nature

https://doi.org/10.1038/s41586-024-08487-6

Accelerated Article Preview

A framework for neural organoids, assembloids and transplantation studies

Received: 26 August 2024

Accepted: 4 December 2024

Accelerated Article Preview

Cite this article as: Pașca, S. P. et al. A framework for neural organoids, assembloids and transplantation studies. *Nature* https://doi.org/10.1038/s41586-024-08487-6 (2024)

Sergiu P. Pașca, Paola Arlotta, Helen S. Bateup, J. Gray Camp, Silvia Cappello, Fred H. Gage, Jürgen A. Knoblich, Arnold R. Kriegstein, Madeline A. Lancaster, Guo-Li Ming, Gaia Novarino, Hideyuki Okano, Malin Parmar, In-Hyun Park, Orly Reiner, Hongjun Song, Lorenz Studer, Jun Takahashi, Sally Temple, Giuseppe Testa, Barbara Treutlein, Flora M. Vaccarino, Pierre Vanderhaeghen & Tracy Young-Pearse

A framework for neural organoids, assembloids and transplantation studies

Sergiu P. Pașca[1,2*], Paola Arlotta[3,4], Helen S. Bateup[5,6], J. Gray Camp[7,8], Silvia Cappello[9,10], Fred H. Gage[11], Jürgen A. Knoblich[12,13], Arnold R. Kriegstein[14], Madeline A. Lancaster[15,16], Guo-Li Ming[17,18,19,20], Gaia Novarino[21], Hideyuki Okano[22], Malin Parmar[23], In-Hyun Park[24], Orly Reiner[25,26], Hongjun Song[17,18,19,27], Lorenz Studer[28], Jun Takahashi[29], Sally Temple[30], Giuseppe Testa[31,32], Barbara Treutlein[33], Flora M. Vaccarino[34,35,36], Pierre Vanderhaeghen[37,38], Tracy Young-Pearse[39,40]

1. Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA.
2. Stanford Brain Organogenesis, Wu Tsai Neurosciences Institute and Bio-X, Stanford University, Stanford, CA, USA.
3. Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA.
4. Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA.
5. Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, CA, USA.
6. Department of Neuroscience, University of California Berkeley, Berkeley, CA, USA.
7. Institute of Human Biology (IHB), Roche Pharma Research and Early Development, Roche Innovation Center Basel, Basel, Switzerland
8. Biozentrum, University of Basel, Basel Switzerland
9. Department of Physiological Genomics, Biomedical Center (BMC), Faculty of Medicine, Ludwig Maximilian University of Munich, Germany.
10. Max Planck Institute of Psychiatry, Munich, Germany
11. Laboratory of Genetics, The Salk Institute for Biological Studies, La Jolla, CA, USA.
12. Institute of Molecular Biotechnology, Austrian Academy of Sciences, Vienna Biocenter, Vienna, Austria.
13. Department of Neurology, Medical University of Vienna, Vienna, Austria
14. Department of Neurology and Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California San Francisco, San Francisco, CA, USA
15. MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, UK
16. Cambridge Stem Cell Institute, University of Cambridge, Cambridge UK
17. Department of Neuroscience and Mahoney Institute for Neurosciences, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA, USA.
18. Department of Cell and Developmental Biology, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA, USA.
19. Institute for Regenerative Medicine, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA, USA.
20. Department of Psychiatry, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA, USA.
21. Institute of Science and Technology of Austria, Klosterneuburg, Austria.
22. Keio University Regenerative Medicine Research Center, Kanagawa, Japan.
23. Department of Experimental Medical Science, Lund Stem Cell Center, Lund University, Lund, Sweden.

24. Department of Genetics, Yale Stem Cell Center, Yale School of Medicine, New Haven, CT, USA.
25. Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel.
26. Department of Molecular Neuroscience, Weizmann Institute of Science, Rehovot, Israel.
27. The Epigenetics Institute, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA, USA.
28. The Center for Stem Cell Biology, Developmental Biology Program, Sloan Kettering Institute for Cancer Research, New York, NY, USA.
29. Department of Clinical Application, Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan.
30. Neural Stem Cell Institute, Rensselaer, NY, USA.
31. Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy.
32. Human Technopole, Viale Rita Levi Montalcini, Milan, Italy.
33. Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland.
34. Child Study Center, Yale University, New Haven, CT, USA.
35. Department of Neuroscience, Yale University, New Haven, CT, USA.
36. Yale Kavli Institute for Neuroscience, New Haven, CT, USA.
37. VIB-KU Leuven Center for Brain & Disease Research, Leuven, Belgium.
38. Department of Neurosciences, Leuven Brain Institute, Leuven, Belgium.
39. Department of Neurology, Ann Romney Center for Neurologic Diseases, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA.
40. Harvard Stem Cell Institute, Harvard University, Cambridge, MA, USA.

*Correspondence: spasca@stanford.edu

Preface

As the field of neural organoids and assembloids rapidly expands, there is an emergent need for guidance and advice on designing, conducting and reporting experiments to increase the reproducibility and utility of these models. Here, our consortium– representing specialized laboratories from around the world– presents a framework for the experimental process that ranges from ensuring the quality and integrity of human pluripotent stem cells to characterizing and manipulating neural cells in vitro, and from transplantation techniques to considerations for modeling human development, evolution, and disease. As with all scientific endeavors, we advocate for rigorous experimental designs tailored to explicit scientific questions, and transparent methodologies and data sharing, to provide useful knowledge for both current research practices and for developing regulatory standards.

Introduction

The field of stem cell-based modeling of human development, evolution and disease using organoids and assembloids has seen a tremendous surge in interest with more than 3,000 articles published annually. Together with two-dimensional (2D) cellular models, these three-dimensional (3D) preparations, whether utilized *in vitro* or *in vivo* after transplantation into animals, leverage the inherent self-organization capabilities of stem and progenitor cells to mimic aspects of physiology and hold the potential to reveal new human tissue biology and pathophysiology[1]. Their application for neuroscience is particularly welcomed as inaccessibility has limited the biological understanding of the human nervous system and the mechanisms underlying neurodevelopmental and neurodegenerative disorders. It is, however, becoming increasingly difficult to navigate the expanding literature and extract what are the experimental standards for building reliable 3D microphysiological systems, especially as *in vitro* studies are often not reported with sufficient experimental detail[2]. Importantly, these complex multi-cellular systems are maintained in long-term cultures (often for hundreds of days) and assessments require increasingly sophisticated assays that range from cellular -omics to functional assays and circuit-probing in vitro and in animals. Consequently, unlike experimental systems with faster turnaround and simpler readouts, adjustments based on successful or unsuccessful outcomes in a human cellular model are often delayed, which impacts experimental progress limiting reproducibility across laboratories. As numerous groups are implementing these multi-cellular systems for development, evolution, or disease applications, and as funding, publishing and regulatory agencies become engaged, it is becoming increasingly important to outline critical experimental variables and challenges for this rapidly expanding field.

Following a similar consortium effort to clarify nomenclature for the field[1], we gathered a group of international researchers to outline an experimental framework that can apply to neural organoids, assembloids and their xeno-transplantation, to highlight some of the limitations and to delineate quality control measures and standards (**Figure 1**). This consensus arises from our experience in modeling the development, evolution, and diseases of the human nervous system. While we encourage methods development, we emphasize that experiments with these models should be thoughtfully designed to answer a specific scientific question as knowledge of the variables can also mitigate the cost of experiments. In other words, considering the variety of experimental models available, including 2D models, with their advantages and caveats, it is critical to match the experimental system to the question. Moreover, we underscore the need for transparency in sharing protocols and experimental details, including how quantifications were performed, and the requirements for depositing raw data in public databases.

Quality control measures can vary widely based on experimental, translational or clinical goals. However, the variables we ~~are~~ describe here should offer sufficient guidance for tailoring experimental designs to a wide array of objectives and needs. We hope that this summary, derived from our collective experience will serve as a guidance to those in-or entering this field, especially trainees, and ultimately will lead to more optimally designed studies. Lastly, we anticipate that this framework for the experimental process may have implications for regulatory agencies such as FDA, MHRA, EMA and PMDA relevant to developing therapeutics in an academic or industry setting, and that many of these guidelines will be relevant to both 2D cellular models and to 3D models of other organs.

A foundational step: quality of hPS cells

One of the most important considerations is the quality of the human pluripotent stem cell (encompassing embryonic and induced, hES or hiPS) lines used to derive neural and other cell lineages. These cells are not only prone to spontaneous differentiation and sensitive to cell

culture conditions and growth factors to maintain pluripotency, but each hiPS cell line can harbor hundreds to thousands of single nucleotide mutations, copy number variants and genomic/epigenomic changes. While some genomic variants can be acquired in culture even after relatively short periods (5–10 passages), others reflect somatic mutations inherited from founder cells[3-6]. Even hPS cell lines considered isogenic may not always be strictly isogenic and reprogramming itself can be mutagenic[4]. Therefore, it is crucial to confirm at the onset of obtaining and then regularly verify and report genome integrity status as well as to authenticate hPS cell lines as suggested by the International Society for Stem Cell Research (ISSCR) guidelines on stem cell research and modeling[7]. It is important to note that classic karyotyping methods or low-pass whole-genome sequencing (WGS) may lack the necessary resolution, and~~/or~~ array-based DNA technologies, such as array-comparative genomic hybridization (array CGH) and single nucleotide polymorphism (SNP)-arrays are recommended. There is no consensus on a minimal number of de novo genetic events in a hPS cell line that warrants exclusion, and finding a genetic variant does not necessarily compromise the results as long as they are described in detail in the manuscript and unless they impact genes known to interfere with cellular function directly[7]. We also caution that erosion of X-chromosome inactivation and abnormal imprinting can impact the reliability of disease models[8], particularly in female-derived cells where it can lead to abnormally high levels of gene expression from the normally inactive X chromosome [9]. Crucially, experiments should be accompanied by detailed information on the cell lines used (including the sex, age, and ancestry background of the donor), the culture conditions, the range of passage numbers used, the genome integrity, and Mycoplasma contamination status following, for instance, the checklist recommendations by the ISSCR[7]. Moreover, hPS cell lines that are compared should be cultured under the same conditions. The use of control and patient cell lines that are not matched demographically or obtained from different sources is not recommended. For example, deriving all control lines in one laboratory and all patient lines in another can introduce a confounding variable. To minimize these issues, gene editing should ideally be performed in the same lab and using the same methods.

With the advent of CRISPR engineering, there has been a rise in the use of genetically engineered hiPS cell lines in addition to the use of patient-derived cell lines. While this approach provides greater control over genetic background, it is important to note that CRISPR/Cas9-modified hiPS cell lines require thorough re-characterization, particularly in terms of pluripotency and genome integrity, if subcloning of hiPS cell lines is used after editing. The consensus is that multiple genetic backgrounds (individuals) should be subjected to editing to generalize findings. Experimental design could include unmodified hiPS cell lines at the targeted locus that went through the process of editing, the parental line, and multiple edited clones to increase the likelihood that phenotypic effects are attributable to the desired on-target editing and not to artifacts of cloning or editing. Another valuable control is to use revertant lines, in which the mutations are corrected in the same cell line. Additionally, orthogonal validation with patient-derived lines or other strategies, such as shRNAs, CRISPRa/i or antisense oligonucleotides, is recommended. When assessing the effect of environmental factors or studying copy number variations (CNVs) that cannot be engineered into control hiPS cell lines, large cohorts of patient-derived lines with demographically matched controls remain the gold standard. Cell villages, chimeroids and mosaic organoids are advancing the ability to achieve the large sample sizes needed for these studies.

Production of hiPS cells requires appropriate donor consent for making and using these lines. Consent restrictions, differences in the international regulatory landscape and institutional policies governing the sharing of human cell lines have sometimes prevented the broader use of a set of lines across laboratories and, overall, present a significant challenge in

comparing experiments. The KOLF2.1J line (of European ancestry), has been proposed as an all-around well-performing hiPS cell line for collaborative studies[10] but more than one line is needed. Several repositories currently provide hiPS cell lines (CIRM, NIH, Allen Institute, SFARI, Coriell, JAX). Moving forward, creating a collection of broadly consented, well-characterized, genetically and ancestrally diverse hiPS cell lines named in a standardized manner, could greatly facilitate protocol benchmarking and accelerate the validation of disease phenotypes across various research settings.

Generating and characterizing neural cells

Human neural development spans hundreds of days, extending decades into the postnatal period. Timing of *in vitro* neural differentiation is largely conserved in most culture conditions and consequently, modeling neurogenesis, astrogenesis, oligodendrogenesis, synapse or circuit formation with human cells will involve long, laborious experiments. Given the complexity and duration of these experiments, it is crucial to carefully design and power them[11]. Due to variability across cell lines, particularly when establishing new protocols and confirming phenotypes, we recommend employing multiple hiPS or hES cell lines and experimental differentiation batches to enable robust statistical analysis (for example, using two different XX and two different XY donor hiPS cell lines ideally in several independent differentiation experiments). It is critical to set a priori criteria for inclusion of lines and avoid removing cell lines post-hoc. Differentiations will inevitably need to be staggered, and incorporating quality control steps at key stages can enhance reliability and save resources. For instance, during the first month of cortical organoid differentiation, a panel of region-specific markers covering domains of the nervous system can be readily checked by qPCR or similar inexpensive methods (e.g., *FOXG1* for forebrain, *EMX1* for dorsal forebrain, *NKX2.1* or *DLX2* for ventral forebrain, *TCF7L2* for diencephalon, *EN1* for midbrain or anterior hindbrain including cerebellum, *HOX* family genes for hindbrain and spinal cord, *TTR* for choroid plexus, as examples). Morphological measures, such as diameter and shape, can sometimes indicate issues of cell identity or viability with some caveats depending on the method: if grown in suspension, organoids need to be separated to avoid fusion, while if embedded in extracellular matrices it can be difficult to run certain assays where access to cells is needed. Cell death and organoid morphology should be monitored and documented over time and across batches. These characteristics can vary depending on the differentiation method and the culture conditions and this variability may interfere with phenotyping. It will be helpful to report the degree of cell death in the core of the organoid and to describe how this was accounted for in experiments or mitigated to reduce cavitation, especially in long term cultures. Immunocytochemistry can also be employed, but while some cytoarchitectural features can reliably be observed (e.g., ventricular-like zones), more mature neuroanatomical features are not generally present in current multi-cellular preparations. For instance, distinguishing subventricular-like zones and outer subventricular-like in cortical organoids or defining multiple cortical layers can often be challenging. Functional quality control experiments should be considered, including characterization using calcium imaging, patch clamp or extracellular recording or neurotransmitter release.

Developing a new differentiation protocol necessitates extensive characterization and validation (**Box 1**). Implementing established protocols should, at a minimum, confirm cell identity in the neural cultures derived from the hPS cell lines used and assess assay variability for the phenotype of interest.

Patch-clamping or extracellular recordings can be complemented by calcium or voltage imaging in intact organoids or assembloids. Immunocytochemical characterization of organoids has often been limited in the literature, with only selected examples or part of organoids

presented. Moving forward, this characterization should include multiple organoid sections, multiple organoids, and multiple experimental batches with appropriate normalization procedures that consider unbiased sampling of cell distribution, quantification of the total number of cells and, when possible, comparison to primary tissue samples. For the development of assembloids, chimeroids and mosaic organoids, it is essential to carefully detail the conditions for integration, including the optimal stages of differentiation for fusion and mixing, respectively, and assays. For instance, when it comes to generating neuro-immune assembloids in which neural organoids are integrated with immune cells, the timing of integration and the compatibility of the cell culture media must be carefully considered. For chimeroids and mosaic organoids, examination of clonal dynamics (hPS or differentiated cells), patterning of neural progenitors used for mixing and the ideal timing is required. Claims of circuit formation in assembloids should include evidence of connectivity by imaging of projections and, ideally, analyses of functional connectivity, such as retrograde or anterograde tracing, synapse formation and electrophysiology.

There should be transparency in reporting experimental details. Methods sections should include information on cell culture media and composition, cell culture dishes, cell density, oxygen levels, concentration and timing of growth factors added, passage number of starting cells and passage method, the efficiency of formation of a particular structure and how the structure is defined, the lots of various reagents (including extracellular matrices), and the presence of serum or other undefined components. Finally, it is important to deposit protocols in public repositories such as protocols.io, provide cell type annotations and be available to address questions and share reagents to facilitate method reproducibility across labs.

We also consider that it is time to broadly implement established best practices for single cell analysis across modalities[12], including detailed reporting of meta-data and the uploading of raw data (if permitted by donor consent) and partially processed data (e.g., count matrix) on public repositories such as NIH GEO, dbGAP, NeMO archive, CellxGene, NIMH Data Archive (NDA), UCSC browser, etc. even when publishing in journals that do not require this. This will be increasingly important as integration across datasets and comparison to primary tissue molecular atlases become standard procedures. Benchmarking to developmental cell atlases, when available, is necessary to substantiate claims about deriving specialized cell types, rather than solely relying on the expression of a few selected markers. Similarly, claims regarding physiological properties should be supported by evidence beyond the mere presence of a specific cell type in an organoid. While dissection of parts of organoids for profiling has been commonly performed, we recommend profiling intact organoids to obtain a comprehensive overview of cell composition. If only subregions are used or organoids are pooled, this should be transparently reported. To expedite the derivation and validation of new cell types, developing quality scores of similarities and user-friendly tools, such as VoxHunt[13], is important.

Significant effort has been dedicated to developing neural differentiation protocols. However, a major challenge arises when implementing these protocols in other laboratories, as they are often subject to modifications or optimizations. While improvements are always welcome, even minor adjustments to a protocol, such as altering the concentration or timing of growth factor application, schedule for media changes, etc., may require revalidation of reliability, as well as the identity and functionality of resulting 3D cultures. This practice also poses challenges for reproducibility efforts, as meta-analyses or comparisons across published studies become exceedingly difficult.

Probing and manipulating neural cells in 3D

Perhaps one of the next major challenges in leveraging the potential of human multi-cellular models is the lag in developing and optimizing tools for capturing the functional properties of these complex 3D cultures. So far, most phenotyping efforts have focused on assessing cell diversity and molecular signatures through single-cell transcriptomics or immunocytochemistry and the morphology of organoids. However, unless there are severe brain structural defects in patients (e.g., microcephaly), the morphology of organoids is rarely a reliable measure for phenotyping. A dramatic reduction in the size of organoids, especially those carrying mutations from patients with only minor or no changes in brain or cortical volume, is often more indicative of technical issues related to genetic engineering, differentiation processes, or variability between different cell lines. In fact, characteristics of hiPS cell lines, their genetic background and the stage of differentiation (timing), contribute to a large fraction of the variability in these models[14,15]. Therefore, it may be necessary to increase the number of lines used in experiments. Even when sample sizes are limited, precise reporting of technical details, including specific timepoints, the number of cell lines, batches, and replicates for each experiment, is crucial for reproducibility and reliability. Special attention should be paid to the research question, the scalability of the assay, and the corresponding sample size. For instance, addressing polygenic inheritance likely requires large numbers of lines and readouts that can be adapted at scale. To substantiate claims about genetic background effects in disease models, a large number of independent lines is likely required, often necessitating orthogonal assays for validation or the examination of multiple pedigrees. In these cases, it is also useful to report data on multiple lines from a single individual. Furthermore, the timing for assessing alterations is a crucial factor and will vary depending on the brain region model, differentiation method or disorder being studied. For the spinal cord, which develops faster, experiments can extend up to 40-50 days *in vitro*. However, for corticogenesis, several months or more are necessary to observe the generation of deep and superficial layer neurons (neurogenesis), unless strategies to accelerate or stop neurogenesis are used. For astrogenesis, assays at early stages in cortical organoids are often confounded by progenitors with astrocytic features, for example radial glia rather than astrocytes, the latter often require over 4-5 months *in vitro* to develop, unless strategies to accelerate or stop gliogenesis are used[16]. Notably, very early stages of development of guided neural organoids (i.e., first few weeks of *in vitro* differentiation) can sometimes exhibit more variability than later stage organoids, likely due to the time needed for various hiPS cell lines and states to reach key stages. Related to timing, assessing the maturation or 'aging' status of cell-based organoids and assembloids is crucial for establishing the relevance of disease models, especially for studying neurodegenerative diseases. The same principles hold for neurotoxicology, where the physiological relevance and the developmental stage of the model used should be described and assessed for effective integration with epidemiological data.

Functional assays in organoids and assembloids provide insights into neural development, maturation, and disease modeling, with fluorescence imaging, neurotransmitter release and electrophysiological methods being central to these efforts. Calcium imaging with genetically encoded indicators, such as gCaMPx, is relatively straightforward to implement and can now be used to monitor activity across large populations of neurons in intact neural organoids or across multi-part assembloids. Combining this imaging readout with reporter labeling facilitates the identification of cell-type-specific effects; however, challenges remain in finding, validating, and delivering cell-type-specific reporters for human neural cells, and these reporters often become silenced over time. In contrast to electrophysiological methods, however calcium imaging has a slower temporal resolution, is susceptible to photobleaching and phototoxicity,

and cannot be equated with spiking or synaptic activity, especially in developing human neural networks.

Patch clamp electrophysiology, sometimes coupled with optogenetic and chemogenetic manipulations, although low-throughput and requiring advanced expertise, remains the gold standard for measuring the activity of individual neurons due to its resolution and precision in capturing the electrical state and membrane properties of a cell. Because this approach is low throughput, it is important to perform power analysis, considering cell type diversity and maturity, to accurately establish the needed sample size. Meanwhile, extracellular recordings provide a less invasive alternative, capturing the spiking activity of neuronal populations and can also be coupled with optogenetic stimulation or neurotransmitter uncaging strategies. These methods are typically used acutely or require the organoids or assembloids to be sectioned, which can disrupt their three-dimensional structure and potentially affect physiological properties. Multielectrode arrays, on the other hand, can be limited by lower signal resolution and can introduce artifacts, especially if 3D organoids require long-term plating on a flat surface. Moving forward, there is need for recording platforms that can seamlessly integrate with and record chronically (for hundreds of days, ideally) from intact neural organoids and assembloids without interfering with their 3D self-organization and differentiation unlike methods that require 2D culture plating for recording. Moreover, human primary tissue recordings are important as a benchmark for neural recording studies involving organoids and assembloids.

Lastly, we emphasize that the complexity of an assay should not preclude the rigor, or the sample size required to obtain robust results. A range of cell lines, biological and technical replicates, and a combination of functional assays are often necessary to ensure that findings are robust.

Transplanting and probing human neural cells in animals

Organoid and assembloid models of the nervous system are constrained by their lack of functional vascularization, incomplete morphological and electrophysiological maturity of neurons and glial cells, and lack of meaningful sensory and other physiological inputs that shape circuit activity during development. Transplantation of human neural cells into host organisms can mitigate some ~~of these~~ limitations by enabling vascularization and integration into living host circuits. Transplantation methods hold significant potential for developing cell therapies and potentially offering a preclinical platform for testing drugs directly on human cells.

Several aspects of experimental design must be considered when developing transplantation models, whether grafting whole organoids or dissociated neural cells derived from 2D or 3D cultures. These experiments are complex, and demand surgical expertise. The developmental stage of the transplanted cells and the host at the time of engraftment—whether prenatal, early postnatal, or adult— could influence outcomes. Quality control measures are important, given the extended duration of these experiments, sometimes exceeding nine months. Ensuring cell quality prior to transplantation, in terms of viability, purity, and cell identity is vital for enhancing reproducibility. Magnetic resonance imaging offers a non-invasive and efficient method to observe graft integration and growth, although at low resolution, and can facilitate the refinement of new transplantation procedures. When possible, two-photon imaging can be useful and informative given its resolution and ability to monitor circuit function, but technically challenging[17]. For graft characterization, we advise verifying survival, proliferation, maturation, and migration within the central nervous system of the host, alongside monitoring changes in cell identity and graft composition over time and assessing possible cell fusion– a process in which individual human and host cells become one by merging their plasma membranes. Identifying human cells may necessitate staining with human-specific antibodies

(e.g., anti-HNA, STEM-121, STEM-123, hNCAM, h.nuc) or orthogonal approaches to distinguish host versus graft cells. Lastly, functional integration into host circuits requires validation, including anatomical reconstruction of projections, labeling of efferent and afferent connections potentially with viral vectors, and electrophysiological characterization, *in vivo* imaging and optogenetic manipulation, ensuring identification of the transplanted cells to be distinguished from the host cells, and, when relevant, assessment of potential impact on animal behavior.

It is crucial to recognize that xenograft transplantation into the rodent nervous system typically occurs against an immunocompromised background or in immunocompetent animals receiving immunosuppressive treatments, constraining the neuro-immune interaction questions that can be asked. Developmental timelines are species-specific; therefore, engraftment is intrinsically heterochronic, and rodents are not always the ideal species for transplantation, and other laboratory animals may have to be considered, although the use of primates as a host is ethically challenging.

The transplantation of human cells into the nervous system of animals introduces various ethical, moral, societal, and legal concerns that are the focus of ongoing discussions[18]. Here, we want to underscore the importance of ethical aspects, including obtaining human donor consent for the use of human cells *in vivo* in animals, ensuring animal welfare, and being mindful of public perceptions regarding the purpose of these experiments and their promise to develop therapeutics. As the field progresses and transplantation techniques improve, it is essential to identify potential ethical tipping points, including emergent features, ensuring that these considerations are integrated into the research process and animal welfare[19].

Considerations on modeling development, evolution and disease

At the core of developing useful, predictive models of biological processes is rigorous benchmarking. To ensure the reliability and relevance of experimental outcomes, it is critical to seek to obtain some sort of 'ground truth' information for normal human brain development and disease models. This is often challenging for neuropsychiatric disorders but can sometimes be achieved by comparison to human postmortem or surgically-resected tissue and clinical or neuroimaging data, using *ex vivo* fetal tissue preparations or relevant animal models (including non-human primates), which remain essential tools for neuroscience. Myriads of disease phenotypes have been described in hiPS cell-derived neural cultures over the past 15 years; some of them involve major changes in cell type composition, electrophysiological function or cell morphology that are in apparent contrast to the clinical presentation in the same patients. Therefore, it is possible that such phenotypes may be artefactual, a partial manifestation of the *in vivo* phenotype, or a manifestation of a phenotype that would be compensated *in vivo* or at a neural circuit level. This is unsurprising, as *in vitro* differentiation is often noisy and can amplify phenotypes, and these defects should be interpreted in the context of the disease observed in humans. Therefore, asserting major claims regarding the pathophysiology of complex and poorly understood psychiatric disorders without corroborating findings through independent validation in alternate experimental systems or with clinical data can be misleading. It is important to consider that these are *models* of aspects of disease and should be interpreted as such. This is especially important for neuropsychiatric disorders (neurodevelopmental and neurodegenerative), as most of them are behaviorally defined.

Ensuring robust statistical and experimental design constitutes another fundamental principle, including the use of randomization, blinding, and *a priori* power calculations It is crucial to define samples: individual cells, individual versus bulk organoids or assembloids, differentiation batches, transplanted animals, hiPS cell lines of the same or different donors. When multiple data points are collected from the same experimental batch without maintaining

independence, there is a significant risk of inflating false positive results. Graphs should show all data collected and points should be colored (or shaped) by the line or batch used; figure legends should include details of what the points represent. A supplementary table could indicate lines and replicates used in each experiment. It is also important to distinguish between technical (assessing variability introduced by the experimental measurement or procedure) and biological replicates (accounting for variability between different biological entities or samples). Additionally, transparency regarding which organoids or assembloids were excluded from analysis and the rationale behind these decisions [2], including the definition of technical failure, is paramount. Such transparency can promote the integrity and reproducibility of research, especially with complex cellular models of the human nervous system.

When modeling evolutionary or developmental features using organoids and assembloids, several experimental considerations and limitations should be considered. The developmental stage of the organoids and assembloids should be considered, especially as comparison across species is challenging, and claims of 'uniquely human' evolutionary features may be related to heterochrony– a phenomenon by which similar developmental processes unfold over different timescales in different species, with neoteny potentially playing a role in retaining early developmental features. Variation in pluripotency states across species adds another layer of complexity, potentially impacting the timing of differentiation *in vitro* and the ability to make accurate phenotypic comparisons. Cell culture conditions under which these models are maintained, particularly the use of high glucose media and hyperoxia, may alter developmental and evolutionary-relevant processes and could influence experimental outcomes. Importantly, claims about modeling extinct hominid brain development or function—are particularly challenging as we have limited information on the cognitive and social functioning of these species. These considerations underscore the need for an experimental design suited to these unique challenges and the acknowledgment of inherent limitations when using models to study complex biological processes.

Lastly, modeling disease using human stem cell-based organoids and assembloids necessitates a deep understanding of the inherent complexities and limitations of these systems. Bridging the gap between cellular models and clinical presentations remains a formidable challenge, and we urge caution in interpreting results and linking cellular phenotypes to complex cognitive and behavioral processes. A primary consideration is the genetic heterogeneity among hiPS cell donors, which is significantly larger than the genetic variance seen in inbred animal models. This variability underscores the critical importance of choosing appropriate controls for these experiments, including obtaining and reporting detailed donor information and, ultimately registering hiPS cell lines to ensure other groups can easily obtain them. For hiPS cell studies, matching clinical features of the cohort, such as sex, and ancestry or using genetically related controls may be an approach to reduce heterogeneity. Unlike in clinical trials where sample size estimation and power analyses are standard prerequisites, such practices are not mandatory for *in vitro* experiments, even when utilizing patient-derived cells. However, the implementation of power analysis to determine the necessary sample size for detecting an expected effect size is essential. In cases where specific power analysis is not feasible or relevant, employing the largest possible sample size is important. The utilization of multiple cell lines from the same individual has been shown to reach a power plateau[11], indicating that incorporating cells from multiple individuals is more likely to increase robustness. Ultimately, unlocking the translational potential of this field will necessitate the adoption of enhanced statistical methodologies; for instance, employing linear and generalized mixed effect models[20] that account for data dependence can provide a more accurate analysis compared to traditional methods like t-tests and ANOVA, which assume

independence among observations. It is advisable to consult statisticians to ensure the appropriate application of these methods.

Final remarks

We are at an exciting time in neuroscience fueled by the potential of human neural cellular models. These human stem cell-based models, combined with xenograft and other experimental approaches, are poised to accelerate our understanding of human development, evolution, and disorders of the nervous system. In conjunction with the prior consensus on nomenclature[1], the framework for the experimental process ~~that~~ we propose here will hopefully accelerate the application of organoid, assembloid, and other multi-cellular models, bringing us closer to realizing their translational potential. This framework and further international collaborative initiatives across laboratories promise to enhance reliability and reproducibility, improve communication within and across disciplines, informing trainees and those entering this field as well as publishing, funding and regulatory agencies, and hopefully sparking further dialogue.

Acknowledgments

We would like to thank members of our labs who provided feedback on earlier version of this manuscript, including A. Jourdon, V. Mariano, TL Li, N. Caporale, E. Villa, M. Sutcliffe.

Contributions

All authors discussed the content of the manuscript and reviewed and edited the entire manuscript. S.P.P. led the writing of the manuscript following extensive discussions and with input from all authors.

Competing interests

The authors declare no competing interests.

Figure 1

Differentiation of human pluripotent stem cells into unguided or guided organoids, assembloids and their transplantation into animals. Boxes indicate assays and other variables that should be taken in consideration when designing experiments.

Box 1

Recommendations for characterization and validation of a new differentiation protocol necessitates extensive (Box 1).

- Morphological characterization, including monitoring and documenting the size and shape of organoids over time, as well as establishing cell health of the 3D culture (e.g., by using dyes for cell viability)
- Validation across multiple, genetically distinct hiPS or hES cell lines.
- Dynamic assessment of expression patterns of region-specific markers through gene expression (qPCR) or spatial assessment of key markers by immunocytochemistry, ideally compared to primary brain tissue samples or published relevant datasets.
- Single-cell or nucleus RNA profiling of individual or pooled organoids at several stages, followed by computational mapping onto human reference atlases (Allen Brain Atlas, Human Cell Atlas, UCSC, etc). After reproducibility is established with single organoid profiling (scRNA-seq, qPCR, etc), organoids can be pooled.
- Assessment of reliability across lines, experimental batches, and individual organoids.
- Functional characterization through imaging and electrophysiological recordings.

References

1 **Pasca, S. P. *et al.* A nomenclature consensus for nervous system organoids and assembloids. *Nature* 609, 907-910 (2022). https://doi.org:10.1038/s41586-022-05219-6**
A consensus article providing a framework for the nomenclature of human multicellular models of nervous system development and disease.

2 RIVER, W. G. Reporting in vitro experiments responsibly– the RIVER recommendations. *MetaArxiv* (2023).

3 Lezmi, E., Jung, J. & Benvenisty, N. High prevalence of acquired cancer-related mutations in 146 human pluripotent stem cell lines and their differentiated derivatives. *Nat Biotechnol* (2024). https://doi.org:10.1038/s41587-023-02090-2

4 Panda, A. *et al.* Clonally Selected Lines After CRISPR-Cas Editing Are Not Isogenic. *CRISPR J* **6**, 176-182 (2023). https://doi.org:10.1089/crispr.2022.0050

5 Abyzov, A. *et al.* Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* **492**, 438-442 (2012). https://doi.org:10.1038/nature11629

6 International Stem Cell, I. *et al.* Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nat Biotechnol* **29**, 1132-1144 (2011). https://doi.org:10.1038/nbt.2051

7 **Ludwig, T. E. *et al.* ISSCR standards for the use of human stem cells in basic research. *Stem cell reports* 18, 1744-1752 (2023). https://doi.org:10.1016/j.stemcr.2023.08.003**
The International Society for Stem Cell Research (ISSCR)'s recommendations and reporting criteria for enhancing the reproducibility and rigor of stem cell research.

8 Mekhoubad, S. *et al.* Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell* **10**, 595-609 (2012). https://doi.org:10.1016/j.stem.2012.02.014

9 Bar, S., Seaton, L. R., Weissbein, U., Eldar-Geva, T. & Benvenisty, N. Global Characterization of X Chromosome Inactivation in Human Pluripotent Stem Cells. *Cell reports* **27**, 20-29 e23 (2019). https://doi.org:10.1016/j.celrep.2019.03.019

10 Pantazis, C. B. *et al.* A reference human induced pluripotent stem cell line for large-scale collaborative studies. *Cell Stem Cell* **29**, 1685-1702 e1622 (2022). https://doi.org:10.1016/j.stem.2022.11.004

11 Brunner, J. W. *et al.* Power and optimal study design in iPSC-based brain disease modelling. *Mol Psychiatry* **28**, 1545-1556 (2023). https://doi.org:10.1038/s41380-022-01866-3

12 Heumos, L. *et al.* Best practices for single-cell analysis across modalities. *Nat Rev Genet* **24**, 550-572 (2023). https://doi.org:10.1038/s41576-023-00586-w

13 Fleck, J. S. *et al.* Resolving organoid brain region identities by mapping single-cell genomic data to reference atlases. *Cell Stem Cell* **28**, 1177-1180 (2021). https://doi.org:10.1016/j.stem.2021.03.015

14 Gordon, A. *et al.* Long-term maturation of human cortical organoids matches key early postnatal transitions. *Nat Neurosci* **24**, 331-342 (2021). https://doi.org:10.1038/s41593-021-00802-y

15 Jourdon, A. *et al.* Modeling idiopathic autism in forebrain organoids reveals an imbalance of excitatory cortical neuron subtypes during early neurogenesis. *Nat Neurosci* **26**, 1505-1515 (2023). https://doi.org:10.1038/s41593-023-01399-0

16 Wang, M. *et al.* Morphological diversification and functional maturation of human astrocytes in glia-enriched cortical organoid transplanted in mouse brain. *Nat Biotechnol* (2024). https://doi.org:10.1038/s41587-024-02157-8

17 Schafer, S. T. *et al.* An in vivo neuroimmune organoid model to study human microglia phenotypes. *Cell* **186**, 2111-2126 e2120 (2023). https://doi.org:10.1016/j.cell.2023.04.022

18 Farahany, N. A. *et al.* The ethics of experimenting with human brain tissue. *Nature* **556**, 429-432 (2018). https://doi.org:10.1038/d41586-018-04813-x

19 **National Academies of Sciences Engineering and Medicine (U.S.). Committee on Ethical Legal and Regulatory Issues Associated with Neural Chimeras and Organoids, National Academies of Sciences Engineering and Medicine (U.S.). Committee on Science Technology and Law & National Academies of Sciences Engineering and Medicine (U.S.). Policy and Global Affairs. *The emerging field of human neural organoids, transplants, and chimeras : science, ethics, and governance*. (The National Academies Press, 2021).**
An overview of ethical and governance issues associated with human neural organoids, assembloids and their transplantation put forward by a committee of scientists, ethicists, and legal scholars conveved by the National Academies of Sciences Engineering and Medicine.

20 Yu, Z. *et al.* Beyond t test and ANOVA: applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron* **110**, 21-35 (2022). https://doi.org:10.1016/j.neuron.2021.10.030