

It's time to incorporate diversity into our basic science and disease models

To address health disparities and facilitate increasingly personalized treatments, we need to develop new models for basic and disease research that reflect diverse ancestral backgrounds and sex, and ensure that diverse populations are included among donors and research participants.

Rick Horwitz, Ekemini A. U. Riley, Maria T. Millan and Ruwanthi N. Gunawardane

The power of science, a long-time cornerstone of disease mitigation, has been on public display during the ongoing COVID-19 pandemic in the form of new strategies for vaccine development, epidemiological methods, therapeutic agents, and diagnostic tests. However, this pandemic also underscores the importance and urgency of tuning the power and reach of science to address disparities in infection and outcome. In addition to cultural, environmental and socioeconomic status, host genetic factors appear to have a role as at least 13 loci have been linked to infection or severe disease outcomes¹. Genetic ancestral relationships also appear to be significant. For example, a COVID-19 risk locus for severe disease is present in 50% of people in South Asia but only in 16% of people in Europe².

The role of genetic background and ancestral relationships is seen in other complex diseases, through the presence of variants associated with certain diseases that differ among ancestries³. Despite this, most human research data derive largely from white populations of European ancestry^{4,5}. Fortunately, this disparity is on the radar of both governmental and non-governmental organizations. For example, the US National Institutes of Health (NIH) has sponsored the All of Us Research Program, which aspires to build "...one of the most diverse health databases in history"⁶. The RESPOND research network is investigating the underlying factors that contribute to the increased risk profile for men of African ancestry for prostate cancer⁷, and the Global Parkinson's Genetics Program is genotyping more than 150,000 people across the world to better understand the genetics of Parkinson's disease, resulting in the most ancestrally diverse Parkinson's population dataset so far^{8,9}.

While progress is being made at the population level, conversations about the need for ancestral and sex diversity in basic human disease research models

are largely absent; yet they are crucial for understanding human biology and actualising the twenty-first century goal of personalized medicine through targeted clinical strategies for the mitigation and prevention of disease. Ostensibly, this omission is due, in part, to the challenges of investigative research on living humans at the molecular, cellular and tissue level.

Human induced pluripotent stem (hiPS) cells and hiPSell-derived organoids are model systems ripe for realizing the need for diversity in ancestral background and sex in human biomedical research. A library of well-characterized, standardized and accessible hiPS cells from diverse ethnic populations would be a game-changer for both investigative and translational research on humans. They would aim to replace the genomically unstable cancer cell lines that are commonly used in research¹⁰ with replicating diploid cells that differentiate into many kinds of cell types and tissues.

We propose that a suite of well-characterized hiPS cell lines with ancestral and sex diversity would become a benchmarked collection of cell lines for basic science and studies of disease mechanisms, and for the introduction of disease mutations and performance of screens for novel therapeutics. Moreover, the inclusion of isogenic hiPS cell lines would provide the reproducibility needed for comparisons among different researchers, as well as between healthy and pathological cells. These libraries should include standardized, reproducible and openly shared protocols for cell growth and differentiation, tools for their use and study, and baseline data on phenotypes, genomics and cell organization. A repository for high-quality data emerging from these lines would also be useful. Finally, tool lines incorporating well-characterized disease mutations could be included in these collections, establishing a baseline against which to compare intrinsically variable patient-derived hiPS cells.

Such resources, while laborious and time-consuming to create, are already in play through nascent efforts. The California Institute for Regenerative Medicine is moving in this direction with a repository of more than 2,600 hiPS cells from individuals of diverse ancestries, including African, Hispanic, Native American, East and South Asian, and European¹¹. These lines were developed using a standardized protocol and available with standardized consents¹² and pre-negotiated licenses for commercial use. In addition, the NIH-sponsored iPSC Neurodegenerative Disease Initiative will model more than 100 different mutations associated with Alzheimer's disease and related dementias, all in an isogenic hiPS cell background¹³. This represents the largest genome engineering project so far — a major step in the right direction and a great benefit to the neurodegeneration field. Finally, the isogenic collection of over 50 hiPS cell lines created by the Allen Institute for Cell Science could serve as a model for what high-quality repositories might look like^{14,15}. These isogenic lines are well-characterized genomically and biologically and are openly available along with tools and protocols for their use; this approach could be expanded to include diversity in donor ethnicity and sex.



Clearly, the science and approaches are in place for initiatives that lead to a diverse suite of well-characterized isogenic hiPS cell lines and promote their use in biomedical research. The next step is the will and financial support to move forward rapidly.

To reach its potential, some challenges will need to be addressed. Donor permissions are now required for the production and distribution of cell lines with supporting genomic data; however, some may still be hesitant to deposit their cells and accompanying genomic information. With the enormous benefits of diverse research models in sight, hopefully donors will emerge to fill the need. Furthermore, hiPS cells often do not differentiate to the

full adult state, hindering research and disease models requiring the adult state, and organoids are still an emerging field and not standardized for all applications. These latter concerns are active areas of investigation that will be buoyed by an initiative similar to this. A natural extension of the hiPS cell work would be a parallel effort to develop tissue and brain resources with donor diversity, further strengthening the model systems used for biomedical research.

If we are to truly understand human biology, address health disparities and personalize our treatments, we need to go beyond our important, ongoing efforts in addressing diversity and inclusion in the workforce and the delivery of healthcare. We need to improve the data we generate by including diverse populations among donors and research participants. This will require new models and tools for basic and disease research that more closely reflect the

diversity of human tissues, across diverse donor backgrounds. In doing so, it will be critical to involve populations that have been historically marginalized in biomedical research. A diverse suite of well-characterized isogenic hiPS cell lines and plans to promote their use in biomedical research could help cell biologists achieve our shared goals of increasing diversity and gender equality, and our knowledge of human biology. □

Rick Horwitz¹✉, Ekemini A. U. Riley² ,
Maria T. Millan³ and
Ruwanthi N. Gunawardane¹ 

¹Allen Institute for Cell Science, Seattle, WA, USA.

²Coalition for Aligning Science, LLC, Chevy Chase, MD, USA. ³California Institute for Regenerative Medicine, Oakland, CA, USA.

✉e-mail: rickh@alleninstitute.org

Published online: 29 November 2021

<https://doi.org/10.1038/s41556-021-00803-w>

References

1. COVID-19 Host Genetics Initiative. *Nature* <https://doi.org/10.1038/s41586-021-03767-x> (2021).
2. Zeberg, H. & Pääbo, S. *Nature* **587**, 610–612 (2020).
3. Wojcik, G. L. et al. *Nature* **570**, 514–518 (2019).
4. Sirugo, G., Williams, S. M. & Tishkoff, S. A. *Cell* **177**, 26–31 (2019).
5. Peterson, R. E. *Cell* **179**, 589–603 (2019).
6. *All of Us Research Program* (NIH, 2021); <https://allofus.nih.gov/>
7. *RESPOND Research Network* (RESPOND, 2021); <https://respondstudy.org/Default.aspx>
8. *Global Parkinson's Genetics Program* (2021); <https://www.gp2.org/about-gp2/>
9. The Global Parkinson's Genetics Program *Mov. Disord.* **36**, 842–851 (2021).
10. Drubin, D. G. & Hyman, A. A. *Mol. Biol. Cell.* **28**, 1409–1411 (2017).
11. *iPSC Repository* (CIRM, 2021); <https://www.cirm.ca.gov/researchers/ipsc-repository>
12. Lin, S. S., DeLaura, S. & Jones, E. M. *Stem Cell Res.* **44**, 101671 (2020).
13. Ramos, D. M. et al. *Neuron Apr.* **109**, 1080–1083 (2021).
14. *Cell Catalog* (Allen Institute for Cell Science, 2021); <https://www.allencell.org/cell-catalog.html>
15. Roberts, B. et al. *Mol. Biol. Cell.* **28**, 2854–2874 (2017).

Competing interests

The authors declare no competing interests.