Review

Synthetic gene circuits as tools for drug discovery

Adam M. Beitz,1 Conrad G. Oakes,2 and Kate E. Galloway1,*

Within mammalian systems, there exists enormous opportunity to use synthetic gene circuits to enhance phenotype-based drug discovery, to map the molecular origins of disease, and to validate therapeutics in complex cellular systems. While drug discovery has relied on marker staining and high-content imaging in cell-based assays, synthetic gene circuits expand the potential for precision and speed. Here we present a vision of how circuits can improve the speed and accuracy of drug discovery by enhancing the efficiency of hit triage, capturing disease-relevant dynamics in cell-based assays, and simplifying validation and readouts from organoids and microphysiological systems (MPS). By tracking events and cellular states across multiple length and time scales, circuits will transform how we decipher the causal link between molecular events and phenotypes to improve the selectivity and sensitivity of cell-based assays.

Enhancing phenotypic drug discovery with synthetic biology

Development of new drugs requires identification and optimization of functional molecules that often proceeds sequentially through high-throughput discovery screening, hit validation, hit expansion and optimization, and preclinical disease modeling (see Glossary) (Figure 1, Key figure) [1]. Target-based drug discovery (TDD) programs aim to identify drug candidates that modulate a defined biological target that is hypothesized to play a causal role in initiating or sustaining a disease. TDD programs offer precision by directly screening for drug candidate efficacy using purified molecular targets [2]. By contrast, phenotypic drug discovery (PDD) programs aim to identify drug candidates that modulate a physiologically-relevant biological system or cellular signaling pathway and screen for drug efficacy using cell-based assays [2]. Cell-based assays are target-agnostic and provide the opportunity to identify therapeutics that resolve disease-associated phenotypic deficits even when the disease etiology remains obscure and targets are undefined. Thus, PDD has recently received renewed interest for its potential to identify first-in-class drugs that act via a novel mechanism of action (MoA) as well as identify drugs for disease with complex or unknown etiology [1–4]. In order to identify drug candidates that translate to effective therapeutics in the clinic, PDD programs require cell-based assays that faithfully recapitulate and report on disease-relevant processes. To efficiently and effectively screen in phenotypic models, assay development for initial discovery screens must balance throughput with the ability to capture these disease processes [1]. As PDD probes a larger biological space than TDD to identify potential therapeutics, the number of hits from an initial discovery screen in a PDD program may be very large [2]. Following initial discovery screens, hit validation and triage are required to produce a small number of high-confidence candidates to proceed to the resource-intensive stages of hit expansion, compound optimization, and higher-complexity preclinical models [4].

Developing cell-based assays for PDD that can identify molecules that translate to effective therapeutics in the clinic remains a substantial challenge [2,3]. Monitoring disease-relevant events and cellular states across multiple length and time scales will transform how we decipher the causal link between molecular events and phenotypes to improve the selectivity and sensitivity of cell-based assays.

Highlights

- Synthetic gene circuits enable real-time monitoring of diverse molecular events in live cells.
- Synthetic circuits provide internal controls to expedite hit validation and facilitate hit prioritization.
- Synthetic circuits enable on-line validation of induced pluripotent stem cell (iPSC)-derived cells.
- Synthetic circuits potentiate longitudinal tracking of neurodegenerative-associated phenotypes.

1Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA
2Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

*Correspondence: katiegal@mit.edu (K.E. Galloway).
biomarkers often requires endpoint assays that preclude longitudinal tracking of cellular populations. Hit validation and triage can require multiple additional assays and extensive molecular profiling to remove hits that act via undesirable MoAs. Higher-complexity in vitro disease models such as organoids and MPS remain difficult to validate and interpret, particularly at high throughput. To address these challenges, synthetic gene circuits may serve as ideal tools for improving PDD by providing live-cell, multiplexed readouts of an intracellular state (Figure 1). Synthetic gene circuits are engineered networks of genes linked through transcriptional and post-transcriptional regulation. Native biological circuits utilize data within the cell to coordinate cellular behaviors in space and time. Synthetic biology aims to develop tools that tap into these native networks and dynamically access information in the cell. In this review, we highlight how synthetic circuits can enhance PDD by reporting diverse molecular events to maximize the information gained through high-content imaging, by validating and monitoring multicellular systems, and by tracking dynamic, disease-relevant molecular events.

From tools to systems: building synthetic gene circuits to assay diverse molecular events

Live cell-based assays utilize reporter genes to sense and translate the activity of pathways and gene regulatory networks (GRNs) into a quantifiable output. Reporter assays have been adapted in both drug candidate screening [5] and pharmaceutical quality control [6]. Reporter-based assays that only detect transcriptional activation in a cellular pathway have had limited success in identifying novel therapeutics, potentially indicating limited predictive validity of these assays [7]. However, transcriptional activation represents only a small portion of the biological space impacted by diseases. Given that transcriptional activation alone has performed poorly, sensors that report on more specific and disease-relevant molecular events may improve the predictive validity of reporter gene assays. Synthetic circuits can incorporate transcriptional, post-transcriptional, and post-translational regulatory elements to more closely mimic disease-associated pathways.

Synthetic biologists have developed a number of tools and circuits that induce gene expression in response to a range of intracellular events from splicing to MAPK signaling (Figure 2A–D). Reporter gene constructs that activate upon alternative splicing of a messenger RNA have been used to develop therapeutics for spinal muscular atrophy [8]. The Mammalian Membrane Two-Hybrid (MaMTH) system reports protein–protein interactions by releasing a synthetic transcriptional activator to the nucleus [9]. A small molecule library screen using MaMTH identified compounds that limited aberrant signaling from a mutant form of epidermal growth factor receptor (EGFR) by reporting its interactions with Shc1, a protein known to facilitate EGFR signaling. One of the identified candidates inhibited EGFR by altering the endosomal trafficking of mutant EGFR rather than directly blocking kinase activity, a mechanism that required a cell-based phenotypic assay to display efficacy [9]. Integrated reporter genes that detect molecular events at different layers of a pathway of interest provide single-cell resolution of molecular events. They may also expedite assay readout by replacing time-consuming biochemical assays for quantifying intracellular protein and RNA levels.

Synthetic circuits can condense signals across regulatory layers and GRNs into a single output. Multi-input synthetic circuits can ‘classify’ cells to report a molecular phenotype through the use of logic decoding promoters [10,11], DNA recombinases [12], split-transcriptional activators (Figure 2E) [13], and microRNAs (Figure 2F) [14]. For instance, various cancer cell lines have a unique molecular phenotype defined by their microRNA profile [14]. A synthetic classifier circuit can distinguish between cell types using the expression of five different microRNAs [14]. In melanoma, increased expression of two genes, EGFR and NGFR, mark cells primed for drug resistance [15]. Multi-input circuits could enable live-identification of these cells, allowing longitudinal tracking of the emergence of this subpopulation and response to drug treatments.

---

**Glossary**

- **Circuit retroactivity**: unintended or emergent interactions between a circuit and native systems or other synthetic modules that often impede designed functions.
- **Disease modeling**: constructing model cellular systems to understand the molecular mechanisms and/or phenotypes that contribute to disease processes.
- **Event recording**: stable modification of DNA in response to a transient signal.
- **GCaMP**: a genetically encoded biosensor that transmits calcium binding into fluorescence.
- **High-content imaging**: high-resolution imaging of cells to identify phenotypes and cellular responses.
- **Mechanism of action (MoA)**: the biochemical process by which a drug interacts with its target and produces its effect.
- **Microphysiological systems (MPS)**: an in vitro model that uses organoids or other 3D constructs to study cells in a framework that is more similar to in vivo than traditional in vitro methods.
- **Molecular phenotype**: the cellular state that characterizes a chemical or genetic perturbation, often defined from a comparative transcriptional profile.
- **Organ-on-a-chip (OoC)**: a type of MPS that uses microfluidic chips that recapitulate the physiological environment of a human organ.
- **Phenotypic drug discovery (PDD)**: process for identifying therapeutics that induce disease-relevant changes in the phenotype of cell-based assays, regardless of the molecular target of the therapeutic.
- **Post-transcriptional regulation**: processing of mRNA transcripts to control their stability and translation into proteins.
- **Promoter**: a region of DNA that is bound by RNA polymerase to initiate transcription of downstream genes.
- **Reporter gene**: a live-cell metric for determining activity of a process or pathway, oftentimes a light-detectable protein or metabolite.
- **Signal-to-noise ratio**: a measure of the activated signal from a reporter compared with the basal activity or background.
- **Single-molecule fluorescence in situ hybridization (smFISH)**: a technique where target RNAs of interest are hybridized with fluorescently labeled oligonucleotides to form subnuclear...
While most reporter assays directly connect the event of interest to transcriptional activation of the reporter gene, synthetic gene circuits can process signals from reporters to clarify the results of a screen [11,16]. For instance, synthetic circuits can increase the dynamic range of reporter assays, which may be limited by the promoter strength of a gene of interest. A recently developed signal amplifier expresses a synthetic transcriptional activator in response to activation of the unfolded protein response (UPR) [16]. The UPR is a stress-associated response triggered by the accumulation of unfolded proteins in the endoplasmic reticulum. UPR activation has been implicated across a wide range of disorders, including cancer and neurodegenerative diseases [17,18]. The synthetic activator amplifies reporter expression, enabling detection of small changes in GRN activity. A second module uses a coregulated repressor to buffer against basal signaling [20]. Similarly, negative regulators of pathway activity can buffer circuits against noise amplification and reduce basal activity [19]. By improving the signal-to-noise ratio, signal amplifier circuits support higher-fidelity candidate drug screens.

Circuits can track dynamic signaling patterns and stably report transient molecular events to increase the information conveyed through a screen. EGFR signaling transmits through the classical mitogen-activated protein kinase (MAPK) cascade via extracellular signal-regulated kinase (ERK). Altered ERK signaling dynamics can lead to divergent cell fates, including the induction of proliferation in cancer [20,21]. Signaling dynamics can be assayed using kinase translocation reporters that transport a reporter protein out of the nucleus in response to kinase activity [21,22]. However, screens for altered dynamics are limited in scale due to the large number of time points needed to track short time scale signaling events and longer time scale phenotypic outputs [21]. Circuits can be used to discriminate between sustained and pulsatile signaling [23,24]. One such circuit, called a dynamic decoder, discriminates a continuous signal from a transient signal pulse. By incorporating native post-transcriptional regulatory elements into a reporter gene, the dynamic decoder recapitulates native circuit feedback on the reporter gene to identify cells with sustained signaling (Figure 2H) [24].

Synthetic circuits with internal controls can maximize the potential of high-content screening

The variance and reproducibility of cell-based assays determine the scale at which drug candidates can be screened. Cellular heterogeneity can obscure insights from cell-based assays. For instance, the rates of global processes such as transcription, translation, and degradation vary with cell-cycle stage and proliferation rate, as well as by cell type [25–27]. These variables can fluctuate extensively through the cell cycle without indicating changes in cellular identity [28]. Thus, controlling for variations in cell cycle-associated processes represents an opportunity and a significant challenge for extracting meaningful phenotypic signals from cell-based assays [29]. Synthetic circuits provide internal controls for validating assays and controlling for cell-to-cell and assay-to-assay variance to refine the quantification of drug candidate efficacy.

Pathway reporters ostensibly provide activity-dependent expression such that low and high outputs can differentiate cell states. Extrinsic variation from the cell cycle and other sources may obscure differences between activity based on a single reporter (Figure 3A, center). Extrinsic variation that impacts systems-level features, such as transcription and translation rate, can be monitored and compensated for through dual reporter-based systems to improve assay resolution (Figure 3A, right) [30–32]. Improved resolution may aid identification of therapies in initial screening libraries that modulate pathway activity through a novel mechanism but require optimization to achieve a large, therapeutic effect. In addition, the control reporter serves as a positive control for reporter function on a cell-by-cell basis. Loss of the control signal may result from cytotoxicity, interference of the drug candidate with assay readout, or nonspecific mechanisms such as protein or RNA synthesis inhibition [1,4].
PDD programs aim to identify therapeutics acting through novel mechanisms; however, previously identified mechanisms often dominate the list of hits identified in a screen [4]. Synthetic circuits may also be used to narrow the list of candidate therapies to those with novel MoAs. Molecular phenotyping via RNA sequencing or other methods can identify compounds that act through ‘frequent hitter’ mechanisms and produce similar transcriptional profiles [4]. However, the low throughput and high costs associated with these techniques prohibits their large-scale application. Multi-input synthetic circuits may be used to condense a molecular phenotype of a frequent hitter mechanism into a single readout that can be included alongside the primary reporter for rapid hit validation and triage (Figure 3B).

Multicolor fluorescent assays permit multiplexed reporters within the same screen (Figure 3C). While one channel is often required to quantify cell numbers, additional fluorescent channels can be dedicated to disease monitoring, internal controls, and undesirable-mechanism reporters to expedite hit triage in high-content screens. Integration of multiple genes on the same construct enables multiple readouts from a screen to quantify molecule efficacy at a single-cell level with
Figure 2. Building synthetic gene circuits to assay diverse molecular events. (A) Reporters of specific splicing isoforms can be designed by inclusion of a reporter gene in frame with an exon of interest. (B) Inclusion of microRNA target sites in an untranslated region of a reporter gene leads to a loss of signal when the microRNA is active. (C) Protein–protein interactions lead to reconstitution of a cleavage signal, such as ubiquitin, releasing a synthetic transcription factor (TF) to the nucleus where it activates the output module [122]. (D) Translocation of an active kinase, such as ERK, to the nucleus causes the kinase translocation reporter (the output module), to translocate to the cytoplasm. (E) Multi-input logic can be realized through multipart assemblies that activate reporter expression. Splitting synthetic transcriptional activators, into activator (teal), DNA-binding (gray), and DNA-targeting (black) components can implement AND logic by placing the expression of each activator component under control of a different transcription factor [13]. (F) microRNAs can implement OR logic by placing different target sites into an untranslated region of a repressor to activate expression of the reporter gene when either of the microRNAs is present. (G) The dynamic range amplifier works via a synthetic transcription factor that activates reporter expression while also repressing negative regulators of the reporter. Range is expanded by inclusion of a transcriptional repressor (dark blue) that limits reporter expression in response to basal-level signaling while the post-translational repressor (orange) degrades the reporter, leading to enhanced resolution of signaling dynamics [16]. (H) Dynamic decoder circuits work by activating in response to a defined dynamic signaling pattern. For example, activation of the MAPK signaling pathway leads to a slow transcriptional activation, fast post-transcriptional repression, and delayed post-translation stabilization of a reporter gene. By differing time scales of regulation, the decoder can be designed to activate the reporter only in response to a defined, sustained input.
minimal added complexity. High-content imaging platforms enable further multiplexing using the subcellular localization of reporters. ORACLE fuses fluorescent reporters to a nuclear rim-localized protein, enabling nuclear segmentation and simultaneous imaging of cell cycle and transcription factor (TF) activity reporters using the same fluorescent channel [33]. Through proper selection of reporter protein fluorescent spectra and localization, circuits can add a condensed, high-dimensional feature for integration into morphology-based assays such as Cell Painting [34]. As methods for parsing cell-based assays based on morphology and other extracted features are rapidly expanding [35,36], we anticipate that circuit-based outputs will complement current assays by providing an information-rich feature for deep learning and other machine learning algorithms. By offering robust internal controls and disease-relevant outputs, circuits offer high-sensitivity metrics for identifying cellular phenotypes and drug-induced responses in cell-based assays.

Validating cells and simplifying readouts in high-complexity, multicellular assays

The predictive validity of cell-based assays relies on in vitro assays capturing processes and mechanisms that are relevant to the disease in vivo. Thus, assays are only as reliable and accurate
as the cells used within them [37]. While cell lines have been the default cells used for drug discovery, transformed cell lines may not accurately represent primary cells affected by disease [38]. Finally, accurately modeling some diseases requires multiple human cell types arranged in specific 3D architectures. In light of these challenges, human induced pluripotent stem cell (iPSC)-derived cells, organoids, and MPS, including organ-on-a-chip (OoC), are increasingly attractive models to enhance the physiological relevance during drug discovery (Figure 4A). However, while these technologies provide exciting avenues for improving the in vitro modeling of complex, multicellular diseases, several barriers remain to broad adoption of these technologies at a commercial scale in standard drug discovery workflows [37]. Challenges include the complexity of the MPS, low reproducibility of organoid systems, and an ever-expanding diversity of cell types that lack live-readout methods for quality control and validation. Synthetic circuits can improve the online validation of derived cells as well as simplify readouts from complex MPS systems.

While iPSC-derived cells offer exciting potential for modeling diseases, validating the in vitro differentiation of iPSCs into disease-affected cells remains a substantial challenge [39–42]. Synthetic gene circuits that report cell type and maturity will improve disease modeling and drug screening using iPSC-derived cells, including organoids. Current protocols for generating organoids suffer from low reproducibility both within and across human iPSC lines [43,44]. Low reproducibility limits well-controlled comparison across batches of organoids and between healthy and diseased donors [44]. Moreover, organoid generation protocols are time-intensive. Derivation of oligodendrocytes and microglia for models of neurodegenerative diseases requires multiple months, which severely limits the speed of protocol iteration [45–48]. Even with prolonged culture, directed differentiation of iPSCs in vitro generates immature cell types reflecting an embryonic state, which may confound modeling of adult-onset diseases [45,47,49–52]. Live-cell assessment of cellular identity and maturity may expedite the development of protocols for robust generation of mature cell types in vitro; however, current live-assessment of cell state remains limited. iPSC lines with integrated reporters provide a live readout of key genetic markers; however, reporters of transcriptional activation of a single genetic locus are limited by low fidelity in vitro [53]. Circuits can enhance live validation of cell state and maturity by reporting the state of several key GRNs in combination with other cell type-specific layers of regulation. As the number of cell types modeled in a single system expands, ensuring assay reproducibility will require robust metrics for evaluating cell identity and the distribution of cell types.

A recently developed framework provides a screening-based approach, allowing screening of cell-state reporters in the absence of known genetic drivers or pathway responses [54]. The Synthetic Promoters with Enhanced Cell-State Specificity (SPECS) library can be screened for activity in two or more distinguishable cell states. Promoter activity is assayed via next generation sequencing of cells sorted by reporter intensity, then a trained regression model uses this data to predict promoter activity in each cell state and identify cell state-specific promoters. Potentially, a SPECS screen of mature and immature cells could be used to identify multiple promoters that together indicate maturity, which can then be combined into a multi-input classifier circuit in iPSCs that reports sufficient differentiation and maturity of derived cells (Figure 4B). MicroRNAs can be used in conjunction with cell state-specific promoters to capture a larger biological space, as needed. MicroRNA classifiers successfully discriminate between cell types derived from iPSCs and have been used to purify cardiomyocytes by inducing apoptosis in cells that lack the desired microRNA profile (Figure 4C) [55].

Development of circuits that specifically report the progress of cell types and organoids generated from iPSCs may provide valuable biological insight into native development. Ultimately, these circuits could guide differentiation. Current protocols for directed differentiation and
organoid generation rely on external signals within the cellular matrix and media to induce differentiation. Consequently, variance in reagent components, including cellular matrix and differentiation-inducing growth factors, induce batch-to-batch variability within and between iPSC lines. Enhanced reproducibility may be achieved by developing genetic circuits that guide differentiation via TF overexpression [56–59]. Recently, TF overexpression has been employed to direct differentiation from iPSCs [60–62] and control multiple cell types in organoids [63]. TF-based control may improve reproducibility by directly activating essential GRNs and reinforcing lineage commitment and maturation [64]. However, TF overexpression leads to super-physiological levels of TFs within the nucleus, which may compromise cell function [65].
may optimally regulate cell fate at specified levels of expression and stoichiometries [66]. Enabling precise, temporal control of gene expression via gene circuits is a hallmark of synthetic biology [19,67–70]. As actuators of cell fate, TF cocktails provide a key component in constructing feedback controllers capable of autonomously guiding cells towards defined cell fates [60,64,71]. A circuit that temporally controls the expression of three TFs enhances the derivation of insulin-secreting beta cells from iPSCs [72]. Thus, in addition to their sensing abilities, genetic circuits provide the ideal system for guiding cells from iPSCs into functionally mature cells. In the future, these circuits may also be used to autonomously guide cells into becoming fully mature cell types for studying the multicellular consortia in a variety of systems from 2D co-culture to iPSC-based MPS.

Organoids and MPS recapitulate the important interactions and 3D architecture of multiple cell types to enhance the validity of in vitro systems [37]. As circuits can enhance the validation of cell state at the single-cell level, they may also enhance the validation of multicellular interactions. Secreted, diffusible signals can be detected by synthetic receptors that activate expression upon ligand binding [73,74]. Receptors may be engineered into the cells of interest, or they may be engineered into theranostic cells that serve as ‘cellular sentinels’ [123]. Rather than using biochemical assays to detect secreted proteins and metabolites from a cell of interest, theranostic cells can be engineered with synthetic receptors for the secreted molecule for in situ detection. While organoids provide improved physiological mimics, their 3D nature makes traditional high-content imaging challenging to perform, analyze, and interpret. Synthetic sender and receiver cells engineered to sense and respond to intercellular signaling molecules may report on mature cell-to-cell interactions. Receiver cells that report organoid and MPS function via secreted molecules would not require 3D culture methods and could be engineered from immortalized cell lines, simplifying the construction and monitoring of these systems. Alternatively, circuits that can record events such as cell state, signaling events, and clonal identity can improve the tracking and analysis of how cells are affected by therapeutics, even in organoids and other systems where live-monitoring may be difficult (Box 1). By simplifying readouts from complex cellular assays into quantifiable metrics, circuits will enable higher throughput validation of drug candidates in preclinical models.

A vision for circuit-based screening in neurodegenerative diseases

Circuits provide a facile mechanism for monitoring the intracellular dynamics that drive phenotypic changes. For neurological disorders, a constellation of factors influences the onset and course of disease. Accurately modeling these diseases and capturing the molecular events that precede death requires access to the distinct, human neural subpopulations affected. iPSCs enable the generation and testing of virtually any human cell type, including many inaccessible neural cells. Relevant disease phenotypes have already been captured in in vitro-derived neural cells, suggesting that these cells could be used as a tool to study neurological diseases [75–79]. By engineering iPSC-derived neurons, circuits can be included that capture systems-wide, disease-associated phenomena, such as elevated signaling, autophagy, and cellular age.

Accumulation of cellular aggregates represents a common hallmark of neurodegenerative diseases in vivo and in iPSC-derived neural models of amyotrophic lateral sclerosis and Huntington’s, Parkinson’s, and Alzheimer’s diseases. In vitro aggregates correlate with disruptions to ubiquitin-mediated proteolysis, UPR, and autophagy [80–82]. However, while aggregates provide markers of degeneration, their role in degeneration and impact on protein processing pathways remain less clear. By monitoring changes in protein processing through the proteolysis and UPR pathways, recently developed circuits can be used to identify drugs that abrogate UPR-associated neurodegeneration [16,81,83,84]. A circuit that reports on aggregate accumulation in
yeast provides one example for how such circuits could be used to identify molecules that limit or accelerate aggregation (Figure 5A) [85]. By monitoring processes preceding cellular death, such as aggregate accumulation and UPR activity, combined phenotype- and circuit-based screening will accelerate identification of on-target compounds.

Changes in proteostasis can lead to the accumulation of surface receptors and hyperexcitability, a common phenotype in neurodegenerative disease models [86–88]. In vitro hyperexcitability induces neuronal death. Currently, fluorescent reporters such as GCaMP, a genetically encoded Ca²⁺ sensor, are used to measure the frequency of glutamate-induced calcium signaling [89]. However, the time scale of neurodegeneration makes connecting GCaMP measurements to cell-fate outcomes within the same assay challenging. Alternatively, circuits that sense neuronal activation frequency and duration may improve long-term tracking and mapping of neuronal signaling events, drug-induced changes, and phenotypic outcomes [24]. Potentially, hyperexcitability could be monitored using engineered receptors in combination with synthetic circuits. For example, we envision that synthetic circuits utilizing engineered receptors for glutamate could be developed to activate a frequency-decoding circuit. By designing the circuit to regulate transcription, mRNA stability, and protein stability over different time scales, reporter expression can be selectively activated in response to higher frequency receptor activation (Figure 5B) [24]. By enabling the simultaneous monitoring of proteostasis and hyperexcitability over the time scale of neurodegeneration, synthetic circuits can be used to identify compounds that selectively affect these processes in the diseased-affected neuronal subtype. As the number of genetically encoded sensors of neurological processes continues to expand [90–92], we envision that circuits may bridge these short time scale, dynamic phenotypes with longer time scale assays. Continued development of circuits that decode dynamic signals will enable simple readouts and/or recording of cellular events across longer time scales to facilitate drug screens that span fast molecular events to longer-term phenotypic outcomes.

Aging represents an important determinant of disease onset in vivo that remains challenging to assess in living cells. Aged cells display defects in metabolism, nuclear structure, and nuclear

---

**Box 1. Event recording integrates cell history and cell state**

Capturing transient events occurring over longer time scales represents a challenge in studying biological systems. Event recording tools that utilize DNA editing enzymes can repurpose DNA into a log of circuit activity. Event recording circuits can preserve transient signals and cellular lineage at specified edit sites in the DNA, which can be read while preserving the spatial organization of cells (Figure A). Events can be later read out through sequencing and high-content imaging to infer cellular lineages and histories.

Recombinases such as Cre activate reporter expression by excising premature stop signals or flipping a gene into an active orientation [113,114]. Event recording may also be implemented using CRISPR-Cas9 [115,116]. Error-prone repair of Cas9-cleaved DNA frequently leads to random insertions or deletions (indels) at the site of damage. The Mammalian Synthetic Cellular Recorders Integrating Biological Events (mSCRIBE) memory system uses a self-targeting single gRNA to continuously generate indels within the gRNA sequence [116]. Continuous recording generates a number of indels that is proportional to the strength and duration of the input signal [116]. mSCRIBE has been used to record tumor necrosis factor-α-induced activation of the NF-κB pathway. The Memory by Engineered Mutagenesis with Optical In Situ Readout (MEMOIR) system employs single-molecule fluorescence in situ hybridization (smFISH) (Figure B, i) to detect indels rather than sequencing to read out DNA edits in situ from specified ‘scratchpad’ sites [115]. Event recording sites placed in an untranslated region of a synthetic transcript enable simultaneous lineage tracing and transcriptomics via single-cell RNA sequencing (Figure B, ii) [117,118].

As an alternative to Cas9, serine integrases can be directed to an array of target recombination sites to enable three DNA states (unedited, flanked, or collapsed) that can be read using in situ readouts [119]. CRISPR-directed base editors have emerged as another event recording tool as they efficiently induce nonrandom changes in the DNA [112,120,121]. To read single-base edits in situ, the Zombie system uses phage RNA polymerases and smFISH to transcribe and read edited DNA barcodes [112]. The DNA-based Ordered Memory and Iteration Network Operator (DOMINO) system uses base editors to induce successive changes in DNA, which translate into increasing numbers of TF binding sites [121]. As the number of binding sites increases, the transcriptional output increases, providing a dynamic, single-cell metric for monitoring the number of successive events (Figure B, iii).
transport, which can inform the development of reporters and circuits [75,93,94]. To assess how aging impacts different cell types, a reporter of nucleocytoplasmic compartmentalization was constructed utilizing fluorescent reporters tethered to either a nuclear localization signal or nuclear export signal. While deficits in nuclear import correlated with aging in neurons, ratios of the reporters displayed large variance, making comparisons between samples of different ages challenging.

Circuits that amplify the ratio of nuclear to cytoplasmic proteins may improve detection of deficits in nucleocytoplasmic compartmentalization (Figure 5C). A negative regulator of circuit activity could also be added to reduce noise amplification and limit basal activity [19].

Concluding remarks and future perspectives
Synthetic gene circuits hold enormous potential for engineering cellular models of disease. Synthetic circuits have the potential to improve the drug discovery process by assaying the specific activity of drug candidates within the requisite cellular environment. Advances in synthetic biology that enable transcriptional reporters of disease-associated molecular events coupled with increasingly complex synthetic circuits massively expand the range of measurable phenotypes in live, single cells. We envision that circuits can both improve the reproducibility of cell-based assays and capture dynamic phenotypes. Despite these advantages, few synthetic circuits have translated into industrial efforts in phenotype-based screening. Key challenges such as limited predictable de novo design, emergent behaviors in primary cells, and poor performance in iPSC-derived cells, have hindered broad adoption of synthetic circuits in cellular engineering and

Outstanding questions
How can synthetic circuits be rapidly integrated into cell-based disease models to improve phenotypic drug discovery and improve translational outcomes of discovered hits?

Can circuit design in mammalian cells be automated to monitor desired pathways and phenotypes via software to produce a robust circuit design?

Can synthetic circuits be designed to limit circuit retroactivity while maintaining a robust output signal?

How can transgenes be engineered into iPSCs to minimize transgene silencing in iPSC-derived cells? What are the molecular determinants of transgene silencing?

Performance of synthetic circuits can vary across cell types, and circuit engineering can require multiple iterations of design–build–test. Can design rules be elucidated that relate circuit performance in iPSCs to performance in iPSC-derived cells to reduce the number of iterations that require differentiation of iPSCs to the desired cell type?

Can synthetic circuits control the differentiation of iPSCs and accelerate the production of mature cell types?
phenotype-based screening. Fully leveraging synthetic circuits as standard tools for phenotypic screening will require improved methods to engineer transgenes in primary cells, including iPSCs and iPSC-derived cells such as organoids.

Figure 5. A vision for circuit-based screening in neurodegenerative diseases. (A) A DNA-binding repressor protein fused to an aggregation-prone protein of interest limits expression of a reporter circuit in the absence of aggregation. Aggregation of the protein of interest sequesters the repressor, inducing expression of the circuit. The signal from this circuit can be amplified using a positive feedback loop by coexpressing a synthetic transcriptional activator (dark blue) with the reporter. (B) Changes in proteostasis lead to accumulation of receptors, making the cell hyperexcitable. By tuning the rate of transcriptional, post-transcriptional, and post-translational regulation, a dynamic decoder circuit could activate when signal frequencies exceed a target level. (C) Aged cells display compromised nucleocytoplasmic compartmentalization (NCC), where nuclear-localized proteins leak into the cytoplasm and nuclear-exported proteins leak into the nucleus [75]. A DNA-binding protein (dark blue), such as a zinc-finger protein, can be fused to nuclear export signal (NES) and one half of a pair of dimerizing peptides. In aged cells with impaired NCC, this protein can leak into the nucleus, dimerize with a nuclear-localized activation domain with the complementary dimerization peptide, and activate transcription of a reporter gene (orange). A second DNA-binding protein (light grey) is fused to a nuclear localization signal (NLS). Dimerization with the nuclear-localized activation domain activates transcription of a second reporter gene (green). Loss of nuclear localization reduces expression of this green reporter. The aging-reporter circuit is designed with positive feedback to amplify detection. To limit basal signaling, the control circuit negatively regulates the aging-reporter circuit.
Developments in components and circuit design have enhanced the tunability and predictability of circuit behavior to enable generation of synthetic circuits that perform complex functions with high fidelity. However, despite this progress, it remains challenging to predictably engineer \textit{de novo} functions into circuits (see \textit{Outstanding questions}) [35]. Developing circuits often requires several iterations of the design–build–test cycle to obtain the desired behavior. Advances in circuit design automation will expedite this process to make circuit-based screens more accessible. As design rules for robust circuit composition are elucidated, algorithms to automate large-scale circuit design are being formalized into accessible software [96–99]. While design in bacterial systems has been the focus of software development, these tools are expanding to eukaryotes [100,101]. Elucidating the rules of mammalian synthetic circuit design will support computational approaches to the rapid design and prototyping of gene circuits. Given the technical challenges, we envision that direct collaboration between academic labs and industrial partners may most efficiently focus the development of circuits suited for specific cell-based assays as well as foundational tools for complex, multicell type assays.

Limiting the burden of genetic circuits represents an important challenge for preserving relevant phenotypes in disease models. Understanding of how circuits integrate into the native cellular context is important for preventing \textit{circuit retroactivity} within the model and within the circuit. Circuit retroactivity, where the system does not function as intended due to coupling to other systems within the cell or to the specific cell states, represents a notable challenge to robust circuit design [95,102,103]. Similarly, examples from introduction of synthetic effectors of the RNAi pathway indicate that synthetic components can sequester resources and systemically alter cellular behavior [104–106]. Consequently, validation of design represents an important hurdle in the translation of circuits to commercial-scale applications.

The usefulness of synthetic genetic programs critically relies on the ability to robustly maintain and control the expression of transgenes. However, simple overexpression of transgenes in primary cells remains nontrivial, with large fractions of transduced cells silencing transgenes [107–109]. Human-specific issues of transgene silencing compound the challenge of translating synthetic biology tools into meaningful therapies [108,110]. Transgene silencing is particularly problematic in primary cells and iPSCs [111]. Despite the existence of “safe-harbor” sites in the genome, transgenes expressed from these sites are highly prone to epigenetic silencing [108]. In fact, the rules for which cell types in combination with which loci promote or resist silencing remain undefined. Moreover, although regulatory elements such as specific promoters, insulators, and other structured elements can reduce silencing, their effects are not uniform across cell types. Improved understanding of how native systems operate and regulate transgenes is needed to reliably program transgene expression and engineer novel cellular behaviors.

Solving the challenges of \textit{de novo} circuit predictability will open up new frontiers of drug exploration that harness the powerful array of synthetic biology tools. With enhanced ability to capture biological information, circuits will facilitate the dissection of the causal mechanisms that link molecular events and phenotypes. By harvesting information from live, human, disease-affected cell types, synthetic gene circuits will improve the reproducibility and reliability of disease modeling, accelerate hit triage, and simplify validation and readouts from complex systems to expedite the drug discovery process.

Acknowledgments
We thank Christopher Johnstone, Sreha Kabaria, Brittany Lende, Kasey Love, Nathan Wang, and Joshua Wilkerson for useful discussions and feedback. A.M.B. is supported by the National Science Foundation Graduate Research Fellowship Program under grant No. 1745302.
Declaration of interests

No interests are declared.

References

22. Akin, T.J. et al. (2020) MAPK activity dynamics regulate non-cell autonomous effects of oncogene expression. eLife 9, e05041
42. de Ruus, D.J. et al. (2021) Current and future applications of induced pluripotent stem cell-based models to study pathological proteins in neurodegenerative disorders. Mol. Psychiatry Published online January 25, 2021. https://doi.org/10.1038/s41380-020-00999-7

224 Trends in Biotechnology, February 2022, Vol. 40, No. 2