Human stem cell-based disease modeling: prospects and challenges
Joshua Z Johnson and Dirk Hockemeyer

Human stem cell-based disease models have great promise to advance our understanding of human disease. These models can be derived from patients with genetic disorders and manipulated with genome editing and myriad differentiation protocols to model pathologies in vitro. However, several challenges have impeded the full potential of stem cell-based in vitro disease modeling. Many genetically predisposed diseases take time to manifest and occur in specific tissue microenvironments, and these parameters are often not adequately modeled using conventional shorter-term monolayer cultures. These challenges must be overcome especially for cases where animal models also incompletely recapitulate the complex pathologies found in humans. As prominent ways to tackle these challenges we discuss here how advanced genome editing tools in human stem cells and human organoid cultures, specifically the example of intestinal organoids, contribute genetically defined models that recapitulate phenotypes of disease.

Address
Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA
Corresponding author: Hockemeyer, Dirk (hockemeyer@berkeley.edu)

A significant challenge in iPSC-based disease modeling lies in the fact that each disease-specific iPSC line is genetically distinct due to the genetic variability among patients [6]. As a consequence, phenotypes of iPSC disease models can show striking variability between individual patient-derived cell lines [7]. Moreover, variability can also be caused by the reprogramming process used to create the iPSCs [8,9]. This variability greatly challenges our ability to model disorders with mild or complex phenotypes. Recently, we and others have overcome this limitation by establishing the use of site-specific nucleases (reviewed in [10,11]) in hPSCs, allowing a level of genetic control previously limited to traditional model systems [12,13–15]. As a result, we can now perform targeted gene knock-outs, generate tissue-specific cell lineage reporters, overexpress genes from defined loci, and introduce and repair point mutations in hPSCs. This genetic amenability of hPSCs allows researchers to generate sets of isogenic cells that differ exclusively at the site of editing. Consequently, the phenotypes identified in these cells can be attributed to the disease-relevant mutation rather than the specific genetic background of a given patient.

A proof of concept for this approach in hPSCs was the genome editing-mediated correction of disease-causing mutations in α-synuclein that cause a familial form of Parkinson’s disease [16]. Comparing isogenic cortical neurons differentiated from these iPSCs identified that α-synuclein mutations caused accumulation of nitrosative and endoplasmic reticulum stresses [17]. Furthermore, comparing these isogenic iPSCs in a similar approach showed that α-synuclein disease-causing mutations predisposed iPSC-derived dopaminergic neurons to mitochondrial stresses from environmental toxins known to be associated with Parkinson’s disease [18].

An elegant approach to increasing the efficiency of gene repair of disease alleles in vitro combines genome editing with the use of piggyBac transposase to correct patient-derived iPSCs for a point mutation in the α(1)-antitrypsin gene, which causes α(1)-antitrypsin deficiency [19]. This gene repair approach utilizes antibiotic selection of a zinc finger nuclease (ZFN)-mediated correction of the disease-causing allele using a selection cassette. Overexpression of piggyBac transposase can later be used to ‘scarelessly excise’ the selection cassette once a corrected hPSC clone is isolated and genotyped. This strategy yields efficient bis-allelic changes in patient-derived iPSCs, and restores enzymatic function of iPSC-derived and transplanted hepatocytes.
A more general translational application of genome editing that increased the versatility of iPSC-based disease modeling has been demonstrated for trisomy 21. Jiang et al. [20] showed that Down syndrome patient-derived iPSCs could be engineered to insert an inducible gene for the Xist IncRNA into chromosome 21. Induction of Xist in the edited iPSCs transcriptionally represses the third copy of chromosome 21 and thereby reverses cellular disease phenotypes in vitro.

Since these initial studies utilizing ZFNs [12*,21**] and transcription activator-like effector nucleases (TALENs) [13,22–24], the advent of the ‘Cas9 revolution’ — the establishment of site specific nucleases based on the bacterial adaptive defense system CRISPR (Clustered Regularly Interspersed Short Palindromic Repeats)/Cas9 (Cas9) — has made genetic engineering of stem cells a widely available and standard tool in human disease modeling. Since the founding work by Jinek et al. [25*], Cas9 has
been used to increase the feasibility of gene targeting, especially in the context of editing human cells [26–28]. The ease of Cas9-mediated genome editing and its ability to be arrayed with multiple guide RNA (gRNA) sequences [29] allow us to simultaneously interrogate several loci. This provides a platform to mechanistically understand the tremendous influx of human disease genetics identified by genome-wide association studies (GWAS). For example, we recently used this approach to elucidate the mechanism by which mutations in the promoter of the catalytic subunit of telomerase (TERT) [30,31], which are the most frequent non-coding mutations in human cancer identified by GWAS [32,33], facilitate tumorigenesis [34]. By introducing these promoter mutations and comparing otherwise isogenic cell lines, we demonstrated that hESCs were unable to silence TERT expression upon differentiation. Thus, we were able to use results of GWAS to guide our hypotheses and then use precise Cas9-mediated genome editing and differentiation protocols to elucidate the mechanism by which these mutations facilitate cancer.

Adaptations and improvements have increased the ease and scope of Cas9-mediated disease modeling in hPSCs. Cas9-mediated genome editing requires a gRNA that bears homology to the genomic locus to be modified. The target locus must have a short sequence motif called a protospacer-adjacent motif (PAM), 5’-NGG-3’ in the case of Cas9 from Streptococcus pyogenes (spCas9), to be present 3 bases away from the cut site, limiting the number of potential cut sites in the genome. One approach for circumventing this limitation is to use CRISPR/CAS-systems from different organisms that respond to different PAM sequences [35–37]. An alternative approach is to engineer spCas9 to respond to alternative PAMs by structure-based engineering of Cas9 [38]. These orthogonal Cas9s are useful in designing editing strategies because of the mechanism of double-strand break repair in human cells [39,40]. Editing strategies that target sites closer to the intended modification are expected to significantly increase the efficiency of targeting [39,40], which is especially important in the correction or introduction of point mutations to generate isogenic cell lines.

Cas9 technologies can be used to introduce functional reporters, which when combined with phenotypes that can be robustly assayed may allow for the rapid determination of complex epistatic networks that govern disease phenotypes. To that end, Cas9 can be efficiently integrated into the AAVS1 safe harbor locus [12,13] as an inducible system in hPSCs [41] to rapidly and robustly generate triple knockout cell lines in one editing step [42]. This platform could be employed to study complex diseases and dissect complex epistasis by the simultaneous modification of several genomic loci. Additionally, Cas9 has been used in genome-wide screens with gRNA libraries targeting multiple loci within all known protein-coding genes [43,44]. Such screening strategies are expected to become a standard tool to test for genome-wide loss-of-function phenotypes within the background of disease-modeling hPSCs.

**Genetically defined model systems of human tissues and microenvironments**

Modeling complex diseases using an *in vitro* stem cell tissue culture system is dependent upon the establishment of robust differentiation protocols that permit the isolation of a large number of a discrete population of cells that can be fully differentiated into functional cells and tissue types of the adult human. While differentiation protocols exist for a limited number of cell types, the lack of protocols for many other cell types has curtailed efforts for modeling human pathologies. Furthermore, established protocols mostly focus on the isolation and purification of individual cell types rather than the derivation of a functional tissue that appropriately recapitulates the disease-relevant microenvironment. Thus, modeling the biology of human disease that manifests because of the complex interplay of different cell types in a tissue becomes a challenge. Consequently, current tissue culture models largely fail to recapitulate the milieu of changes that cells acquire over extended periods of disease progression.

One approach to overcome the challenges of monolayer cultures is to develop three-dimensional organoid cultures to study human disease. Organoids are a cellular culture system that generates multiple cell types of the corresponding organ spatially organized as they are *in vivo* and that recapitulates aspects of the tissue’s function (reviewed in [45]). Such cultures can be derived from hPSCs or human tissue samples, and these model systems can be manipulated using the genome editing technologies described above. Organoid systems have so far been established to model a variety of organs such as the intestine [46], stomach [47,48], pancreas [49], mammary tissue [50], brain [51], liver [52], and lung [53]. Here we will review how these organoid technologies combined with genome editing have been used to model human disease with a focus on intestinal organoids.

**Genome editing in human organoids to elucidate intestinal biology**

Intestinal biology significantly differs between mice and humans making it necessary to develop novel approaches to understand intestinal pathology in a human-derived *in vitro* context. Human intestinal organoids were first derived from adult intestinal stem cells [46], but later work showed that they could also be derived from hPSCs [54]. An alternative approach that our lab has recently developed used an intestinal stem cell marker, LGR5, labeled with GFP by genome editing [55]. Edited hESCs were injected into immune-compromised mice, allowed to
differentiate in the *in vivo* environment, and the labeled intestinal stem cell marker was used to isolate those cells that were then cultured to produce intestinal organoids in a 3D matrix.

Differentiation of hPSCs into intestinal tissue results in an immature transcriptional profile that resembles stages of early development and fetal tissue despite the ability to generate all of the major cell types of the epithelium [55, 56]. Similar observations were made upon analysis of pancreatic beta cells derived from hPSCs [57] and in differentiation protocols that differentiate cells into hepatocytes [58, 59]. Failure to differentiate into mature adult tissue poses a general challenge for disease modeling in hPSC-derived tissues. In the case of intestinal organoids, recent progress was made to overcome these challenges by developing xenograft transplantation protocols of growing human iPSC-derived intestinal organoids in the mouse kidney capsule [60]. Here the human cells matured and adopted a transcriptional profile more similar to that of adult tissue, suggesting that functional xenograft transplantation could be a general approach that allows for the necessary maturation and differentiation of hPSCs to model pathologies that manifest in adulthood.

An alternative to obviate the issue of ‘maturity’ in the *in vitro* systems is to use primary tissue-derived organoid cultures isolated from patients for disease modeling, as these already possess the developmental tissue properties of an adult donor. This has been done successfully with the intestine in the context of cystic fibrosis (CF) [61]. In intestinal organoid cultures isolated from CF patient samples the disease-causing allele in the CF transmembrane conductor receptor can be repaired using Cas9-mediated genome editing, restoring the function of the adult primary tissue-based intestinal organoids. Primary tissue-derived intestinal organoids [47] have also been used to understand the genetics of different stages of colorectal cancer [62, 63]. The most common mutations of colorectal cancer found in *APC, P53, KRA* and *SMAD4* were combinatorially introduced to model the different stages of colorectal cancer. Mutations in *APC* and *P53* were sufficient to show hallmarks of tumor progression, but all four mutations allowed the tissue to show the most severe and physiologic cancer phenotype in xenografts [63]. Similar approaches might be used in the future to parse GWAS data sets for less well understood cancer mutations to determine which genes are important in the driving of disease, and which may be only passengers or important for later stages.

Addressing the challenge of modeling the tissue microenvironment involves introducing the *in vivo* complexity into a tissue culture system. Aspects of such complexity can be seen by culturing human intestinal organoids with bacterial pathogens in order to model host–pathogen interactions. Interaction of *Clostridium difficile* [64] and *Salmonella enterica* serovar Typhimurium [65] with human intestinal organoids results in clinically relevant responses and phenotypes such as an appropriate inflammatory cytokine response [65].

Although these genetically amenable models offer much potential for understanding intestinal biology, they do not fully recapitulate the intestine, as it is a complex organ with essential links to the vasculature, enteric nervous system, immune system, and the microbiota. A potential solution is engraftment of intestinal organoids into a mammalian intestine, which would provide these missing physiological components. Initial studies already demonstrate allotransplantation of mouse organoid tissue [66], and these experiments were recently translated to the xenograftment of fetal human colon-derived organoids into the colon of mice [67]. Despite challenges of xenograftment, such as in many cases the need for immune-compromised hosts and potentially aberrant interactions of host factors with human tissue, we predict that it will become a powerful system to understand disease and the activity of putative drugs in a complex human tissue context.

**Conclusions**

In summary, there have been tremendous advancements in the modeling of human disease. These include the still ongoing discovery of disease alleles from GWAS, genome editing and its implications in stem cell genetics, and the ability to induce patients’ cells to pluripotency. These advancements are of interest because the mechanistic insight gleaned from studies using these approaches can directly influence drug design and clinical strategies. Applying these tools to organoid cultures that more closely recapitulate the *in vivo* microenvironment and that ideally can functionally engraft into mammalian model organisms will greatly facilitate our ability to understand and treat human disease.

**Acknowledgements**

We thank Ryan Forster and Franziska Lorbeer and all other members of the Hockemeyer lab for helpful discussion. Dirk Hockemeyer is a New Scholar in Aging of the Ellison Medical Foundation and is supported by the Glenn Foundation as well as The Shurl and Kay Curci Foundations and the NIH (R01/HL12884A).

**References and recommended reading**

Papers of particular interest, published within the period of review, have been highlighted as:  
- of special interest  
- of outstanding interest


This paper establishes that use of site-specific nucleases to edit human pluripotent stem cells.


These studies provide the proof of concept for the generation of isogenic induced pluripotent stem cell disease models using genome editing.


These studies provide the proof of concept for the generation of isogenic induced pluripotent stem cell disease models using genome editing.


This paper elucidates the mechanism of the most common non-coding mutations in cancer, as described by GWAS, by comparing Cas9-edited isogenic pluripotent stem cells.


