Published in final edited form as:

Trends Mol Med. 2014 February ; 20(2): 91–104. doi:10.1016/j.molmed.2013.11.004.

iPSC-derived neurons as a higher-throughput readout for autism: Promises and pitfalls

Daria Prilutsky¹, Nathan P. Palmer¹, Niklas Smedemark-Margulies², Thorsten M. Schlaeger³, David M. Margulies^{1,4}, and Isaac S. Kohane^{1,5}

¹Center for Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA

²Amherst College Class of 2013 Undergraduate Neuroscience Program, Amherst, MA, 01002, USA

³Stem Cell Program, Boston Children's Hospital, Boston, MA 02115, USA

⁴Divisions of Genetics and Developmental Medicine, Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA

⁵Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA

Abstract

The elucidation of disease etiologies and establishment of robust, scalable, high-throughput screening assays for autism spectrum disorders (ASDs) have been impeded by both inaccessibility of disease-relevant neuronal tissue and the genetic heterogeneity of the disorder. Neuronal cells derived from induced pluripotent stem cells (iPSCs) from autism patients may circumvent these obstacles and serve as relevant cell models. To date, derived cells are characterized and screened by assessing their neuronal phenotypes. These characterizations are often etiology-specific or lack reproducibility and stability. In this manuscript, we present an overview of efforts to study iPSC-derived neurons as a model for autism, and we explore the plausibility of gene expression profiling as a reproducible and stable disease marker.

Keywords

Autism; iPSC; gene expression; high-throughput assay

Cellular models to unravel complexity behind autism

Autism Spectrum Disorders (ASDs) (see Glossary) are a group of neurodevelopmental disorders characterized by deficits in social cognition, communication, and behavioral flexibility. There are no specific, approved pharmacologic treatments for ASDs, and there

© 2013 Elsevier Ltd. All rights reserved.

Corresponding author: Kohane, I.S. (Isaac_Kohane@hms.harvard.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

are very few targets for drug development, although clinical trials with metabotropic glutamate receptor (mGluR) 5 antagonists, for example, AFQ056 (ClinicalTrials.gov, NCT01357239) to treat Fragile X [1, 2]; and recombinant human Insulin Growth Factor-1 (rhIGF-1) (ClinicalTrials.gov, NCT01777542) to treat Rett syndrome [3] are underway. To date, primary treatment for ASD patients has consisted mainly of behavioral therapy [4]. Recent studies have demonstrated that ASD has a strong and complex genetic basis, and hundreds of candidate genes have been identified, each with many different putatively disruptive variants [5, 6]. Recent large scale sequencing efforts have revealed that many individuals with ASD have multiple point mutations and copy number variants encompassing likely candidate genes. Current efforts focus on identifying a set of pathways that may underlie the connection between these candidate genes. Some sequencing studies have identified disruptive variants in glutamatergic pathway genes as associated with, and potentially causal of, syndromic and non-syndromic autism [7, 8].

Model systems for the study of ASDs have been elusive. There are several monogenic syndromes (e.g., Rett's, Timothy, Fragile X, and tuberous sclerosis), which may present with symptoms of ASD, and a more complete understanding of these disorders may help unravel the molecular and cellular complexity behind other ASDs. There exist promising but limited animal models for ASDs. Murine model systems, for example, can provide genetic homogeneity and allow the study of behavioral phenotypes, but may have inherently limited applicability to understanding effects in human neocortical regions [9]. Direct tissue assays from humans are desirable, but such tissue is rarely obtained, especially during the crucial early developmental period when the disorder is first manifest [10]. Viable tissue samples are also unavailable, since brain biopsies from patients with ASD for establishing an *in vitro* neuronal cell line are considered strictly unethical.

To circumvent these obstacles, some are attempting to use patient-derived induced pluripotent stem cells (iPSCs) to characterize these neurodevelopmental abnormalities. These cells have the properties of self-renewal and pluripotency, making them a promising resource for the modeling of pathogenesis, drug screening, and cell-based therapies [11, 12]. Somatic cells such as fibroblasts or mononuclear cells can be reprogrammed into a pluripotent state and subsequently differentiated into the tissue of interest, such as neurons. Since the resulting neural cells presumably retain all of that individual's genetic predispositions, this approach has tremendous potential as an *in vitro* screening assay for ASDs.

To date, several groups have generated disease-specific lines from patients with monogenic ASDs including Rett syndrome (RTT) [13–19], Fragile X syndrome (FXS) [20, 21] and Timothy syndrome (TS) [22]. The main objectives of these studies have been to identify a disease-specific cellular phenotype and then to examine whether this phenotype can be rescued by therapeutic intervention. Several studies have found phenotypic differences between diseased lines and controls, and work on Rett syndrome in particular has yielded a highly reproducible morphological phenotype. While cellular phenotypes have been demonstrated in the Rett-derived cells [13–16, 19], these phenotypes are not immediately amenable to high-throughput drug screening, and the phenotypes are very sensitive to both culture conditions and lineage-specific variables.

Another approach to understanding the molecular changes that occur in autism is through gene expression profiling. Gene expression profiling has been applied to postmortem brain tissue from human donors [23] and murine models [24, 25], as well as peripheral blood [26] and lymphoblastoid lines [27–29] to compare expression signatures of diseased and wild-type tissues. Gene expression profiling has repeatedly identified several common broad pathways that consistently show differential regulation between ASD samples and neurotypical controls, including long-term potentiation (LTP), gap-junction, chemokine-immune and Wnt signaling [23, 27–29]. These early results support the notion that the ASD phenotype may stem from a smaller set of pathways than the number of genes or variants that have been implicated. Therefore, an individual with ASD may have a gene expression signature that can serve as a marker of disease.

In this review, we discuss the plausibility of evaluating gene expression in iPSC-derived cell lines as a high-throughput disease marker in ASD and, eventually, as a predictor for responses to pharmacological intervention. By comprehensively comparing the transcriptomic signature of a specific sample to thousands of other samples in the national Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>), we are able to determine [30, 31] whether developing cells follow consistent trajectories in gene expression space during maturation, and whether projecting measurements of transcriptional activity onto a larger transcriptomic analysis framework can reliably establish cell identity, similarity to tissue of origin or to target tissue, and stage of maturity. Applying these methods of analysis to iPSCs and derived cell lines holds promise for the creation of higher-throughput screening systems. We discuss our perspective on the future use of iPSCs and their derivatives as a disease model, and compare the plausibility of phenotypic and transcriptomic measures as screening metrics. We anticipate that a combination of these methods will be effective for identifying therapeutic molecules, stratifying populations with a specific pattern of disease, and illuminating the underlying mechanisms of ASDs.

History of iPSCs and promises of neurons-on-a-dish to model autism

Derivation of iPSCs from human somatic cells

Embryonic stem cells (ESCs) possess a virtually unlimited capacity for self-renewal and differentiation and are easy to culture and manipulate, making them invaluable for modeling defined genetic disorders and for generating disease-relevant cell types. However, ESCs can only be obtained from embryos, a fact that severely limits their value as a tool for understanding genetically complex pathologies. In a landmark paper published in 2006 [12], Takahashi and Yamanaka described how differentiated somatic cells from adult mice could be turned into cells that closely resemble ESCs. They showed that retroviral transfer of four genes encoding the transcription factors Oct4, Sox2, Klf4 and c-Myc could cause reprogramming of skin fibroblasts into ESC-like “induced pluripotent stem cells” (iPSCs). Since then, similar methods have been shown to allow generation of iPSCs from human skin-derived fibroblasts and other cell types [32, 33].

One limitation of retrovirally-derived iPSCs as a disease model is the persistence of the introduced reprogramming agents as foreign proviral sequences in the iPSC genome. These sequences may affect nearby endogenous genes, and any expression of the transgenes

beyond the reprogramming phase may alter cellular behavior and disease phenotypes. These considerations have led to the development of a variety of reprogramming methods that do not result in permanent integration of exogenous DNA sequences. Transgene-free human iPSCs have now been generated by transient transfection with episomal vectors [34], mini-circle DNA vectors [35], synthetic mRNAs [36], fully-excisable piggyBac transposons [37], direct protein delivery [38], and transduction with non-integrating viral vectors based on adenovirus [39] or Sendai virus (SeV) [40]. For studying complex pathologies in disorders such as ASD, a non-integrating technology for generating iPSCs is preferable.

The high phenotypic variability among iPSC clones [41, 42] is problematic in the use of these cells for disease modeling and for assaying therapeutic response. Even from the same individuals, one can get divergent iPSC lines. For this reason, many groups have started to characterize these cells with quantitative measures rather than qualitative ones [43–45]. Some groups have tried to address this issue using “scorecards” [43]. Bock et al. developed one such scoring system to correlate differentiation potential with DNA methylation and gene expression in the undifferentiated state. This tool yields a quick and comprehensive assessment of pluripotent cell lines and can predict whether a specific clone is suitable for differentiation. For monogenic disorders, the challenges posed by the variability between cell lines can be overcome by using nuclease-assisted homologous recombination to engineer isogenic sets of iPSCs with various disease genotypes (e.g., wild type vs. heterozygous vs. homozygous). Conditional (e.g., doxycycline- or tamoxifen-inducible) gene rescue or knock-down systems can also help overcome some of the challenges posed by line-to-line variability. However, for the study of genetically complex pathologies, no simple solutions exist. Instead, such studies require much larger patient cohorts and a greater number of independent iPSC lines per patient to be compared in carefully designed experiments in order to distinguish genuine disease effects from system-intrinsic phenotypic noise among lines.

Some of the observed variability among derived lines may result from incomplete reprogramming where a few gene loci are observed to be in an aberrant epigenetic state that does not resemble that of bona fide pluripotent stem cells [46]. This form of incomplete reprogramming tends to bias differentiation towards cellular lineages that are developmentally related to that of the donor cell [47]. For example, iPSCs obtained from mouse fibroblasts, hematopoietic and myogenic cells exhibit distinct transcriptional and epigenetic patterns [48]. This epigenetic memory of the donor cell of origin can be erased by differentiation and serial reprogramming, by treatment with chromatin-modifying drugs, and by nuclear transfer into an oocyte [49]. Epigenetic differences among iPSC lines can sometimes be attenuated by simple long-term culture and repeated passaging of cells [50]. However, driving the cells through many extra doublings is cumbersome, costly, and increases the chance of chromosomal abnormalities, affecting the validity of *in vitro* cellular model [51].

Protocols to derive neurons from iPSCs

When designing a differentiation protocol to generate neuronal cell lines relevant to autism, it is important to consider the choice of cell type, brain regional specificity, recapitulation of

neurodevelopmental program and functional connectivity (synaptogenesis). A variety of protocols have been established to generate many iPSC-derived neural subtypes, spanning the entire neural tube. Three of the more frequently studied neural cell types are: (i) spinal motor neurons; (ii) midbrain dopaminergic neurons, and; (iii) cortical neurons. Protocols for producing each cell type differ with regard to a number of parameters, including media composition, the timing, concentration and nature of exogenous signaling factors, cell density, and culturing substrate and environment [52, 53]. The process of neural differentiation begins with the modulation of cellular morphogens such as Wingless/Int proteins (Wnt), Sonic hedgehog (SHH), bone morphogenetic proteins (BMP), retinoic acid (RA) and fibroblast growth factor (FGF), which are deposited in gradients to establish the main subdivisions of the central nervous system [52, 54]. In the study of ASD, it remains unclear which neuronal cell type is best to study. Several investigators suggest focusing on generation of cortical neurons, since the underlying disease mechanisms are believed to cause defects in cortical connectivity [53, 55].

There are two major classes of protocols to initiate iPSC-derived neuronal differentiation: (i) floating embryoid bodies (EBs) aggregate formation; and (ii) monolayer adherent cell culture [53]. Most of the protocols involve an initial aggregation of the cells into EBs, followed by the generation of “neural rosettes”. It was demonstrated that floating aggregates of mouse ES cells generate naive telencephalic precursors that acquire sub-regional identities by responding to extracellular patterning signals. In this study, the authors used serum-free medium referred to as ‘SFEB’ (serum-free, floating culture of embryoid body–like aggregates) [56]. In their next study, the authors improved 3D aggregate culture, allowing highly efficient differentiation of ESCs into cortical progenitors and functional projection neurons [57]. Their approach is advantageous because it generates 3D structure which may recapitulate structural phenomena such as cortical layer formation [58]. However, there are other protocols that rely on monolayer cultures rather than EBs to allow consistent and homogenous cell–cell interactions. A protocol for neural differentiation in adherent culture by inhibiting dual Sma and Mothers Against Decapentaplegic (SMAD) signaling (BMP and transforming growth factor β TGF β inhibition) was developed [59]. In most protocols, the cells can be directed into different fates (forebrain, midbrain, spinal cord) by modulating use of specific morphogens and culture conditions [33, 32, 60].

Cerebral cortex is mostly composed of two classes of neurons, which could be of interest: (i) ~20% GABAergic inhibitory interneurons; and (ii) ~80% glutamatergic excitatory projection neurons. To date, iPSC work has focused on programming iPSCs differentiation into excitatory neurons [53] and resulted in heterogeneous populations of both types. The most accepted doctrine for the generation of a specific neuronal type is to block alternative fates. To generate cortical pyramidal neurons, the action of ventralizing Shh should be inhibited [61], with additional inhibition of BMP/Wnt dorsalizing signaling [52, 62]. Recently introduced examples of such protocols include the use of defined medium with no morphogen and presence of the Shh inhibitor cyclopamine [63]; co-culture with stromal cells and addition of FGF2 and RA [64] in rodent and human ESCs systems. In contrast, to generate cortical interneurons originating from ventral telencephalon, addition of SHH is required along with the addition of differentiation factors such as FGF2 and IGF [65–67]; inhibition of Wnt signaling [68]; or addition of activin as a potent neurotrophic factor [69].

Recent studies suggest that retinoids are required for the development of cortical neurons in rodent systems [70] and for the differentiation of cortical neurons from human iPSCs [71], if administered at a specific time-frame and concentration. Examples of these protocols are summarized in Figure 1.

When designing a differentiation protocol for an *in vitro* model of cortical disease, one should also consider recapitulating the trajectory of *in vivo* development, and monitor the cells for the expression of markers for various developmental stages, and functional benchmarks. It is useful to recapitulate functional synaptic network formation, because of the presumed tight link between synaptic function and pathophysiology of ASDs. Neuronal lineage can be validated by analyzing the cells for the expression of timing- and region-specific markers, morphology, and electrophysiological properties. Important characteristics indicating successful recapitulation of cortical development include: (i) a period of cortical neurogenesis followed by astrocyte genesis measured by the pattern and timing of expression of transcription factor combinations unique to a layer (deep/upper) cortical identity; (ii) apico-basal cell polarity of neuroprogenitors; (iii) terminal neuronal differentiation and acquisition of mature electrophysiological properties; (iv) formation of functional synapses; and (v) synchronized network activity. Some studies reproduced human cerebral cortex development from pluripotent stem cells, addressing key milestones in neurodevelopment [71, 72]. Elegant work by Lancaster et al. reproduced neurodevelopment by generating cerebral organoids in a 3D culture system using Matrigel droplets as scaffolds differentiating in a spinning bioreactor [73].

It is also possible to cause the trans-differentiation of mature fibroblasts directly into induced neurons [74, 75] by forced expression of key transcription factors, skipping the intermediate pluripotency stage. These induced neurons can form synapses and generate action potentials. In this approach, the source tissue has a finite capacity to replicate. However, trans-differentiation may skip a developmental window during which the molecular deficit underlying autism is manifest, and the overexpression of neuronal genes in converting somatic tissue may obscure the expression abnormality that causes disease [76].

Current non-transcriptomic phenotyping of iPSC-derived neurons and sources of variation

For an iPS-derived cell line to be useful as both a disease model and a drug-screening platform, it must manifest a disease-specific phenotype that is stable across replicates and individuals. The cellular phenotype must be measurable in a rapid and scalable manner, and it must normalize with treatment. Only a few useful examples of iPS-derived neurodevelopmental ‘disease-in-a-dish’ models [10, 77] have been studied *in vitro* and evaluated for: (i) morphological changes; (ii) impaired neurogenesis and migration; (iii) synaptic dysfunction; (iv) electrophysiological impairment; (v) toxic protein accumulation; and (vi) neuronal death. These measures can be addressed and used as a screening platform to test the efficiency of a candidate drug. Cellular function including cellular morphology, such as soma size or dendritic spine density, can be measured by cellular imaging and High-Content Screening (HCS) technology [78]. HCS uses high-resolution microscopy to screen for phenotypic changes, allowing the assay of only those cells which express a marker of interest. It distinguishes cell types from one another based on amounts of proteins

synthesized and morphology, and may thereby measure the drug effects on features in each cell type. Therefore, this technology is well suited for high-throughput screening and quantitative observation of phenotypes in some systems. It complements the transcriptomic measure of perturbation of each gene in a pathway which may underlie the observed phenotypic changes. At this time, HCS requires sophisticated image processing algorithms which are particularly challenging to apply to neuronal cell types with complex cellular morphology. Electrophysiological recording as a phenotypic measure has the disadvantage of providing measurements late in cellular development. Additionally, it may take up to nine months [79], even when iPSC-derived progenitors are implanted into mouse brain, to reach the functional synaptically active stage of neuronal development.

As previously noted, the phenotypic variability of iPSCs can be considerable and this is certainly true of iPSC-derived neurons. Furthermore, undifferentiated iPSCs are co-mingled with undefined constituents in the culture (such as feeder cells), which may further add noise to the phenotypic characterization. Therefore, many studies prefer the use of defined-media and environment to decrease variability. Studies have also demonstrated that there is considerable variation in differentiation potential among pluripotent cell lines [45, 80, 81]. Some lines are difficult to neuralize and therefore erroneous assignment of disease-relevant phenotype may ensue [45]. Consequently, quantitative measures are preferable when choosing a line for further analysis [43, 80]. However, specifically for neurodevelopmental disorders, this approach may lead to masking of a disease-relevant phenotype due to developmental defect [45]. Another source of variation is the heterogeneous neuronal cultures that are generated with various types of neurons and various states of functional maturation [45]. Given these challenges, can gene expression read-outs work in concert with cellular phenotyping and become a complementary method of characterization?

Studies of autism in derived neurons

iPSC-based studies in monogenic and syndromic autism

To date, numerous structural variations and candidate genes (e.g., neurexin 1 (*NRXN1*), neuroligin 3/4X (*NLGN3/4X*), SH3 and multiple ankyrin repeat domains 2/3 (*SHANK2/3*), synaptic Ras GTPase activating protein 1 (*SYNGAP1*), ubiquitin protein ligase E3A (*UBE3A*) and discs, large (Drosophila) homolog-associated protein 2 (*DLGAP2*) have been shown to be associated with ASDs [5, 53]. Known functions of these genes indicate that alteration of synaptic homeostasis could be a common biological process associated with ASD [82]. These genes, with the exception of the recently reported knockdown of *NRXN1* [83], have not yet been explored using iPSCs, and it is unknown whether iPSC-derived neurons from patients with idiopathic autism will consistently recapitulate the phenotype. Encouraging initial results have emerged from using iPSCs to model monogenic diseases that share symptoms with ASD, such as Rett syndrome [13–17, 19], Fragile X syndrome [20, 21, 84], Timothy syndrome [22, 85], and sporadic form of Schizophrenia (SZD) [86–89]. The disease-related phenotype in these studies was characterized mainly based on neuronal morphology and neurite analysis, calcium imaging, electrophysiology, immunocytochemistry, synaptic protein expression, and neuronal connectivity. In some studies gene expression profiling using microarrays and transcriptome sequencing was

applied as well. In RTT, TS and SZD, a clear disease-relevant phenotype was manifested and rescued by treatment, indicating that iPSC generation technology can recapitulate aspects of genetic disease. Other neurologic disorders modeled using iPSCs include Angelman syndrome [90], Prader-Willi syndrome [90, 91], and branched chain ketoacid dehydrogenase kinase (*BCKDK*) associated mutations [92]. However, the search for a stable end-point neuronal phenotype that may be pharmacologically modified continues. Therefore, further analysis and studies will be necessary to determine more robust differences between diseased neurons and those derived from normal iPSCs. Examples of human iPSC-derived neural cell lines and disease-related phenotypes from various neurodevelopmental disorders are summarized in Table 1.

iPSC-based studies in idiopathic and non-syndromic autism

DeRosa et al. studied iPSC-derived neurons from patients with idiopathic autism generated from peripheral blood mononuclear cells using lentiviral transduction [93]. The studied individuals displayed no evidence of known neurogenetic syndromes or confounding medical conditions which could give rise to an ASD phenotype. The resulting iPSCs efficiently formed embryoid bodies expressing markers from the three different germ layers and could be differentiated into GABAergic neurons. This study, however, did not demonstrate a cellular phenotype that distinguishes between their derived neurons and those derived from iPSCs from neurotypic individuals. Regardless, extending the iPSC-modeling technology beyond monogenic ASDs to the study of non-syndromic autism could help to expose cellular pathways that are common to many forms of autism [94].

One challenge when studying neurodevelopmental disorders, especially autism, is to account for the contribution of genetic background to the manifestation of the phenotype in individuals. Controls from unaffected proximate relatives will subtract out some of the genetic background effects, but conversely, related individuals who appear to be unaffected may in fact be manifesting an intermediate phenotype. Unrelated controls can be used to assess the latter possibility. If control cell lines are from well-phenotyped unaffected relatives, then confounding genetic heterogeneity can be limited [53]. Another scenario is to model loss-of-function monogenic disorders by mimicking down-regulation of disease-associated loci in wild type cells or by rescuing the phenotype using overexpression of wild-type alleles in diseased cells [32]. Additional approaches to generate isogenic control lines are to take advantage of X chromosome inactivation to model X-linked disorders such as RTT [14–16, 95] and FXS [20, 21], or to use targeted gene-editing or gene-engineering techniques such as zinc-finger nucleases (ZFN) [96, 97] or transcription activator-like effector nucleases (TALEN) [98].

Can transcriptome analysis of induced neurons be used as a reliable marker of disease and response to therapy?

Existing methods of phenotyping iPS-derived cells are not yet sufficiently reliable, affordable, and scalable to permit the creation of a high-throughput screening assay for autism. However, there are promising early studies of a robust transcriptional signal of pluripotency, tissue and cell specificity [30, 31, 99–101] that may allow these challenges to

be overcome. The first issue to address is whether scalable transcriptomic measurements can be found which are capable of distinguishing between induced neurons from neurotypic and autistic patients. The next step is to analyze whether such measurements and analysis will be more stable than other phenotypic measurements of these cells. Finally, measurements of the transcriptome must be validated as a tool to screen candidate drugs for preliminary signals of efficacy. The question here is whether the underlying gene expression program of a derived neuron can serve as a more robust and recognizable measure than other cellular phenotypes.

Several high-throughput technologies have been developed that enable evaluation of the coordinated expression levels of tens of thousands of genes [102], evaluate hundreds of thousands of single-nucleotide polymorphisms, and sequence individual genomes easily and inexpensively. The data produced by these assays have provided the research and commercial communities with an opportunity to define improved clinical prognostic indicators and develop a molecular understanding of the systemic underpinnings of a variety of diseases. The standard gene expression microarray is one of the most popular techniques for measuring the relative expression intensities of tens of thousands of genes simultaneously. Early acceptance of this “high-throughput” technique was limited based on several high-profile studies citing reproducibility problems [103]. Subsequently, however, many of these inconsistencies were explained by differences in the cited array technologies, post-processing normalization, and statistical analyses [104, 105]. Following this initial uncertainty, a number of studies have successfully demonstrated biological consistency among expression signatures from different high-throughput array technologies [106].

Several groups have studied the transcriptomic variability of various stages of differentiation in iPSC-derived models of neurodevelopment (Table 2) [13, 22, 72, 83, 86, 87, 92, 107]. In some studies, gene expression characteristics of specific differentiation stages could be segregated into meaningful biological and clinical subgroups [22], though the small number of samples in these studies may limit the generalizability of their results. The simplest way to expand on these results is to project gene expression data from different clinical states and differentiation stages onto a more extended platform comprising diverse tissues and disease phenotypes [31]. Typical expression analyses compare expression level across two states (e.g., cases versus controls) or a limited number of phenotypic classes. Such comparative analyses impose subjective decisions about what constitutes an appropriate control population, limiting the analysis to a specific phenotype and again reducing generalizability. To circumvent this limitation, a data-rich analysis environment for a more comprehensive approach to gene expression analysis, in which phenotypes can be characterized in the context of tissues and diseases, is required. A scalable method, referred to as Concordia (Box 1), has been put forward, that associates expression patterns with phenotypes in order to assign phenotype labels to new samples and identify phenotypically meaningful gene signatures in the context of a comprehensive transcriptomic space [31].

Box 1**Definition of comprehensive transcriptomic map: Concordia**

Concordia [31] analyzes a specific phenotype in the context of data-rich transcriptomic space, avoiding the need for predefined control groups and presupposed relationships between phenotypes. Concordia has proved to be a replicable method of characterizing a cell's lineage and state of development. It has produced a comprehensive gene expression analysis that reveals a multidimensional continuum from ESC and iPSCs to fully differentiated tissues, and identified transcription patterns associated with pluripotent stem cells [30]. This method identified genes with expression levels that are highly specific to the stem cell samples as compared to non-stem cell samples. In particular, the stem cell gene 189 set (SCGS) was identified as representative of a tightly conserved core of transcriptional programming among stem cells. This gene set was capable of differentiating between the pluripotent, multipotent, progenitor, malignant and normal samples, retaining the tissue specific features. Based on the SCGS, the authors defined an index to compare relative stemness, which allowed the differentiation between various grades of tumors, indicating that there is a high degree of stem cell-specific gene expression which differs between heterogeneous cancers.

Gene expression profiling allows a quantitative view into a case's response to the environment as prescribed by its DNA-coded instructions at any given point in time. We refer to such a profile as the cell's position in the state space of gene expression. Any underlying genetic lesion that causes an aberrant response may potentially be observable as an alteration in transcriptional activity (e.g., a change in absolute quantity of expression or alternative isoform expression), thus shifting the cell's position within this space. Concordia provides early evidence that gene expression can be a valid and stable readout, even in the data generated from various laboratories using different protocols and experimental conditions [31]. Concordia-based analysis and earlier studies show that: (i) cell identity is manifest by transcriptional activity [99–101]; (ii) developing cells follow consistent trajectories during maturation; and (iii) it is possible to measure similarity of tissue of origin and stage of maturity between cells in transcriptional space [30, 31].

The potential applications of comprehensive Concordia-type analysis include generation of a data-platform or transcriptional landscape constructed from available studies of cellular lineages and relevant primary tissues as represented in example in Box 2 (Figure I). The variance captured by such a map will further allow to frame statistically significant classification boundaries to define a phenotypes (phenotype concept enrichment), and extract the phenotype-specific marker gene signature characteristic to a cluster. This analysis will allow characterization of a specific phenotype in the context of broad transcriptomic landscape rather than in the context of dichotomous classes [31]. Precedence for Concordia-type analyses includes a recently introduced well-controlled isogenic system using human ESCs-derived Rett syndrome neurons to identify disease-related phenotype and pathways that are affected in mutant neurons [108]. Conventional profiling studies may define pathways relevant to small number of patients, however the value of Concordia will be subsequent to these discoveries by parsing multiple pathways down to reveal the most

commonly affected pathways in idiopathic ASD in the context of a comprehensive platform. Additional potential application of Concordia include the estimation of a quantitative measure of pluripotency or “stemness” index [30], previously introduced as set of genes indicative of pluripotency (Box 2, Figure II).

Box 2

Examples of potential application of Concordia-type analysis

1. **Generation of comprehensive gene expression map.** An example use of Concordia-type comprehensive transcriptomic analysis constructed from publicly available studies of human primary neuronal, stem cell-derived neuronal cultures and brain tissues, is represented on Figure I. This example database is comprised of 899 gene expression samples belonging to 25 series performed on the Illumina HumanRef-8 v3.0 expression beadchips that were obtained from NCBI’s Gene Expression Omnibus (GEO) [22, 23, 110–127]. The map uncovers gene expression alterations that result from the reprogramming of somatic tissue (fibroblasts) into pluripotent stem cells, which are then differentiated into neuronal cultures. These induced neurons may then be compared to various regions of brain and to primary neuronal cultures. As it is demonstrated in Figure I, the first two principal components (PCs) of the expression level of 17,596 genes across the database provide a representation of the phenotypic relationships and a specific signature characteristic of a differentiation stage. This map captures variance in the data and provides evidence that system is “pre-tuned” for Concordia-type analysis.
2. **Estimation of “stemness” index or pluripotency score [30]:** An example estimation of pluripotency score represented on Figure II, reflects a quantitative transcriptional measure of reprogramming and neuronal differentiation by capturing gene expression at four time points: fibroblasts, iPSCs, neural progenitors, and neurons (Fig. II (A) and (B)) [22]. Moreover, this gene expression signature collocates the four time points’ samples and clearly separates the early and late stages of differentiation in primary neuronal cultures as shown in Figure II (C) [121], although each experimental design is an independent representation and the scale is relative. Consequently, the pluripotency score between different experiments cannot be quantitatively compared and therefore a comprehensive analysis that includes tissue type, differentiation state, protocol, pluripotency and other sources of variance is required. Further work will be required to determine the relative contribution of measures such as PluriTest [44] to cell fate commitment, as compared to more comprehensive approaches.

Figure I. Example of comprehensive transcriptomic map for reprogramming and neuronal differentiation

Principal component analysis of whole-transcriptome profiles for blood, lymphoblast cell lines, brain tissue, fibroblasts, induced pluripotent stem cells, embryonic stem cells, primary neural progenitors and derived neurons showing clustering of cell types based on the first two principal components (PC1 and PC2). This database is comprised of 899 gene expression samples belonging to 25 series performed on the Illumina HumanRef-8 v3.0 expression beadchips that were obtained from NCBI's Gene Expression Omnibus (GEO)[22, 23, 110–127]. Interestingly, the gene expression signature of primary neuronal cultures (neural progenitor cells (NPCs) at 0, 2, 4 and 8 weeks) consistently shifts towards brain tissue over the course of days in culture and neural differentiation.

Figure II. Reprogramming and neuronal differentiation distribution over stemness index or pluripotency score

Each curve represents the distribution of pluripotency score values for a particular differentiation stage. The differentiation process is represented from the extreme right curve to the extreme left, with decreasing pluripotency score, clearly separating between early and late stages of differentiation. The spectrum of pluripotency in fibroblasts, induced pluripotent stem cells (iPSCs), iPS-derived neural progenitor cells (NPCs) and iPSC-derived neurons (neurons), which were derived from individuals with Timothy syndrome (A) and from control individuals (B). iPSCs were generated from dermal

fibroblasts through retroviral reprogramming and then differentiated into NPCs and neurons using conditions favoring generation of cortical neurons [22]; (C) Spectrum of pluripotency in differentiating primary normal neural progenitor cells at 0, 2, 4 and 8 weeks of differentiation *in vitro* [121].

To facilitate disease-related phenotype stable discovery, comprehensive transcriptomic analysis may help to reduce different sources of variation in iPSCs-derived systems. There may be benefits for (i) stratifying influence of genetic background in cells of origin; (ii) identifying optimal pluripotent cell line with good differentiation potential using pluripotency score; (iii) measuring the progress of derived precursor cells en route to becoming fully developed neurons, and selecting neuronal cell types and states of functional maturation which are sensitive to a specific disease. After establishing disease-related phenotype in a transcriptional space, this phenotype can be perturbed by a drug and the trajectories that result from perturbation can be evaluated. Concordia's potential is to identify the trajectory change in a transcriptional space following the influence of candidate drug, and thereby identifying candidates capable of restoring normative transcriptional behavior. A priori the "right" trajectory for the transcriptome is not obvious. A Concordia-type analysis for drug screening first requires establishing a comprehensive landscape of gene expression and phenotype concept enrichment for relevant cell types and tissues. Within this landscape, expression patterns indicative of healthy neurotypic phenotype, as well as archetypal profiles of the various axes of disease are identified. Then, in the context of this broad well-characterized transcriptomic landscape, one can identify marker genes that characterize a specific phenotype. After derivation of a gene set, whose transcriptional activity has strong predictive power, measurement on these selected targets (e.g., through qRT-PCR) in previously unseen iPSC-derived neurons assayed in the drug screen may be performed.

The importance of existing phenotypic characterization methods, such as morphological, synaptogenic, and electrophysiological phenotypes, should not be underestimated. Generation of a comprehensive neurodevelopmental landscape for further analysis requires the integration of phenotypic, clinical and transcriptional measurements in a multi-dimensional space for an accurate and consistent annotation. This requires profiling of cells derived from neurotypical individuals as well as from individuals with ASDs. Given such a landscape/space, we can seek to identify transcriptomically characteristic disease clusters. Within each disease cluster, transcriptional measurements can be used to provide a quantitative measure of perturbation of the position of the individual, as represented by their iPSCs, in this landscape for each therapeutic. As each such perturbation potentially represents a change across the entire transcriptome, standard pathway enrichment techniques, for example GSEA [109], can be used to quantify which pathways are most affected and those candidate pathways that may be perturbed to treat autism. The most important and open question suggested here is: Which range of therapeutic perturbations and which class of therapeutics will most reliably result in a shift within the gene expression state space toward the neurotypical cluster?

Concluding remarks and future perspectives

The recent introduction of iPSC strategies provides a very promising tool for creating *ex vivo* models of neurodevelopmental disorders. This approach is particularly relevant for ASDs because ASD is patient-specific and the patient's neuronal tissue is inaccessible. iPSC-derived neurons provide an ideal alternative to *in situ* brain tissue. However, the phenotypic screening approaches used today are not robust and scalable enough for high-throughput measurements of pharmacological stressors. In order to model a drug response, clear and stable phenotypic patient versus control differences should be identified and then perturbed by a stressor. In this manuscript, the use of transcriptomic analysis is discussed in various neurodevelopmental disorders as a phenotypic marker, and comprehensive gene expression analysis is proposed, based on which the trajectory of cellular development and morphological end-points can be defined for normal and disease states.

If iPSC-derived cells are to serve as clinically relevant *ex vivo* models of neurodevelopmental disease, they must be demonstrated to be robust and reproducible. The important directions of the field include deriving cellular systems from properly stratified populations with minimized genetic variation, establishing a stable characterization method for resulting cells that should be validated with other phenotypic measurements, and selecting an appropriate line, cell type and maturation level to model ASD (Box 3). Currently, no one modality of measurement can claim to provide a gold-standard perspective, therefore establishment of a characterization tool will explicitly involve a multidimensional space and integration of clinical, phenotypic and transcriptional information. Neuronal differentiation protocols will improve and comprehensive transcriptomic analysis will complement existing phenotyping techniques. Integrating information across these various modalities will buffer idiosyncrasies of specific measurement types, and thereby enable investigators to identify stable biomarkers for the dynamics of the nervous system in neurodevelopmental diseases, and provide useful end-points for future high-throughput screening using human iPSC-derived models.

Box 3

Outstanding Questions

- What is the optimal and least variable stage of differentiation to result in disease-specific phenotype?
- What is the correlation between phenotyping methods and transcriptomic characterization?
- Does the neuron-specific difference dominate the disease-specific signal?
- How does the protocol of iPSCs derivation affect stability of the end-phenotype?
- How many cell lines are needed to detect various reasonable effect size levels of change between patient lines and control lines, especially given the high variability?

Particular experiments that would be important to pursue include: (i) define the cellular system and identify critical factors and events that drive iPSCs toward “autism”-relevant neurons; (ii) evaluate expression patterns as potential markers differentiating these neural derivatives in cases and controls; (iii) evaluate the relationship among expression signatures in brain regions and in various iPSC-derived neuronal lines, as this will allow further assessment of the extent to which the transcriptional activity of iPSC-derived neurons resembles that of neurons *in vivo*; (iv) examine whether pluripotency, differentiation programs and pathways are consistent across various tissues and diseases; and (v) compare the effects on putative therapeutics on the transcriptional program of the “autistic” state versus a “normal” state determine if they correctly predict short and long-term clinical outcomes. If these results are confirmed and expanded, it will be possible to obtain fibroblasts from a patient, reprogram them into iPSCs, differentiate these cells into neurons, and characterize the neurons to screen them for the selection of a precisely individualized therapy.

Acknowledgments

We thank the members of the laboratories of Dr. Isaac Kohane (especially Dr. Patrick Schmid), Dr Louis Kunkel, Dr. Michael Greenberg, Dr. Thorsten Schlaeger, Dr. Alvin Kho and Dr. Kevin Eggan for helpful discussions and valuable advice. We also thank Dr. Mustafa Sahin for comments on the manuscript. We are particularly grateful to Dr. Bulent Ataman for his invaluable comments and scientific guidance. I.S.K. and D.P. acknowledge the support of Conte Center for Computational Neuropsychiatric Genomics (NIH P50MH94267). We apologize to all whose work we could not cite due to journal specified space limitations.

Glossary

Autism spectrum disorders (ASDs)	a heritable group of neurodevelopmental disorders associated with a complex neuropsychiatric phenotype. ASDs are characterized by deficits in social cognition, disordered communication, and the presence of stereotyped, repetitive behaviors in early childhood. Rett syndrome (RTT), Fragile X syndrome (FXS), Timothy syndrome (TS) and tuberous sclerosis are monogenic forms of neurodevelopmental disorders belonging to ASDs
Apico-basal cell polarity	asymmetric distribution of proteins and other molecules in cells forming apical and basolateral surfaces, which is implicated in differentiation, proliferation and morphogenesis
Behavioral therapy	wide range of methods used to treat psychological problems, by systematically changing behavior. One of the most supported examples of such an approach to treat autism is an applied behavior analysis (ABA)
Concordia	comprehensive transcriptomic map incorporating data from various tissues and disorders. This method allows analyzing a specific phenotype in the context of data-rich transcriptomic space [31]
Cortical neuron	main functional cell type of the brain’s cerebral cortex, enabling most of the complex activity of the brain. Various types of cortical

	neurons classified broadly by their ability to excite or inhibit neuronal activity
Dopaminergic neurons	nerve cell whose primary neurotransmitter is dopamine, and whose effect is modulatory
Embryoid bodies (EBs)	three-dimensional multicellular aggregates of differentiating pluripotent stem cells that often contain derivatives of the three germ layers
Embryonic stem cells (ESCs)	are derived from the pluripotent cells of the inner cell mass or epiblast cell layer of early embryos. ESCs possess an unlimited capacity for <i>in vitro</i> self-renewal while retaining pluripotent differentiation potential
GABAergic neuron	nerve cell whose primary neurotransmitter is γ -aminobutyric acid (GABA), and whose effect is inhibitory
Gene-editing or gene-engineering	method which allows precise genetic modification by targeted genome cleavage (double stranded break) using engineered nuclease, followed by the generation of modification during subsequent DNA repair
Glutamatergic neuron	nerve cell whose primary neurotransmitter is glutamate, and whose effect is excitatory
High-Content Screening (HCS)	a quantitative cellular imaging method, that uses high-resolution fluorescence microscopy and imaging processing tools to screen for phenotypic changes in cells, therefore referred as a phenotypic screen. It distinguishes cell types from one another based on changes in production of cellular products such as proteins and/or changes in cellular morphology
Induced pluripotent stem cells (iPSCs)	pluripotent stem cells that were generated from somatic cells by a process called reprogramming. Reprogramming can be induced by transient over-expression of genes (such as the classical Yamanaka set of reprogramming factors, Oct4, Sox2, Klf4 and c-Myc) that erases the somatic cell's epigenetic marks and activates an endogenous, self-sustaining ESC-like transcriptional regulatory circuitry. The transcriptome and developmental potential of iPSCs closely resembles that of ESCs
Morphogen	a signaling molecule that influences the organization of cells by generating concentration gradients during process of morphogenesis
Motor neuron	an efferent nerve cell, located in central nervous system (CNS), which projects axons outside CNS to innervate muscle

Neural rosette	radially organized columnar neuroepithelial-like cells, which are formed during neural differentiation of stem cells and represent developmental signature of neuroprogenitors
Neurite	neuronal branch or any projection from neuronal cell body (axon or dendrite)
Neurite analysis	a measurement of a neurite outgrowth, such as neurite length, density, number of roots and number of branching points
Pluripotency	the potential of a stem cell to differentiate into any of the three germ layers: endoderm, mesoderm or ectoderm
Principal component analysis (PCA)	a statistical technique, in which orthogonal transformation is used to convert a number of possibly correlated variables into a smaller number of linearly uncorrelated variables, referred as principal components (PCs). The first principal component has the largest possible variability in the data, with each succeeding component accounting for as much of the remaining variability as possible. PCA allows dimensionality reduction of high-dimensional data, identifying patterns, and expressing the data in a way which helps to highlight similarities and differences without losing too much information
Stem cell gene set (SCGS)	set of genes with expression levels that are highly specific to the stem cell samples as compared to non-stem cell samples, representing a tightly conserved core of transcriptional programming among stem cells [30]
Stemness index or pluripotency score	a relative measure of 'stemness'; an index or score defined over the SCGS, identifying the dominant differences between the samples within the context of the stem cell genes [30]
Synaptogenesis	the process of generation of functional synapses between neurons in the nervous system. It is a critical step in neural network generation
Telencephalon	anterior part of the embryonic forebrain or the corresponding part of the adult forebrain derived from it, including cerebral hemispheres
Transcriptome	total set of RNA molecules transcribed from the genome (mRNA, tRNA, rRNA), including all non-coding RNA, in a specific biological system

References

1. Dolen G, et al. Mechanism-based approaches to treating fragile X. *Pharmacology & therapeutics*. 2010; 127:78–93. [PubMed: 20303363]
2. Sourial M, et al. Progress toward therapeutic potential for AFQ056 in Fragile X syndrome. *Journal of Experimental Pharmacology*. 2013; 5:45–54.

3. Tropea D, et al. Partial reversal of Rett Syndrome-like symptoms in *MeCP2* mutant mice. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106:2029–2034. [PubMed: 19208815]
4. Blenner S, et al. Diagnosis and management of autism in childhood. *BMJ*. 2011; 343:d6238. [PubMed: 22021467]
5. State MW, Levitt P. The conundrums of understanding genetic risks for autism spectrum disorders. *Nature neuroscience*. 2011; 14:1499–1506.
6. Geschwind DH. Genetics of autism spectrum disorders. *Trends in cognitive sciences*. 2011; 15:409–416. [PubMed: 21855394]
7. Auerbach BD, et al. Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature*. 2011; 480:63–68. [PubMed: 22113615]
8. Kelleher RJ 3rd, et al. High-throughput sequencing of mGluR signaling pathway genes reveals enrichment of rare variants in autism. *Plos One*. 2012; 7:e35003. [PubMed: 22558107]
9. Rakic P. Evolution of the neocortex: a perspective from developmental biology. *Nature reviews. Neuroscience*. 2009; 10:724–735.
10. Chailangkarn T, et al. Modeling neurodevelopmental disorders using human neurons. *Current opinion in neurobiology*. 2012
11. Takahashi K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007; 131:861–872. [PubMed: 18035408]
12. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006; 126:663–676. [PubMed: 16904174]
13. Marchetto MCN, et al. A Model for Neural Development and Treatment of Rett Syndrome Using Human Induced Pluripotent Stem Cells. *Cell*. 2010; 143:527–539. [PubMed: 21074045]
14. Ananiev G, et al. Isogenic pairs of wild type and mutant induced pluripotent stem cell (iPSC) lines from Rett syndrome patients as in vitro disease model. *Plos One*. 2011; 6:e25255. [PubMed: 21966470]
15. Cheung AY, et al. Isolation of *MECP2*-null Rett Syndrome patient hiPS cells and isogenic controls through X-chromosome inactivation. *Human molecular genetics*. 2011; 20:2103–2115. [PubMed: 21372149]
16. Kim KY, et al. Neuronal maturation defect in induced pluripotent stem cells from patients with Rett syndrome. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108:14169–14174. [PubMed: 21807996]
17. Amenduni M, et al. iPS cells to model *CDKL5*-related disorders. *European journal of human genetics: EJHG*. 2011; 19:1246–1255. [PubMed: 21750574]
18. Hotta A, et al. Isolation of human iPS cells using EOS lentiviral vectors to select for pluripotency. *Nature methods*. 2009; 6:370–376. [PubMed: 19404254]
19. Ricciardi S, et al. *CDKL5* ensures excitatory synapse stability by reinforcing NGL-1-PSD95 interaction in the postsynaptic compartment and is impaired in patient iPSC-derived neurons. *Nature cell biology*. 2012; 14:911–923.
20. Sheridan SD, et al. Epigenetic Characterization of the *FMRI* Gene and Aberrant Neurodevelopment in Human Induced Pluripotent Stem Cell Models of Fragile X Syndrome. *Plos One*. 2011; 6
21. Liu J, et al. Signaling defects in iPSC-derived fragile X premutation neurons. *Human molecular genetics*. 2012; 21:3795–3805. [PubMed: 22641815]
22. Pasca SP, et al. Using iPSC-derived neurons to uncover cellular phenotypes associated with Timothy syndrome. *Nat Med*. 2011; 17:1657–U1176. [PubMed: 22120178]
23. Voineagu I, et al. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*. 2011; 474:380–384. [PubMed: 21614001]
24. Moy SS, Nadler JJ. Advances in behavioral genetics: mouse models of autism. *Molecular psychiatry*. 2008; 13:4–26. [PubMed: 17848915]
25. Robertson HR, Feng G. Annual Research Review: Transgenic mouse models of childhood-onset psychiatric disorders. *Journal of child psychology and psychiatry, and allied disciplines*. 2011; 52:442–475.

26. Kong SW, et al. Characteristics and predictive value of blood transcriptome signature in males with autism spectrum disorders. *Plos One*. 2012; 7:e49475. [PubMed: 23227143]
27. Sarachana T, et al. Investigation of post-transcriptional gene regulatory networks associated with autism spectrum disorders by microRNA expression profiling of lymphoblastoid cell lines. *Genome medicine*. 2010; 2:23. [PubMed: 20374639]
28. Nishimura Y, et al. Genome-wide expression profiling of lymphoblastoid cell lines distinguishes different forms of autism and reveals shared pathways. *Human molecular genetics*. 2007; 16:1682–1698. [PubMed: 17519220]
29. Hu VW, et al. Gene expression profiling of lymphoblastoid cell lines from monozygotic twins discordant in severity of autism reveals differential regulation of neurologically relevant genes. *BMC genomics*. 2006; 7:118. [PubMed: 16709250]
30. Palmer N, et al. A gene expression profile of stem cell pluripotentiality and differentiation is conserved across diverse solid and hematopoietic cancers. *Genome biology*. 2012; 13:R71. [PubMed: 22909066]
31. Schmid PR, et al. Making sense out of massive data by going beyond differential expression. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:5594–5599. [PubMed: 22447773]
32. Han SS, et al. Constructing and deconstructing stem cell models of neurological disease. *Neuron*. 2011; 70:626–644. [PubMed: 21609821]
33. Vaccarino FM, et al. Induced pluripotent stem cells: a new tool to confront the challenge of neuropsychiatric disorders. *Neuropharmacology*. 2011; 60:1355–1363. [PubMed: 21371482]
34. Okita K, et al. A more efficient method to generate integration-free human iPS cells. *Nature methods*. 2011; 8:409–412. [PubMed: 21460823]
35. Jia F, et al. A nonviral minicircle vector for deriving human iPS cells. *Nature methods*. 2010; 7:197–199. [PubMed: 20139967]
36. Warren L, et al. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell stem cell*. 2010; 7:618–630. [PubMed: 20888316]
37. Woltjen K, et al. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature*. 2009; 458:766–770. [PubMed: 19252478]
38. Kim D, et al. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell stem cell*. 2009; 4:472–476. [PubMed: 19481515]
39. Zhou W, Freed CR. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells*. 2009; 27:2667–2674. [PubMed: 19697349]
40. Fusaki N, et al. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proceedings of the Japan Academy. Series B, Physical and biological sciences*. 2009; 85:348–362.
41. Hu BY, et al. Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107:4335–4340. [PubMed: 20160098]
42. Yamanaka S. Induced pluripotent stem cells: past, present, and future. *Cell stem cell*. 2012; 10:678–684. [PubMed: 22704507]
43. Bock C, et al. Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell*. 2011; 144:439–452. [PubMed: 21295703]
44. Muller FJ, et al. A bioinformatic assay for pluripotency in human cells. *Nature methods*. 2011; 8:315–317. [PubMed: 21378979]
45. Sandoe J, Egan K. Opportunities and challenges of pluripotent stem cell neurodegenerative disease models. *Nature neuroscience*. 2013; 16:780–789.
46. Lister R, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*. 2011; 471:68–73. [PubMed: 21289626]
47. Kim K, et al. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nature biotechnology*. 2011; 29:1117–1119.

48. Polo JM, et al. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature biotechnology*. 2010; 28:848–855.
49. Kim K, et al. Epigenetic memory in induced pluripotent stem cells. *Nature*. 2010; 467:285–290. [PubMed: 20644535]
50. Sullivan GJ, et al. Induced pluripotent stem cells: epigenetic memories and practical implications. *Molecular human reproduction*. 2010; 16:880–885. [PubMed: 21059705]
51. Martins-Taylor K, Xu RH. Concise review: Genomic stability of human induced pluripotent stem cells. *Stem Cells*. 2012; 30:22–27. [PubMed: 21823210]
52. Petros TJ, et al. Pluripotent stem cells for the study of CNS development. *Frontiers in molecular neuroscience*. 2011; 4:30. [PubMed: 22016722]
53. Kim KY, et al. Cellular reprogramming: a novel tool for investigating autism spectrum disorders. *Trends in molecular medicine*. 2012; 18:463–471. [PubMed: 22771169]
54. Nat, R., et al. Human Pluripotent Stem Cells Modeling Neurodegenerative Diseases. In: Bhartiya, D.; Lenka, N., editors. *Pluripotent Stem Cells*. 2013. p. 477-516.
55. Just MA, et al. Functional and anatomical cortical underconnectivity in autism: evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cereb Cortex*. 2007; 17:951–961. [PubMed: 16772313]
56. Watanabe K, et al. Directed differentiation of telencephalic precursors from embryonic stem cells. *Nature neuroscience*. 2005; 8:288–296.
57. Eiraku M, et al. Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell stem cell*. 2008; 3:519–532. [PubMed: 18983967]
58. Eiraku M, Sasai Y. Self-formation of layered neural structures in three-dimensional culture of ES cells. *Current opinion in neurobiology*. 2012
59. Chambers SM, et al. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature biotechnology*. 2009; 27:275–280.
60. Marchetto MCN, et al. Pluripotent stem cells in neurodegenerative and neurodevelopmental diseases. *Hum Mol Genet*. 2010; 19:R71–R76. [PubMed: 20418487]
61. Gaspard N, et al. An intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature*. 2008; 455:351–357. [PubMed: 18716623]
62. Rash BG, Grove EA. Patterning the dorsal telencephalon: a role for sonic hedgehog? *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2007; 27:11595–11603. [PubMed: 17959802]
63. Gaspard N, et al. Generation of cortical neurons from mouse embryonic stem cells. *Nature protocols*. 2009; 4:1454–1463.
64. Ideguchi M, et al. Murine embryonic stem cell-derived pyramidal neurons integrate into the cerebral cortex and appropriately project axons to subcortical targets. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2010; 30:894–904. [PubMed: 20089898]
65. Maroof AM, et al. Prospective isolation of cortical interneuron precursors from mouse embryonic stem cells. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2010; 30:4667–4675. [PubMed: 20357117]
66. Danjo T, et al. Subregional specification of embryonic stem cell-derived ventral telencephalic tissues by timed and combinatorial treatment with extrinsic signals. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2011; 31:1919–1933. [PubMed: 21289201]
67. Goulburn AL, et al. Generating GABAergic cerebral cortical interneurons from mouse and human embryonic stem cells. *Stem cell research*. 2012; 8:416–426. [PubMed: 22280980]
68. Maroof AM, et al. Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. *Cell stem cell*. 2013; 12:559–572. [PubMed: 23642365]
69. Cambray S, et al. Activin induces cortical interneuron identity and differentiation in embryonic stem cell-derived telencephalic neural precursors. *Nature communications*. 2012; 3:841.
70. Siegenthaler JA, et al. Retinoic acid from the meninges regulates cortical neuron generation. *Cell*. 2009; 139:597–609. [PubMed: 19879845]
71. Shi Y, et al. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nature neuroscience*. 2012; 15:477–486. S471.

72. Mariani J, et al. Modeling human cortical development in vitro using induced pluripotent stem cells. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:12770–12775. [PubMed: 22761314]
73. Lancaster MA, et al. Cerebral organoids model human brain development and microcephaly. *Nature*. 2013; 501:373–379. [PubMed: 23995685]
74. Pfisterer U, et al. Direct conversion of human fibroblasts to dopaminergic neurons. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108:10343–10348. [PubMed: 21646515]
75. Vierbuchen T, et al. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*. 2010; 463:1035–1041. [PubMed: 20107439]
76. Brennand KJ, Gage FH. Modeling psychiatric disorders through reprogramming. *Disease models & mechanisms*. 2012; 5:26–32. [PubMed: 21954066]
77. Marchetto MC, et al. Induced pluripotent stem cells (iPSCs) and neurological disease modeling: progress and promises. *Human molecular genetics*. 2011; 20:R109–115. [PubMed: 21828073]
78. Wu C, et al. Automatic robust neurite detection and morphological analysis of neuronal cell cultures in high-content screening. *Neuroinformatics*. 2010; 8:83–100. [PubMed: 20405243]
79. Espuny-Camacho I, et al. Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo. *Neuron*. 2013; 77:440–456. [PubMed: 23395372]
80. Boulting GL, et al. A functionally characterized test set of human induced pluripotent stem cells. *Nature biotechnology*. 2011; 29:279–286.
81. Osafune K, et al. Marked differences in differentiation propensity among human embryonic stem cell lines. *Nature biotechnology*. 2008; 26:313–315.
82. Toro R, et al. Key role for gene dosage and synaptic homeostasis in autism spectrum disorders. *Trends in genetics: TIG*. 2010; 26:363–372. [PubMed: 20609491]
83. Zeng L, et al. Functional Impacts of *NRXN1* Knockdown on Neurodevelopment in Stem Cell Models. *Plos One*. 2013; 8:e59685. [PubMed: 23536886]
84. Urbach A, et al. Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells. *Cell stem cell*. 2010; 6:407–411. [PubMed: 20452313]
85. Krey JF, et al. Timothy syndrome is associated with activity-dependent dendritic retraction in rodent and human neurons. *Nature neuroscience*. 2012; 16:201–209.
86. Brennand KJ, et al. Modelling schizophrenia using human induced pluripotent stem cells. *Nature*. 2011; 473:221–225. [PubMed: 21490598]
87. Pedrosa E, et al. Development of patient-specific neurons in schizophrenia using induced pluripotent stem cells. *Journal of neurogenetics*. 2011; 25:88–103. [PubMed: 21797804]
88. Chiang CH, et al. Integration-free induced pluripotent stem cells derived from schizophrenia patients with a *DISC1* mutation. *Molecular psychiatry*. 2011; 16:358–360. [PubMed: 21339753]
89. Lin M, et al. Allele-biased expression in differentiating human neurons: implications for neuropsychiatric disorders. *Plos One*. 2012; 7:e44017. [PubMed: 22952857]
90. Chamberlain SJ, et al. Induced pluripotent stem cell models of the genomic imprinting disorders Angelman and Prader-Willi syndromes. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107:17668–17673. [PubMed: 20876107]
91. Yang J, et al. Induced pluripotent stem cells can be used to model the genomic imprinting disorder Prader-Willi syndrome. *The Journal of biological chemistry*. 2010; 285:40303–40311. [PubMed: 20956530]
92. Novarino G, et al. Mutations in *BCKD*-kinase lead to a potentially treatable form of autism with epilepsy. *Science*. 2012; 338:394–397. [PubMed: 22956686]
93. DeRosa BA, et al. Derivation of autism spectrum disorder-specific induced pluripotent stem cells from peripheral blood mononuclear cells. *Neuroscience letters*. 2012; 516:9–14. [PubMed: 22405972]
94. Dolmetsch R, Geschwind DH. The human brain in a dish: the promise of iPSC-derived neurons. *Cell*. 2011; 145:831–834. [PubMed: 21663789]

95. Pomp O, et al. Unexpected X chromosome skewing during culture and reprogramming of human somatic cells can be alleviated by exogenous telomerase. *Cell stem cell*. 2011; 9:156–165. [PubMed: 21816366]
96. Urnov FD, et al. Genome editing with engineered zinc finger nucleases. *Nature reviews. Genetics*. 2010; 11:636–646.
97. Hockemeyer D, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nature biotechnology*. 2009; 27:851–857.
98. Hockemeyer D, et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nature biotechnology*. 2011; 29:731–734.
99. Su AI, et al. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101:6062–6067. [PubMed: 15075390]
100. Butte AJ, Kohane IS. Creation and implications of a phenome-genome network. *Nature biotechnology*. 2006; 24:55–62.
101. Zheng-Bradley X, et al. Large scale comparison of global gene expression patterns in human and mouse. *Genome biology*. 2010; 11:R124. [PubMed: 21182765]
102. Emilsson V, et al. Genetics of gene expression and its effect on disease. *Nature*. 2008; 452:423–428. [PubMed: 18344981]
103. Shi L, et al. The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. *Nature biotechnology*. 2010; 28:827–838.
104. Kothapalli R, et al. Microarray results: how accurate are they? *BMC bioinformatics*. 2002; 3:22. [PubMed: 12194703]
105. Tan PK, et al. Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic acids research*. 2003; 31:5676–5684. [PubMed: 14500831]
106. Sotiriou C, Piccart MJ. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? *Nature reviews. Cancer*. 2007; 7:545–553.
107. Lin M, et al. RNA-Seq of human neurons derived from iPS cells reveals candidate long non-coding RNAs involved in neurogenesis and neuropsychiatric disorders. *Plos One*. 2011; 6:e23356. [PubMed: 21915259]
108. Li Y, et al. Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons. *Cell stem cell*. 2013; 13:446–458. [PubMed: 24094325]
109. Subramanian A, et al. 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*. 2005; 102:15545–15550. [PubMed: 16199517]
110. Chen AE, et al. Optimal timing of inner cell mass isolation increases the efficiency of human embryonic stem cell derivation and allows generation of sibling cell lines. *Cell stem cell*. 2009; 4:103–106. [PubMed: 19200798]
111. Kim JB, et al. Direct reprogramming of human neural stem cells by *OCT4*. *Nature*. 2009; 461:649–653. [PubMed: 19718018]
112. Konopka G, et al. Human-specific transcriptional regulation of CNS development genes by *FOXP2*. *Nature*. 2009; 462:213–217. [PubMed: 19907493]
113. Zaehres H, et al. Induction of pluripotency in human cord blood unrestricted somatic stem cells. *Experimental hematology*. 2010; 38:809–818. 818 e801–802. [PubMed: 20541586]
114. Pena OM, et al. Endotoxin tolerance represents a distinctive state of alternative polarization (M2) in human mononuclear cells. *J Immunol*. 2011; 186:7243–7254. [PubMed: 21576504]
115. Peixoto L, et al. Integrative genomic approaches highlight a family of parasite-specific kinases that regulate host responses. *Cell host & microbe*. 2010; 8:208–218. [PubMed: 20709297]
116. Zhu S, et al. Reprogramming of human primary somatic cells by OCT4 and chemical compounds. *Cell stem cell*. 2010; 7:651–655. [PubMed: 21112560]
117. Chen TS, et al. Enabling a robust scalable manufacturing process for therapeutic exosomes through oncogenic immortalization of human ESC-derived MSCs. *Journal of translational medicine*. 2011; 9:47. [PubMed: 21513579]

118. Cole SW, et al. Transcript origin analysis identifies antigen-presenting cells as primary targets of socially regulated gene expression in leukocytes. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108:3080–3085. [PubMed: 21300872]
119. Zhang J, et al. A human iPSC model of Hutchinson Gilford Progeria reveals vascular smooth muscle and mesenchymal stem cell defects. *Cell stem cell*. 2011; 8:31–45. [PubMed: 21185252]
120. Noggle S, et al. Human oocytes reprogram somatic cells to a pluripotent state. *Nature*. 2011; 478:70–75. [PubMed: 21979046]
121. Konopka G, et al. Modeling the functional genomics of autism using human neurons. *Molecular psychiatry*. 2012; 17:202–214. [PubMed: 21647150]
122. Chow ML, et al. Preprocessing and Quality Control Strategies for Illumina DASL Assay-Based Brain Gene Expression Studies with Semi-Degraded Samples. *Frontiers in genetics*. 2012; 3:11. [PubMed: 22375143]
123. Li W, et al. Rapid induction and long-term self-renewal of primitive neural precursors from human embryonic stem cells by small molecule inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108:8299–8304. [PubMed: 21525408]
124. Ray M, et al. Comparison of low and high dose ionising radiation using topological analysis of gene coexpression networks. *BMC genomics*. 2012; 13:190. [PubMed: 22594378]
125. Santer DM, et al. Plasmacytoid dendritic cells and C1q differentially regulate inflammatory gene induction by lupus immune complexes. *J Immunol*. 2012; 188:902–915. [PubMed: 22147767]
126. Luo R, et al. Genome-wide transcriptome profiling reveals the functional impact of rare de novo and recurrent CNVs in autism spectrum disorders. *American journal of human genetics*. 2012; 91:38–55. [PubMed: 22726847]
127. de Jong S, et al. A gene co-expression network in whole blood of schizophrenia patients is independent of antipsychotic-use and enriched for brain-expressed genes. *Plos One*. 2012; 7:e39498. [PubMed: 22761806]

Highlights

- Neurons derived from induced pluripotent stem cells represent a promising resource to model autism
- Phenotypic characterization methods for derived neurons lack reproducibility and stability
- Gene expression profiling may represent a plausible reproducible disease readout
- Projection onto comprehensive transcriptome space may assign phenotype labels and trajectories

Figure 1. Overview of a pipeline to generate iPSCs and cortical neurons for modeling ASDs

Somatic tissue such as skin is obtained from the patient and expanded into fibroblasts. These fibroblasts are reprogrammed into iPSCs, which are later differentiated to a neural forebrain fate. Cortical neural fate may be achieved using several examples of protocols. Forebrain neural progenitors may be directed into a cortical excitatory projection fate or an inhibitory interneuron fate [52]. The resulting neurons are validated for their structural and functional properties in culture. Next, these cells may be transplanted into mouse cortex to evaluate their ability to generate networks, fire action potentials and form synapses. Sources of variation throughout this pipeline are summarized [45]. Abbreviations: BMP, bone morphogenetic protein; FGF, fibroblast growth factor; GABA, γ -aminobutyric acid; IGF, insulin growth factor; iPSCs, induced pluripotent stem cells; RA, retinoic acid; SHH, sonic hedgehog; SMAD, Sma and Mothers Against Decapentaplegic signaling; TGF β transforming growth factor β Wnt, Wntless/Int proteins.

Table 1

iPSC-derived neurons and disease-related phenotypes for neurodevelopmental disorders (compiled from [10,53]).^{a,b,c,d}

Disease	Disease-related phenotype in iPSC-derived neurons	Methods to characterize disease-related phenotype	Time from harvest of fibroblasts to phenotype of neurons ^d	Refs
Rett syndrome	Reduced spine density and number of synapses, smaller soma size, altered calcium signaling and electrophysiological defects	Immunocytochemistry and neuronal morphology quantification (cell soma size, neuronal dendrites and spines, synapse quantification); Cell cycle analysis (at level of neuro-progenitors) Calcium imaging; Electrophysiology	Total: ~8–11 weeks Differentiation: 4–5 weeks	[13]
	Reduced nuclear size	Neuronal nuclei measurement	Total: ~7–10 weeks Differentiation: 30days	[14]
	Reduced neuronal soma size	Morphology (soma size)	Total: ~12–15 weeks Differentiation: ~8–9 weeks	[15]
	Defects in neuronal maturation: reduced expression of cellular markers at mature neuronal stage	Immunocytochemistry; Gene expression analysis: quantitative RT-PCR; Determination of apoptotic cells	Total: ~7–10 weeks Differentiation: 25 days	[16]
Fragile X syndrome	Fewer and shorter neurites	Immunocytochemistry; Morphology	Total: ~6–9 weeks Differentiation: (from progenitors to neurons) 19 days	[20]
	Synaptic proteins: decreased PSD95 expression; Synaptic density: decreased PSD95 puncta density; Neurite outgrowth: decreased neurite length; Calcium imaging: functionally abnormal neurons, increased amplitude/frequency and altered response to glutamate uptake	Immunocytochemistry; Synaptic protein expression: immunoblot; Calcium imaging; Neurite analysis (neurite outgrowth, number of roots, and number of branch points); Synaptic density	Total: ~7–12 weeks Differentiation: ~4–6 weeks	[21]
	Impaired calcium signaling and electrophysiology, Defect in activity-dependent gene expression; Abnormality in differentiation, decreased SATB2 expression Higher expression of tyrosine hydroxylase (TH) and increased production of catecholamines (norepinephrine and dopamine)	Number of neurons, proliferation and migration; Calcium imaging; Electrophysiology; Gene-expression profiling (microarrays); Single cell quantitative RT-PCR; Immunocytochemistry; High-performance liquid chromatography measurements (catecholamines levels)	Total: ~9 weeks Differentiation: 43days	[22]
Schizophrenia (SZD)	Diminished neuronal connectivity and decreased neurite number, reduced PSD95 protein levels and glutamate receptor expression Altered gene expression profiles	Rabies virus trans-neuronal tracing (neuronal connectivity); Neurite analysis (imaging and counting); Immunocytochemistry; Electrophysiology; Calcium imaging; Synaptic protein staining analysis (imaging), synapse density; Gene expression profiling (microarrays)	Total: ~16weeks Differentiation: 4–12 weeks	[86]
CDKL5-related disease (atypical form of Rett syndrome)	Aberrant dendritic spines; Significantly reduced number of synaptic contacts (decreased PSD95/VGLUT1 puncta density)	Immunocytochemistry Dendritic protrusion analysis (imaging); Infection with lentivirus/GFP (highlight spine morphology)	Total: ~14–16 weeks Differentiation: 55–60days	[19]

Disease	Disease-related phenotype in iPSC-derived neurons	Methods to characterize disease-related phenotype	Time from harvest of fibroblasts to phenotype of neurons ^d	Refs
BCKDK associated mutations	No disease-related phenotype reported	Cell survival and proliferation; Cell morphology (neuron body size, number of dendrites per cell), quantification of cell density and presynaptic marker density; Immunocytochemistry; Gene expression profiling	Total: ~6–12 weeks Differentiation: ~6 weeks	[92]

^a Abbreviations: (iPSCs), induced pluripotent stem cells; *BCKDK*, branched chain ketoacid dehydrogenase kinase; *CDKL5*, cyclin-dependent kinase-like 5; PSD95, postsynaptic density protein 95; RT-PCR, reverse transcription polymerase chain reaction; *SATB2*, special AT-rich sequence-binding homeobox 2; SZD, schizophrenia; TH, tyrosine hydroxylase; VGLUT1, vesicular glutamate transporter 1.

^b All periods of time for reprogramming and differentiation were calculated based on the assumption that a well-established protocol already exists. The time required for reprogramming is an estimate based on manuscripts discussed; time window for differentiation represents time required (as mentioned in the relevant manuscript) for differentiation of iPSCs to mature/intermediate characterized stage described in the paper (neural progenitors or neurons).

^c We are mentioning the studies in which authors attempted to perform phenotypic comparison between disease model to control.

^d Time estimation is including 3–6 weeks reprogramming to produce iPSC intermediate.

Table 2

Gene expression data of iPSCs and derived neurons. ^{a,b}

Original phenotype/Disease	GSE #	Source Tissue	Target Tissue	iPSC reprogramming	Gene expression data	Platform	Ref
Timothy syndrome	25542	Fibroblasts	Cortical neurons (glutamatergic, GABAergic)	Retroviral	1 Fibroblasts	Illumina HumanRef-8 v3.0 expression beadchip	[22]
					2 iPSCs		
					3 iPSC-derived neurospheres (day 7)		
					4 iPSC-derived neurons at rest (day 45)		
					5 iPSC-derived neurons kept in KCl		
					6 ESCs		
					7 ESC-derived neurospheres		
					8 ESC-derived neurons at rest		
Rett syndrome	21037	Fibroblasts	Cortical neurons (GABAergic, glutamatergic)	Retroviral	1 Fibroblasts	Affymetrix Human Gene 1.0 ST Array	[13]
					2 iPSCs		
					3 ESCs		
BCKDK mutation	39447	Fibroblasts	Tuj1+/Map2+ neurons	Episomal	1 Fibroblasts	Nimblegen Homosapiens HG18 Expression Array	[92]
					2 iPSCs		
					3 iPSC-derived neural progenitors		
					4 iPSC-derived neurons		
Schizophrenia	25673	Fibroblasts	Glutamatergic (most, ~60%) GABAergic (30%) Dopaminergic (<10%)	Lentiviral	1 iPSC-derived neurons (6weeks old)	Affymetrix Human Gene 1.0 ST Array	[86]
	26629	Skin fibroblasts, liver fibroblasts	Glutamatergic neurons	Retroviral	1 iPSCs 2 iPSC-derived neurons (day 10) 3 iPSC-derived neurons (day 32)	Affymetrix Human Gene 1.0 ST Array	[87]
NRXN1 knockdown	-	Fibroblasts	Neural stem cells, Tuj1+ neurons, astrocytes, oligodendrocytes	PiggyBac transposon	1 iPSC-derived neural stem cells (0 weeks)	Illumina HiSeq 2000	[83]

Original phenotype/Disease	GSE #	Source Tissue	Target Tissue	iPSC reprogramming	Gene expression data	Platform	Ref
Additional studies (no defined phenotype)					2 iPSC-derived neural stem cells (4 weeks)		
	32625	Fibroblasts	Glutamatergic neurons	Retroviral	1 iPSCs 2 iPSC-derived neurons	Illumina HiSeq 2000	[107]
	41565	Fibroblasts	Cortical neurons: early forebrain neurons, glia, glutamatergic, GABAergic	Retroviral	1 iPSCs 2 ESCs 3 iPSC-derived neurons (50 days)	Illumina HumanHT-12 V4.0 expression beadchip	[72]

^a Abbreviations: iPSCs, induced pluripotent stem cells; *BCKDK*, branched chain ketoacid dehydrogenase kinase; ESCs, embryonic stem cells; GEO, Gene Expression Omnibus; GSE, GEO series; KCl, potassium chloride; *NRXN*, neuroligin.

^b We have excluded the transcriptomic data from iPSC-derived neurons from Parkinson's, Alzheimer's, spinal muscular atrophy fronto-temporal dementia and other neurological disorders despite its existence in the database, to focus on neurodevelopmental space.