Martinez, I., Bonder, M. J., Garcia-Bernardo, J., … Stegle, O. (2019). Single-cell RNA-sequencing of differentiating iPS cells reveals dynamic genetic effects on gene expression. *BioRxiv*. https://doi.org/10.1101/630996

Data & Statistics on Autism Spectrum Disorder | CDC. (n.d.). Retrieved September 9, 2019, from https://www.cdc.gov/ncbddd/autism/data.html

Datlinger, P., Rendeiro, A. F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., … Bock, C. (2017). Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods*, *14*(3), 297–301. https://doi.org/10.1038/nmeth.4177

dbGaP PsychENCODE Consortium. (n.d.). Retrieved from https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001220.v1.p1&phv=310069&phd=&pha=&pht=6752&phvf=&phdf=&phaf=&phtf=&dssp=1&consent=&temp=1

De Rubeis, S., He, X., Goldberg, A. P., Poultney, C. S., Samocha, K., Cicek, A. E.,

… Buxbaum, J. D. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*, *515*(7526), 209–215. https://doi.org/10.1038/nature13772

Deciphering Developmental Disorders Study. (2015). Large-scale discovery of novel genetic causes of developmental disorders. *Nature*, *519*(7542), 223–228. https://doi.org/10.1038/nature14135

Derksen, F., Bensing, J., & Lagro-Janssen, A. (2013). Effectiveness of empathy in general practice: a systematic review. *The British Journal of General Practice : The Journal of the Royal College of General Practitioners*, *63*(606), e76-84. https://doi.org/10.3399/bjgp13X660814

Di Lullo, E., & Kriegstein, A. R. (2017). The use of brain organoids to investigate neural development and disease. *Nature Reviews Neuroscience*, Vol. 18, pp. 573–584. https://doi.org/10.1038/nrn.2017.107

Dick, D. M., Agrawal, A., Keller, M. C., Adkins, A., Aliev, F., Monroe, S., … Sher, K. J. (2015). Candidate Gene–Environment Interaction Research. *Perspectives on Psychological Science*, *10*(1), 37–59. https://doi.org/10.1177/1745691614556682

Dillon, J. J., DeSimone, C. V., Sapir, Y., Somers, V. K., Dugan, J. L., Bruce, C. J., … Friedman, P. A. (2015). Noninvasive potassium determination using a mathematically processed ECG: Proof of concept for a novel "blood-less, blood test." *Journal of Electrocardiology*, *48*(1), 12–18. https://doi.org/10.1016/j.jelectrocard.2014.10.002

Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C. P., Jerby-Arnon, L., … Regev, A. (2016). Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell*, *167*(7), 1853-1866.e17. https://doi.org/10.1016/j.cell.2016.11.038

Dominguez-Salas, P., Moore, S. E., Baker, M. S., Bergen, A. W., Cox, S. E., Dyer, R. A., … Hennig, B. J. (2014). Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nature Communications*, *5*(1), 3746. https://doi.org/10.1038/ncomms4746

DSM-5. (2016). https://doi.org/10.1176/appi.pn.2016.7a7

Dull, T., Zufferey, R., Kelly, M., Mandel, R. J., Nguyen, M., Trono, D., & Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. *Journal of Virology*, *72*(11), 8463–8471.

Ecker, J. R., Geschwind, D. H., Kriegstein, A. R., Ngai, J., Osten, P., Polioudakis, D., … Zeng, H. (2017). The BRAIN Initiative Cell Census Consortium: Lessons Learned toward Generating a Comprehensive Brain Cell Atlas. *Neuron*, *96*(3), 542–557. https://doi.org/10.1016/j.neuron.2017.10.007

Ehrlich, M., Mozafari, S., Glatza, M., Starost, L., Velychko, S., Hallmann, A.-L., … Kuhlmann, T. (2017). Rapid and efficient generation of oligodendrocytes from human induced pluripotent stem cells using transcription factors. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1614412114

Elkabetz, Y., Panagiotakos, G., Al Shamy, G., Socci, N. D., Tabar, V., & Studer, L. (2008). Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. *Genes and Development*. https://doi.org/10.1101/gad.1616208

Engel, S. M., Miodovnik, A., Canfield, R. L., Zhu, C., Silva, M. J., Calafat, A. M., & Wolff, M. S. (2010). Prenatal Phthalate Exposure Is Associated with Childhood Behavior and Executive Functioning. *Environmental Health Perspectives*, *118*(4), 565–571. https://doi.org/10.1289/ehp.0901470

Eraslan, G., Avsec, Ž., Gagneur, J., & Theis, F. J. (2019). Deep learning: new computational modelling techniques for genomics. *Nature Reviews Genetics*, *20*(7), 389–403. https://doi.org/10.1038/s41576-019-0122-6

Esposito, G., Azhari, A., & Borelli, J. L. (2018). Gene × Environment Interaction in Developmental Disorders: Where Do We Stand and What's Next? *Frontiers in Psychology*, *9*, 2036. https://doi.org/10.3389/fpsyg.2018.02036

Fini, J.-B., Mughal, B. B., Le Mével, S., Leemans, M., Lettmann, M., Spirhanzlova, P., … Demeneix, B. A. (2017). Human amniotic fluid contaminants alter thyroid hormone signalling and early brain development in Xenopus embryos. *Sci. Rep.*, *7*, 43786. https://doi.org/10.1038/srep43786

Fisher, R. A. (1919). XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh*, *52*(2), 399–433. https://doi.org/10.1017/S0080456800012163

Florio, M., & Huttner, W. B. (2014). Neural progenitors, neurogenesis and the evolution of the neocortex. *Development*, *141*(11), 2182–2194. https://doi.org/10.1242/dev.090571

Florio, Marta, Borrell, V., & Huttner, W. B. (2017). Human-specific genomic signatures of neocortical expansion. *Current Opinion in Neurobiology*, *42*, 33–44. https://doi.org/10.1016/j.conb.2016.11.004

Frisch, M., Klocke, B., Haltmeier, M., & Frech, K. (2009). LitInspector: literature and signal transduction pathway mining in PubMed abstracts. *Nucleic Acids Res.*, *37*(Web Server issue), W135-40. https://doi.org/10.1093/nar/gkp303

Gabriele, M., Lopez Tobon, A., D'Agostino, G., & Testa, G. (2018). The chromatin basis of neurodevelopmental disorders: Rethinking dysfunction along the molecular and temporal axes. *Prog. Neuropsychopharmacol. Biol. Psychiatry*. https://doi.org/10.1016/j.pnpbp.2017.12.013

Gabriele, M., Vulto-van Silfhout, A. T., Germain, P.-L., Vitriolo, A., Kumar, R., Douglas, E., … de Vries, B. B. A. (2017). YY1 Haploinsufficiency Causes an Intellectual Disability Syndrome Featuring Transcriptional and Chromatin Dysfunction. *The American Journal of Human Genetics*, *100*(6), 907–925. https://doi.org/10.1016/j.ajhg.2017.05.006

Gamazon, E. R., Segrè, A. V., van de Bunt, M., Wen, X., Xi, H. S., Hormozdiari, F., … Ardlie, K. G. (2018). Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. *Nature Genetics*, *50*(7), 956–967. https://doi.org/10.1038/s41588-018-0154-4

Gameiro, G. R., Sinkunas, V., Liguori, G. R., & Auler-Júnior, J. O. C. (2018). Precision Medicine: Changing the way we think about healthcare. *Clinics (Sao Paulo, Brazil)*, *73*, e723. https://doi.org/10.6061/clinics/2017/e723

Gandal, M. J., Haney, J. R., Parikshak, N. N., Leppa, V., Ramaswami, G., Hartl, C., … Geschwind, D. H. (2018). Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science (New York, N.Y.)*, *359*(6376), 693–697. https://doi.org/10.1126/science.aad6469

Ganna, A., Verweij, K. J. H., Nivard, M. G., Maier, R., Wedow, R., Busch, A. S., … Zietsch, B. P. (2019). Large-scale GWAS reveals insights into the genetic architecture of same-sex sexual behavior. *Science*, *365*(6456), eaat7693. https://doi.org/10.1126/SCIENCE.AAT7693

Garcez, P. P., Loiola, E. C., Da Costa, R. M., Higa, L. M., Trindade, P., Delvecchio, R., … Rehen, S. K. (2016). Zika virus: Zika virus impairs growth in human neurospheres and brain organoids. *Science*, *352*(6287), 816–818. https://doi.org/10.1126/science.aaf6116

Gaugler, T., Klei, L., Sanders, S. J., Bodea, C. A., Goldberg, A. P., Lee, A. B., … Buxbaum, J. D. (2014). Most genetic risk for autism resides with common variation. *Nature Genetics*, *46*(8), 881–885. https://doi.org/10.1038/ng.3039

GENEX-FB2. (n.d.). Retrieved October 8, 2019, from https://github.com/hobrien/GENEX-FB2

Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. (2007). *Nature*, *447*(7145), 661–678. https://doi.org/10.1038/nature05911

Geschwind, D. H. (2011). Genetics of autism spectrum disorders. *Trends in Cognitive Sciences*, *15*(9), 409–416. https://doi.org/10.1016/j.tics.2011.07.003

Giandomenico, S. L., & Lancaster, M. A. (2017). Probing human brain evolution and development in organoids. *Curr. Opin. Cell Biol.*, *44*, 36–43. https://doi.org/10.1016/j.ceb.2017.01.001

Giandomenico, S. L., Mierau, S. B., Gibbons, G. M., Wenger, L. M. D., Masullo, L., Sit, T., … Lancaster, M. A. (2019). Cerebral organoids at the air–liquid interface generate diverse nerve tracts with functional output. *Nature Neuroscience*, *22*(4), 669–679. https://doi.org/10.1038/s41593-019-0350-2

Gjorevski, Nikolce, Sachs, N., Manfrin, A., Giger, S., Bragina, M. E., Ordóñez-Morán, P., … Lutolf, M. P. (2016). Designer matrices for intestinal stem cell and organoid culture. *Nature*, *539*(7630), 560–564. https://doi.org/10.1038/nature20168

Gjorevski, Nikolche, Ranga, A., & Lutolf, M. P. (2014). Bioengineering approaches to guide stem cell-based organogenesis. *Development*, *141*(9), 1794–1804. https://doi.org/10.1242/dev.101048

Global Research on Developmental Disabilities Collaborators, B. O., Davis, A. C., Wertlieb, D., Boo, N.-Y., Nair, M. K. C., Halpern, R., … Kassebaum, N. J. (2018). Developmental disabilities among children younger than 5 years in 195

countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet. Global Health*, *6*(10), e1100–e1121. https://doi.org/10.1016/S2214-109X(18)30309-7

Gorin, G., Svensson, V., & Pachter, L. (2019). RNA velocity and protein acceleration from single-cell multiomics experiments. *BioRxiv*, 658401. https://doi.org/10.1101/658401

Grimes, C. A., & Jope, R. S. (2001). The multifaceted roles of glycogen synthase kinase 3beta in cellular signaling. *Progress in Neurobiology*, *65*(4), 391–426. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11527574

Grishin, D., Obbad, K., & Church, G. M. (2019). Data privacy in the age of personal genomics. *Nature Biotechnology*, 1–3. https://doi.org/10.1038/s41587-019-0271-3

Grishin, D., Obbad, K., Estep, P., Quinn, K., Zaranek, S. W., Zaranek, A. W., … Church, G. (2018). Accelerating Genomic Data Generation and Facilitating Genomic Data Access Using Decentralization, Privacy-Preserving Technologies and Equitable Compensation. *1*. https://doi.org/10.30953/bhty.v1.34

GTEx Portal - eQTL dashboard. (n.d.). Retrieved October 8, 2019, from https://gtexportal.org/home/eqtlDashboardPage

Gunz, P., Neubauer, S., Maureille, B., & Hublin, J.-J. (2010). Brain development after birth differs between Neanderthals and modern humans. *Current Biology : CB*, *20*(21), R921-2. https://doi.org/10.1016/j.cub.2010.10.018

Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F., & Theis, F. J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nature Methods*, *13*(10), 845–848. https://doi.org/10.1038/nmeth.3971

Health, T. L. P. (2019). Next generation public health: towards precision and fairness. *The Lancet Public Health*, *4*(5), e209. https://doi.org/10.1016/S2468-2667(19)30064-7

Heaton, H., Talman, A. M., Knights, A., Imaz, M., Durbin, R., Hemberg, M., & Lawniczak, M. (2019). souporcell: Robust clustering of single cell RNAseq by genotype and ambient RNA inference without reference genotypes. *BioRxiv*, 699637. https://doi.org/10.1101/699637

Hegarty, J. P., Pegoraro, L. F. L., Lazzeroni, L. C., Raman, M. M., Hallmayer, J. F., Monterrey, J. C., … Hardan, A. Y. (2019). Genetic and environmental influences on structural brain measures in twins with autism spectrum disorder. *Molecular Psychiatry*, 1–11. https://doi.org/10.1038/s41380-018-0330-z

Homberg, J. R., Kyzar, E. J., Scattoni, M. L., Norton, W. H., Pittman, J., Gaikwad, S., … Kalueff, A. V. (2016). Genetic and environmental modulation of neurodevelopmental disorders: Translational insights from labs to beds. *Brain Research Bulletin*, *125*, 79–91. https://doi.org/10.1016/J.BRAINRESBULL.2016.04.015

Homberg, J. R., Kyzar, E. J., Stewart, A. M., Nguyen, M., Poudel, M. K., Echevarria, D. J., … Kalueff, A. V. (2016). Improving treatment of neurodevelopmental

disorders: recommendations based on preclinical studies. *Expert Opinion on Drug Discovery*, *11*(1), 11–25. https://doi.org/10.1517/17460441.2016.1115834

Horton, R. (2018). Offline: In defence of precision public health. *The Lancet*, *392*(10157), 1504. https://doi.org/10.1016/S0140-6736(18)32741-7

Huang, Y., McCarthy, D. J., & Stegle, O. (2019). Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. *BioRxiv*, 598748. https://doi.org/10.1101/598748

Hublin, J.-J., Neubauer, S., & Gunz, P. (2015). Brain ontogeny and life history in Pleistocene hominins. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1663), 20140062. https://doi.org/10.1098/rstb.2014.0062

Iakoucheva, L. M., Muotri, A. R., & Sebat, J. (2019). Getting to the Cores of Autism. *Cell*, *178*(6), 1287–1298. https://doi.org/10.1016/j.cell.2019.07.037

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., … Wigler, M. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, *515*(7526), 216–221. https://doi.org/10.1038/nature13908

Izumi, N., Fumoto, K., Izumi, S., & Kikuchi, A. (2008). GSK-3beta regulates proper mitotic spindle formation in cooperation with a component of the gamma-tubulin ring complex, GCP5. *The Journal of Biological Chemistry*, *283*(19), 12981–12991. https://doi.org/10.1074/jbc.M710282200

Jo, J., Xiao, Y., Sun, A. X., Cukuroglu, E., Tran, H.-D., Göke, J., … Ng, H.-H. (2016). Midbrain-like Organoids from Human Pluripotent Stem Cells Contain Functional Dopaminergic and Neuromelanin-Producing Neurons. *Cell Stem Cell*, *19*(2), 248–257. https://doi.org/10.1016/j.stem.2016.07.005

John, R. M., & Rougeulle, C. (2018). Developmental Epigenetics: Phenotype and the Flexible Epigenome. *Front Cell Dev Biol*, *6*, 130. https://doi.org/10.3389/fcell.2018.00130

Kain, R., Kahn, S., Thompson, D., Lewis, D., Barker, D., Bustamante, C., … Barry, D. (2019). Database shares that transform research subjects into partners. *Nature Biotechnology*, *37*(10), 1112–1115. https://doi.org/10.1038/s41587-019-0278-9

Kang, H. M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., McCarthy, E., … Ye, C. J. (2018a). Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.*, *36*(1), 89–94. https://doi.org/10.1038/nbt.4042

Kang, H. M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., McCarthy, E., … Ye, C. J. (2018b). Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nature Biotechnology*, *36*(1), 89–94. https://doi.org/10.1038/nbt.4042

Kawasaki, H., Mizuseki, K., Nishikawa, S., Kaneko, S., Kuwana, Y., Nakanishi, S., … Sasai, Y. (2000). Induction of midbrain dopaminergic neurons from ES cells

by stromal cell-derived inducing activity. *Neuron*. https://doi.org/10.1016/S0896-6273(00)00083-0

Kelava, I., & Lancaster, M. A. (2016). Stem Cell Models of Human Brain Development. *Cell Stem Cell*, *18*(6), 736–748. https://doi.org/10.1016/j.stem.2016.05.022

Kelsey, G., Stegle, O., & Reik, W. (2017). Single-cell epigenomics: Recording the past and predicting the future. *Science*, *358*(6359), 69–75. https://doi.org/10.1126/science.aan6826

Kester, M. H. A., Martinez de Mena, R., Obregon, M. J., Marinkovic, D., Howatson, A., Visser, T. J., … Morreale de Escobar, G. (2004). Iodothyronine Levels in the Human Developing Brain: Major Regulatory Roles of Iodothyronine Deiodinases in Different Areas. *The Journal of Clinical Endocrinology & Metabolism*, *89*(7), 3117–3128. https://doi.org/10.1210/jc.2003-031832

Kevin Mitchell. (2019, September 14). Beyond reductionism – systems biology gets dynamic. Retrieved September 16, 2019, from http://www.wiringthebrain.com/2019/09/beyond-reductionism-systems-biology.html?spref=tw

Kim, S., Lee, H., Chung, M., & Jeon, N. L. (2013). Engineering of functional, perfusable 3D microvascular networks on a chip. *Lab Chip*, *13*(8), 1489–1500. https://doi.org/10.1039/c3lc41320a

Kim, W.-Y., Wang, X., Wu, Y., Doble, B. W., Patel, S., Woodgett, J. R., & Snider, W. D. (2009). GSK-3 is a master regulator of neural progenitor homeostasis. *Nature Neuroscience*, *12*(11), 1390–1397. https://doi.org/10.1038/nn.2408

Klei, L., Sanders, S. J., Murtha, M. T., Hus, V., Lowe, J. K., Willsey, A., … Devlin, B. (2012). Common genetic variants, acting additively, are a major source of risk for autism. *Molecular Autism*, *3*(1), 9. https://doi.org/10.1186/2040-2392-3-9

Kopec, G., Shekhawat, P. S., & Mhanna, M. J. (2017). Prevalence of diabetes and obesity in association with prematurity and growth restriction. *Diabetes Metab. Syndr. Obes.*, *10*, 285–295. https://doi.org/10.2147/DMSO.S115890

Kortenkamp, A. (2014). Low dose mixture effects of endocrine disrupters and their implications for regulatory thresholds in chemical risk assessment. *Current Opinion in Pharmacology*, *19*, 105–111. https://doi.org/10.1016/j.coph.2014.08.006

Kyrousi, C., & Cappello, S. (2019). Using brain organoids to study human neurodevelopment, evolution and disease. *Wiley Interdisciplinary Reviews: Developmental Biology*, e347. https://doi.org/10.1002/wdev.347

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., … Kharchenko, P. V. (2018). RNA velocity of single cells. *Nature*, *560*(7719), 494–498. https://doi.org/10.1038/s41586-018-0414-6

Lancaster, M. A., Corsini, N. S., Wolfinger, S., Gustafson, E. H., Phillips, A. W., Burkard, T. R., … Knoblich, J. A. (2017). Guided self-organization and cortical

plate formation in human brain organoids. *Nature Biotechnology*, *35*(7), 659–666. https://doi.org/10.1038/nbt.3906

Lancaster, M. A., Renner, M., Martin, C.-A., Wenzel, D., Bicknell, L. S., Hurles, M. E., … Knoblich, J. A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature*, *501*(7467), 373–379. https://doi.org/10.1038/nature12517

Lee, D. H. (2018). Evidence of the Possible Harm of Endocrine-Disrupting Chemicals in Humans: Ongoing Debates and Key Issues. *Endocrinology and Metabolism (Seoul, Korea)*, *33*(1), 44–52. https://doi.org/10.3803/EnM.2018.33.1.44

Li, M., Pevny, L., Lovell-Badge, R., & Smith, A. (1998). Generation of purified neural precursors from embryonic stem cells by lineage selection. *Current Biology*. https://doi.org/10.1016/S0960-9822(98)70399-9

Li, Y. I., van de Geijn, B., Raj, A., Knowles, D. A., Petti, A. A., Golan, D., … Pritchard, J. K. (2016). RNA splicing is a primary link between genetic variation and disease. *Science*, *352*(6285), 600–604. https://doi.org/10.1126/science.aad9417

LIBD eQTL Browser. (n.d.). Retrieved October 8, 2019, from http://eqtl.brainseq.org/phase1/eqtl/

Liew, Z., Ritz, B., von Ehrenstein, O. S., Bech, B. H., Nohr, E. A., Fei, C., … Olsen, J. (2015). Attention Deficit/Hyperactivity Disorder and Childhood Autism in Association with Prenatal Exposure to Perfluoroalkyl Substances: A Nested Case–Control Study in the Danish National Birth Cohort. *Environmental Health Perspectives*, *123*(4), 367–373. https://doi.org/10.1289/ehp.1408412

Lifetime. (n.d.). Retrieved from https://lifetime-fetflagship.eu/

Linné, C. von. (2013). Caroli a Linné ... Systema naturae : per regna tria natura, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis /. In *Caroli a Linné ... Systema naturae : per regna tria natura, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis /*. https://doi.org/10.5962/bhl.title.68927

Loo, K. M. . van, & Martens, G. J. . (2007). Genetic and Environmental Factors in Complex Neurodevelopmental Disorders. *Current Genomics*, *8*(7), 429. https://doi.org/10.2174/138920207783591717

Loomans-Kropp, H. A., & Umar, A. (2019). Cancer prevention and screening: the next step in the era of precision medicine. *Npj Precision Oncology*, *3*(1), 3. https://doi.org/10.1038/s41698-018-0075-9

López-Tobón, A., Villa, C. E., Cheroni, C., Trattaro, S., Caporale, N., Conforti, P., … Testa, G. (2019). Human Cortical Organoids Expose a Differential Function of GSK3 on Cortical Neurogenesis. *Stem Cell Reports*, *0*(0). https://doi.org/10.1016/j.stemcr.2019.09.005

Lui, J. H., Hansen, D. V., & Kriegstein, A. R. (2011). Development and evolution of the human neocortex. *Cell*, Vol. 146, pp. 18–36.

https://doi.org/10.1016/j.cell.2011.06.030

Luo, C., Lancaster, M. A., Castanon, R., Nery, J. R., Knoblich, J. A., & Ecker, J. R. (2016). Cerebral Organoids Recapitulate Epigenomic Signatures of the Human Fetal Brain. *Cell Rep.*, *17*(12), 3369–3384. https://doi.org/10.1016/j.celrep.2016.12.001

Ma, Y.-C., Song, M.-R., Park, J. P., Henry Ho, H.-Y., Hu, L., Kurtev, M. V., … Greenberg, M. E. (2008). Regulation of Motor Neuron Specification by Phosphorylation of Neurogenin 2. *Neuron*, *58*(1), 65–77. https://doi.org/10.1016/j.neuron.2008.01.037

Maccari, S., Krugers, H. J., Morley-Fletcher, S., Szyf, M., & Brunton, P. J. (2014). The Consequences of Early-Life Adversity: Neurobiological, Behavioural and Epigenetic Adaptations. *Journal of Neuroendocrinology*, *26*(10), 707–723. https://doi.org/10.1111/jne.12175

Machines Learn a Biased View of Women | WIRED. (n.d.). Retrieved September 8, 2019, from https://www.wired.com/story/machines-taught-by-photos-learn-a-sexist-view-of-women/

Mackey, T. K., Kuo, T.-T., Gummadi, B., Clauson, K. A., Church, G., Grishin, D., … Palombini, M. (2019). 'Fit-for-purpose?' – challenges and opportunities for applications of blockchain technology in the future of healthcare. *BMC Medicine*, Vol. 17. https://doi.org/10.1186/s12916-019-1296-7

Madhavan, M., Nevin, Z. S., Shick, H. E., Garrison, E., Clarkson-Paredes, C., Karl, M., … Tesar, P. J. (2018). Induction of myelinating oligodendrocytes in human cortical spheroids. *Nature Methods*, *15*(9), 700–706. https://doi.org/10.1038/s41592-018-0081-4

Mandric, I., Schwarz, T., Majumdar, A., Perez, R., Subramaniam, M., Ye, C. J., … Halperin, E. (2019). Optimal design of single-cell RNA sequencing experiments for cell-type-specific eQTL analysis. *BioRxiv*, 766972. https://doi.org/10.1101/766972

Manolio, T. A. (2010). Genomewide Association Studies and Assessment of the Risk of Disease. *New England Journal of Medicine*, *363*(2), 166–176. https://doi.org/10.1056/NEJMra0905980

Mansour, A. A., Gonçalves, J. T., Bloyd, C. W., Li, H., Fernandes, S., Quang, D., … Gage, F. H. (2018). An in vivo model of functional and vascularized human brain organoids. *Nature Biotechnology*, *36*(5), 432–441. https://doi.org/10.1038/nbt.4127

Marchetto, M. C., Belinson, H., Tian, Y., Freitas, B. C., Fu, C., Vadodaria, K. C., … Muotri, A. R. (2017). Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. *Molecular Psychiatry*, *22*(6), 820–835. https://doi.org/10.1038/mp.2016.95

Marczylo, E. L., Jacobs, M. N., & Gant, T. W. (2016). Environmentally induced epigenetic toxicity: potential public health concerns. *Crit. Rev. Toxicol.*, *46*(8), 676–700. https://doi.org/10.1080/10408444.2016.1175417

Mariani, J., Coppola, G., Zhang, P., Abyzov, A., Provini, L., Tomasini, L., … Vaccarino, F. M. (2015). FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. *Cell*, *162*(2), 375–390. https://doi.org/10.1016/j.cell.2015.06.034

Martinez Arias, A., & Steventon, B. (2018). On the nature and function of organizers. *Development (Cambridge)*. https://doi.org/10.1242/dev.159525

Marton, R. M., Miura, Y., Sloan, S. A., Li, Q., Revah, O., Levy, R. J., … Pașca, S. P. (2019). Differentiation and maturation of oligodendrocytes in human three-dimensional neural cultures. *Nature Neuroscience*, *22*(3), 484–491. https://doi.org/10.1038/s41593-018-0316-9

May, P. A., Baete, A., Russo, J., Elliott, A. J., Blankenship, J., Kalberg, W. O., … Hoyme, H. E. (2014). Prevalence and Characteristics of Fetal Alcohol Spectrum Disorders. *PEDIATRICS*, *134*(5), 855–866. https://doi.org/10.1542/peds.2013-3319

McCartney, D. L., Hillary, R. F., Stevenson, A. J., Ritchie, S. J., Walker, R. M., Zhang, Q., … Marioni, R. E. (2018). Epigenetic prediction of complex traits and death. *Genome Biology*, *19*(1), 136. https://doi.org/10.1186/s13059-018-1514-1

McCubrey, J. A., Rakus, D., Gizak, A., Steelman, L. S., Abrams, S. L., Lertpiriyapong, K., … Martelli, A. M. (2016). Effects of mutations in Wnt/β-catenin, hedgehog, Notch and PI3K pathways on GSK-3 activity—Diverse effects on cell growth, metabolism and cancer. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1863*(12), 2942–2976. https://doi.org/10.1016/J.BBAMCR.2016.09.004

McCulloch, W. S. (Warren S. (1970). *Embodiments of mind*. M.I.T. Press.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115–133. https://doi.org/10.1007/BF02478259

McGinnis, C. S., Patterson, D. M., Winkler, J., Conrad, D. N., Hein, M. Y., Srivastava, V., … Gartner, Z. J. (2019). MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nature Methods*, *16*(7), 619–626. https://doi.org/10.1038/s41592-019-0433-8

McKenzie, K., Milton, M., Smith, G., & Ouellette-Kuntz, H. (2016). Systematic Review of the Prevalence and Incidence of Intellectual Disabilities: Current Trends and Issues. *Current Developmental Disorders Reports*, *3*(2), 104–115. https://doi.org/10.1007/s40474-016-0085-7

Meloni, M., & Testa, G. (2014). Scrutinizing the epigenetics revolution. *Biosocieties*, *9*(4), 431–456. https://doi.org/10.1057/biosoc.2014.22

Menendez, L., Kulik, M. J., Page, A. T., Park, S. S., Lauderdale, J. D., Cunningham, M. L., & Dalton, S. (2013). Directed differentiation of human pluripotent cells to neural crest stem cells. *Nat. Protoc.*, *8*(1), 203–212. https://doi.org/10.1038/nprot.2012.156

Miller, D. J., Bhaduri, A., Sestan, N., & Kriegstein, A. (2019). Shared and derived features of cellular diversity in the human cerebral cortex. *Current Opinion in Neurobiology*, *56*, 117–124. https://doi.org/10.1016/J.CONB.2018.12.005

Milligan, S. R., Balasubramanian, A. V, & Kalita, J. C. (1998). Relative potency of xenobiotic estrogens in an acute in vivo mammalian assay. *Environmental Health Perspectives*, *106*(1), 23–26. https://doi.org/10.1289/ehp.9810623

Mills, M. C., & Rahal, C. (2019). A scientometric review of genome-wide association studies. *Communications Biology*, *2*(1), 9. https://doi.org/10.1038/s42003-018-0261-x

Mimitou, E. P., Cheng, A., Montalbano, A., Hao, S., Stoeckius, M., Legut, M., … Smibert, P. (2019). Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nature Methods*, *16*(5), 409–412. https://doi.org/10.1038/s41592-019-0392-0

Miniscalco, C., Nygren, G., Hagberg, B., Kadesjö, B., & Gillberg, C. (2006). Neuropsychiatric and neurodevelopmental outcome of children at age 6 and 7 years who screened positive for language problems at 30 months. *Developmental Medicine & Child Neurology*, *48*(5), 361–366. https://doi.org/10.1017/S0012162206000788

Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Scientific Reports*, *6*(1), 26094. https://doi.org/10.1038/srep26094

Mitchell, K. J. (2015). The Genetic Architecture of Neurodevelopmental Disorders. In *The Genetics of Neurodevelopmental Disorders* (pp. 1–28). https://doi.org/10.1002/9781118524947.ch1

Molnár, Z., Clowry, G. J., Šestan, N., Alzu'bi, A., Bakken, T., Hevner, R. F., … Kriegstein, A. (2019). New insights into the development of the human cerebral cortex. *Journal of Anatomy*, *235*(3), 432–451. https://doi.org/10.1111/joa.13055

Moosa, A., Shu, H., Sarachana, T., & Hu, V. W. (2018). Are endocrine disrupting compounds environmental risk factors for autism spectrum disorder? *Hormones and Behavior*, *101*, 13–21. https://doi.org/10.1016/j.yhbeh.2017.10.003

Morgenstern, R., Whyatt, R. M., Insel, B. J., Calafat, A. M., Liu, X., Rauh, V. A., … Factor-Litvak, P. (2017). Phthalates and thyroid function in preschool age children: Sex specific associations. *Environ. Int.*, *106*, 11–18. https://doi.org/10.1016/j.envint.2017.05.007

Mullur, R., Liu, Y.-Y., & Brent, G. A. (2014). Thyroid hormone regulation of metabolism. *Physiol. Rev.*, *94*(2), 355–382. https://doi.org/10.1152/physrev.00030.2013

Nabais Sá, M. J., Gabriele, M., Testa, G., & de Vries, B. B. (2019). Gabriele-de Vries Syndrome. In *GeneReviews®*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/31145572

Namba, T., & Huttner, W. B. (2017). Neural progenitor cells and their role in the

development and evolutionary expansion of the neocortex. *Wiley Interdisciplinary Reviews: Developmental Biology*, *6*(1), e256. https://doi.org/10.1002/wdev.256

Nica, A. C., & Dermitzakis, E. T. (2013). Expression quantitative trait loci: present and future. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1620), 20120362. https://doi.org/10.1098/rstb.2012.0362

Noctor, S. C., Martinez-Cerdeño, V., & Kriegstein, A. R. (2007). Neural stem and progenitor cells in cortical development. *Novartis Foundation Symposium*, *288*, 59–73; discussion 73-8, 96–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18494252

Nowakowski, T. J., Bhaduri, A., Pollen, A. A., Alvarado, B., Mostajo-Radji, M. A., Di Lullo, E., … Kriegstein, A. R. (2017). Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science (New York, N.Y.)*, *358*(6368), 1318–1323. https://doi.org/10.1126/science.aap8809

Nowotny, H., & Testa, G. (2010). *Naked genes : reinventing the human in the molecular age*. MIT Press.

Ode, A., Källén, K., Gustafsson, P., Rylander, L., Jönsson, B. A. G., Olofsson, P., … Rignell-Hydbom, A. (2014). Fetal Exposure to Perfluorinated Compounds and Attention Deficit Hyperactivity Disorder in Childhood. *PLoS ONE*, *9*(4), e95891. https://doi.org/10.1371/journal.pone.0095891

Ozone, C., Suga, H., Eiraku, M., Kadoshima, T., Yonemura, S., Takata, N., … Sasai, Y. (2016). Functional anterior pituitary generated in self-organizing culture of human embryonic stem cells. *Nat. Commun.*, *7*, 10351. https://doi.org/10.1038/ncomms10351

Parker, H. S., Leek, J. T., Favorov, A. V., Considine, M., Xia, X., Chavan, S., … Fertig, E. J. (2014). Preserving biological heterogeneity with a permuted surrogate variable analysis for genomics batch correction. *Bioinformatics*, *30*(19), 2757–2763. https://doi.org/10.1093/bioinformatics/btu375

Paşca, A. M., Sloan, S. A., Clarke, L. E., Tian, Y., Makinson, C. D., Huber, N., … Paşca, S. P. (2015). Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. *Nat. Methods*, *12*(7), 671–678. https://doi.org/10.1038/nmeth.3415

PEABODY, F. W. (1927). THE CARE OF THE PATIENT. *JAMA: The Journal of the American Medical Association*, *88*(12), 877. https://doi.org/10.1001/jama.1927.02680380001001

Pelch, K. E., Bolden, A. L., & Kwiatkowski, C. F. (2019). Environmental Chemicals and Autism: A Scoping Review of the Human and Animal Research. *Environmental Health Perspectives*, *127*(4), 046001. https://doi.org/10.1289/EHP4386

Petrucelli, N., Daly, M. B., & Pal, T. (2016). *BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer*. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK1247/

Platzer, K., Cogné, B., Hague, J., Marcelis, C. L., Mitter, D., Oberndorff, K., … Jamra, R. (2018). Haploinsufficiency of *CUX1* Causes Nonsyndromic Global Developmental Delay With Possible Catch-up Development. *Annals of Neurology*, *84*(2), 200–207. https://doi.org/10.1002/ana.25278

Pletikos, M., Sousa, A. M. M., Sedmak, G., Meyer, K. A., Zhu, Y., Cheng, F., … Sestan, N. (2014). Temporal specification and bilaterality of human neocortical topographic gene expression. *Neuron*, *81*(2), 321–332. https://doi.org/10.1016/j.neuron.2013.11.018

Pollen, A. A., Bhaduri, A., Andrews, M. G., Nowakowski, T. J., Meyerson, O. S., Mostajo-Radji, M. A., … Kriegstein, A. R. (2019). Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution. *Cell*, *176*(4), 743-756.e17. https://doi.org/10.1016/j.cell.2019.01.017

Pollen, A. A., Nowakowski, T. J., Chen, J., Retallack, H., Sandoval-Espinosa, C., Nicholas, C. R., … Kriegstein, A. R. (2015). Molecular identity of human outer radial glia during cortical development. *Cell*, *163*(1), 55–67. https://doi.org/10.1016/j.cell.2015.09.004

Process to set scientific criteria to identify endocrine disruptors | Public Health. (n.d.). Retrieved September 9, 2019, from https://ec.europa.eu/health/endocrine_disruptors/process_en

Qi, Y., Zhang, X. J., Renier, N., Wu, Z., Atkin, T., Sun, Z., … Studer, L. (2017). Combined small-molecule inhibition accelerates the derivation of functional cortical neurons from human pluripotent stem cells. *Nature Biotechnology*. https://doi.org/10.1038/nbt.3777

Qian, X., Nguyen, H. N., Song, M. M., Hadiono, C., Ogden, S. C., Hammack, C., … Ming, G.-L. (2016). Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. *Cell*, *165*(5), 1238–1254. https://doi.org/10.1016/j.cell.2016.04.032

Quadrato, G., Brown, J., & Arlotta, P. (2016). The promises and challenges of human brain organoids as models of neuropsychiatric disease. *Nat. Med.*, *22*(11), 1220–1228. https://doi.org/10.1038/nm.4214

Quadrato, G., Nguyen, T., Macosko, E. Z., Sherwood, J. L., Min Yang, S., Berger, D. R., … Arlotta, P. (2017). Cell diversity and network dynamics in photosensitive human brain organoids. *Nature*, *545*(7652), 48–53. https://doi.org/10.1038/nature22047

Quesnel-Vallières, M., Weatheritt, R. J., Cordes, S. P., & Blencowe, B. J. (2019). Autism spectrum disorder: insights into convergent mechanisms from transcriptomics. *Nature Reviews Genetics*, *20*(1), 51–63. https://doi.org/10.1038/s41576-018-0066-2

Raja, W. K., Mungenast, A. E., Lin, Y.-T., Ko, T., Abdurrob, F., Seo, J., & Tsai, L.-H. (2016). Self-Organizing 3D Human Neural Tissue Derived from Induced Pluripotent Stem Cells Recapitulate Alzheimer's Disease Phenotypes. *PLoS One*, *11*(9), e0161969. https://doi.org/10.1371/journal.pone.0161969

Rajarajan, P., Borrman, T., Liao, W., Schrode, N., Flaherty, E., Casiño, C., …

Akbarian, S. (2018). Neuron-specific signatures in the chromosomal connectome associated with schizophrenia risk. *Science (New York, N.Y.)*, *362*(6420), eaat4311. https://doi.org/10.1126/science.aat4311

Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., … Yosef, N. (2017). The Human Cell Atlas. *ELife*, *6*. https://doi.org/10.7554/eLife.27041

Rigby, M. J. (2019). Ethical Dimensions of Using Artificial Intelligence in Health Care. *AMA Journal of Ethics*, *21*(2), E121-124. https://doi.org/10.1001/amajethics.2019.121

Robinson, B. M. (2012). Malignant pleural mesothelioma: an epidemiological perspective. *Annals of Cardiothoracic Surgery*, *1*(4), 491–496. https://doi.org/10.3978/j.issn.2225-319X.2012.11.04

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Rodriques, S. G., Stickels, R. R., Goeva, A., Martin, C. A., Murray, E., Vanderburg, C. R., … Macosko, E. Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science (New York, N.Y.)*, *363*(6434), 1463–1467. https://doi.org/10.1126/science.aaw1219

Romano, M. E., Webster, G. M., Vuong, A. M., Thomas Zoeller, R., Chen, A., Hoofnagle, A. N., … Braun, J. M. (2015). Gestational urinary bisphenol A and maternal and newborn thyroid hormone concentrations: The HOME Study. *Environmental Research*, *138*, 453–460. https://doi.org/10.1016/j.envres.2015.03.003

Ruderfer, D. M., Ripke, S., McQuillin, A., Boocock, J., Stahl, E. A., Pavlides, J. M. W., … Kendler, K. S. (2018). Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell*, *173*(7), 1705-1715.e16. https://doi.org/10.1016/j.cell.2018.05.046

Ruzzo, E. K., Pérez-Cano, L., Jung, J.-Y., Wang, L., Kashef-Haghighi, D., Hartl, C., … Wall, D. P. (2019). Inherited and De Novo Genetic Risk for Autism Impacts Shared Networks. *Cell*, *178*(4), 850-866.e26. https://doi.org/10.1016/j.cell.2019.07.015

Saffen, D. (2015). The genetic architecture of autism spectrum disorders (ASDs) and the potential importance of common regulatory genetic variants. *Science China Life Sciences*, *58*(10), 968–975. https://doi.org/10.1007/s11427-012-4336-5

Sakaguchi, H., Kadoshima, T., Soen, M., Narii, N., Ishida, Y., Ohgushi, M., … Sasai, Y. (2015). Generation of functional hippocampal neurons from self-organizing human embryonic stem cell-derived dorsomedial telencephalic tissue. *Nat. Commun.*, *6*, 8896. https://doi.org/10.1038/ncomms9896

Sanders, S. J., He, X., Willsey, A. J., Ercan-Sencicek, A. G., Samocha, K. E., Cicek, A. E., … State, M. W. (2015). Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron*, *87*(6), 1215–1233. https://doi.org/10.1016/j.neuron.2015.09.016

Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*(7510), 421–427. https://doi.org/10.1038/nature13595

Schug, T. T., Blawas, A. M., Gray, K., Heindel, J. J., & Lawler, C. P. (2015). Elucidating the Links Between Endocrine Disruptors and Neurodevelopment. *Endocrinology*, *156*(6), 1941–1951. https://doi.org/10.1210/en.2014-1734

scVelo. (n.d.).

Sestan, N., & State, M. W. (2018). Lost in Translation: Traversing the Complex Path from Genomics to Therapeutics in Autism Spectrum Disorder. *Neuron*, *100*(2), 406–423. https://doi.org/10.1016/j.neuron.2018.10.015

Shabalin, A. A. (2012). Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics*, *28*(10), 1353–1358. https://doi.org/10.1093/bioinformatics/bts163

Shah, S., Bonder, M. J., Marioni, R. E., Zhu, Z., McRae, A. F., Zhernakova, A., … van Zwet, E. W. (2015). Improving Phenotypic Prediction by Combining Genetic and Epigenetic Associations. *The American Journal of Human Genetics*, *97*(1), 75–85. https://doi.org/10.1016/j.ajhg.2015.05.014

Shang, L., Cho, M. T., Retterer, K., Folk, L., Humberson, J., Rohena, L., … Chung, W. K. (2015). Mutations in ARID2 are associated with intellectual disabilities. *Neurogenetics*, *16*(4), 307–314. https://doi.org/10.1007/s10048-015-0454-0

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., … Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303

Shapiro, J. A. (2011). *Evolution: a view from the 21st century*. Retrieved from http://dannyreviews.com/h/Evolution_21st_Century.html

Shi, H., Kichaev, G., & Pasaniuc, B. (2016). Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. *The American Journal of Human Genetics*, *99*(1), 139–153. https://doi.org/10.1016/j.ajhg.2016.05.013

Singh, A. R., Lawrence, W. H., & Autianx, J. (1975). Maternal-Fetal Transfer of 14C-Di-2-ethylhexyl Phthalate and 14C-Diethyl Phthalate in Rats. *Journal of Pharmaceutical Sciences*, *64*(8), 1347–1350. https://doi.org/10.1002/jps.2600640819

Singh, B. K., Sinha, R. A., & Yen, P. M. (2018). Novel Transcriptional Mechanisms for Regulating Metabolism by Thyroid Hormone. *International Journal of Molecular Sciences*, *19*(10), 3284. https://doi.org/10.3390/ijms19103284

Skinner, M. K., Manikkam, M., & Guerrero-Bosagna, C. (2010). Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends in Endocrinology and Metabolism: TEM*, *21*(4), 214–222. https://doi.org/10.1016/j.tem.2009.12.007

Sloan, S. A., Darmanis, S., Huber, N., Khan, T. A., Birey, F., Caneda, C., … Paşca,

S. P. (2017). Human Astrocyte Maturation Captured in 3D Cerebral Cortical Spheroids Derived from Pluripotent Stem Cells. *Neuron*, *95*(4), 779-790.e6. https://doi.org/10.1016/j.neuron.2017.07.035

Spearow, J. L., Doemeny, P., Sera, R., Leffler, R., & Barkley, M. (1999). Genetic variation in susceptibility to endocrine disruption by estrogen in mice. *Science (New York, N.Y.)*, *285*(5431), 1259–1261. https://doi.org/10.1126/science.285.5431.1259

Stachowiak, E. K., Benson, C. A., Narla, S. T., Dimitri, A., Bayona Chuye, L. E., Dhiman, S., … Stachowiak, M. K. (2017). Cerebral organoids reveal early cortical maldevelopment in schizophrenia—computational anatomy and genomics, role of FGFR1. *Transl. Psychiatry*, *7*(11), 6. https://doi.org/10.1038/s41398-017-0054-x

Statistics - SFARI Gene. (n.d.). Retrieved September 8, 2019, from https://gene.sfari.org/about-human-gene/statistics/

Stenz, L., Rahban, R., Prados, J., Nef, S., & Paoloni-Giacobino, A. (2019). Genetic resistance to DEHP-induced transgenerational endocrine disruption. *PLOS ONE*, *14*(6), e0208371. https://doi.org/10.1371/journal.pone.0208371

Stewart, A. M., & Kalueff, A. V. (2015). Developing better and more valid animal models of brain disorders. *Behavioural Brain Research*, *276*, 28–31. https://doi.org/10.1016/j.bbr.2013.12.024

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P. K., Swerdlow, H., … Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. *Nature Methods*, *14*(9), 865–868. https://doi.org/10.1038/nmeth.4380

Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B. Z., Mauck, W. M., … Satija, R. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biology*, *19*(1), 224. https://doi.org/10.1186/s13059-018-1603-1

Stuart, T., & Satija, R. (2019). Integrative single-cell analysis. *Nature Reviews Genetics*, *20*(5), 257–272. https://doi.org/10.1038/s41576-019-0093-7

Šubelj, L., & Bajec, M. (2014). Group detection in complex networks: An algorithm and comparison of the state of the art. *Physica A: Statistical Mechanics and Its Applications*, *397*, 144–156. https://doi.org/10.1016/J.PHYSA.2013.12.003

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., … Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

System, B. R. F. S., & Others. (2014). Atlanta (GA): US Department of Health and Human Services, Centers for Disease Control and Prevention. *National Center for Chronic Disease Prevention and Health Promotion*.

Systems Medicine. (2016). https://doi.org/10.1007/978-1-4939-3283-2

Tanay, A., & Regev, A. (2017). Scaling single-cell genomics from phenomenology to mechanism. *Nature*, *541*(7637), 331–338. https://doi.org/10.1038/nature21350

Taverna, E., Götz, M., & Huttner, W. B. (2014). The Cell Biology of Neurogenesis: Toward an Understanding of the Development and Evolution of the Neocortex. *Annual Review of Cell and Developmental Biology*, *30*(1), 465–502. https://doi.org/10.1146/annurev-cellbio-101011-155801

Taylor, D. M., Aronow, B. J., Tan, K., Bernt, K., Salomonis, N., Greene, C. S., … White, P. S. (2019). The Pediatric Cell Atlas: Defining the Growth Phase of Human Development at Single-Cell Resolution. *Developmental Cell*, *49*(1), 10–29. https://doi.org/10.1016/j.devcel.2019.03.001

Thapar, A., Cooper, M., & Rutter, M. (2017). Neurodevelopmental disorders. *The Lancet. Psychiatry*, *4*(4), 339–346. https://doi.org/10.1016/S2215-0366(16)30376-5

The Dark Secret at the Heart of AI - MIT Technology Review. (n.d.). Retrieved September 8, 2019, from https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/

The Human Exposome Project. (n.d.). Retrieved October 15, 2019, from https://humanexposomeproject.com/

The LifeTime FET Flagship. (n.d.). Retrieved September 8, 2019, from https://lifetime-fetflagship.eu/index.php/the-initiative/

Topol, E. J. (2019). *Deep medicine : how artificial intelligence can make healthcare human again* (2019th ed.).

Torkamani, A., Andersen, K. G., Steinhubl, S. R., & Topol, E. J. (2017). High-Definition Medicine. *Cell*, *170*(5), 828–843. https://doi.org/10.1016/j.cell.2017.08.007

Totaro, A., Castellan, M., Battilana, G., Zanconato, F., Azzolin, L., Giulitti, S., … Piccolo, S. (2017). YAP/TAZ link cell mechanics to Notch signalling to control epidermal stem cell fate. *Nat. Commun.*, *8*, 15206. https://doi.org/10.1038/ncomms15206

Tracy, S., Yuan, G.-C., & Dries, R. (2019). RESCUE: imputing dropout events in single-cell RNA-sequencing data. *BMC Bioinformatics*, *20*(1), 388. https://doi.org/10.1186/s12859-019-2977-0

Tran, N. Q. V., & Miyake, K. (2017). Neurodevelopmental Disorders and Environmental Toxicants: Epigenetics as an Underlying Mechanism. *International Journal of Genomics*, *2017*, 7526592. https://doi.org/10.1155/2017/7526592

Tritschler, S., Büttner, M., Fischer, D. S., Lange, M., Bergen, V., Lickert, H., & Theis, F. J. (2019). Concepts and limitations for learning developmental trajectories from single cell genomics. *Development*, *146*(12), dev170506. https://doi.org/10.1242/DEV.170506

Trujillo, C. A., Gao, R., Negraes, P. D., Gu, J., Buchanan, J., Preissl, S., … Muotri, A. R. (2019). Complex Oscillatory Waves Emerging from Cortical Organoids Model Early Human Brain Network Development. *Cell Stem Cell*, *0*(0). https://doi.org/10.1016/j.stem.2019.08.002

Turakhia, M. P., Desai, M., Hedlin, H., Rajmane, A., Talati, N., Ferris, T., … Perez, M. V. (2019). Rationale and design of a large-scale, app-based study to identify cardiac arrhythmias using a smartwatch: The Apple Heart Study. *American Heart Journal*, *207*, 66–75. https://doi.org/10.1016/j.ahj.2018.09.002

van der Wijst, M. G. P., Brugge, H., de Vries, D. H., Deelen, P., Swertz, M. A., & Franke, L. (2018). Single-cell RNA sequencing identifies celltype-specific cis-eQTLs and co-expression QTLs. *Nature Genetics*, *50*(4), 493–497. https://doi.org/10.1038/s41588-018-0089-9

Vanden Heuvel, J. P., Thompson, J. T., Frame, S. R., & Gillies, P. J. (2006). Differential Activation of Nuclear Receptors by Perfluorinated Fatty Acid Analogs and Natural Fatty Acids: A Comparison of Human, Mouse, and Rat Peroxisome Proliferator-Activated Receptor-α, -β, and -γ, Liver X Receptor-β, and Retinoid X Receptor-α. *Toxicological Sciences*, *92*(2), 476–489. https://doi.org/10.1093/toxsci/kfl014

Vandenberg, L. N., Colborn, T., Hayes, T. B., Heindel, J. J., Jacobs, D. R., Lee, D.-H., … Myers, J. P. (2012). Hormones and Endocrine-Disrupting Chemicals: Low-Dose Effects and Nonmonotonic Dose Responses. *Endocrine Reviews*, *33*(3), 378–455. https://doi.org/10.1210/er.2011-1050

Vegas, N., Cavallin, M., Maillard, C., Boddaert, N., Toulouse, J., Schaefer, E., … Bahi-Buisson, N. (2018). Delineating *FOXG1* syndrome. *Neurology Genetics*, *4*(6), e281. https://doi.org/10.1212/NXG.0000000000000281

Velasco, S., Kedaigle, A. J., Simmons, S. K., Nash, A., Rocha, M., Quadrato, G., … Arlotta, P. (2019). Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature*, *570*(7762), 523–527. https://doi.org/10.1038/s41586-019-1289-x

Vigna, P., & Casey, M. J. (2016). *The age of cryptocurrency: how bitcoin and the blockchain are challenging the global economic order*. Retrieved from http://www.theinvestorspodcast.com/wp-content/uploads/2018/03/TheAgeofCryptoCurrency-NEW.pdf

Vitriolo, A., Gabriele, M., & Testa, G. (2019). From enhanceropathies to the epigenetic manifold underlying human cognition. *Human Molecular Genetics*. https://doi.org/10.1093/hmg/ddz196

Wang, D., Liu, S., Warrell, J., Won, H., Shi, X., Navarro, F. C. P., … Gerstein, M. B. (2018). Comprehensive functional genomic resource and integrative model for the human brain. *Science*, *362*(6420), eaat8464. https://doi.org/10.1126/SCIENCE.AAT8464

Wang, M. H., Cordell, H. J., & Van Steen, K. (2019). Statistical methods for genome-wide association studies. *Seminars in Cancer Biology*, *55*, 53–60. https://doi.org/10.1016/J.SEMCANCER.2018.04.008

Watanabe, K., Kamiya, D., Nishiyama, A., Katayama, T., Nozaki, S., Kawasaki, H., … Sasai, Y. (2005). Directed differentiation of telencephalic precursors from embryonic stem cells. *Nature Neuroscience*, *8*(3), 288–296. https://doi.org/10.1038/nn1402

Weiner, D. J., Wigdor, E. M., Ripke, S., Walters, R. K., Kosmicki, J. A., Grove, J., … Robinson, E. B. (2017). Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nature Genetics*, *49*(7), 978–985. https://doi.org/10.1038/ng.3863

Weinstein, J. A., Regev, A., & Zhang, F. (2019). DNA Microscopy: Optics-free Spatio-genetic Imaging by a Stand-Alone Chemical Reaction. *Cell*, *178*(1), 229-241.e16. https://doi.org/10.1016/j.cell.2019.05.019

When It Comes to Gorillas, Google Photos Remains Blind | WIRED. (n.d.). Retrieved September 8, 2019, from https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/

WHO | Global assessment of the state-of-the-science of endocrine disruptors. (2013). *WHO*. Retrieved from https://www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en/

Whyatt, R. M., Liu, X., Rauh, V. A., Calafat, A. M., Just, A. C., Hoepner, L., … Factor-Litvak, P. (2012). Maternal Prenatal Urinary Phthalate Metabolite Concentrations and Child Mental, Psychomotor, and Behavioral Development at 3 Years of Age. *Environmental Health Perspectives*, *120*(2), 290–295. https://doi.org/10.1289/ehp.1103705

Williams, G. R. (2008). Neurodevelopmental and neurophysiological actions of thyroid hormone. *J. Neuroendocrinol.*, *20*(6), 784–794. https://doi.org/10.1111/j.1365-2826.2008.01733.x

Wills, Q. F., Livak, K. J., Tipping, A. J., Enver, T., Goldson, A. J., Sexton, D. W., & Holmes, C. (2013). Single-cell gene expression analysis reveals genetic associations masked in whole-tissue experiments. *Nature Biotechnology*, *31*(8), 748–752. https://doi.org/10.1038/nbt.2642

Willsey, A. J., Morris, M. T., Wang, S., Willsey, H. R., Sun, N., Teerikorpi, N., … Krogan, N. J. (2018). The Psychiatric Cell Map Initiative: A Convergent Systems Biological Approach to Illuminating Key Molecular Pathways in Neuropsychiatric Disorders. *Cell*, *174*(3), 505–520. https://doi.org/10.1016/j.cell.2018.06.016

Wilsch-Bräuninger, M., Florio, M., & Huttner, W. B. (2016). Neocortex expansion in development and evolution — from cell biology to single genes. *Current Opinion in Neurobiology*, *39*, 122–132. https://doi.org/10.1016/j.conb.2016.05.004

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., … Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, *50*(5), 668–681. https://doi.org/10.1038/s41588-018-0090-3

Xiang, Y., Tanaka, Y., Cakir, B., Patterson, B., Kim, K.-Y., Sun, P., … Park, I.-H. (2019). hESC-Derived Thalamic Organoids Form Reciprocal Projections When Fused with Cortical Organoids. *Cell Stem Cell*, *24*(3), 487-497.e7. https://doi.org/10.1016/j.stem.2018.12.015

Xie, S., Duan, J., Li, B., Zhou, P., & Hon, G. C. (2017a). Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells. *Mol. Cell*, *66*(2), 285-299.e5. https://doi.org/10.1016/j.molcel.2017.03.007

Xie, S., Duan, J., Li, B., Zhou, P., & Hon, G. C. (2017b). Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells. *Mol. Cell*, *66*(2), 285-299.e5. https://doi.org/10.1016/j.molcel.2017.03.007

Xu, W., Xin, C., Lin, Q., Ding, F., Gong, W., Zhou, Y., … Hu, S. (2014). Adolescent Mouse Takes on An Active Transcriptomic Expression During Postnatal Cerebral Development. *Genomics, Proteomics & Bioinformatics*, *12*(3), 111–119. https://doi.org/10.1016/j.gpb.2014.04.004

Ye, L., Guo, J., & Ge, R.-S. (2014). Environmental Pollutants and Hydroxysteroid Dehydrogenases. In *Vitamins and hormones* (Vol. 94, pp. 349–390). https://doi.org/10.1016/B978-0-12-800095-3.00013-4

Yin, X., Mead, B. E., Safaee, H., Langer, R., Karp, J. M., & Levy, O. (2016). Engineering Stem Cell Organoids. *Cell Stem Cell*, *18*(1), 25–38. https://doi.org/10.1016/j.stem.2015.12.005

Yokota, Y., Eom, T.-Y., Stanco, A., Kim, W.-Y., Rao, S., Snider, W. D., & Anton, E. S. (2010). Cdc42 and Gsk3 modulate the dynamics of radial glial growth, inter-radial glial interactions and polarity in the developing cerebral cortex. *Development*, *137*(23), 4101–4110. https://doi.org/10.1242/dev.048637

Yoo, H. (2015). Genetics of Autism Spectrum Disorder: Current Status and Possible Clinical Applications. *Experimental Neurobiology*, *24*(4), 257–272. https://doi.org/10.5607/en.2015.24.4.257

Yui, S., Azzolin, L., Maimets, M., Pedersen, M. T., Fordham, R. P., Hansen, S. L., … Jensen, K. B. (2018). YAP/TAZ-Dependent Reprogramming of Colonic Epithelium Links ECM Remodeling to Tissue Regeneration. *Cell Stem Cell*, *22*(1), 35-49.e7. https://doi.org/10.1016/j.stem.2017.11.001

Zanella, M., Vitriolo, A., Andirko, A., Martins, P. T., Sturm, S., O'Rourke, T., … Testa, G. (2019). 7q11.23 Syndromes Reveal BAZ1B as a Master Regulator of the Modern Human Face and Validate the Self-Domestication Hypothesis. *BioRxiv*, 570036. https://doi.org/10.1101/570036

Zhang, Ye, Sloan, S. A., Clarke, L. E., Caneda, C., Plaza, C. A., Blumenthal, P. D., … Barres, B. A. (2016). Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron*. https://doi.org/10.1016/j.neuron.2015.11.013

Zhang, Yingsha, Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., … Südhof, T. C. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron*, *78*(5), 785–798. https://doi.org/10.1016/j.neuron.2013.05.029

Zhang, Yonggang, & Hu, W. (2012). NFκB signaling regulates embryonic and adult neurogenesis. *Frontiers in Biology*, *7*(4). https://doi.org/10.1007/s11515-012-1233-z

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.*, *8*, 14049. https://doi.org/10.1038/ncomms14049

Zoeller, R. T., Brown, T. R., Doan, L. L., Gore, A. C., Skakkebaek, N. E., Soto, A. M., … Vom Saal, F. S. (2012). Endocrine-Disrupting Chemicals and Public Health Protection: A Statement of Principles from The Endocrine Society. *Endocrinology*, *153*(9), 4097–4110. https://doi.org/10.1210/en.2012-1422