Title
Investigating mechanisms of endogenous, systems-level coordination in stress adaptation and homeostasis

Author
Bonny, Alain

Publication Date
2020

Peer reviewed|Thesis/dissertation

Investigating Mechanisms of Endogenous, Systems-Level Coordination in Stress Adaptation and Homeostasis

by
Alain Bonny

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Hana El-Samad
52F8E320792C4C4...

Hana El-Samad
Chair

DocuSigned by:
Sophie Dumont
DocuSigned by:
DC...

Sophie Dumont

Joe DeRisi
50AE848EEDE2449...

Committee Members

Acknowledgments

To attempt to capture all for whom I am deeply indebted to is both humbling and inspiring. I have been so fortunate to be surrounded by the unconditional love and support of family, friends, mentors, mentees and colleagues over the years. I genuinely owe all that I've become to the many individuals that guided me along the way.

To my brother, Jean-Paul, I've always tried to be a role model that you could look up to. I'm beyond proud of all of your accomplishments, and glad to call you one of my best friends. To my parents, Gardy and Brigitte, who have continuously been my biggest source of inspiration, support, and guidance my entire life. You fanned the flames of my curiosity at a young age and never let anyone or anything put out that flame. I remember at a young age stealing little potions and gadgets (much to their annoyance) and pretending I was a scientist by concocting something contraptions. Rather than get mad, which I'm sure they were, they never discouraged me from pursing my interests. This unwavering support continued as I developed aspirations to be a professional athlete, musician, physician and finally an actual scientist. While my intellectual curiosity was sustained by them, I would be remiss if I didn't acknowledge the tremendous sacrifices they made to ensure that my brother and I received every opportunity from traveling the world, to attending the best schools possible. I will probably never know the full extent of the sacrifices you made for JP and I, but from the bottom of my heart, I love you, am grateful for you, and appreciate all that you've done.

A favorite quote of mine is "You become the average of your five closest friends." If the same group of friends incorporated me into that calculation, it would be clear that I bring

down the average. I'm so lucky to have friends that challenge my biases, open my eyes to new perspectives, and accompany me on the many journeys life has offered. My closest friends growing up, Sean, Ryan, Jerry and Joe, despite my move across the country, you've all been there for me in good times and bad. My 707 Prospect Fam, we all came into that house as friends and left as family. Those three years were the most memorable of my life, and I have each of you to thank for helping me grow during that time, especially in a new place with an incredibly stressful job. To that end, to my former students and colleagues at Da Vinci Design, I couldn't have asked for a better work environment. The staff was a collection of devoted, passionate and brilliant people that put the students above all, and the students bring me more faith than ever before for the future. Lastly, my ride or dies, Joey, Ashley, Ricardo, Jenna and Sandy, we've been all over the world from the hills of Haiti (and an unexpected detour to Panama), the lakes of Wisconsin, to the mountains of New Zealand. I love you guys, and I don't know where I would be without your love and friendship.

Reflecting back on all that I've learned in graduate school, it would not have been accomplished without the help of my brilliant friends and colleagues in the El-Samad Lab. I've been so fortunate to be able to learn and grow alongside you all over the past six years. I knew virtually nothing about systems and synthetic biology when I rotated, and through your patience, support and encouragement, I at least feel a little more competent than when I began. Specifically, Elham, Patrick, Ben, Susan, Taylor, Andrew, João, Lindsey - you all made lab an inclusive, fun and rigorous place to learn how to do science.

I decided to pursue a PhD for the purpose of learning how to think. This was not a trivial task, and it could not have been accomplished without the support, engagement and mentorship of my many scientific collaborators. To start, my graduate adviser, Hana, gave me the freedom to pursue my interests, but also intervened when necessary to focus and hone in on pursuing questions with rigor. Importantly, the thriving lab ecosystem is truly a testament to the time and effort Hana has invested so that we constantly push each other to do good science. In addition to Hana, I've been fortunate to have mentors and advocates

both inside and outside the lab as well. Specifically, my friend and frequent collaborator, João, whom I latched onto early on in graduate school and learned many of the tissue culture techniques that were integral to much of the work we would later publish together. Additionally, my thesis committee of Joe DeRisi and Sophie Dumont, thank you for your guidance to let me see the big picture, telling me what I needed to hear, and supporting me throughout my maturation as a scientist. My undergraduate mentor, Scott Showalter, was the first person who trusted me to do science in a lab. Importantly, he also never pressured me to pursue a PhD, but instead gave me the space to let me find my way, and was always a persistent supporter of me. Lastly, my mentor, De'Broski Herbert, was a scientific role model who showed me that a career as an academic is possible as a black man. He has always been a source of inspiration and confidence that helped me navigate some of the hills and valleys that come with academia.

Finally, I would like to dedicate this work to my single biggest advocate and supporter, my grandmother, Ma Titie. Thank you for everything. Love and miss you.

Contributions

Several chapters in this work have been adapted from manuscripts that are either published in a journal or as a preprint. These chapters do not necessarily match their final published form.

Chapter 2 of this thesis is a prereprint of the following manuscript:

Bonny AR, Kochanowski, K, Diether, M, El-Samad H. Stress-Induced Transient Cell Cycle Arrest Coordinates Metabolic Resource Allocation to Balance Adaptive Tradeoffs. BioRxiv.

Chapter 3 of this thesis is a reprint of the following accepted paper:

Fonseca, JP*, Bonny, AR*, Kumar, GR, Ng, AH, Town, J, Wu, QC, Aslankoohi, E, Chen, SY, Dods, G, Harrigan, P, Osimiri, LC, Kistler, AL, El-Samad, H. A Toolkit for Rapid Modular Construction of Biological Circuits in Mammalian Cells. ACS Syn Biol (Accepted).

Investigating mechanisms of endogenous, systems-level coordination in stress adaptation and homeostasis

Alain R Bonny

Abstract

Cells routinely deploy a spectrum of specialized counter-measures to deal with stress. However, a near constant feature of response to stress is a down regulation or arrest of the cell cycle. It has been widely assumed that this cell cycle modulation facilitates a faster cellular adaptation, a hypothesis that has been posited but still awaits strong support. Here, we used the *S. cerevisiae* hyperosmotic shock response as a model system to pursue the role of cell cycle arrest in stress adaptation. We found that decoupling the stress response program from the cell cycle led to an adaptive program that was accelerated compared to the wild type strain. Furthermore, we uncovered that stress-induced cell cycle

arrest restricted the utilization of metabolic resources, whereas unabated cell cycle progression diverted a surplus into the stress adaptation response. To probe why a wild type cell would ration its metabolic resources for an ostensibly suboptimal response, we identified an adaptive tradeoff that makes this rationing of resources advantageous under repeated stress. Given the universality of cell cycle regulation in stress responses, we suggest that this study might present a new paradigm in which the coordinated action of stress responses, cell cycle and metabolism help cells navigate the tradeoffs between rapid stress adaptation and long-term robustness.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

A paramount task that cells must address how to survive and thrive in a changing environment. Cells have evolved to maintain physiological processes within a range of homeostatic environmental conditions (ref); however, when conditions push a cell beyond the range of its homeostasis, the cell must adapt to stress. Previous work has elucidated the cognate pathways that are primarily responsible for mitigating the effect of the stress, demanding cellular resources. The sudden activation of these stress responsive pathways is at odds with other resource-intensive pathways that are continuously active, such as the cell cycle. To alleviate this conflict over resources, it's been observed that the activation of stress response pathways usually coincides with the arrest of the cell cycle, this allowing the cell to divert resources. It's now appreciated that cell cycle arrest during adaptation to stress adaptation is a fundamental aspect of stress response (ref), but the precise role of cell cycle arrest in the reallocation of resources remains elusive.

In chapter one of this work, I used the *S. cerevisiae* hyperosmotic shock response as a model system to pursue this question, where I found a surprising answer that seems to defy current hypotheses. In short, I found that decoupling the stress response program from the

cell cycle machinery led to an adaptive stress response that was substantially accelerated compared to the wild type strain. I pursued this observation mechanistically to uncover that stress-induced cell cycle arrest tapered the flux of metabolic resources, whereas aberrant cell cycle progression diverted a surplus into the stress adaptation program. I then went further, asking why a wild type cell would ration its metabolic resources, mediated through cell cycle control, even under acute stress. I uncovered a fitness tradeoff that makes this rationing of resources beneficial under repeated stress. Given the universality of cell cycle regulation in stress responses, we argue that this work provides a mechanism for how cells sacrifice faster adaptation for long-term robustness.

In addition to hypothesis-driven work, in chapter two I demonstrate the ability to build new tools to enable new lines of questioning. Here, I present a mammalian cloning toolkit library to enable the rapid assembly of gene circuits. This toolkit leverages the wide availability of tools, but introduces these disparate tools into a centralized, modular cloning system that we believe will become an invaluable resource as synthetic biology is used across multiple disciplines in engineering biology.

Chapter 2

Stress-Induced Transient Cell Cycle Arrest Coordinates Metabolic Resource Allocation to Balance Adaptive Tradeoffs

2.1 Abstract

The ability of a cell to mount a robust response to an environmental perturbation is paramount to its survival. While cells deploy a spectrum of specialized counter-measures to deal with stress, a near constant feature of these responses is a down regulation or arrest of the cell cycle. It has been widely assumed that this modulation of the cell cycle is instrumental in facilitating a faster response towards cellular adaptation. Here, we directly investigate the role of cell cycle arrest in the hyperosmotic shock response of the model organism *S. cerevisiae* by deleting the osmoshock-stabilized cell cycle inhibitor Sic1, thus

enabling concurrent stress response activation and cell cycle progression. Contrary to expectation, we found that removal of stress-induced cell cycle arrest accelerated the adaptive response to osmotic shock instead of delaying it. Using a combination of time time-lapse microscopy, genetic perturbations and quantitative mass spectrometry, we discovered that unabated cell cycle progression during stress enables the liquidation of internal glycogen stores, which are then shunted into the osmotic shock response to fuel a faster adaptation. Therefore, osmo-adaptation in wild type cells is delayed because cell cycle arrest diminishes the ability of the cell to tap its glycogen stores. However, acceleration of osmo-adaptation in mutant cells that do not arrest comes at the cost of acute sensitivity to a subsequent osmo-stress. This indicates that despite the ostensible advantage faster adaptation poses, there is a trade-off between the short-term benefit of faster adaptation and the vulnerability it poses to subsequent insults. We suggest that cell cycle arrest acts as a carbon flux valve to regulate the amount of material that is devoted to osmotic shock, balancing short term adaptation with long-term robustness.

2.2 Introduction

Cells and organisms are constantly challenged in their environment with insults that vary in origin, magnitude and duration. In order to respond to these insults, cells have evolved a large battery of adaptive stress responses that allow them to survive and maintain their homeostasis. Different stress responses show a remarkable diversity in their sensing, regulation, and logic[1–5]. However, an almost constant feature of any stress response is the involvement of cell cycle slow-down or arrest[6–10]. It is widely assumed that this is because it is advantageous for a cell not to divide during stressful conditions in order to safeguard its own fitness and that of its future progeny[11,12]. Furthermore, by arresting division, resources and energy can be diverted from the replication and division program to the

stress response program, allowing it to proceed faster or more efficiently[13]. Despite the near universality of these assumptions, the precise contribution of cell cycle arrest to adaptation remains poorly understood.

In the model organism *Saccharomyces cerevisiae*, two models for the interaction between cell cycle arrest and stress pathways have been proposed. In the first model, the cell cycle is deemed to be a passive recipient of upstream signals from the stress-sensing pathway, with several studies showing that the onset of a stress condition coincides with a downregulation of the cell cycle[11,14]. In the second model, the cell cycle is considered to be an important factor in the stress response, acting as the arbiter of the stress-handling capacity of the cell. To support this model, studies have demonstrated that the rate of cell cycle progression (i.e. growth rate) is a large contributor to survival under a variety of stresses, with slower growing cells being more fit in nutrient poor media and heat shock as well as other environmental insults[15,16]. Evidently, these models are not incompatible and strong data exist to support both. For example, growth rate is undeniably a contributing factor in mounting a robust stress response, and its effect has been demonstrated from bacteria to higher eukaryotes[17]. At the same time, yeast strains absent stress-sensing effectors display poor survival in the face of cognate insults irrespective of growth rate, further supporting the importance of frontline stress sentinels[2,18–20]. It is likely that a more applicable model is one in which a stress pathway can modulate cell cycle progression and changes in the cell cycle can feedback to actively modulate the stress response.

To explore this paradigm, we used the hyperosmotic glycerol (HOG) response as a convenient framework. The HOG program is a canonical stress response activated by the presence of excess osmolytes in the extracellular environment. The increase in osmotic pressure difference between the inside and outside of the cell drives water out, causing the cellular volume to decrease. At the onset of a step input of hyperosmotic shock, the central HOG mediator, Hog1, rapidly translocates from the cytoplasm to the nucleus where it

interacts with a variety of targets to initiate the production and accumulation of glycerol[21]. The accumulation of glycerol in the cytoplasm re-establishes the osmotic pressure gradient to its basal level, and once volume has been corrected, Hog1 exits the nucleus[22]. Importantly, in addition to the initiation of glycerol synthesis, Hog1 stabilizes the stoichiometric inhibitor of b-type cyclins 5 and 6, Sic1, to transiently arrest the cell in the G1 phase of the cell cycle[11]. Volume restoration and exit of Hog1 from the nucleus also coincides with resumption of cell cycle progression[23]. The adaptive translocation pattern of Hog1 has been the subject of many studies for its robust, reproducible and stereotyped pattern, which acts as a real-time reporter of hyperosmotic adaptation[24,25].

Using HOG as a model system, we investigated the role of transient cell cycle arrest in the adaptive response to hyperosmotic shock. Our approach was to decouple the HOG response program from the canonical cell cycle machinery such that both processes proceed simultaneously during osmotic shock. By following Hog1 translocation as a reporter of HOG adaptation, we were able to quantify deviations from Hog1's stereotyped translocation pattern as an indication of an altered stress response. Surprisingly, we found that unabated cell cycle progression during osmoshock accelerated osmo-adaptation as measured by Hog1 translocation. Remarkably, other canonical markers of adaptation such as glycerol production and volume recovery also proceeded faster. These data indicated that cell cycle arrest impedes, rather than facilitates, adaptation to stress. To pinpoint the mechanistic roots of this phenotype, we used mass spectrometry ^{13}C isotope tracing to probe the differences in metabolic flux between wild type and cell cycle arrest-disabled cells. We discovered that progression in the cell cycle during osmostress initiated catabolism of internal glycogen that was mediated by the enzyme Gph1. Breakdown of glycogen fueled faster glycerol synthesis in the mutant cells, giving them the ability to restore turgor pressure faster than wild type. Therefore, cell cycle seems to be the guardian of a metabolic valve that remains closed when cell cycle is arrested. To investigate what vulnerabilities arise from opening of this valve, and rationalize why the wild type cells still

implement cell cycle arrest despite the delay it imposes on stress adaptation, we subjected cells to repeated osmostress pulses. Under these conditions, wild type cells maintain substantial population viability while cells lacking arrest display worse survival, which keeps diminishing upon repeated stress pulses. Therefore, while adaptation to an osmotic stress proceeds faster when there is no cell cycle arrest, this phenotype leaves the cell particularly susceptible to subsequent osmoshocks. We believe that these data dispute the notion that cell cycle arrest or slow-down simply underlies rapid recovery from one instance of stress, instead suggesting that these relationships are multifactorial with complexities that we are yet to understand. Our findings also pinpoint an example where connection between three important cellular networks - a stress pathway, cell cycle regulation, and metabolic control - seem to collaborate in order to strike a balance between mounting a rapid adaptive response to acute threats and ensuring viability in the face of future insults.

2.3 Removal of Hog1-mediated cell cycle arrest accelerates adaptation to hyperosmotic shock

To assess the role of cell cycle arrest in adaptation to osmotic shock, we removed the ability of Hog1 to initiate G1-phase cell cycle arrest by generating a *sic1*Δ knockout strain (Figure 2.1A). In this strain, we tagged Hog1 with mVenus at its endogenous locus to allow visualization of its nuclear translocation by microscopy. We also incorporated the same Hog1-mVenus construct in a wild type (WT) strain to allow for the comparison of its Hog1 dynamics with those of the *sic1*Δ mutant. For precise temporal control in applying a step input of osmotic shock, we used a commercially available microfluidic platform that allowed us to quickly induce osmoshock and monitor Hog1 dynamics in the *sic1*Δ and WT strains by time lapse microscopy. We chose the osmolyte sorbitol as the input to induce hyper-osmotic stress because it is an inert sugar in the presence of glucose[26].

Using this experimental setup, we subjected WT and *sic1*Δ mutant cells to a step input of 1.2 M sorbitol osmoshock and monitored Hog1 nuclear translocation dynamics. Following the sorbitol input, Hog1 rapidly translocated to the nucleus in both strains with similar nuclear influx dynamics (Figure 2.1B). The degree of maximum nuclear enrichment was virtually indistinguishable in both strains, suggesting that the mutant maintains the ability to sense and respond to acute osmotic shock. In the WT strain, Hog1 exited the nucleus and its cytoplasmic levels adapted to the pre-stimulus values within 45 minutes on average, consistent with previous reports[25]. Surprisingly, however, the return of Hog1 to the cytoplasm was much faster in the *sic1*Δ cells, occurring on average within 33 minutes (Figure 2.1C). This constitutes a significant 30% speed-up compared to WT (Figure 2.1C, right). This acceleration of Hog1 adaptation in non-arresting cells was not a sorbitol-specific effect, since experiments carried out with NaCl also showed the same phenotype (Supplemental Figure 2.5A, 2.5B). Additionally, despite the protein Sic1 being typically described as G1-specific regulator, we observe this phenotype averaged over the entire asynchronous population, likely underestimating the effect of its deletion. Furthermore, prior work has shown strongly correlated transcriptional response between G1- and G2-arrested cells suggesting that the phenotype we observe is the direct result of HOG activated during concomitant cell cycle progression[13].

The faster Hog1 response in *sic1*Δ cells can be the result of a breakdown of coordination between the regulatory osmotic response and turgor pressure of the cell. If this were the case, then Hog1 would recover its cytoplasmic localization without full recovery in other physiological parameters such as cellular volume and glycerol accumulation necessary for this recovery. On the other hand, if the fast Hog1 adaptation were the result of an acceleration within an intact recovery program, then the profile of volume recovery and glycerol accumulation should also be accelerated. We therefore investigated these critical phenotypes to see if the integrity of the adaptive program is maintained despite accelerated Hog1 dynamics in the *sic1*Δ strain.

First, we measured internal glycerol content using a commercial colorimetric assay following a 1.2M sorbitol input over 60 minutes in both the WT and *sic1*Δ strains (Figure 2.1D). Using this assay, we determined that both WT and mutant cells upregulated their glycerol production by 5-fold at the end of the 60 minutes. However, while the WT cells did not begin to dramatically upregulate glycerol production until 30 minutes after the onset of stress, *sic1*Δ began increasing glycerol synthesis only 15 minutes after stress (Figure 2.1D, inset). In agreement with this finding, quantification of cellular surface area as a surrogate for volume also showed a faster recovery for the mutant relative to the WT (Figure 2.1E, top). While this volume recovery phenotype was reproducible, its extent was slightly lower than the Hog1 and glycerol phenotype, showing a 15% difference between WT and mutant (Figure 2.1E, right). This could be due to the difficulty in cell tracking and quantification of surface area, or to a strict upper limit on the expansion properties on the cell wall[27]. Taken together, the three phenotypes strongly support the hypothesis that the mutant has the same coordinated osmotic stress response as the WT type, but with faster dynamics that ensue from the inability of these cells to arrest their cell cycle during osmotic shock.

2.4 Glycerol production is accelerated using internal sources in mutant that lacks cell cycle arrest

Because Hog1 translocation, glycerol production and volume all correlate with faster recovery dynamics in the mutant strain, we next sought to investigate if another cellular process was fueling the accelerated production of glycerol, the catalyst for osmotic shock recovery. We hypothesized that a surplus of carbon material from central glycolysis could be shunted into glycerol production, thus resulting in heightened glycerol synthesis. One scenario for this to happen is one in which the deletion of Sic1 augments the ability of cells to import extracellular glucose, resulting in a greater amount of carbon material entering

glycolysis to be directed towards glycerol production. To test this hypothesis we performed a mass spectrometry ^{13}C isotope tracing experiment to compare the glucose uptake rate between the mutant and WT strains. In this experiment, cells were grown in ^{12}C glucose media. At time zero, an aliquot of cells was transferred onto filter paper over a vacuum manifold and continuously perfused with fully-labeled ^{13}C glucose media with and without 1.2 M sorbitol for various durations before quenching the sample (Figure 2.2A, Supplemental Figure 2.6A). A pilot experiment (not shown) suggested that turnover rates of glycolytic intermediates occur on the order of seconds, with intermediates reaching steady-state after 1 minute. Therefore, we chose to quench samples at 10, 20, 30, 45 and 60 seconds in order to capture the rate at which the internally ^{12}C-enriched glycolysis intermediates are degraded and newly synthesized metabolites incorporate the perfused ^{13}C. We measured ^{13}C incorporation in a number of metabolites representative of central glycolysis and the glycerol branch. Consistent with a fast turn-over of these metabolites, we observed a rapid decay of 12C enrichment among the glycolysis intermediates within seconds. In the examples of Glucose-6-Phosphate (G6P) and fructose bisphosphate (FBP), the *sic1*Δ strain incorporated ^{13}C with a slower rate than the WT in the presence and absence of osmotic shock (Figure 2.2B, Supplemental Figure 2.6B). Using the decay rate of ^{12}C enrichment as a surrogate for glucose uptake, we derived glucose import rates of 0.117 s-1 and 0.04 s-1 for WT and *sic1*Δ, respectively, during osmotic shock for G6P. Similarly, FBP incorporated ^{13}C at rates of 0.12 s-1 and 0.064 s-1 for WT and *sic1*Δ, respectively. The approximately 2-to-3-fold slower glucose uptake rate in the *sic1*Δ mutant was consistent for all glycolysis intermediates we targeted from upper glycolysis in glucose-6-phosphate (G6P) to phosphoenolpyruvate (PEP) in lower glycolysis (Supplemental Figure 2.6B, 2.6C). Therefore, it is clear that the faster recovery of the *sic1*Δ mutant cannot be ascribed to uptake of external glucose, since the mutant is visibly slower at glucose uptake. Slower glucose uptake often has the implication of a reduced doubling rate, which has been suggested to confer an advantage during stress

adaptation[16,28]. We measured the growth rate of the *sic1*Δ mutant and observed that it has a doubling time of 3.4 hours compared to 2.1 hours of the WT in defined media (Supplemental Figure 2.7A). To assess whether slower growth could be correlated to the accelerated osmoshock adaptation, we compared the growth rate of all strains used in this study (see subsequent sections for different strains, including those with different deletions in metabolic genes) against their respective adaptation time (Supplemental Figure 2.7B). We found a weak correlation (R =0.34, P =0.51) between the growth rate of the strains tested and their adaptation time to osmotic stress, suggesting that while growth may contribute to the accelerated phenotype, it is unlikely to be the only factor. The difference in glucose uptake rates between WT and *sic1*Δ suggest that the mutant is not importing more extracellular carbon material into glycolysis. To further investigate the source of its accelerated glycerol production, we reasoned that if the excess carbon material was not from extracellular sources, then it was conceivable that the glycerol synthesis in *sic1*Δ could be assisted by diverted intracellular carbon stores. To investigate this possibility, we again turned to ^{13}C mass spectrometry isotope tracing to test whether, and to what extent, intracellular carbon was used for glycerol production in both strains. In order to selectively enrich internal stores with a unique carbon isotope, we grew the cells on ^{12}C glucose, and 5 minutes before time zero, resuspended the cells in fully-labeled ^{13}C glucose. We then continued this treatment with and without osmoshock (1.2M sorbitol) for the remainder of the experiment, collecting samples until 30 minutes after osmoshock (Figure 2.2C). Based on the rapid turn-over of glycolytic intermediates (Supplemental Figure 2.6B, 2.6C), we hypothesized that within 5 minutes all glycolysis intermediates will be ^{13}C enriched, but macromolecules such as storage carbohydrates would remain enriched in ^{12}C due to their slower turn-over[29]. Given that the cell is only provided ^{13}C carbon at the onset of osmotic shock, and glycolysis intermediates likely enriched ^{13}C in the 5 minutes before stress, any detection of ^{12}C after HOG activation would implicate the liquidation of an internal macromolecule. By monitoring the panel of targeted metabolites in glycolysis and glycerol

precursors, we detected ^{12}C enrichment along central glycolysis, suggesting that both strains liquidated internal stores during osmoshock. The dynamic pattern of ^{12}C enrichment of metabolites for the WT cells showed pattern of low and unchanged level (FBP and PEP) or a level that declined as a function of time (G6P and DHAP) (Figure 2.2D, Supplemental Figure 2.8B). The temporal enrichment dynamics of the *sic1*Δ mutant, however, was markedly different showing a transient peak of ^{12}C enrichment followed by a decrease for all metabolites (Figure 2.2D, Supplemental Figure 2.8B). In the first step of glycolysis, the metabolite G6P started with a low ^{12}C enrichment, which increased and peaked between 1 and 5 minutes and subsequently decreased. The same pattern was present for glycolysis intermediates FBP and PEP. More importantly, this same transient ^{12}C signature was present in the glycerol production branch represented by dihydroxyacetone phosphate (DHAP). This transient ^{12}C enrichment pattern in the *sic1*Δ mutant suggests that ^{12}C in glycolysis intermediates is temporarily replaced by ^{12}C from carbohydrate stores in a flux that traverses the metabolic route to glycerol production. However, after this wave, external ^{12}C is incorporated again in metabolites, underlying the subsequent decline in ^{12}C enrichment. It is worth noting that in the absence of stress, the WT had a higher basal ^{12}C enrichment likely due to a higher basal turnover rate of macromolecules, consistent with its faster growth rate than the *sic1*Δ strain[29]. Only during the onset of stress did the *sic1*Δ mutant have a brief, higher ^{12}C enrichment with the aforementioned peak throughout (Figure 2.2C, Supplemental Figure 2.8B). Collectively, these data strongly suggest that, unlike the WT, the *sic1*Δ briefly shunts internal carbon stores as extra flux into glycerol production at the beginning of osmotic shock. Interestingly, the immediate precursor to glycerol, glycerol-3-phosphate (Gl3P), had an order of magnitude greater proportion of ^{12}C compared to the other metabolites tested (Figure 2.2D). The increase in ^{12}C could be attributed to back-flux from existing ^{12}C-enriched glycerol by way of a futile cycle to degrade excess ATP during severe stress[30]. Despite the discrepancy between Gl3P and DHAP, the *sic1*Δ strain was still enriched with

a greater amount of ^{12}C than the WT, likely reflecting the convergence of the aforementioned futile cycle and excess internal carbon liquidation.

2.5 Internal glycogen is liquidated using the Gph1 enzyme and shunted into glycerol production in mutant that lacks cell cycle arrest

Previous studies have established links between cell cycle progression and glycolysis as processes that are integral to cellular physiology[31?,32]. These links are mediated mechanistically through biochemical interactions between cell cycle regulators such as CDK1 and storage catabolism enzymes for trehalose and glycogen (Nth1 and Gph1, respectively)[33,34]. Given these data, we hypothesized that during osmotic shock, the *sic1*Δ mutant could activate storage catabolism enzymes through unabated cell cycle progression, resulting in a burst of glycolytic flux that was then shunted into excess glycerol production (Figure 2.3A). Further, we predicted that by coupling either a Nth1 or Gph1 knockout with *sic1*Δ, we could rescue the mutant to the WT phenotype, for example as measured by Hog1 localization dynamics. Cellular trehalose levels have been widely established as mediators of stress recovery[35]. However, surprisingly, in a *sic1*Δ*nth1*Δ mutant, Hog1 nuclear levels still adapted significantly faster than in WT following osmotic stress (Supplemental Figure 2.3A, 2.9B), suggesting that trehalose is not the main liquidated internal carbon source. However, when we coupled the *sic1*Δ deletion with a knockout of Gph1 (*sic1*Δ*gph1*Δ mutant), Hog1 adapted nearly 30% slower, closely resembling WT recovery time (Figure 2.3B, 2.3C). When we combined these genetic perturbations in a *sic1*Δ*nth1*Δ*gph1*Δ strain, Hog1 again adapted in time comparable to WT, supporting the notion that the Gph1-mediated breakdown of glycogen is the main driver of the accelerated

*sic1*Δ phenotype (Supplemental Figure 2.3C). To ensure that the *gph1*Δ rescue is specific to mitigate the effect of *sic1*Δ, and not a broad glycolytic flux perturbation irrespective of genetic background, we predicted that Gph1-mediated glycogen breakdown would not affect an otherwise WT Hog1 response, because glycogen catabolism is halted due to normal cell cycle arrest. Indeed, in support that *gph1*Δ is a specific rescue to *sic1*Δ phenotype, Hog1 adaptation time following osmostress in *gph1*Δ cells is nearly indistinguishable from the WT response (Supplemental Fig 2.9B, 2.9C). Consistent with the rescue in Hog1 dynamics, we observed that the deletion of Gph1 counter-acted the *sic1*Δ effect and reduced the glycerol accumulation in a manner commensurate to the WT rate (Figure 2.3D).

2.6 Accelerated recovery due to glycogen storage liquidation during osmotic shock prioritizes faster adaptation over robustness to repeated insults

Bypassing cell cycle arrest during osmotic shock results in accelerated recovery due to cell cycle-mediated carbon flux shunted into glycerol production. This suggests that the *sic1*Δ mutant might have an advantage upon one instance of osmotic stress. Conceivably, multiple instances of osmotic shock can amplify the advantage that *sic1*Δ cells have over WT cells. Alternatively, the faster recovery advantage of *sic1*Δ cells might come at the cost of other vulnerabilities that are only revealed dynamically[36]. To test the endurance of *sic1*Δ mutant upon a series of osmotic shocks, we subjected cells to the same 90 minute step input of 1.2 M sorbitol as before, but followed by three 45 minute shocks separated by 5 minutes (Figure 2.4A). At the end of each step input, we calculated the relative adaptation time using Hog1 nuclear residence as a metric and also assessed the viability of cells. After the first osmotic shock, both strains adapt as previously shown, with the *sic1*Δ

mutant recovering faster than its WT counterpart (Figure 2.4B). Following the first step input, adaptation to subsequent osmostress inputs proceeded faster (around 50% faster in this case) in both WT and mutant strains due to accumulation of glycerol from the previous cycle and consistent with previous reports[36] (Figure 2.4C). However, in the subsequent pulses, many *sic1*Δ mutant cells started exhibiting morphological differences from their WT counterparts. These cells displayed a deflated phenotype with visible material accumulated in the nearby vicinity, suggesting the cell lysed and released intracellular debris. This morphology is consistent with a breakdown of cell wall integrity, which is a common mechanism of death in serial osmotic shock perturbations[27]. Algorithmic quantification of this phenotype revealed a marked increase in cells that have a breakdown in their cell wall integrity for each subsequent pulse. By the end of the fourth step input of osmotic shock, 25% of *sic1*Δ cells displayed a cell wall-compromised morphology, while only 5% of WT cells displayed a similar phenotype (Figure 2.4D, 2.4E). Interestingly, the *sic1*Δ*gph1*Δ genetic background shares the same acute vulnerability as *sic1*Δ to repeated osmotic pulses despite its adaptation that resembles that of the WT. However, the triple mutant *sic1*Δ*nth1*Δ*gph1*Δ is able to survive the sequential osmoshocks in a manner that is more similar to the WT (Supplemental Figure 2.10D). This difference in vulnerability suggests unique roles for the breakdown of trehalose and glycogen that warrant further investigation. Nonetheless, it is clear that while the accelerated adaptation ostensibly provides an advantage in the face of a single step input of osmotic shock, the *sic1*Δ mutant is severely ill-suited for repeated insults.

2.7 Discussion

Cell cycle dynamics have a fundamental presence in many adaptations to stress, but a mechanistic understanding of its role has long been outstanding. Here we attempted to

understand some aspects of the role of cell cycle arrest in the context of the well-studied HOG pathway and associated osmotic stress. Our main experiment was to decouple the HOG stress program from the cell cycle machinery and monitor the nuclear localization dynamics of HOG master effector (Hog1). This experiment revealed that the normally stereotyped behavior of Hog1 accelerated in cells where cell cycle cannot arrest upon osmoshock. From there, we confirmed that the HOG program was still competent by looking to other canonical hallmarks of osmoshock recovery - internal glycerol content and cellular volume. Following the hypothesis that an upstream process was dictating the faster dynamics, we utilized quantitative mass spectrometry to implicate faster internal glycogen liquidation by the mutant as a route by which glycerol synthesis increases and mediates faster adaptation to the stress. In strong agreement with this insight, deletion of glycogen catabolism enzyme Gph1 rescues the *sic1*Δ Hog1 translocation and glycerol accumulation phenotypes. Puzzled by the observation that the WT is not optimal with respect to the speed of its recovery, we hypothesized that the mutant might have vulnerabilities that the WT can circumvent. Following this reasoning, we identified a critical failure mode of the mutant by subjecting it to multiple step inputs of osmotic shock. The mutant adapts faster and maintains viability after the first osmotic shock, but trades its faster initial adaptation for susceptibility to subsequent osmoshocks, as nearly a quarter of the population is dead by the last insult. Meanwhile, the WT adapts slower after the initial osmotic shock, but maintains nearly 95% viability throughout the experiment, thus highlighting the dichotomy between apparent short-term gain versus long-term resilience against a dynamic environment. We believe that the main contribution of this work is two-fold. First, our investigations provide a higher resolution dissection of the interconnection between three crucial cellular pathways: the cell cycle, the HOG stress response, and carbon metabolism. Contrary to expectation, this connection is not perfectly tuned to maximize the speed of adaptation to stress. In fact, the connection of the HOG pathway to the cell cycle diminishes the ability of the cell to recover rapidly following osmostress, and ablation of

this connection allows the cell to recover a substantial 30% faster. The cell cycle seems to be the gatekeeper of a metabolic valve that can augment carbon flow into glycerol from internal resources, but this valve remains shut in WT type cells. Opening of this valve in mutant (*sic1*Δ) cells seems to mediate their faster recovery, as evidenced by a *gph1*Δ mutant in which deletion of the enzyme presiding over the internal flux from glycogen to glycerol abolishes the fast osmostress recovery. Therefore, our data provide additional mechanistic details to an intricate interplay of pathways that together set the cellular recovery tempo. The second contribution of this work is to formulate an instance in which cells seem to navigate a delicate functional balance, sacrificing the brief advantage of faster recovery from an insult for robustness to future environmental changes (Figure 2.4F). The hypothesis that they evolved to do so prompts us to revise our view of how to interpret the measured dynamics of stress responses and our assumptions about how cells mobilize their resources to combat stress. It is clear that at least in the example of osmostress, *S. cerevisiae* cells do not maximally mobilize their carbon, and hence sacrifice substantial speed in their recovery. It is also evident that the cell cycle serves as an arbiter and enforcer of this suboptimal performance. Since the mutant that circumvents the speed limitation shows tremendous vulnerability to repeated stress, one is compelled to at least hypothesize that this cell cycle control has evolved to alleviate such vulnerability in an environment where repeated or oscillating stress might be more probable. What could be the nature of the vulnerability that the cell cycle is trying to arbitrate? Prior work demonstrated complementary roles between the HOG and cell wall integrity (CWI) pathways, where HOG is activated during the initial shrinkage of the cell, and the CWI is activated during the subsequent volume expansion and growth[37]. It is plausible to suggest that the redirected carbon flux of the sic1 mutant diverts resources away from cell wall remodeling, potentially explaining why *sic1*Δ is selectively vulnerable to a series of osmotic shocks. Evidently, more studies would be necessary to substantiate this hypothesis, but it is tempting to speculate further on the complex allocation of resources between these core

cellular processes. The work presented here reframes cell cycle arrest in a mechanistic light as being a mediator of a slower adaptation response to hyperosmotic shock. In future investigations, it would be interesting to use a similar approach with conditional or inducible mutants to test for the role of cell cycle arrest in other stresses, potentially discovering similar metabolic flux control mechanisms or roles more tailored to specific stresses. Alternatively, expanding the scope of this question to higher eukaryotes could further illuminate the complex relationship between cell cycle, metabolism and stress response, which has been implicated in several pathologies[38]. More broadly, it is clear that as we begin to explore how multiple pathways collaborate to allow a cell to navigate its complex environment, we need to revisit statements about functional allocations and re-explore plausible but exceedingly simple assignment of roles and assumptions of unifunctional optimality of any one pathway. We hope that the data presented in this work form a basis for such investigation, initiated by our more nuanced exploration of the ubiquitous role that cell cycle arrest plays in stress adaptation.

2.8 Materials and Methods

Strain Construction

The base *S. cerevisiae* strain used in this study is w303. Hog1-mVenus at the endogenous locus was generated by ordering oligos of 40 bp homology 5' upstream of the stop codon, PCR amplifying the mVenus-HIS3 cassette, and transforming as previously described40. To knockout genes, 80 bp of homology 5' to the start codon and 3' of the stop codon was used to PCR amplify a selection marker cassette and transformed as described above. PCR products using oligos in the 5'UTR and internal to the selection cassette were used to verify knockouts and insertions.

Growth Conditions

Single colonies were picked and inoculated from auxotrophic SD (6.7 g/L Bacto-yeast nitrogen base without amino acids, 2 g/L complete amino acid mix, and 20 g/L dextrose) agar plates. SDC liquid media used throughout the study consisted of 6.7 g/L Bacto-yeast nitrogen base, 2 g/L complete amino acid mix, and 20 g/L dextrose. **Thermocycler protocols** The GG protocol is primarily used for assembly reactions. The reaction temperature is initially held at 45°C for 2 min to digest the plasmids followed by 20°C for 4 min to anneal constituent parts together. After repeating these first two steps 24 times, the temperature is increased to 60°C for 10 min to digest remaining recognition sites and inactivate the ligase. Then the temperature is held at 80°C for 10 min to inactivate the enzyme. Lastly, the reaction is held at 12°C indefinitely. The "GG End-On" protocol is used when BsaI or BsmBI sites need to be retained in the final product. The temperature is initially held at 45°C for 2 min to digest the plasmid followed by 20°C for 5 min to anneal and ligate the resulting plasmid. These steps are cycled 24 times and then held at 16°C indefinitely.

Microscopy

Time-lapse microscopy was collected on a Nikon Ti inverted scope 40x air objective, with Sutter XL lamp illumination and a Hamamatsu Flash 4.0 camera. YFP (515 ex/528 em) channel was collected using a Chroma CFP/YFP filter set with an exposure time of 300 ms. Automated image acquisition was controlled by Nikon NIS Elements proprietary software. The CellAsics Onix2 Microfludic platform was used to control the changing of media. Pressure of the media perfusion was held constant at 10.8 psi in a microfluidic plate designed to trap haploid yeast (Millipore). To ensure that the yeast cells adapted to conditions within the microfluidic chamber, cells were perfused with normal SDC media for

90 minutes prior to the osmotic shock in all experiments.

Image Processing

The nucleus of each cell was defined as the mean pixel intensity of the brightest 5% of pixels over the segmented cell in the YFP channel. The remainder of the cell outside the brightest 5% was defined as the cytoplasm. Cell tracking and quantification of nuclear/cytoplasmic enrichment was done using automated yeast cell tracking software implemented in Matlab41. Nuclear enrichment is plotted as the population average nuclear to cytoplasmic ratio divided by the average three time points before the onset of osmotic shock minus one, as previously described23. Cellular volume was calculated from segmentation of an out of focus brightfield image using the Nucleaizer web interface (www.nucleaizer.com). The surface area of each cell in pixels was then converted into an approximate volume as described previously[23]. This volume was normalized by the three time points before the onset of osmotic shock.

Mass Spectrometry

Samples were grown overnight to saturation and diluted in SDC media. Cultures were grown to mid-log and 1 mL was transferred to a 0.45 µm PVDF membrane filter paper (Millipore), fixed atop a vacuum manifold. Cells were continuously perfused with either normal SDC media or 1.2 M sorbitol in SDC media for the indicated durations (Figure 2A). At the end of the perfusion period, the filter paper containing cells was immediately transferred to a 2 mL quenching solution of 40:40:20 Methanol:Acetonitrile:H2O chilled to -20 °C. After 2 hours incubation at -20 °C, quenching solution plus cells was transferred to a conical tube and dried for approximately 7 hours in vacuo and resuspended in 45 µL H2O. In the experiment described in Figure 2C, five minutes prior to time zero 50 mL

culture were transferred to a conical tube, spun down at 2000 RPM for 2 minutes and resuspended in 25 mL fully labeled 13C glucose SDC. After time zero, at the indicated time points, 1 mL of culture was transferred to the same filter paper vacuum manifold described above, and after media washed through the sample was quenched as described above. Collected compounds were analyzed using an LC-MS/MS mass spectrometer system consisting of a 1290 Infinity LC (Agilent Technologies) coupled to a 5500 QTRAP triple quadrupole mass spectrometer (AB Sciex) in negative mode and with multiple reaction monitoring (MRM) scan type. Five µL of metabolite extracts were injected on an Agilent PoroShell 120 HILIC-Z column (150 x 2.1 mm, 2.7 µm; Agilent, Santa Clara, CA) using a mobile phase A (water, 10 mM ammonium acetate, 5 µM medronic acid, pH 9) and mobile phase B (90% acetonitrile, 10% water, 10 mM ammonium acetate, 5 µM medronic acid, pH 9) at a constant flow rate of 250 µl/min; Initial conditions: 10% A, 2 min: 10% A, 12 min: 40% A, 15 min: 40% A, 16 min: 10% A, 24 min: 10% A. The MRM settings were adapted from Yuan et al[39]. The raw data were processed and analyzed using custom software in Matlab (Mathworks).

Intracellular glycerol assay

Each strain was inoculated into SDC media, grown overnight at 30 °C, split and diluted into six 600 uL 0.1 OD600 cultures in a 96 well 2 mL plate. Once the cells were in log phase growth, 600 uL of 2.4 M sorbitol was added to one well at time points of 60 minutes, 30 minutes, 15 minutes, 10 minutes, 5 minutes and 0 minutes. After time zero, 200 µL of each culture was transferred to a separate Corning 3904 96-well assay plate plate for an OD600 reading, and the remaining 1 mL immediately spun down for 2 minutes at 2000 RPM. Cells were washed in 400 µL H2O, and spun down again for 2 minutes at 2000 RPM. The culture was then resuspended in 150 µL H2O, and left to incubate for 15 minutes at 95 °C. Following incubation at 95 °C, cells were vortexed for 2 minutes and promptly spun

down for 10 minutes at 4000 RPM. After the pelleting of cell debris, 100 µL of supernatant was carefully removed and transferred to a separate plate and kept at 4 °C. Colorimetric glycerol assays were acquired using a commercial kit (Sigma) where the provided assay powder was resuspended in 40 mL of distilled H2O. For each sample, 5 µL of supernatant was added to 400 µL of glycerol free reagent solution, and left to incubate at room temperature for 15 minutes hidden from light. After 15 minutes, 200 µL of the glycerol free reagent solution-sample mixture was transferred to a separate plate and the OD540 was acquired for each sample on a Tecan Spark 10M plate reader. To account for differences in cell density across samples, the 540 nm readings were normalized by their OD 600 nm reading values.

Growth Assay

Each strain was inoculated into SDC media overnight, reached saturation and diluted the following morning to an OD600 of 0.1. After dilution, 200 µL of culture was transferred to a Corning 3904 96-well assay plate and grown at 30 °C while shaking. Optical density readings were collected at 600 nm every 20 minutes until saturation on a Tecan Spark 10M plate reader.

Cell Viability

At the end of each 45 minute step input of osmotic shock, each cell was manually assessed for cell cycle re-entry and return to basal morphology. Cells that either did not show continued cell cycle progression, or displayed a visible change in refractive index reflective of a change in morphology were labeled with altered morphology.

2.9 Acknowledgements

The authors thank the members of the El-Samad lab, Joe DeRisi (UCSF) and Sophie Dumont (UCSF) for helpful feedback and discussions. H.E-.S is an investigator in the Chan Zuckerberg Biohub. This work was also supported by the National Defense Science Engineering Graduate (NDSEG) Fellowship awarded to A.R.B.

2.10 Contributions

A.R.B. and H.E-.S. conceived of the study. A.R.B., K.K, and M.D. collected and processed data. A.R.B., K.K., and H.E-.S. interpreted results. A.R.B. and H.E-.S. wrote and edited the manuscript with input from all authors.

2.11 Figures

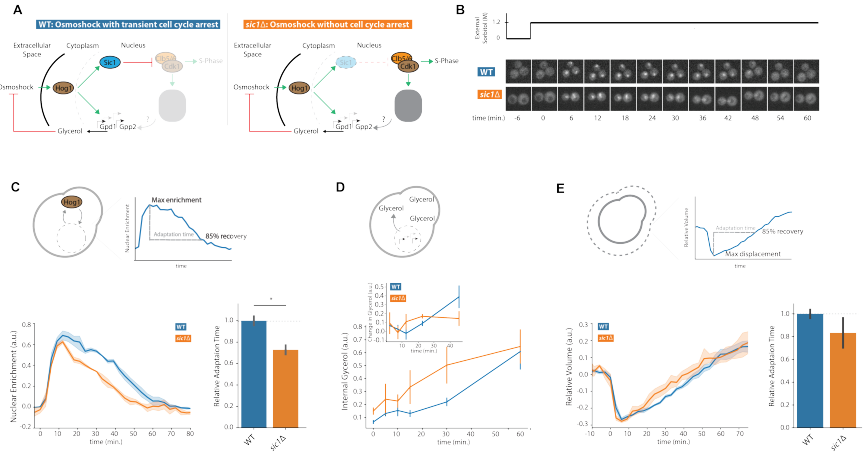


Figure 2.1: Removing cell cycle arrest by deletion of Sic1 accelerates the HOG adaptation program during osmotic shock.

a. A simplified schematic of the HOG pathway showing its coupling to cell cycle arrest and glycerol production. In a cell cycle mutant strain *sic1*Δ, we asked whether the removal of stress-induced cell cycle arrest affects the adaptive response. **b**. Representative time lapse images of endogenously-tagged Hog1-mVenus following step input of 1.2 M sorbitol osmotic shock to WT (top) and *sic1*Δ (bottom) cells. Quantification of all cells presented in Panel C. **c**.Top: Cartoon depicting the translocation dynamics and quantification of adaptation time. Bottom: Quantification of the WT (blue) and *sic1*Δ (orange) Hog1-mVenus nuclear enrichment in the experiment described in Panel B. Shaded regions represent the standard error of the mean (SEM) of n=3 biological replicates. Right: Quantification of Hog1 adaptation for WT (blue) and *sic1*Δ (orange). Values are normalized to the average of WT. Error bars represent the SEM of n=3 biological replicates. *P-value¡0.05; two-sided Student's t-test. **d**. Top: Cartoon schematic depicting the intracellular accumulation of glycerol. Bottom: Quantification of internal glycerol as a function of time for WT (blue) and *sic1*Δ (orange) to a step input of 1.2 M sorbitol osmotic shock. Measurements are taken using a colorimetric assay. Error bars represent the standard deviation for n=3 biological replicates. Inset: the change in glycerol, calculated as the difference between two time points for data in Panel D, is plotted as a function of time.**e**. Top: Cartoon schematic depicting volume recovery and quantification of its adaptation time. Bottom: Change in volume of the WT (blue) and *sic1*Δ (orange) strains in response to a step input of 1.2 M sorbitol osmotic shock. Shaded regions represent the SEM of n=4 biological replicates. Right: Quantification of volume adaptation time of WT (blue) and *sic1*Δ (orange) volume. Values are normalized to the average of WT. Error bars represent the SEM of n=4 biological replicates.

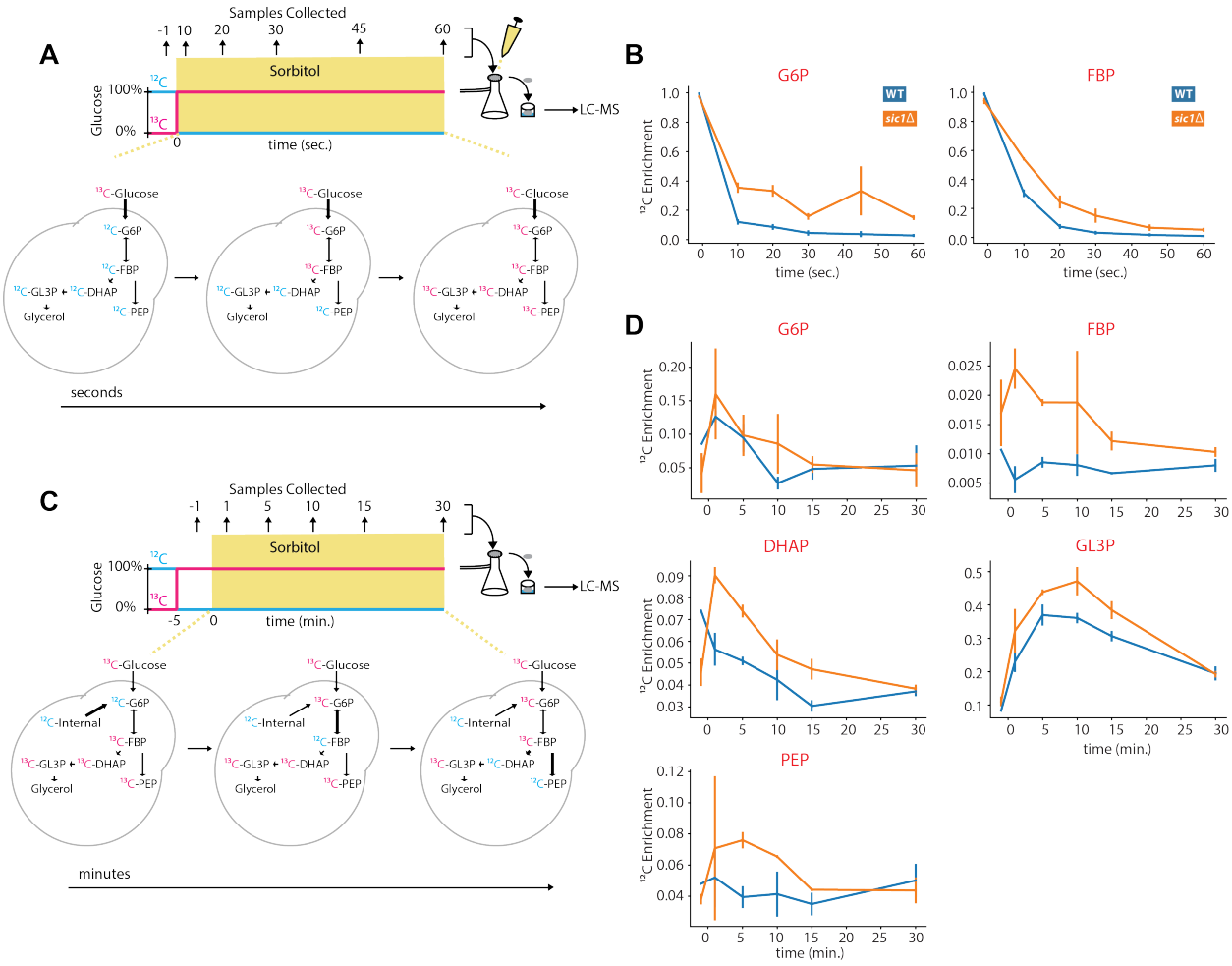


Figure 2.2: An internal carbon store is shunted towards excess glycerol production during osmotic shock in the *sic1*Δ mutant.

a. Cartoon schematic of the experiment to measure extracellular glucose uptake rates. Cells were inoculated overnight, diluted and outgrown in ^{12}C glucose. At time zero a 1 mL sample of cells was transferred to filter paper above a vacuum manifold and continuously perfused with fully-labeled ^{13}C glucose media. A 1.2 M sorbitol input was also administered at time 0. Samples were taken at 10 s, 20 s, 30 s, 45 s, and 60 s and transferred to quenching solution. **b** ^{12}C enrichment over time of central glycolysis metabolite Glucose-6-Phosphate (G6P) (Left) and fructose bisphosphate (FBP) (Right). Traces shown are WT (blue) and *sic1*Δ (orange) strains for experiment described in Panel A. Error bars represent the standard deviation of n=2 technical replicates. **c** Cartoon schematic of experiment to test internal carbon enrichment of targeted metabolites. Cells were inoculated overnight, diluted and outgrown in ^{12}C glucose. Five minutes prior to time zero, cells were resuspended in fully-labeled ^{13}C glucose. At time zero the culture was diluted 1:1 with 2.4 M sorbitol in fully-labeled in ^{13}C glucose. At the indicated time points, 1 mL of culture was placed on filter paper above a vacuum manifold for the media to wash through, transferred to quenching solution and then measured. **d** ^{12}C enrichment over time for a panel of select metabolites in glycolysis and glycerol production. Traces shown are the

WT (blue) and *sic1*Δ (orange) strains for experiment described in Panel C. Error bars represent the standard deviation of n=2 technical replicates.

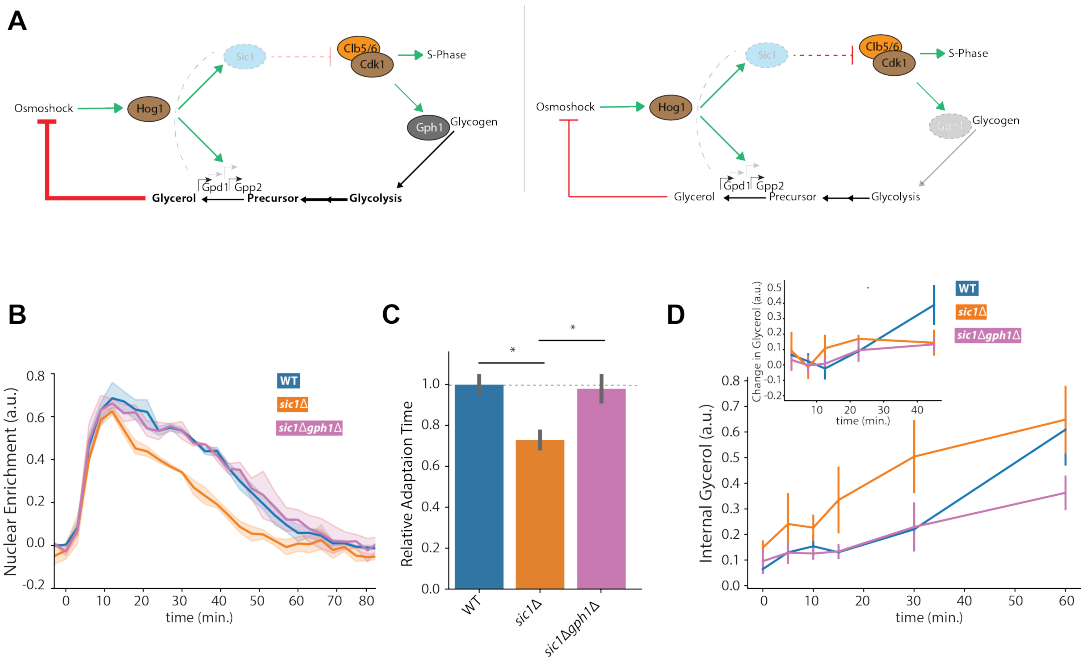


Figure 2.3: Glycogen catabolism enzyme Gph1 mediates glycerol synthesis to fuel acceleration phenotype in *sic1*Δ mutant.

a. Schematic depicting the hypothesis that liquidation of glycogen by activation of Gph1 accelerates the Hog1-mediated glycerol production. **b** Left: Traces of Hog1 nuclear enrichment over time following 1.2 M sorbitol osmotic shock in the WT (blue), *sic1*Δ (orange), *sic1*Δ*gph1*Δ (purple) cells. Shaded regions represent the SEM of n=3 biological replicates. **c** Quantification of adaptation time of Hog1 nuclear enrichment computed as in Figure 1B. Values are normalized to the average WT. Error bars represent the SEM of n=3 biological replicates. *P¡0.05; two-sided Student's t-test. **d** Measurement of internal glycerol over time for the strains shown in Panel B in response to a step input of 1.2 M sorbitol osmotic shock. Measurements are taken using a colorimetric assay. Inset: the change in glycerol, calculated as the difference between two time points for data in Panel D, is plotted as a function of time. Error bars represent the standard deviation for n=3 biological replicates.

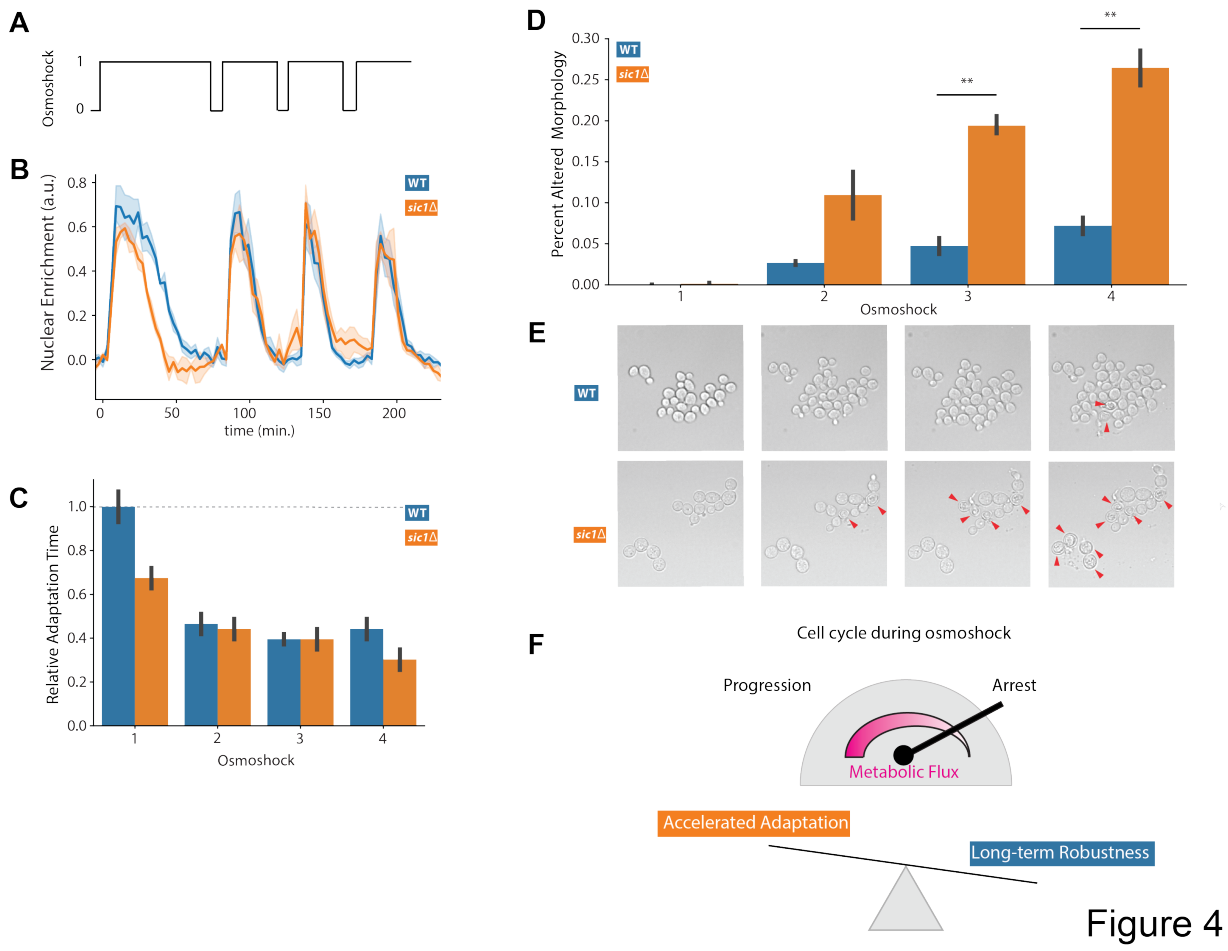


Figure 2.4: Cell cycle arrest mediates tradeoffs between fast recovery and resilience to multiple instances of osmotic shock.

a Top: experiment schematic representing a series of 1.2 M sorbitol osmotic shock step inputs. The first input lasts for 90 minutes and subsequent inputs last 45 minutes, and are separated by 5 minutes. **b** Time traces of Hog1 nuclear enrichment of WT (blue), *sic1*Δ (orange). Shaded regions represent the SEM of n=3 biological replicates. **c** Quantification of adaptation time of Hog1 nuclear enrichment for WT and *sic1*Δ strains for data presented in Panel B. Values are normalized to the average first response for the WT strain. Error bars represent the SEM of n=3 biological replicates. **d** The percent of cells with altered morphological phenotypes at the end of each step input of 1.2 M sorbitol. Error bars represent the SEM of n=3 biological replicates. **P¡0.005; two-sided Student's t-test. **e** Representative brightfield images corresponding to the timepoints quantified in Panel D. Images depict the compromised cell wall morphology indicated by red arrows in the WT (top) and *sic1*Δ (bottom) strains. **f** Conceptual model of the role of cell cycle arrest in hyperosmotic shock response.

2.12 Supplemental Data Figures

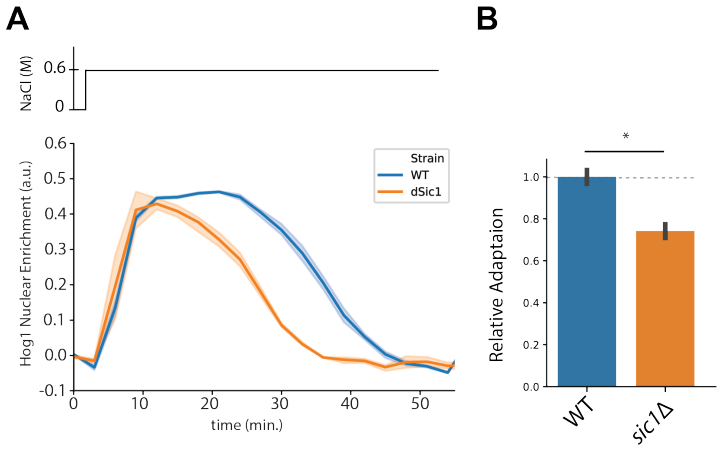


Figure 2.5: The *sic1*Δ mutant recovers faster following osmotic shock induced by 0.6 M NaCl.

a Time traces of Hog1-mVenus nuclear enrichment for WT (blue) and *sic1*Δ (orange) in response to a 0.6 M NaCl step input. Shaded regions represent the SEM of n=3 biological replicates. **b** Quantification of Hog1 adaptation for WT (blue) and *sic1*Δ (orange). Values are normalized to the average WT. Error bars represent the SEM of n=3 biological replicates. *P¡0.05; two-sided Student's t-test.

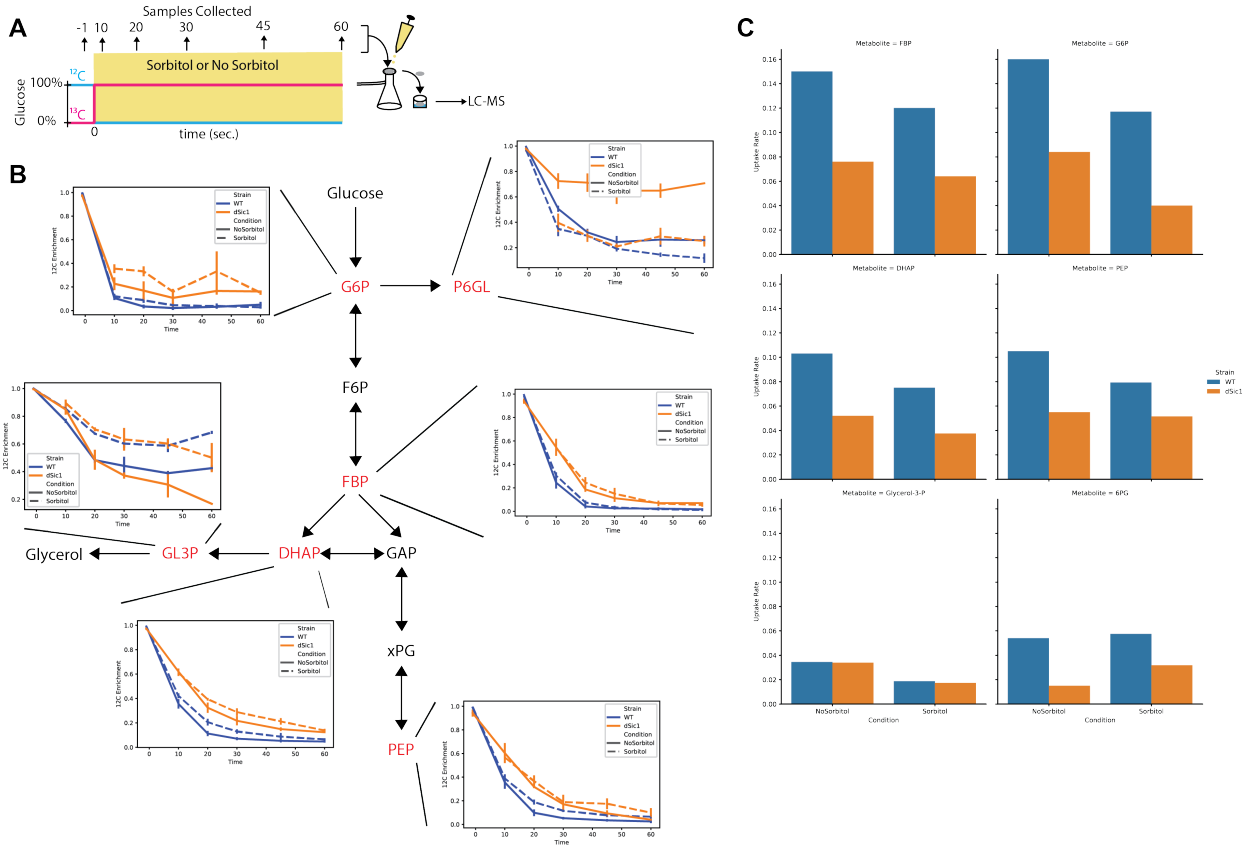


Figure 2.6: WT glucose uptake rate is greater than *sic1*Δ by nearly three-fold in the presence and absence of osmotic shock.

a Cartoon schematic of experiment to infer extracellular glucose uptake rates. At time zero a 1 mL sample of cells was transferred to filter paper above a vacuum manifold and continuously perfused with fully-labeled ^{13}C glucose media with and without 1.2 M sorbitol for durations of 10 s, 20 s, 30 s, 45 s, and 60 s and transferred to quenching solution. **b** A schematic of central glycolysis and the glycerol branch of central metabolism with the targeted metabolites featured. WT (blue) and *sic1*Δ (orange); solid lines represent 1.2 M sorbitol osmotic shock, and dashed lines represent normal defined media. Error bars represent the standard deviation of n=2 technical replicates. **c** Fitting an exponential decay to each ^{12}C enrichment over time trace, bar plots show the average decay constant, which serves as a proxy for ^{13}C incorporation rate, for each metabolite and condition. The plotted constant represents the average of the n=2 technical replicates

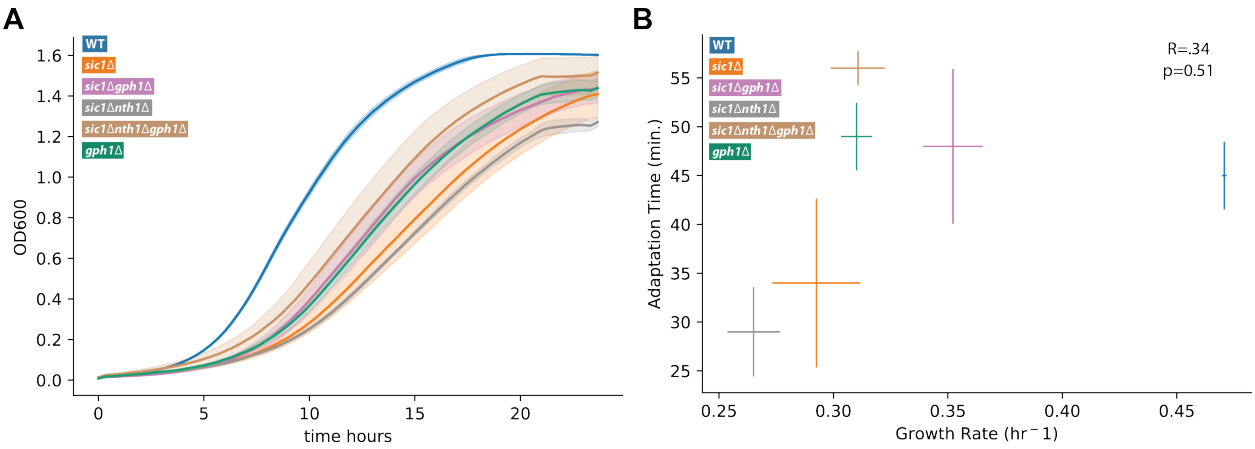


Figure 2.7: Average growth rate bears a weak correlation to average Hog1 adaptation time.

a OD600 readings of WT (blue), *sic1*Δ (orange), *sic1*Δ*nth1*Δ (grey), *sic1*Δ*gph*Δ (purple) ,*sic1*Δ*nth1*Δ*gph1*Δ (gold) and *gph1*Δ (green) strains with readings taken every 20 minutes over 24 hours. Shaded regions represent the SEM of at least n=2 biological replicates. **b** The average growth rate of each strain measured in Panel A plotted against its average adaptation time later shown in Supplemental Figure 5B. Error bars along each axis represent the standard deviation of at least n=2 biological replicates. R, Pearson's correlation coefficient. P-value; Student's t-test.

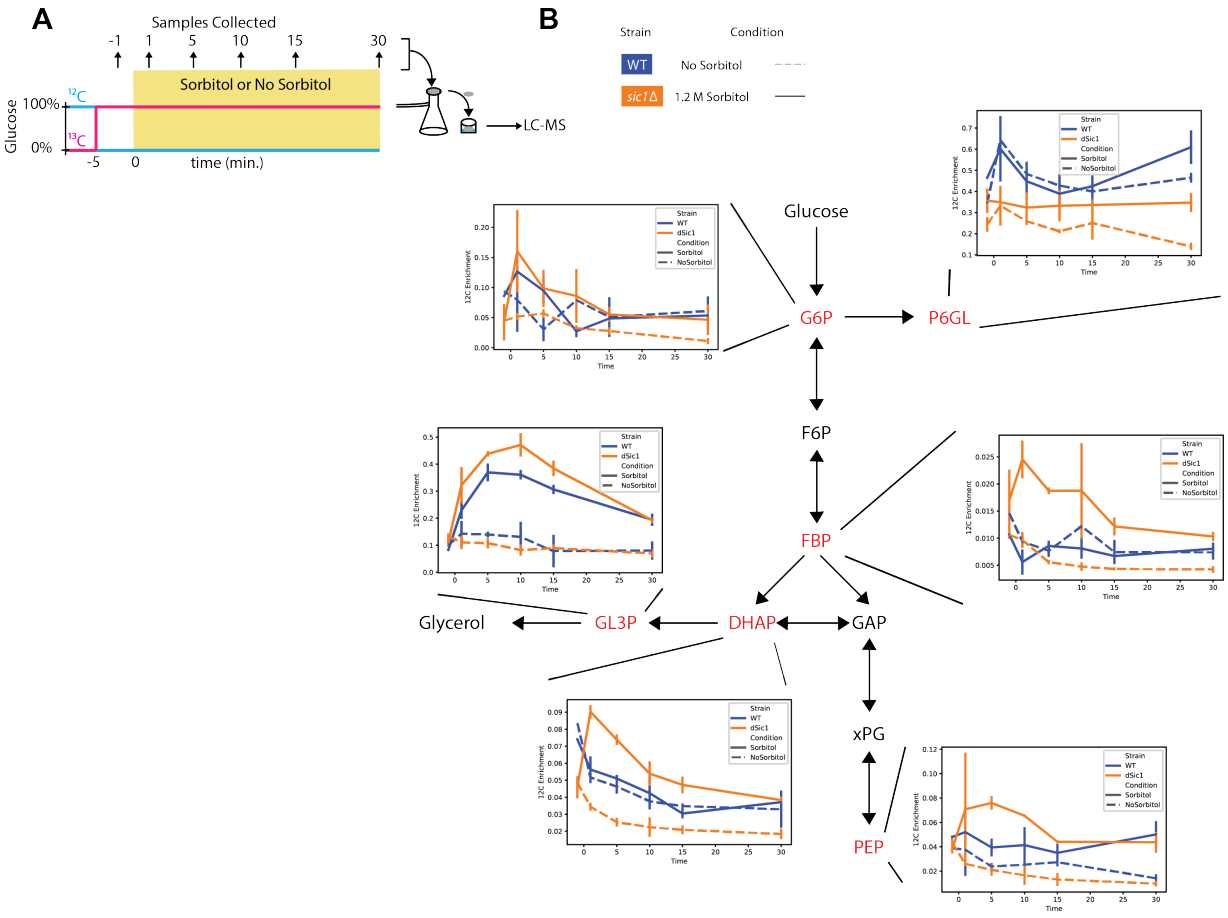


Figure 2.8: Stress-induced mobilization of an internal carbon macromolecule is shunted into central glycolysis in the *sic1*Δ mutant.

a Cartoon schematic of experiment to test internal carbon enrichment of targeted metabolites. Five minutes prior to time zero, cells were resuspended in fully-labeled ^{13}C glucose. At time zero the culture of cells were diluted 1:1 in fully-labeled in ^{13}C glucose with either 2.4 M sorbitol or normal defined media. At the indicated time points, 1 mL of culture was placed on filter paper above a vacuum manifold for the media to wash through, and transferred to quenching solution. **b**. A schematic of central glycolysis and the glycerol branch of central metabolism with the targeted metabolites featured. WT (blue) and *sic1*Δ (orange); solid lines represent 1.2 M sorbitol osmotic shock, and dashed lines represent normal defined media. Error bars represent the standard deviation of n=2 technical replicates.

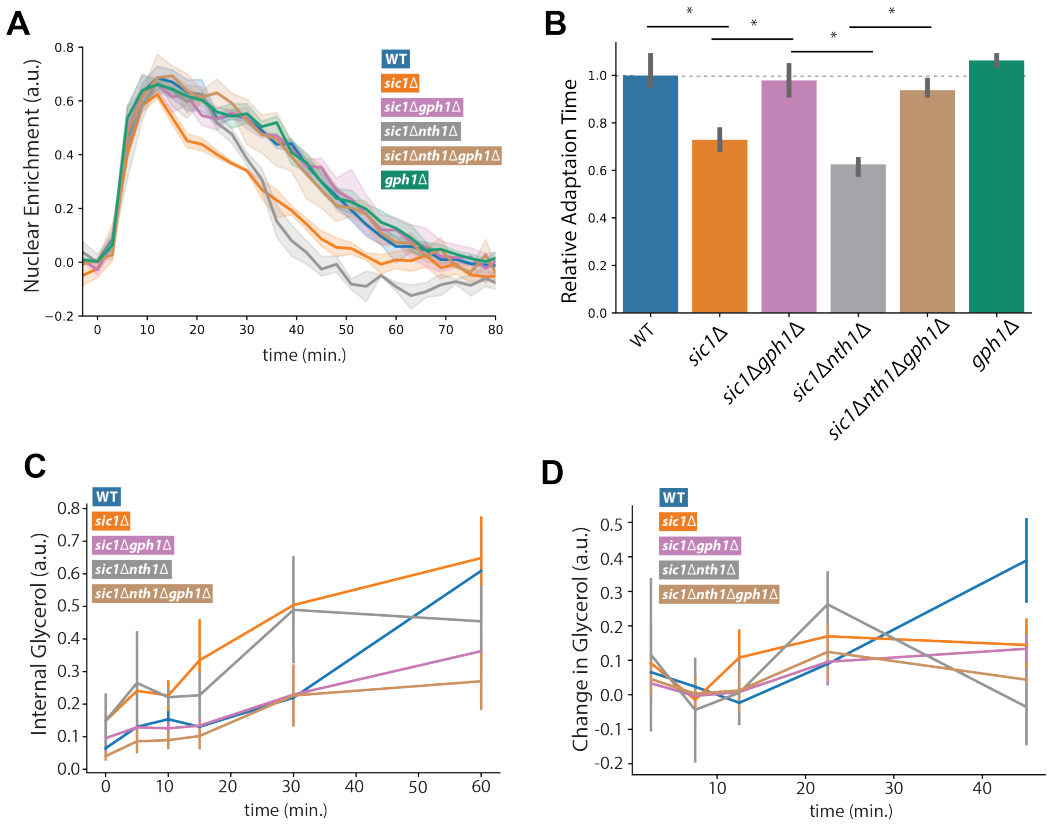
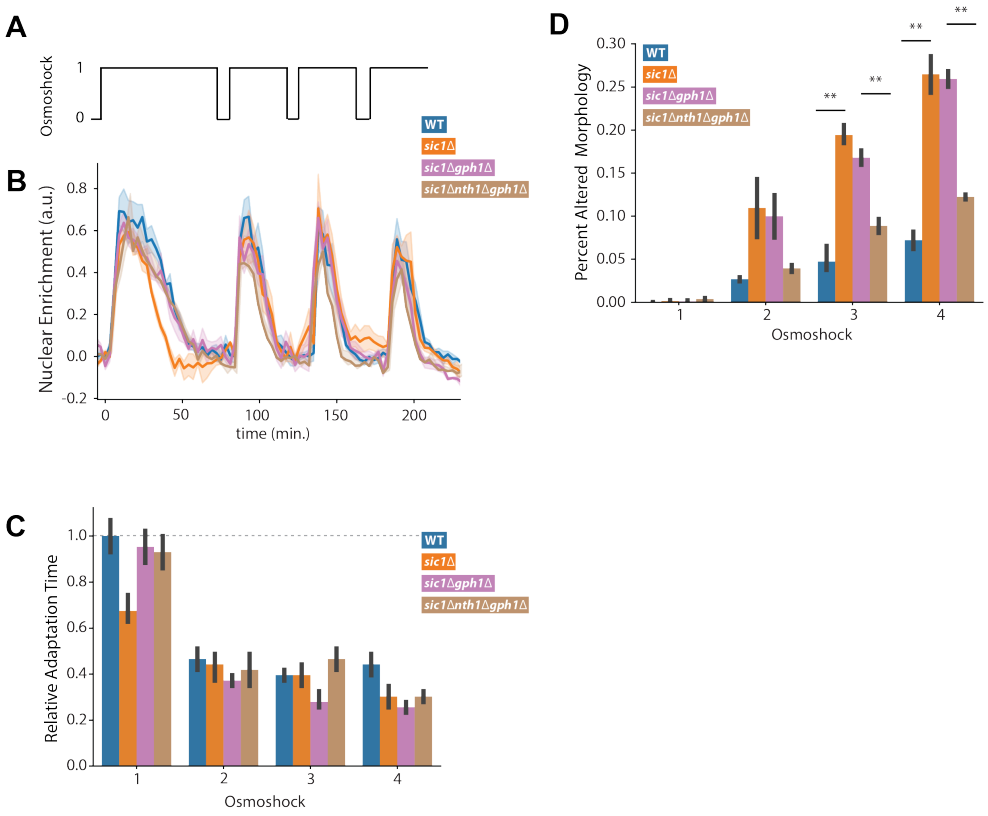


Figure 2.9: Knockout of glycogen catabolism enzyme, Gph1, is sufficient and necessary to rescue accelerated phenotype conferred with removal of Sic1.

a Traces of Hog1 nuclear enrichment over time following 1.2 M sorbitol osmotic shock in the WT (blue), *sic1*Δ (orange), *sic1*Δ*nth1*Δ (grey), *sic1*Δ*gph1*Δ (purple) ,*sic1*Δ*nth1*Δ*gph1*Δ (gold) and *gph1*Δ (green) cells. Shaded regions represent the SEM of n=3 biological replicates. **b**. Quantification of adaptation time of Hog1 nuclear enrichment computed as in Figure 1B. Values are normalized to the average WT. Shaded regions represent the SEM of n=3 biological replicates. *P¡0.05; two-sided Student's t-test.**c** Quantification of internal glycerol as a function of time for WT (blue), *sic1*Δ (orange), *sic1*Δ*nth1*Δ (grey), *sic1*Δ*gph1*Δ (purple) ,*sic1*Δ*nth1*Δ*gph1*Δ (gold) to a step input of 1.2 M sorbitol osmotic shock. Measurements are taken using a colorimetric assay. Error bars represent the standard deviation for n=3 biological replicates. **d**The change in glycerol, calculated as the difference between two time points for data in Panel C, is plotted as a function of time. Error bars represent the standard deviation for n=3 biological replicates.



Supplemental Figure 6

Figure 2.10: Multiple step inputs of osmotic shock reveal different susceptibilities between rescue strains.

a Top: experiment schematic representing a series of 1.2 M sorbitol osmotic shock step inputs. The first input lasts for 90 minutes and subsequent inputs last 45 minutes, and are separated by 5 minutes. **b**. Time traces of Hog1 nuclear enrichment of WT (blue), *sic1*Δ (orange), *sic1*Δ*gph1*Δ (purple), *sic1*Δ*nth1*Δ*gph1*Δ (gold). Shaded regions represent the SEM of n=3 biological replicates.**c** Quantification of adaptation time of Hog1 nuclear enrichment for WT and *sic1*Δ strains for data presented in Panel B. Values are normalized to the average first response for the WT strain. Error bars represent the SEM of n=3 biological replicates. **dThe percent of cells with altered morphological phenotypes at the end of each step input of 1.2 M sorbitol. Error bars represent the SEM of n=3 biological replicates. **P¡0.005; two-sided Student's t-test.

2.12.1 Supplementary Information

Table 2.1: *S. cerevisiae* strain list

Name	Background	Genotype	Description
yARB001	w303a	*HOG1-mVenus-HIS3*	"Wild type" strain
yARB002	w303a	*HOG1-mVenus-HIS3, SIC1::TRP1*	*sic1*Δ strain
yARB003	w303a	*HOG1-mVenus-HIS3, SIC1::TRP1, NTH1::NAT*	*sic1*Δ*nth1*Δ
yARB004	w303a	*HOG1-mVenus-HIS3, SIC1::TRP1, GPH1::LEU2*	*sic1*Δ*gph1*Δ strain
yARB005	w303a	*HOG1-mVenus-HIS3, SIC1::TRP1, GPH1::LEU2, NTH1::NAT*	*sic1*Δ*nth1*Δ*gph1*Δ strain
yARB006	w303a	*HOG1-mVenus-HIS3, GPH1::LEU2*	*gph1*Δ strain

Chapter 3

A Toolkit for Rapid Modular Construction of Biological Circuits in Mammalian Cells

3.1 Abstract

The ability to rapidly assemble and prototype cellular circuits is vital for biological research and its applications in biotechnology and medicine. Current methods for the assembly of mammalian DNA circuits are laborious, slow and expensive. Here we present the Mammalian ToolKit (MTK), a Golden Gate-based cloning toolkit for fast, reproducible and versatile assembly of large DNA vectors and their implementation in mammalian models. The MTK consists of a curated library of characterized, modular parts that can be assembled into transcriptional units and further weaved into complex circuits. We showcase the capabilities of the MTK by using it to generate single-integration landing pads, create and deliver libraries of protein variants and sgRNAs, and iterate through

dCas9-based prototype circuits. As a biological proof of concept, we demonstrate how the MTK can speed the generation of non-infectious viral circuits to enable rapid testing of pharmacological inhibitors of emerging BSL4 viruses that pose a major threat to human health.

3.2 Introduction

Molecular cloning is the cornerstone of modern biological research, enabling the generation of DNA vectors that encode a wide variety of molecules, which can further be organized into genetic circuits with applications ranging from basic discovery to medical platforms. To transform molecular cloning into a fast and reproducible engineering discipline, circuit parts need to be modular, vetted for their function and easy to share. Furthermore, these circuits need to be rapidly assembled, amenable to fast prototyping, and easily deliverable. Conventional approaches to produce DNA vectors involve PCR amplification of DNA fragments, which are ligated following digestion with restriction enzymes. This procedure is time consuming, laborious, error-prone and requires sequencing control steps. Commercial DNA synthesis provides a potential avenue to reliably build libraries of DNA vectors. However, current pricing, size limitation, and turnaround time remain a significant bottleneck. While in yeast and plant systems toolkits for rapid assembly of genetic circuits from libraries of well described modular parts exist[40–45], no similar comprehensive resource exists for mammalian systems. Existing resources such as the Gibson modular assembly platform (GMAP)[46], and more recently, the Mammalian Modular Cloning (mMoClo)[47] and Extensible Mammalian Modular Assembly kit (EMMA)[48] provide a starting point for developing this functionality. However, while these methods use sets of modular parts that can be assembled directionally into transcription units (TUs), they have important remaining shortcomings. GMAP requires sequencing in between all cloning steps, which for

large DNA elements (such as a 7 kb viral polymerase) makes the method slow and expensive since PCR reactions of such elements are difficult and prone to mutations, and require, in the viral polymerase example, at least 7 sequencing reactions for every tested clone. EMMA and mMoClo provide a method to hierarchically assemble DNA vectors that encode large circuits, without the need of additional PCR and sequencing reactions. However they both lack several characteristics that would make them useful for general use. mMoClo, is not available in Addgene or upon request, making its adoption unrealizable. EMMA, while available to users, lacks a library of tested parts for facile use and integration of sgRNA cloning. EMMA also provides good flexibility in terms of part building, but it has not been shown to be able to build circuits with more than 3 transcriptional units. Finally, because EMMA necessitates assembly of 25 parts to build functional plasmids, a substantial energy barrier for its adoption in the community at large seems to exist. In this work, we present a toolkit that capitalizes on the strengths of previous efforts[42,49], but substantially extends their capabilities for mammalian molecular cloning, providing a first platform that can be immediately adopted by the community. The Mammalian ToolKit (MTK) is a library of over 300 parts including vetted promoters, 3'UTRs, fluorescent proteins, insulator and P2A elements that can be rapidly combined to build complex genetic circuits. To facilitate the MTK's ease of use, users can consolidate parts into ready-made, customizable destination vectors to bypass the need to re-assemble multiple plasmids and often-used elements such as selection markers. These vectors can be delivered to a wide range of cell types through viral, recombinase and CRISPR/Cas9 methods. Importantly, the MTK requires the assembly of only 8 unique parts for a functional transcriptional unit, which in turn can be immediately used to build plasmids carrying up to 9 transcriptional units that can encode the expression of a wide variety of coding DNA sequences as well as of sgRNAs for targeting of Cas9 proteins of two species. As a proof of concept, we built a hAAVS1 landing pad for single integration of genetic circuits, created and delivered libraries of protein variants and sgRNAs and rapidly

compared multi-species Cas9-based genetic circuits. Finally, using Ebola virus (EBOV) as an example, we demonstrate how the MTK can simplify and speed the *de novo* generation of non-infectious BSL2 stable cell line expression systems commonly deployed to facilitate analyses of emerging BSL4 viruses.

3.3 An expansive, modular cloning toolkit for rapid prototyping in mammalian cells

The basis for the MTK is a library of parts that are "domesticated" from source DNA using a Golden Gate (GG) reaction10 into a standard vector. This requires designing oligos that anneal to the source DNA, appending restriction sites that enable the resulting PCR product to be digested and ligated into the base parts vector with chloramphenicol resistance (MTK0_027). The resulting plasmid is then sequenced once to ensure fidelity to the source sequence, and can later be reused as a validated modular part in a variety of genetic constructs. Similar to previously published work[42,47–49], we utilize the type-IIS restriction enzymes BsaI and BsmBI with "reach over" endonuclease activity that leaves arbitrary four base overhangs adjacent to the recognition site. The defining feature for a part vector is a unique four base overhang that categorizes it to similar parts, and ensures an ordered 5' to 3' ligation of parts into a transcriptional unit (TU) of multiple parts. We used the same overhangs as defined in the Yeast Toolkit (YTK)[42]. These overhangs have been shown to enable highly efficient assemblies and, because the YTK has reached a wide community of users, parts that have been built for the YTK can be reused in the MTK. With a library of sequenced-verified part vectors, only diagnostic restriction digests are necessary to verify correct assembly in subsequent GG reactions. We employ BsmBI and BsaI GG reactions in an alternating manner to assemble a library of part vectors into libraries of TU plasmids, and ultimately into multi-transcriptional unit (multi-TU)

plasmids. Like part vectors, TUs have unique overhangs that define their position in a multi-TU plasmid. The time from initial source DNA to final plasmid product is four days, assuming the use of fast growing E. coli and 18h for sequencing turnaround(Fig.3.1a). However, once MTK plasmid parts are constructed, they can be reused to assemble new configurations in only two days. This is a substantial gain for a general, non-specialist, user who can achieve many complex circuits already from the large library of parts that the MTK provides.

The MTK encompasses part vector categories 1-8 that are sufficient to build and deliver a vast combinatorial library of genetic constructs to cells (Fig.3.1b). Parts 2, 3 and 4 form the core of a TU, specifying the 5' UTR, coding, and 3' UTR sequences, respectively. Part 2 corresponds to promoter sequences that can be specified to drive constitutive or inducible expression or recruit diverse polymerases. Part 3 vectors are canonical coding sequences that are typically proteins of interest. Part 4 corresponds to 3' UTR sequences that can encode polyadenylation (pA) sequences that terminate transcription or spacer sequences that couple transcription to the downstream TU. Parts 1 and 5 correspond to connector sequences that enable the sequential ordering of TUs into a multi-TU plasmid, with many versions provided in the MTK. For example, Part 1 implements connectors with insulator sequences[50] to minimize polymerase read-through between TUs or connectors with cis-acting P2A elements to enable strongly correlated expression from one promoter driving up to five downstream TUs[51,52]. Overall, the connectors included in the MTK allow the construction of multi-TU plasmids encoding 9 TUs, and hence large genetic circuits. Further nested subdivision can be achieved — for example Part 3 can be replaced with Part 3a and 3b, and still connect to a Part 4, allowing for combinations that implement tethering of localization tags, protein domains or any desired coding sequence both N- and C-terminally with an innocuous linker sequence in between (Fig. 1b). Additionally, coupled Part 234 vectors accommodate rapid cloning of small guide RNA (sgRNA) expression by oligo annealing for CRISPR/Cas9-related genetic constructs. Parts 6, 7 and

8 generally flank a typical TU and can encode the method of delivery to cells, such as homology arms for a locus of integration. Lastly, the MTK contains kanamycin resistant Part 0 destination vectors that allow delivery via viral transduction, transposase transfection[53,54] and homologous recombination from the same, recyclable collection of TUs (Fig.3.1c). Only Part 0 in the final GG reaction needs to be changed in order to accommodate different delivery methods into cells. Overall, the MTK combines intuitive organization with an expansive library (345 parts listed in Supp. Table 1 and publicly available on Addgene) for rapid facile construction of genetic constructs that can be integrated into cells without redesign or re-sequencing, irrespective of the delivery method.

3.4 MTK enables facile construction of independent or multicistronic TUs with different levels of expression

The MTK contains seventeen characterized constitutive promoters derived from a mix of human, mouse and viral[55] origin, and two inducible promoters whose use is illustrated in Fig.5. This set provides a range of expression levels, and a balance of native and transgene promoters for the two conventional cell lines tested. To enable matching expression levels of different proteins without re-using the same promoter, this panel also includes select promoters that closely match each other. To characterize the relative strengths of these promoters we assembled a panel of 14 TUs with varied promoter parts driving mAzamiGreen expression, flanked by the bovine growth hormone (Bgh) pA 3' UTR. An insulated downstream TU, where a CAG promoter expresses mScarlet with the rabbit beta-globin (Rgl) pA 3' UTR (Fig.3.2a), was used for normalization. Transient transfection in HEK293T cells demonstrated that this suite of promoters spanned a smooth continuum

of 2 orders of magnitude over background, with the strongest promoter, CMV, more than 300-fold greater than a promoter-less mAzamiGreen (Fig.3.2a). The relative strengths of these promoters was largely maintained across human and mouse cell lines (Supplementary Fig.3.7a,3.7b), suggesting that they are portable across commonly used cell lines. Additionally, the rank of promoter expression was consistent between transient and stably integrated expression (Supplementary Fig.3.7c,3.7d). The MTK provides options for further fine-tuning of TU expression using a collection of five different 3' UTR sequences. Similar to the promoter comparison, we generated a circuit with the same constitutive promoter (EF1a) and varied the 3' UTR sequence to compare normalized mAzamiGreen expression (Fig.3.2b). We observed a range of 4-fold change in expression of mAzamiGreen among the three conventional 3' UTR sequences. While the Bgh pA and Rgl pA signals have the same effect on gene expression, the simian virus 40 (SV40) pA signal had a nearly 1.5-fold greater effect (Fig.3.2c). Moreover, replacing a canonical 3' UTR sequence with either a spacer sequence with a stop codon (used for lentiviral delivery) or without a stop codon (used for multicistronic read-through) diminished expression of its upstream coding sequence. While we explored here only a few combinations related to one promoter (EF1a), libraries of fine-tuned expression of a protein of interest can be easily generated using this platform by the interested user. Finally, to enable concomitant expression of up to 6 TUs in a multicistronic vector, we incorporated ribosome skipping P2A elements[51,52] into the MTK part 1. We used this part to build a vector with one CAG promoter driving membrane-tethered iRFP713, cytoplasmic mAzamiGreen, mCerulean-tagged p38KTR[56], and histone H2B fused to mScarlet separated by P2A elements, and compared expression from this construct to a parallel multi-TU construct (Fig.3.2c). The multicistronic construct conferred a 35% reduction in vector size (15540 bp to 10020 bp) compared to the multi-TU construct yet maintained the correct localization of each of the proteins in transiently transfected HEK293T cells (Fig.3.2d). These data demonstrate the opportunity to deploy P2A elements to generate multicistronic expression constructs to circumvent the

size limitations of conventional viral delivery vectors.

3.5 MTK contains a landing pad system and accompanying destination vector

When delivering synthetic genetic circuits, it can be essential to have site-specific and reliable single copy integration. Therefore, the MTK includes a BxBI-dependent landing pad (LP) system for integrating synthetic circuits in a locus of choice. This system is divided into two parts: an LP cell line and an LP transfer vector. The LP transfer vector can be used with MTK generated LP cell lines, as well as cell lines carrying the SBI pinpoint system landing pad, or that were generated using other methods and that contain a BxBI attB site[47]. We used a CRISPR/Cas9 approach to build a HEK293T cell line with a BxBI landing site in the well-characterized adeno-associated virus integration site 1 locus[57,58] (hAAVS1 LP) and tested its ability for site-specific recombination. The vector used encoded a multicistronic cassette with hygromycin resistance and nuclear-localized mRuby2. The BxBI attP site was located between the promoter and the first gene of this cassette (Fig.3.3a). We verified the correct integration of the LP cassette by PCR in 8 clones (Supplementary Fig.3.93) and continued its characterization in clone 8 with a monoallelic LP. We further confirmed the presence of the LP in the cell line by mRuby2 expression (Fig.3.3b). In order to verify that the transfer vector was correctly integrated in the LP, we positioned a promoter-less multicistronic cassette encoding blasticidin resistance and nuclear localized tagBFP downstream of the BxBI attB site. When site-specific integration is accomplished, the cell line switches fluorophore (from mRuby2 to tagBFP) and resistance (hygromycin to blasticidin). To test the landing pad, we integrated a transcriptional unit that expresses H2B mAzamiGreen-fused histone H2B from the CAG promoter (Fig.3.3, Supplementary Fig.3.3b). Upon integration of the transfer vector into

the hAAVS1 LP, we noted the decrease in mRuby2 expression together with expression of mAzamiGreen and tagBFP in most cells (Fig.3.3b, Supplementary Fig.3.3b), with a small fraction of cells showing no or incorrect integration of the transfer vector. Finally, while we chose to integrate the landing pad in one genomic locus, others can be easily targeted by replacing the homology arms in the CRISPR/Cas9 vector.

3.6 MTK allows rapid, one step combinatorial construction of libraries

The inherent modularity of the MTK workflow facilitates parallelization of large libraries of vectors in one-step combinatorial reactions. Moreover, such libraries can be reused and delivered repeatedly and in multiple ways. These features are particularly relevant for optimization of a large number of protein or circuit variants or CRISPR/Cas9 screens where efficient throughput and variant representation are critical. Here, we provide two examples of MTK combinatorial library construction that illustrate how it can be applied to streamline the generation of large libraries of TU variants. First, a library of fluorescent proteins that vary in their localization, and which can be integrated as a single-copy into the BxBI landing pad site; second, a library of viral-delivered sgRNAs that target Cas9 to two fluorescent proteins leading to their disruption. In the first library, we combined one of three localization tags (NLS (nucleus), NES (cytoplasm) and Lyn (plasma membrane)) with four fluorescent proteins (tagBFP, mAzamiGreen, mScarlet and iRFP713) (Fig.3.4a). Performed in parallel, this reaction generated 4 fluorescent variants per localization tag. These libraries of variants were further pooled in equimolar amounts with a LP destination vector. This single reaction was predicted to create a library of 64 distinct variants, where each vector in the library encodes three random fluorescent proteins that are localized in the nucleus, cytoplasm and plasma membrane. The final library was delivered to the

hAAVS1-LP HEK293T cell line described before (Fig.3.4). Visualization of expression and localization of the fluorescent proteins after selection with blasticidin showed a qualitative assortment of random, single fluorescent markers located in the three subcellular compartments, (Fig.3.4b) as predicted from the cloning strategy. While other Golden Gate-based toolkits have the ability to generate combinatorial libraries of proteins, they cannot generate combinatorial libraries of sgRNAs. Such libraries are useful for targeting a combination of genes, either with Cas9 or with a wide variety of modified dCas9 variants. To address this need, we showcased the ability of the MTK to generate sgRNA combinatorial libraries by building a library of sgRNAs that target two fluorescent proteins. We first assembled TUs that contain sgRNAs that target tagBFP, mScarlet, or that are non-targeting[59]. We combined those in a multi-TU unit so that a final Lentiviral delivery plasmid contained 2 guide RNAs, ensuring that we had any one of four possible outcomes (knockout of tagBFP, mScarlet, tagBFP and mScarlet, or no knockout) in the final library (Fig.3.4c). Each multi-TU encoded iRFP713 to facilitate the identification of cells that express sgRNAs. Since sgRNA targeting has variable efficiency[60], we concurrently generated four individual lentiviral plasmids for the four outcomes of the library as a control to compare the results of the combinatorial assembly. We independently transduced the four control viruses in HEK293T cells expressing tagBFP, mScarlet and spCas9 (3C cell line) and quantified tagBFP and mScarlet expression two weeks after transduction. Cells that contained the non-targeting sgRNAs expressed high levels of both tagBFP and mScarlet (Fig.3.4d, first panel). Cells expressing the tagBFP-targeting sgRNA showed reduced expression of tagBFP in 86% of the population (Fig.3.4d, second panel), while only 27% of cells showed reduced expression of mScarlet when its guide was expressed (Fig.3.4d, third panel). Accordingly, when both guides were co-expressed, only 31% cells show reduced expression for both proteins, while most cells had reduced tagBFP (Fig.3.4d, fourth panel). We used these results to build a linear classifier, trained on the individual targeted variants of sgRNA to determine the likely sgRNAs that each cell received in a

population transduced with the full sgRNA library. Due to the inefficient targeting of the mScarlet guide, the linear classifier had an average precision of 0.62 and an average recall of 0.58 (Fig.3.4e, Supplementary Fig.3.10) and predicted higher than expected proportions of non-targeting and tagBFP sgRNA-expressing cells when presented with an equal proportion of each variant of the library (Fig.3.4f). To test the combinatorial assembly approach, we transduced the 3C cells with 4 biological replicates of the sgRNA library. After 14 days of selection, we measured tagBFP and mScarlet expression and identified cells belonging to the four possible outcomes of the viral library, suggesting that all combinatorial possibilities are achievable through this method (Fig.3.4g). Importantly, the fraction of each outcome was in accordance with the predictions of the classifier, indicating that the MTK is able to rapidly generate libraries for combinatorial exploration of proteins variants or sgRNAs that maintain the correct ratio of variants.

3.7 MTK facilitates optimization of combinatorial gene circuits for synthetic biology applications

To capitalize on the potential of the MTK for parallel testing of combinatorial circuits, we constitutively expressed deactivated Streptococcus pyogenes and Streptococcus aureus CRISPR Cas9 (dSpCas9 and dSaCas9, respectively) C-terminally fused to either a fluorescent protein (FP), repressor domain (Krüppel associated box, KRAB), or an activator domain (VP64-p65-Rta, VPR)[61–65]. These protein-effector combinations were targeted to two inducible promoters included in the MTK: the GAL4 Upstream Activating Sequence (UAS) or the Tet Responsive Element (TRE) (Fig.3.6a). Each gene circuit was assembled upstream of a constant TU control with constitutive EF1a expression of iRFP713. In total, we transiently transfected twelve different circuits into HEK293T cells and normalized changes in mAzamiGreen expression by iRFP713. The plasmids encoding

these circuits have an average size of 16kbp. While our parallel building protocol required only 2 test restriction enzyme digestions to verify the correct assembly of the circuits, methods based on gibson assembly would have required about 200 sequencing reactions (12 circuits, sequencing 14kbp of each one) for sequence verification of the 12 circuits assembled from 9 PCR products. As expected, in each of the circuit iterations, expression of mAzamiGreen relative to the iRFP713 control was below basal with the repressor domain, or above basal with the activator domain (Fig.3.6b). dSpCas9 achieved approximately 3-fold reduction in fluorescence through the KRAB effector and 30-fold induction in both the UAS and TRE promoters. In contrast, the dSaCas9 fused to VPR was an effective activator, increasing expression 30-fold, but when fused to a KRAB domain, dSaCas9-mediated repression was comparable to background FP expression levels. The TRE promoter exhibited higher basal activity due to its overall greater repression, and lower activation when compared to the UAS promoter. This information expedites the rational design of dCas9-based gene circuits and can be easily scaled to screen and customize various combinations of genetic circuits.

3.8 MTK streamlines the generation of endogenous viral circuits

Viruses represent a class of naturally occurring genetic circuits that are particularly amenable to MTK construction due to their modular genome organization. As such, the MTK provides virologists the opportunity to rapidly "boot up" the genes, replicons, or the complete life cycles of emerging viral agents as soon as their sequences are available. A key example corresponds to the filoviruses, a family of emerging highly pathogenic (BSL4) RNA viruses that includes the Ebola viruses[66]. Global efforts to contain filovirus outbreaks have resulted in several candidate vaccines and antivirals; however, there is currently no

approved preventive or therapeutic treatment. A challenge to advancing our understanding of EBOV was its highly pathogenic nature and the need for BSL4 containment. In the late 1990's a BSL2 system was developed for Zaire Ebola virus (ZEBOV) to facilitate research on its transcription and genome replication sub-lifecycle and antiviral discovery efforts[67,68]. However, this system requires transient transfection of 6 plasmids (four viral proteins (NP, VP35, VP30, and Lpol), T7 polymerase, and a T7-driven minigenome reporter construct[67]) (Fig.**??**a). The complexity and variability of this system has limited feasibility to perform industrial scale compound screening campaigns to identify effective inhibitors of EBOV. More recently, a robust stable ZEBOV RNP cell line system was developed via sequential integration of RNP viral proteins[69]; however, this approach entailed a time-intensive, multi-step approach that took 1 month. This timescale is undesirable if the goal is to enable rapid investigation of emerging viruses or variants of the same virus. Given the urgent need and time sensitive nature that outbreaks of highly pathogenic emerging viral agents pose to human health, we explored how the MTK workflow can streamline and improve the de novo generation of BSL2 tools like the EBOV RNP minigenome replicon system. We used the MTK system to design a single 4-cistronic construct of Zaire ebolavirus (ZEBOV-4cis, Fig.**??**a) directly and simultaneously into five different Part 0 destination vectors (Supplementary Fig. S5a). In transient transfections, all variants of the ZEBOV-4cis constructs displayed levels of minigenome activity similar to the 6-plasmid system, with 40-60-fold higher activity than cells lacking the viral polymerase Lpol. All five of these systems showed robust cell viability (Fig.**??**b, Supplementary Fig.3.11b). Stable cell lines harboring the ZEBOV-4cis construct were generated using PiggyBac transposon-mediated integration in 10 days followed by isolation of clonal cells (Fig.**??**c, Supplementary Fig.3.11c). Comparable viability and minigenome reporter activity were observed in two independent ZEBOV-4cis stable populations and clonal cell lines (Fig.**??**d). The P2A ribosome skipping sites encoded in the ZEBOV-4cis construct confer 2A peptide tags at the C-terminus of each of the NP, VP35, and VP30

proteins, enabling the simultaneous detection of all 3 proteins with a 2A peptide antibody (Fig.**??**e). Larger multi-ORF fusion protein products were not detected with the 2A peptide antibodies or NP and VP35 antibodies, indicating efficient "self-cleavage" occurs at each of the P2A sites. As expected, VP35 showed diminished expression, probably due to the known effect of P2A elements on downstream peptides. This is a trait that may be desirable in certain systems such as viruses that regulate expression levels of downstream genes[70]. Parallel immunofluorescence analysis showed a punctate cytoplasmic localization pattern (Fig.**??**f) consistent with previously described ZEBOV inclusion bodies that correspond to sites of viral replication[71–73]. A lack of available specific antibodies or 2A tag encoded in the Lpol gene made parallel protein expression analyses infeasible for Lpol; however the Lpol dependence of minigenome activity in this system (Fig.**??**b and **??**c) provides confirmation of functional Lpol expression. We also examined the susceptibility of the ZEBOV-4cis stable cells to previously described small molecule inhibitors: Gedunin, an inhibitor of heat shock protein 90, and 6-Azauridine, a nucleoside analog[74,75]. Treatment of ZEBOV-4cis stable cells with 5 µM of each compound revealed ¿50% inhibition of minigenome activity, with minimal impact on cell viability (Fig.**??**g). These data indicate that the ZEBOV-4cis stable cell lines generated here are similarly susceptible to known inhibitors identified in transient RNP minigenome systems or recombinant virus systems[74,75]. Taken together, these data demonstrate that the MTK workflow facilitates rapid generation of BSL2 systems to study filoviruses and other viral agents as they emerge, providing a critical opportunity to increase the throughput and the rate at which we can screen for and identify candidate antiviral compounds.

3.9 Discussion

A key element of cellular engineering is the ability to quickly build, test, and iterate on designs. This is currently infeasible in mammalian cells using conventional methods in molecular cloning. Here, we presented the MTK, a platform that takes an important step to remove bottlenecks in mammalian cellular engineering. A major asset of the MTK is a large, characterized suite of modular parts to build versatile TUs, which we showcased by driving different levels of gene expression for single TUs and creating multicistronic constructs. Additionally, we included lentiviral, recombinase and Cas9 delivery vectors to maximize the number of contexts where circuits can be implemented. Due to the modularity of the MTK, it is straightforward to generate new variants of delivery vectors, which we demonstrated by creating BxBI Landing Pad for the hAAVS1 locus. The set of validated, interchangeable parts of MTK constitute fundamental tools for the rapid and facile assembly of genetic circuits. We illustrated this point in two ways: first, we assembled a combinatorial library of transcriptional units that encode different proteins or guide RNAs; second, we combinatorially assembled new circuits that use different MTK parts. Such libraries are generated through the MTK in a straightforward way, reproducing all variants included in the combinatorial assembly process and maintaining their ratios through all the cloning steps. While these studies tackled a small set of applications, many new combinatorial libraries can be built with any parts in the MTK as it currently exists, or as users add to its parts. For example, introducing a panel of C-terminal fluorescent protein tags into the MTK would only require a one-time PCR and sequence verification for each tag to "domesticate" the series as 4a Parts. Additionally, the facile assembly of sgRNAs expedites the production of expression vectors that are easily amenable to CRISPR screens of all scales. With the many reagents encoded in the MTK library, diverse TUs with multiple functions can be rapidly assembled, screened and repurposed to generate a wide array of final circuit vectors. The capabilities provided in the MTK are particularly

relevant in the context of infectious disease outbreaks where time is of the essence. Responding to and containing such outbreaks is an important public health challenge that demands rapid production and iteration of viral circuits in mammalian cells to enable discovery of inhibitors and analysis of their basic biology. Using ZEBOV as an example, we demonstrated how the MTK can reduce the lead time needed to generate functional reagents and respond to outbreaks of highly pathogenic viruses. We generated functional cell lines stably expressing the ZEBOV replication complex components in days, while maintaining a library of these components for diverse future applications. Furthermore, given that discovery or availability of a virus sequence does not necessarily correlate with successful culturing of the virus in a lab[76], the MTK system provides a powerful and parallelizable method to rapidly test multiple viral strains and primary clinical isolates, as well as engineered variants of each of these that are optimized for expression in cell culture. While this presents a scaling challenge for conventional cloning approaches, the MTK workflow enables faster iteration and identification of optimal viral circuits for downstream analyses. Thus, the MTK presents a tool for virologists to begin to scale functional experimental studies apace with the recent explosive growth in viral genome sequences[77,78]. Finally, while not explored in this work, the MTK constitutes a launching platform for additional exciting capabilities, including the incorporation of barcoding capabilities for bulk and single-cell sequencing technologies[79–81], enhanced flexibility in genetic circuit design[82], and automation to streamline assembly. These advances are poised to position the MTK as a major catalyst for biological research and biotechnology.

3.10 Methods

Bacterial Cell Culture

Commercial MachI and XL10 strains (QB3 MacroLab) were used to transform plasmid vectors. A typical transformation mixture consists of 2 µL of the Golden Gate reaction product, 48 µL bacteria, incubated on ice for 30 minutes, heat shocked at 42°C for 1 minute, recovered on ice for 5 minutes, reaction mixture plated onto selective agar and incubated overnight at 37°C. In the case of multi-TU transformations, cells recovered in LB media for 30 min after heat shock at 37°C before plating reaction onto kanamycin selective agar plates. Cells were cultured in antibiotic concentrations of 100µg/mL chloramphenicol, 25µg/mL carbenicillin and 100µg/mL kanamycin.

Golden Gate Reactions The general reaction mixture follows 0.5 µL per PCR product, annealed oligos, geneblock, or plasmid (50 fmol.µL-1); 1 µL T4 DNA Ligase Buffer (10x) (NEB) with 0.25% PEG; 0.5 µL T4 DNA Ligase (NEB) diluted with water to a total volume of 9.5 µL. For BsaI GG reactions, we added 0.5 µL BsaI-HFv2 (NEB R3733). For BsmBI GG reactions we used 0.5 µL of either BsmBI (NEB) or FastDigest Esp3I (Thermo Scientific FD0454 NEB) (both 10,000 U/mL).

Thermocycler protocols The GG protocol is primarily used for assembly reactions. The reaction temperature is initially held at 45°C for 2 min to digest the plasmids followed by 20°C for 4 min to anneal constituent parts together. After repeating these first two steps 24 times, the temperature is increased to 60°C for 10 min to digest remaining recognition sites and inactivate the ligase. Then the temperature is held at 80°C for 10 min to inactivate the enzyme. Lastly, the reaction is held at 12°C indefinitely. The "GG End-On" protocol is used when BsaI or BsmBI sites need to be retained in the final product. The temperature is initially held at 45°C for 2 min to digest the plasmid followed

by 20°C for 5 min to anneal and ligate the resulting plasmid. These steps are cycled 24 times and then held at 16°C indefinitely.

Domestication of Parts Forward and reverse primers ordered from IDT (www.idtdna.com) were manually designed to anneal to source DNA. See Supplementary Table 3 for part-specific design of domestication primers. In summary, internal BsaI and BsmBI sites were removed and tandem BsaI and BsmBI sites were appended to both the 5' and 3' ends of the sequence (See Supplemental Tables 3 and 4). PCR was performed with the general reaction mixture of 10 µL Q5 polymerase master mix (NEB MO492S) using 1 µL each of the forward and reverse primers (10 µM) and 0.5 µL of template DNA. The desired PCR product was gel extracted (ThermoFisher Gel Extraction K0691) and 1 µL of the final elution added to a BsmBI-mediated GG reaction (described above) with the MTK0_027 domestication vector. The resulting product was transformed into bacteria and grown in selective LB overnight. Plasmid DNA was extracted (ThermoFisher MiniPrep K0503) and sequenced verified.

Removal of Internal BsaI and BsmBI Sites To remove an internal BsaI or BsmBI site when domesticating a coding DNA sequence part, overlapping forward and reverse primers are designed at the restriction site with a synonymous mutation made to ablate the internal site while preserving the coded amino acid. For non coding parts, substitutions were made to maintain CG content when possible. Flanking the site are complementary BsmBI overhangs such that upon digestion of the resulting PCR products, the components anneal together.

Oligo Annealing Oligos were designed such that the desired sequence was at least 15bp long, and that when complemented the oligos generated the overhangs associated with the part number. To anneal oligos into double stranded DNA, we prepared a reaction mixture as follows: 1 µL of each oligo (100 µM), 1 µL T4 Ligase Buffer (NEB), 1 µL T4 PNK (NEB), 6 µL water. Mixture was incubated at 37°C for 1 hr and then diluted

to a volume of 200 µL. To anneal the oligos, a thermocycler protocol was prepared to hold the temperature of the reaction at 96°C for 6 min, and ramp down 0.1°C per second to 23°C. The reaction is then held at 23°C indefinitely. In the case of sgRNA design, this final mix is then added to a BsmbI-mediated GG reaction into its corresponding sgRNA destination vector. In the case of multi-oligo assembly, each oligo is added to the reaction mixture with MTK0_027 in a BsmBI-mediated GG reaction as described above.

Geneblocks Gene fragments were ordered from IDT as either whole constructs or partial constructs with complementary overhangs to ensure proper domestication.

TU assembly Part 1-5 plasmids were pooled together with a recipient Part 678 plasmid following a 2:1 molar ratio in a BsaI-mediated GG reaction as described above.

MTU assembly Constituent transcriptional unit plasmids were pooled in a 2:1 molar ratio with the destination vector in a BsmBI-mediated GG reaction as described above.

Mammalian Cell Culture and Transfection

HEK293T and 3T3 cells were maintained in DMEM (Dulbecco's Modified Eagle Medium, Gibco) supplemented with 10% Fetal Calf Serum (SAFC) and passaged every 3 days. Clonal cell lines of 293T cells carrying BxBI landing pad were obtained by single cell sorting (FACS Aria2) of cells expressing mRuby2, following transfection of parental cells with plasmids carrying CAS9 and 3 guide RNAs for hAAVS1(JPF0432) and landing pad (MTK0_057). Transfections for transient expression were done in quadruplicate, in 96 well (5.104 cells, 300ng total plasmid DNA) or 6 well plate (1.105 cells, 4ug total plasmid DNA) format with Lipofectamine 2000 (Invitrogen), according to manufacturer's instructions. CAS9, Piggybac and BxBI transfections were performed in triplicate with Lipofectamine

2000, according to the manufacturer's instructions. In summary, transfections were performed with 1:1 ratio of CAS9/transposase/recombinase (JPF0432 / pCMV-hyPBase / pCAG-NLS-HA-Bxb1) to transfer vector, for a total of 800ng of DNA per well in 24 well plates seeded with 5.104 cells. The next day following transfection, media was replaced with normal growth media with the appropriate antibiotic selection. Hygromycin (100ug/mL, Invivogen) and Blasticidin (10ug/mL, Neta Scientific) selection was performed for at least 7 days, with media changes every 48h. pCMV-hyPBase was a gift from Wendell Lim and pCAG-NLS-HA-Bxb1 was a gift from Pawel Pelczar (Addgene plasmid 51271 ; http://n2t.net/addgene:51271 ; RRID:Addgene_51271).

Transfections for Zaire ebolavirus RNP assays were performed in duplicate or triplicate using standard calcium phosphate transfection methods. Briefly, complexes were generated in 100 µL of 1X Hepes Buffered Saline (HBS) (Thermo AAJ62623AK) and 12.5 µL of 2M calcium chloride (Fisher Scientific 50995817) per 1 µg of DNA for 15 mins at room temperature prior to transfection of cells.

Lentiviral Production

For lentiviral generation, 24h before transfection, 5E5 HEK293T cells were plated in a 6 well plate containing 2mL of growth media. To prepare virus, transfections of 4ug total DNA of equimolar amounts of of pCMV-dR8.91, pCMV-VSV-G, and transfer vector were performed using Lipofectamine 2000, and following the manufacturer's instructions. Media was changed after 16h and virus were collected 48h after transfection, by filtering the supernatant through a 0.45µm filter (Millipore SLHV033RS). For transductions, 24 h before addition of virus target cells were seeded at a density of 1E5 in a 6 well plate and transduced with 1 mL, 100 µL or 10 µL of viral supernatant supplemented with 4ug/mL of polybrene (SCBT sc-134220). pMD2.G was a gift from Didier Trono (Addgene plasmid

12259 ; http://n2t.net/addgene:12259 ; RRID:Addgene_12259) and pCMV-dR8.91 was a gift from Wendell Lim

Flow Cytometry and Data Analysis

For the analysis of promoter, 3'UTR expression and hAAVS1 LP expression, cells were collected in 96 well plates (Corning) and measured using an LSR2 flow cytometer (BD) with the four laser configuration (488 nm, 635 nm, 355 nm, 405 nm). mAzamiGreen (excitation at 488 nm, emission at 530 nm), mRuby2 or mScarlet (excitation at 561 nm, emission between 610 and 620 nm) and tagBFP (excitation at 355 nm, emission at 450 nm) fluorescence levels were recorded for 10000 events. Gating of single cells, normalization of fluorescence levels, and statistical analysis was performed with custom python scripts (https://github.com/jpfon/MTK).

Microscopy and Image Processing

For the imaging of combinatorial assembly of the fluorescent protein and localization tags library, 5E4 HEK293T cells were plated in an 8-Well µ-Slide (Ibidi) that contains 200µL of growth media. After 24h and before imaging, growth media was replaced with 200µL of Fluorobrite DMEM (Gibco). Imaging was performed in a temperature and atmosphere controlled chamber on a Zeiss microscope equipped with a Yokagawa CSUX1-A1N-E confocal spinning disk. Images were collected with a 40 x 1.1 NA water immersion objective and Photometrics Evolve 512 EMCCD camera. Images were stitched (ZEN, Zeiss) and gamma-corrected (FIJI) for perception enhancement. For the imaging of multicistronic fluorescent proteins, 45E4 HEK293T cells were plated in 8-Well µ-Slide (Ibidi) containing 200µL of growth media. Imaging was performed in a temperature and atmosphere controlled chamber on a Nikon Ti Microscope equipped with a Andor Borealis

CSU-W1 confocal spinning disk. Images were collected through a 20x 0.75NA air objective, using an Andor 4 Laser Launch for tagBFP (excitation at 405nm, collection between 425 and 475nm), mAzamiGreen (excitation at 488nm, collection between 500 and 550nm), mScarlet (excitation at 561nm, collection between 590 and 650nm) and iRFP713 (excitation at 640nm, collection between 665 and 736nm). An Andor Zyla 4.2 sCMOS was used to detect the images and pixel size was 325nm.

Viral DNA sequences and Plasmids

Synthetic cDNA sequences for viral NP, VP35, VP30, and Lpol from Zaire ebolavirus Mayinga 1976 isolate (Accession: NC_002549, H.sapiens-tc/COD/1976/Yambuku-Mayinga) were codon optimized (NP, VP35, VP30) and synthesized by IDT, followed by Gibson assembly into the pCAGGs vector backbone. These yielded pCAGGs-NP, pCAGGs-VP35, pCAGGs-VP30, and pCAGGs-Lpol expression plasmids. In order to generate the 4-cistronic (ZEBOV-4cis), viral proteins were PCR amplified from pCAGGs expression vectors and domesticated into MTK0_027 entry vector using the BsmbI site followed by verification with Sanger sequencing. The T7-driven viral minigenome construct p2.0_3E5E_eGFP was a gift from Elke Mühlberger (Addgene plasmid 69359 ; http://n2t.net/addgene:69359; RRID_Addgene_69359), which was modified via NdeI and NotI restriction sites to generate p2.0-3E5E-nLuc. T7opt in pCAGGS was a gift from Benhur Lee (Addgene plasmid 65974; http://n2t.net/addgene:65974 ; RRID_Addgene_65974).

Compounds

Gedunin (CAS 2753-30-2) (Fisher Scientific 33871) and 6-Azauridine (Sigma A1882-1G) were resuspended in DMSO to generate a 1mM stock solution. Stock solutions were serially

diluted in DMSO and media to treat cells with a final concentration of 5µM in 1% DMSO.

Generation of ZEBOV-4cis stable cell populations and clones

HEK293T cells were seeded in 6-well plates at a density of 0.4E6 cells/well for 24 hours and then transfected with 2 µg of total DNA (1 µg of MTK043-ZEBOV-4cis construct and 1 µg of PiggyBac transposon expression construct (pCMV-hyPBase)) using calcium phosphate transfection. After 2 days cells were transferred to a 10-cm dish to begin selection with hygromycin B (0.4 mg/mL) (Fisher Scientific MIR5930). Complete selection was observed in 7 days resulting in ZEBOV-4cis expressing stable populations from which clones were isolated using limited dilution plating. Briefly, cells were plated into 96-well plates at a density of 0.5 cells/well in the presence of hygromycin for 10 days and subsequently expanded to yield stable clones.

3.10.1 Western blot and Immunofluorescence analyses

Cell lysates were prepared in RIPA lysis and extraction buffer (Thermo 89900) containing protease inhibitors (Sigma P8340) for Western blotting. Lysates were resolved by SDS-polyacrylamide gel electrophoresis (PAGE), transferred to a polyvinylidene difluoride (PVDF) membrane, and subjected to Western blotting using primary antibodies: rabbit polyclonal Zaire ebolavirus NP (IBT Bioservices 0301-045) (1:1000 dilution), mouse monoclonal Zaire ebolavirus VP35 (Kerafast EMS702) (1:1000 dilution), mouse monoclonal 2A peptide (Novus Biologicals NBP2-59627) (1:2000 dilution), Rabbit polyclonal GAPDH (Thermo Fisher Scientific PA1-987) (1:5000 dilution), and secondary antibodies: goat anti-rabbit or goat anti-mouse polyclonal IRDye-800CW antibodies (VWR, 1:5000 dilution). For immunofluorescence analysis, cells seeded in 12-well tissue culture plates for 24 h were fixed in 4% formaldehyde for 20 min, incubated in

permeabilization buffer (1% [vol/vol] Triton X-100 and 0.1% [wt/vol] sodium citrate in PBS) for 10 min and then in blocking buffer (1% [vol/vol] Triton X-100, 0.5% [vol/vol] Tween 20, and 3% bovine serum albumin in PBS) for 30 min. Cells were then incubated overnight with primary antibodies for Zaire ebolavirus VP35 or 2A peptide, secondary antibodies goat anti-mouse or goat anti-rabbit Alexa-488 for 30 min, and stained with DAPI solution (GeneTex GTX16206) for 10 mins. Cells were imaged at 10X (Leica light microscope).

3.10.2 Minigenome reporter assays (luciferase, GFP)

ZEBOV minigenome reporter construct (p2.0-3E5E-nLuc, 250 ng) and T7-expression plasmid (pCAGGs-T7opt, 250 ng) were co-transfected with four ZEBOV expression plasmids (pCAGGs-ZEBOV-NP, pCAGGs-ZEBOV-VP35, pCAGGs-ZEBOV-VP30, pCAGGs-ZEBOV-Lpol; 250 ng each) or multicistronic constructs (ZEBOV-4cis, 1000 ng), and pCAGGs empty vector for a total of 1500 ng DNA per 0.5E6 HEK293T cells using calcium phosphate transfection (see method above). Suspension transfected cells were seeded at a density of 20,000 cells/well into 8 wells of 96-well plates for 2 days in duplicates. Nano luciferase assays for minigenome function and cell titer glo assays for cell viability were performed as per manufacturer's instructions (Promega). For minigenome assays in ZEBOV-4cis stable cells, 500 ng each of p2.0-T7-3E5E-nLuc or p2.0-T7-3E5E-eGFP and 500ng of T7-expression plasmid were co-transfected. Nano luciferase levels were assayed as above and eGFP signal was captured via microscopy (Leica, 4X magnification).

3.10.3 Statistical analysis

Bar plots show mean of biological replicates and, when shown, error bars denote 95% confidence intervals of mean. Statistical analysis was done in Python (https://github.com/jpfon/MTK).

3.10.4 Data availability

All supporting data for the findings and plasmids presented in this study are available on github (https://github.com/jpfon/MTK).

3.11 Acknowledgements

The authors thank the members of the El-Samad lab and Joe DeRisi (UCSF) for helpful feedback, and John Dueber (UC Berkeley) for the gift of constructs used in this study. This work was supported by the National Science Foundation Award 1715108 to H.E-.S and the Defense Advanced Research Projects Agency [grant number HR0011-16-2-0045 to H.E-.S]. The content and information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred. H.E-.S is an investigator in the Chan Zuckerberg Biohub. This work was also supported by the National Defense Science Engineering Graduate (NDSEG) Fellowship awarded to A.R.B., A.H.N. and L.C.O. L.C.O. is also supported by the Paul and Daisy Soros Fellowship for New Americans and National Institute for General Medical Sciences (NIGMS) Initiative for Maximizing Student Development (IMSD) Fellowship. Research performed by G.R.K. and A.L.K. is funded by the Chan Zuckerberg Biohub.

3.12 Contributions

J.P.F., A.R.B., and H.E-.S. conceived of the study. A.L.K. and G.R.K. conceived of the virology application. J.P.F., A.R.B., G.R.K., A.H.N., J.T., Q.C.W., E.A., S.Y.C., G.D., P.H., and L.C.O. constructed parts library. J.P.F., A.R.B., and G.R.K. collected and processed data. J.P.F., A.R.B., G.R.K., A.L.K., and H.E-.S. interpreted results, wrote and edited the manuscript.

3.13 Figures

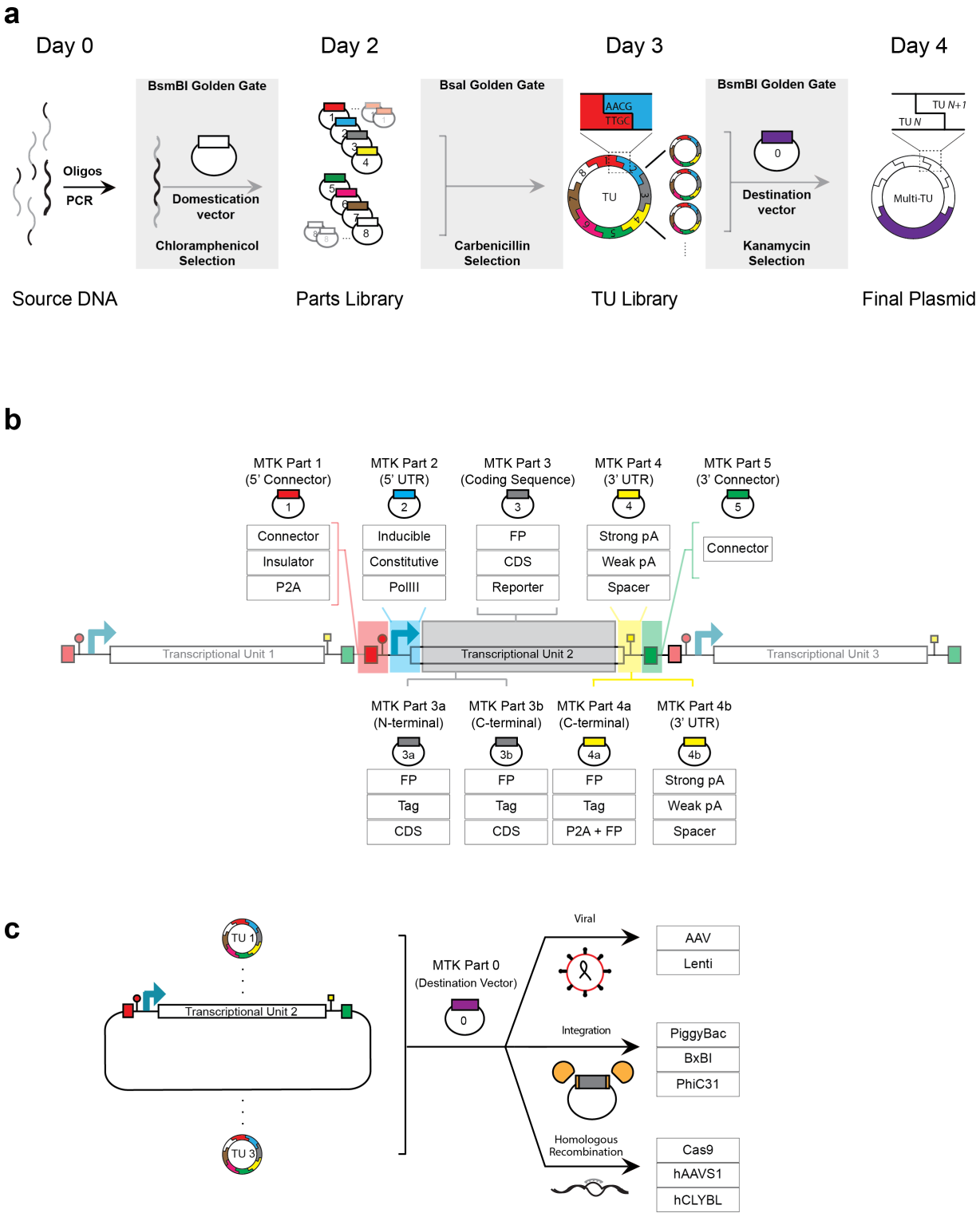


Figure 3.1: Schematic and definition of parts of the Mammalian Toolkit (MTK).

a. Workflow of the MTK starting with a BsmBI part domestication, BsaI transcriptional unit assembly, and a final BsmBI assembly of a multi-transcriptional unit plasmid. Top text indicates approximate time from initial PCR with primers to receiving final plasmid. **b**. Schematic of a standard transcriptional unit (TU). Tables below part definition summarize variations of that part that are present in the MTK. Supplemental table 1 contains a comprehensive list of current parts. **c**. Library of TUs can be re-used with no further modification and delivered to cells in three methods already present in the MTK using a dedicated destination vector (MTK part 0).

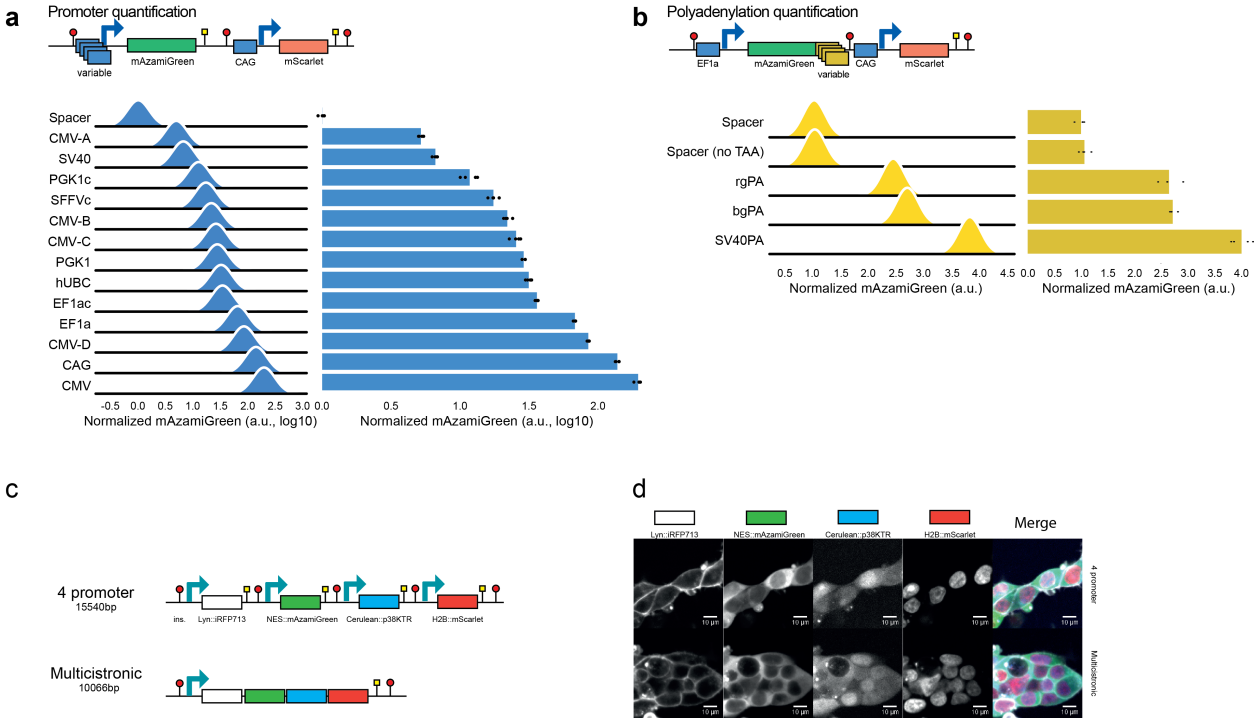


Figure 3.2: Characterization of constitutive promoters,3' UTR and multicistronic connectors provided in the MTK.

a. Top: Schematic of a TU used to quantify promoter expression. Bottom: Ranked expression of mAzamiGreen (normalized to a constitutive mScarlet) driven by different promoters. Distributions show one of four biological replicates, and bar plots represent the mean of all four biological replicates. The mean of each replicate is also shown as a black dot. The letter "c" following the name of a promoter (e.g PGK1 versus PGK1c) is used to designate a "crippled" promoter where the Kozak sequence is disrupted. **b** Top: Schematic of TU used to quantify the effect of the 3' UTR on mAzamiGreen expression driven by the EF1a promoter. Data representation is as described above. **c** Schematic of expression strategies of 4 fluorescent proteins targeted to 4 different cellular locations. In the first strategy, a plasmid contains 4 TUs, each containing a CAG promoter. In the first TU, CAG drives expression of membrane-targeted iRFP713 (Lyn::iRFP713), in the second expression of cytoplasmic mAzamiGreen (NES::mAzamiGreen), in the third expression of p38 kinase translocation reporter fused to mCerulean (Cerulean::p38KTR), and in the fourth histone H2B fused to mScarlet (H2B::mScarlet). In the second strategy, a multicistronic plasmid encodes the same proteins, but all are produced from a single transcript driven by a CAG promoter with 3 P2A sequences separating the 4 peptides. The size in bp of each plasmid is also indicated. **d**Confocal images of HEK293T cells expressing the 4 promoter (top) or multicistronic (bottom) plasmid. Merge panel shows Lyn::iRFP713 in white, NES::mAzamiGreen in green, Cerulean::p38KTR in blue and H2B::mScarlet in red.

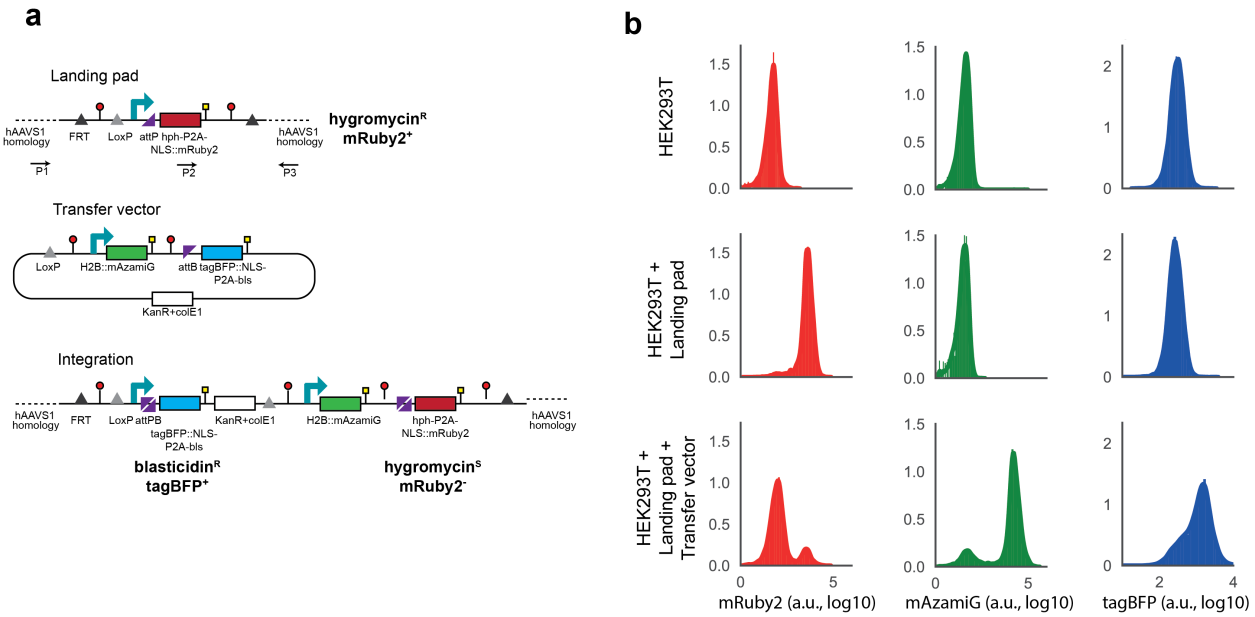


Figure 3.3: Generation and testing of landing pads for HEK293T cells using the MTK.

a. Schematic of landing pad, transfer vector and integrated vector. Also shown are genotyping primers. FRT, flippase recognition target; ins., Insulator; LoxP; LoxP site; attP, BxBI phage attachment site; hph, hygromycin resistance gene, bls, blasticidin resistance gene; NLS, nuclear localization signal; KanR, kanamycin resistance gene; colE1, colE1 origin of replication; attB, BxBI bacterial attachment site; mAzamiG, mAzamiGreen. **b** mRuby2, mAzamiGreen and tagBFP expression in populations of parental, Landing pad and Landing Pad with Transfer vector HEK293T cells. mRuby2 expression indicates presence of hAAVS1 landing pad. mAzamiGreen and tagBFP expression indicates precise integration of transfer vector in hAAVS1 landing pad.

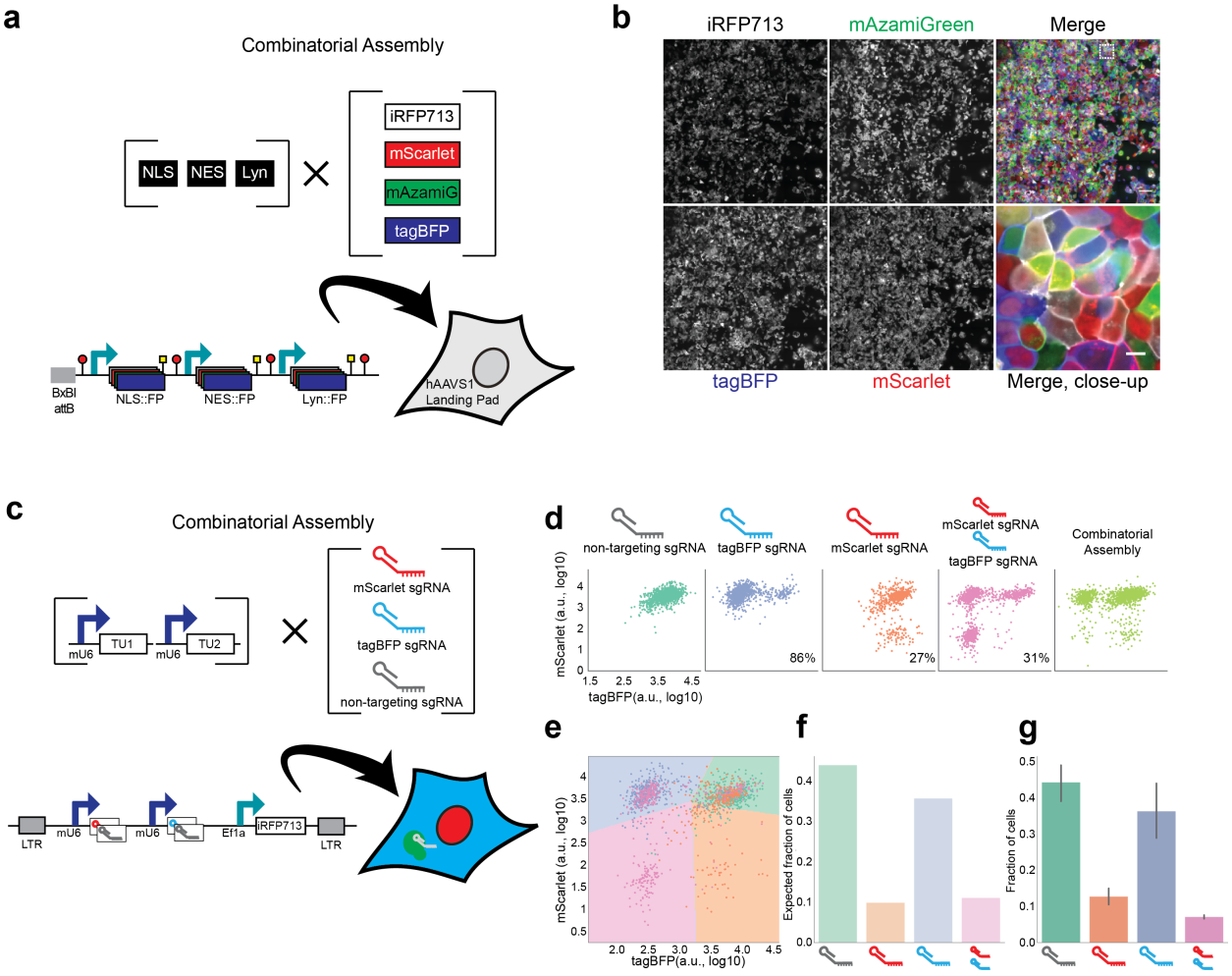


Figure 3.4: Combinatorial assembly of complex libraries using the MTK.

a Overview of the MTK strategy for a combinatorial, one pot, generation of a library containing 12 combinations of 4 different fluorescent proteins (FP) targeted to three cellular locations. The final multi-TU construct is composed of 3 TUs: EF1a driving the nuclear localization of an FP; CMV expressing the cytoplasmic localization of an FP; and CAG producing a membrane-tethered FP. Each FP is one of 4 variable fluorophores (tagBFP, mAzamiGreen, mScarlet and iRFP713), giving a total of 64 possible variants. Constructs are delivered to the landing pad in HEK293T cells. **b** Confocal image of HEK293T+LP, transfected with pooled library in panel (a). iRFP713 shown in white, mScarlet shown in red, mAzamiGreen shown in green and tagBFP shown in blue. Merge and close-up show cells expressing 3 FPs in three subcellular locations displaying the variety of expected combinations. **c** Overview of the MTK strategy for a combinatorial, one pot, library of lentiviral vectors carrying sgRNAs targeting tagBFP, mScarlet, both or non-targeting. The final multi-TU construct is composed of three TUs: mU6 promoter driving the expression of mScarlet or non-targeting sgRNA; mU6 promoter driving the expression of tagBFP or non-targeting sgRNA; EF1a driving the expression of iRFP713 for identification of cells that have integrated the construct. Library was produced and transduced to HEK293T cells expressing tagBFP, mScarlet and Cas9 fused to

mAzamiGreen (3C cells). **d** Scatter plots of tagBFP and mScarlet fluorescence in populations of 3C cells where sgRNAs were individually expressed to target tagBFP, mScarlet, tagBFP and mScarlet, or with non-targeting sgRNA. Last panel shows the scatter plots of tagBFP and mScarlet fluorescence in populations transfected with combinatorially assembled library of sgRNAs. **e** Scatter plot of tagBFP and mScarlet fluorescence in populations of 3C cells, where an equal number of cells expresses sgRNAs that target tagBFP, mScarlet, tagBFP and mScarlet, or with non-targeting sgRNA. Shaded areas correspond to 4 classes identified by a linear classifier (green, non-targeting; blue, tagBFP; orange, mScarlet; pink, mScarlet and tagBFP). **f** Expected fraction of cells expressing different guide combinations as identified by linear classifier in panel (e) where 4 sgRNA combinations have equal ratios. **g** Measured fraction of cells expressing different guide combinations in combinatorial assembly is similar to (f). Bar plot and error bars represents mean and 95% CI of four biological repeats. NLS, Nuclear localization signal; NES, nuclear export signal; Lyn, plasma membrane tag; mAzamiG, mAzamiGreen.

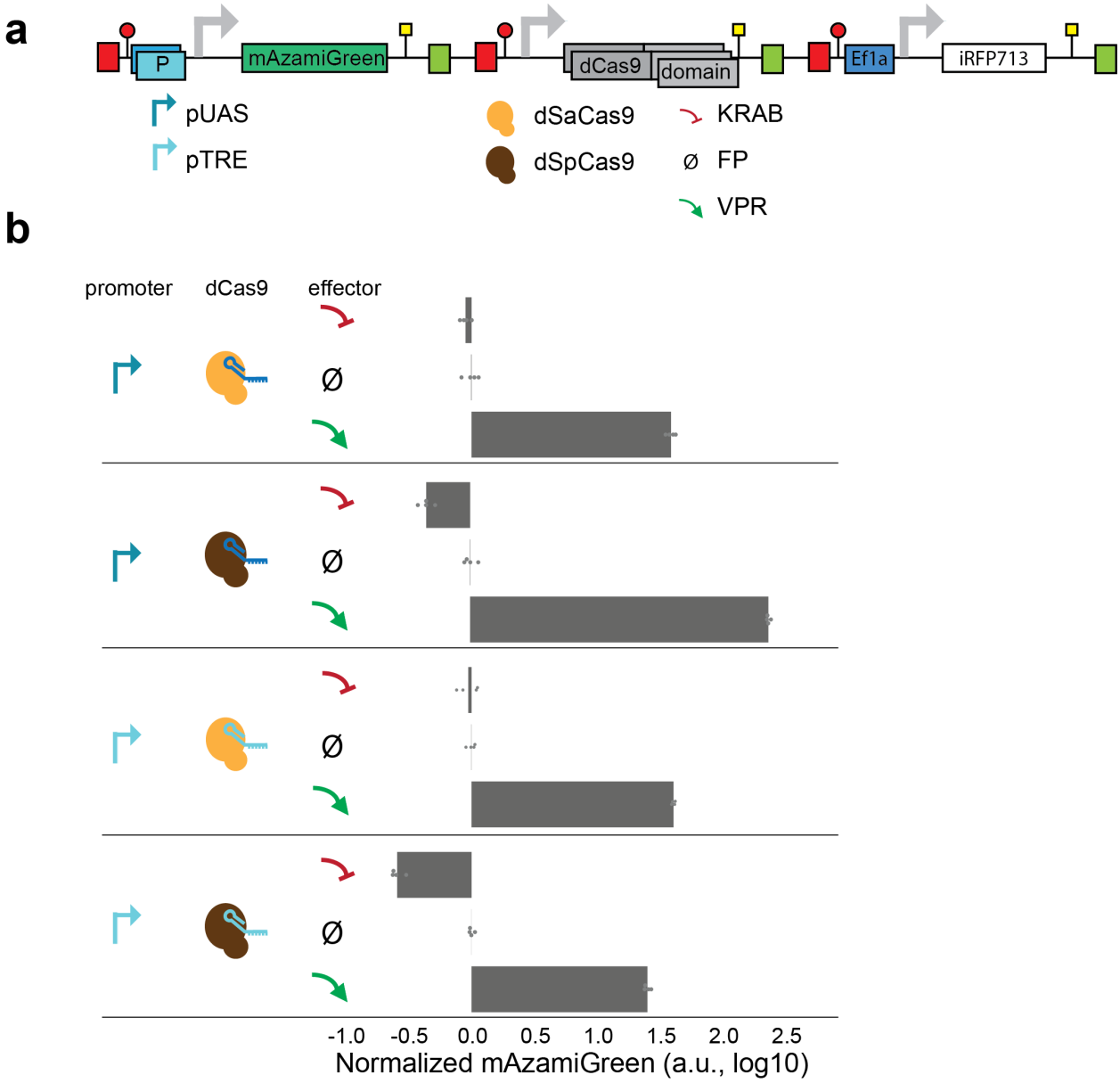


Figure 3.5: Parallelization of dCas9 circuit prototyping using the MTK.

a Schematic of 12 dCas9 circuit variations assembled in parallel. Each variation is a multi-TU plasmid with either a TRE or UAS inducible promoter controlling the expression of mAzamiGreen, followed by EF1a driving either dSpCas9 or dSaCas9 fused to a KRAB domain, fluorophore, or VPR domain. UAS, Upstream Activating Sequence.TRE, Tetracycline Responsive Element. dSpCas9 and dSaCas9, deactivated S. pyogenes and S. aureus Cas9, respectively. KRAB, Krueppel-associated box. FP, fluorescent protein. VPR, VP64-p65-Rta. **b** mAzamiGreen expression (normalized by iRFP713) for different configurations of the circuit. Every row is a different configuration, corresponding to a variation of the promoter, dCas9 used, and effector used. Configurations are grouped by their promoter-dCas9 pairing (different groups are separated by solid lines). Bar plot represents mean of 4 biological replicates, the mean of each is shown as a black dot.

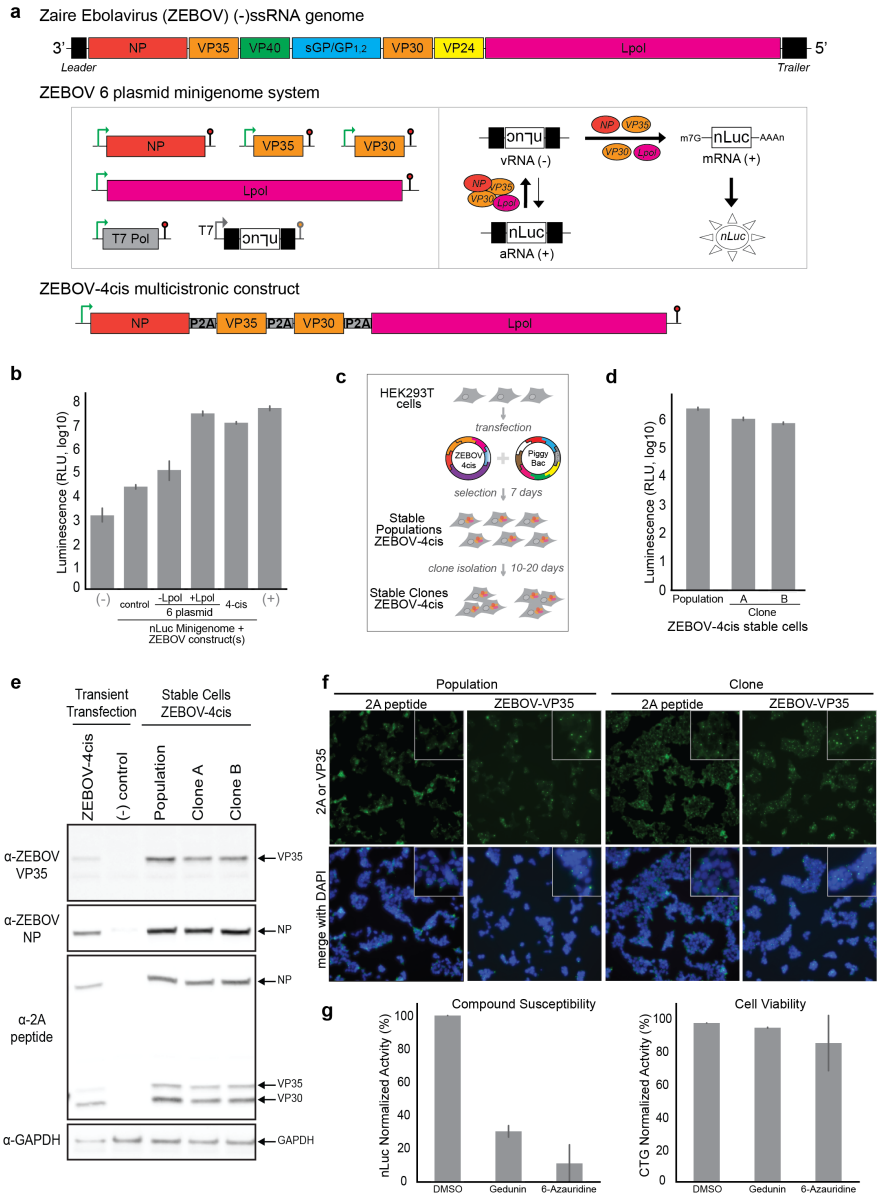


Figure 3.6: Generating multicistronic constructs for Zaire ebolavirus ribonucleoproteins in a mammalian cell host using the MTK.

a Schematic of the different constructs. Panel 1: Schematic of the Zaire ebolavirus (ZEBOV) negative sense single stranded RNA (-ssRNA) genome. Panel 2: Functional minigenome assay for ZEBOV using the 6 plasmid system that includes expression constructs for NP, VP35, VP30, and Lpol viral proteins as well as the T7 polymerase along with a T7-driven minigenome reporter construct flanked by viral non coding 3' and 5' UTRs, the leader and trailer, respectively. Transfection of the 6 plasmids into mammalian cells results in expression of NP, VP35, VP30, Lpol, and T7 and transcription of the minigenome viral RNA (vRNA(-)). The vRNA(-) is replicated to generate antigenomic RNA (aRNA(+)) and transcribed into mRNA by the viral proteins to yield reporter gene expression. Panel 3: Schematic of the Zaire ebolavirus multicistronic construct

(ZEBOV-4cis) with NP, VP35, VP30, and Lpol separated by P2A ribosomal skipping site elements. **b** Luminescence measurements of ZEBOV minigenome activity. HEK293T cells were transfected with ZEBOV nano luciferase (nLuc) minigenome in combination with pCAGGs empty control plasmid (control), the ZEBOV 6 plasmids (with or without Lpol: +Lpol, -Lpol), or with ZEBOV-4cis in PiggyBac part 0 vector. Positive (+) and negative (-) controls for nLuc expression levels include transfection of only pCAGGS-nLuc plasmid or pCAGGs empty plasmid, respectively. Nano luciferase activity was measured two days post transfection. Bar plots represent the mean of biological replicates (n=2). **c** Schematic of ZEBOV-4cis stable cell line generation using PiggyBac transposon. HEK293T cells were co-transfected with MTK043-ZEBOV-4cis and a PiggyBac expression construct for 3 days and selected with hygromycin for 7 days to generate stable cell populations expressing ZEBOV RNP complex proteins. Clones were isolated from these populations via limited dilution plating in additional 10-20 days. **d** Luminescence measurements of minigenome activity in ZEBOV-4cis stable population. ZEBOV-4cis stable cells and clones were transfected with a T7-driven ZEBOV minigenome construct encoding an nLuc reporter along with T7 polymerase for 2 days followed by nano luciferase assay. Bar plots represent mean of technical replicates (n=10). **e** Western blot confirmation of protein expression. ZEBOV-4cis stable population and clones as well as HEK293T cells transfected with MTK0-43-ZEBOV-4cis were processed for Western blot analysis with mouse anti-ZEBOV VP35, rabbit anti-ZEBOV-NP, mouse anti-2A peptide, and rabbit anti-GAPDH antibodies. **f** Immunofluorescence analysis of viral protein localization. ZEBOV-4cis population and clone were stained with with anti-2A peptide and anti-ZEBOV-VP35 primary antibodies and Alexa-488 secondary antibody. DAPI stained nuclei are shown in the merged images. Insets represent 2X magnified fields. **g** Effect of chemical compound inhibitors on minigenome activity in ZEBOV-4cis stable cells. ZEBOV-4cis stable population was transfected with a ZEBOV minigenome construct encoding a nano luciferase reporter and treated with DMSO (1%), Gedunin (5uM), or 6-Azauridine (5uM). After 2 days cells were processed for nano luciferase assay for functional minigenome activity and cell titer glo assays for cell viability. Levels of signal relative to DMSO control is plotted. Bar plots represent mean of biological replicates (n=2).

3.14 Supplemental Data Figures

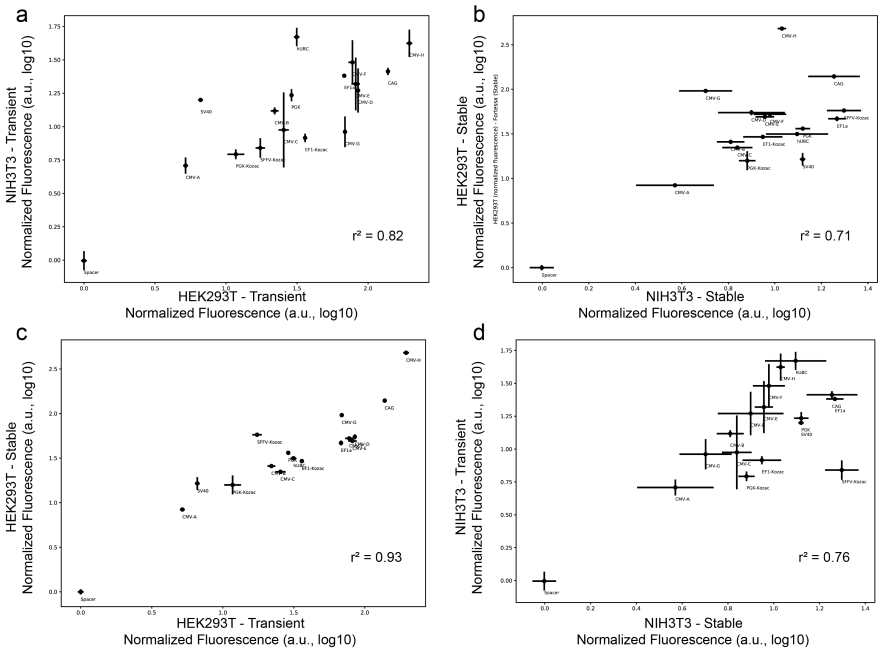


Figure 3.7: Comparison of constitutive promoters across cell lines and delivery methods.

a The mAzamiGreen expression from transient transfection and PiggyBac integration of each promoter was assessed in HEK293T and **b** NIH3T3 cells. **c** mAzamiGreen expression of each promoter was compared between HEK293T and NIN3T3 in transient expression and **d** PiggyBac integration. Each point represents the mean of four biological replicates and error bars represent the standard deviation across replicates.

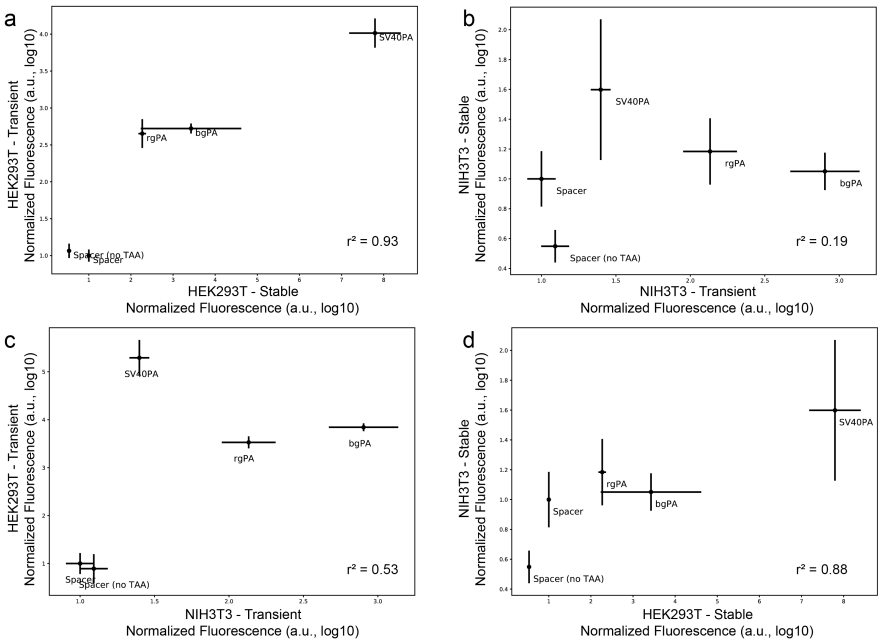


Figure 3.8: Comparison of impact of 3' UTRs across cell lines and delivery methods.

a The mAzamiGreen expression from transient transfection and PiggyBac integration delivery of each 3' UTR was assessed in HEK293T and **b** NIH3T3 cells. **c** mAzamiGreen expression of each 3' UTR was compared between HEK293T and NIN3T3 in transient expression and **d** PiggyBac integration. Each point represents the mean of four biological replicates and error bars represent the standard deviation across replicates.

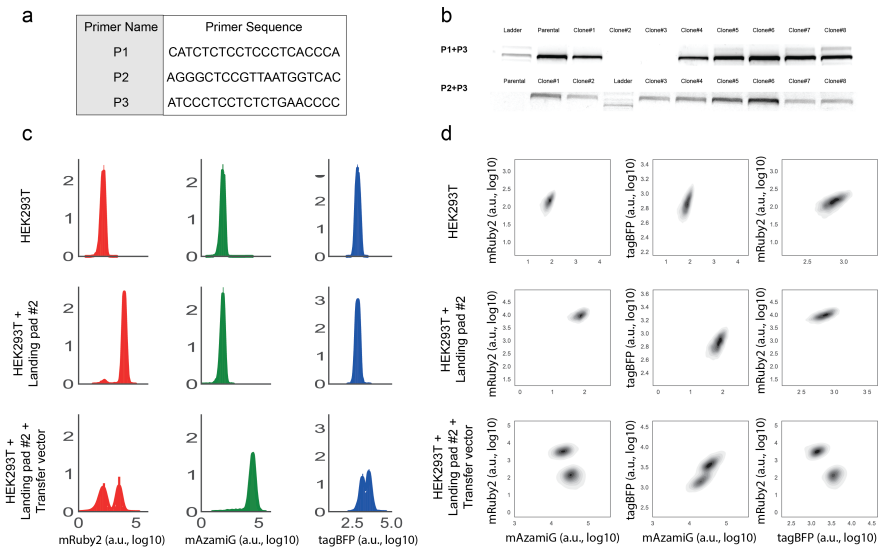


Figure 3.9: Generation of landing pads for human cell lines.

a PCR products from landing pad genotyping. P1+P3 indicate presence of WT hAVVS1 locus and P2+P3 indicate presence of BxBI landing pad in hAAVS1 locus. mRuby2, mAzamiGreen and tagBFP expression in populations of parental, Landing pad and Landing Pad with Transfer vector HEK293T cells. **b** mRuby2, mAzamiGreen and tagBFP expression in populations of parental, Landing pad 2 and Landing Pad 2 with Transfer vector HEK293T cells. In this clone, both wild type alleles of hAAVS1 locus were replaced by the landing pad construct, showing two populations as measured by fluorescence upon integration of the transfer vector and suggesting these two populations had one or two copies that integrated into the genome. mRuby2 expression indicates presence of hAAVS1 landing pad. mAzamiGreen and tagBFP expression indicates precise integration of transfer vector in hAAVS1 landing pad.

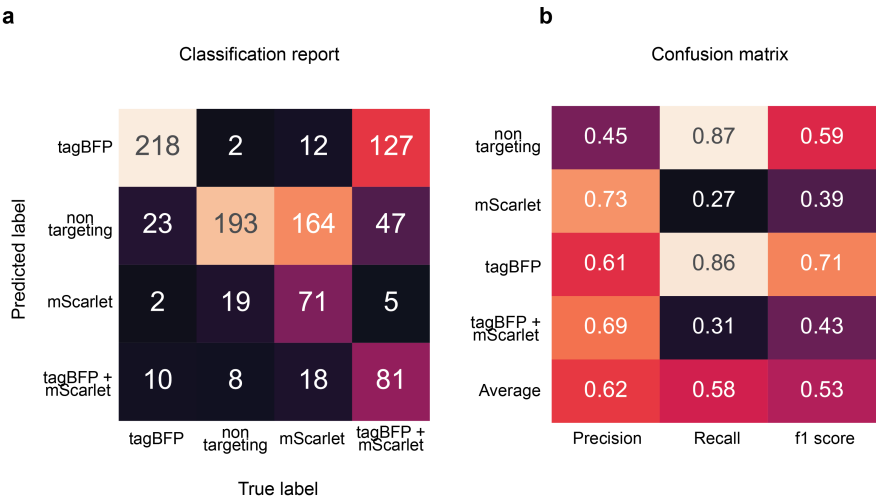


Figure 3.10: Building a linear classifier to distinguish target sgRNA knockdown populations. **a** Classification report of linear classifier. **b**. Confusion matrix of linear classifier.

a

Donor Vector	Donor Name	ZEBOV-4cis Colonies Screened	ZEBOV-4cis Colonies Correct
MTK0-43	PiggyBac Transposon	9	9
MTK0-13	PhiC31 Integrase	4	4
MTK0-14	BxBI Integrase	4	3
MTK0-16	PhiC31 tagBFP	4	4
MTK0-17	BxBI tagBFP	4	4

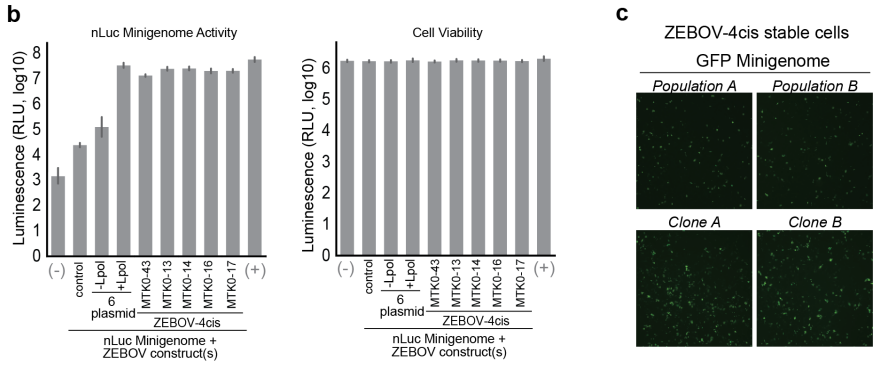


Figure 3.11: Generation and quality control of multicistronic constructs for Zaire ebolavirus ribonucleoproteins.

a Multicistronic construct cloning efficiency. ZEBOV multicistronic construct containing 4 viral ORFs separated by P2A elements (ZEBOV-4cis) was BsmBI assembled directly into donor vectors for genome engineering. Specifically, ZEBOV-4cis was generated in donor vectors for PiggyBac transposon (MTK0-43), PhiC31 Integrase (MTK0-13, MTK0-16 tagBFP), and BxBI Integrase (MTK0-14, MTK0-17 tagBFP). The number of bacterial colonies screened and positive for correct construct by size (NotI digestions) is indicated. **b**. Luminescence measurements of ZEBOV minigenome activity and cell viability. HEK293T cells were transfected with ZEBOV nLuc minigenome in combination with pCAGGs empty control plasmid (control), the ZEBOV 6 plasmids (with or without Lpol: +Lpol, -Lpol), or with ZEBOV-4cis in various part 0 donor vectors. Positive (+) and negative (-) controls include transfection of only pCAGGS-nLuc plasmid or pCAGGs empty plasmid, respectively. Nano luciferase activity was measured two days post transfection. Bar plots represent the mean of biological replicates (n=2).**c** GFP analysis of minigenome activity in stable cells. ZEBOV-4cis stable populations and clones were transfected with a T7-driven ZEBOV minigenome construct encoding the eGFP reporter along with T7 polymerase. After 2 days cells were imaged for GFP detection (Leica, 4X).

3.14.1 Supplementary Information

Table 3.1: Cas9 sgRNAs use in this study

Figure	Target Locus	Protospacer Sequence	Cas9 Species (Part Plasmid)	PAM Sequence	Source
3	hAAVS1 (site 1)	GAGCCACATTAACCGGCCCT	S. pyogenes (MTK3_025)	GGG	This study
3	hAAVS1 (site 2)	ATTCCCAGGGCCGGTTAATG	S. pyogenes (MTK3_025)	TGG	This study
3	hAAVS1 (site 3)	GGGGCCACTAGGGACAGGAT	S. pyogenes (MTK3_025)	TGG	This study
4	tagBFP	CTACAACGTCAAGATCAGAG	S. pyogenes (MTK3_025)	GGG	This study
4	mScarlet	CCACAACGAAGATTATACCG	S. pyogenes (MTK3_025)	TGG	This study
4	human non-targeting	ACGGAGGCTAAGCGTCGCAA	S. pyogenes (MTK3_025)		Doench, J. G. et al. (2016)
5	TRE	TACGTTCTCTATCACTGATA	S. pyogenes (MTK3_025)	GGG	Gao, Y. et al. (2016)
5	TRE	GTTACTCCCTATCAGTGATA	S. aureus (MTK3b_019)	AGGAGT	Gao, Y. et al. (2016)
5	UAS	GAGCACTGTCCTCCGAACGT	S. pyogenes (MTK3_025)	CGG	This study
5	UAS	GAACGTCGGAGCACTGTCCT	S. aureus (MTK3b_019)	CCGAAC	This study

Table 3.2: Cas9 sgRNAs use in this study (continued)

Target Locus	Protospacer Sequence	PAM Sequence	Forward Oligo	Reverse Oligo	Backbone Vector
hAAVS1 (site 1)	GAGCCACATTAACCGGCCCT	GGG	5' TGTTTGGAGCCACATTAACCGGCCCTG	5' TAAACAGGGCCGGTTAATGTGGCTCCA	MTK234_002
hAAVS1 (site 2)	ATTCCCAGGGCCGGTTAATG	TGG	5' TGTTTGATTCCCAGGGCCGGTTAATGG	5' TAAACCATTAACCGGCCCTGGGAATCA	MTK234_002
hAAVS1 (site 3)	GGGGCCACTAGGGACAGGAT	TGG	5' TGTTTGGGGGCCACTAGGGACAGGATG	5' TAAACATCCTGTCCCTAGTGGCCCCCA	MTK234_002
tagBFP	CTACAACGTCAAGATCAGAG	GGG	5' TGTTTGCTACAACGTCAAGATCAGAGG	5' TAAACCTCTGATCTTGACGTTGTAGCA	MTK234_002
mScarlet	CCACAACGAAGATTATACCG	TGG	5' TGTTTGCCACAACGAAGATTATACCGG	5' TAAACCGGTATAATCTTCGTTGTGGCA	MTK234_002
human non-targeting	ACGGAGGCTAAGCGTCGCAA		5' TGTTTGACGGAGGCTAAGCGTCGCAAG	5' TAAACTTGCGACGCTTAGCCTCCGTCA	MTK234_002
TRE	TACGTTCTCTATCACTGATA	GGG	5' TGTTTGTACGTTCTCTATCACTGATAG	5' TAAACTATCAGTGATAGAGAACGTACA	MTK234_002
TRE	GTTACTCCCTATCAGTGATA	AGGAGT	5' TGTTTGGTTACTCCCTATCAGTGATAG	5' ATAACTATCACTGATAGGGAGTAACCA	MTK234_050
UAS	GAGCACTGTCCTCCGAACGT	CGG	5' TGTTTGGAGCACTGTCCTCCGAACGTG	5' TAAACACGTTCGGAGGACAGTGCTCCA	MTK234_002
UAS	GAACGTCGGAGCACTGTCCT	CCGAAC	5' TGTTTGGAACGTCGGAGCACTGTCCTG	5' ATAACAGGACAGTGCTCCGACGTTCCA	MTK234_050

Table 3.3: Primers for MTK part domestication

Part	Forward Primer	Reverse Primer	Destination Vector
1	5' gcatCGTCTCatcGGTCTCaCCCTCGTCTCaNNNN	5' gcatCGTCTCaGGTCTCaCGTT	MTK0_027
1 (reverse connectors)	5' gcatCGTCTCatcGGTCTCaCCCT	5' gcatCGTCTCaGGTCTCaCGTTCGTCTCtNNNN	$MTK0_0 27$
2	5' gcatCGTCTCatcGGTCTCaAACG	5' gcatCGTCTCaGGTCTCaCATA	MTK0_027
3	5' gcatCGTCTCatcGGTCTCaTATG	5' gcatCGTCTCaGGTCTCaGGAT	MTK0_027
3a	5' gcatCGTCTCatcGGTCTCaTATG	5' gcatCGTCTCaGGTCTCaAGAA	MTK0_027
3b	5' gcatCGTCTCatcGGTCTCaTTCT	5' gcatCGTCTCaGGTCTCaGGAT	MTK0_027
4	5' gcatCGTCTCatcGGTCTCaATCCtaa	5' gcatCGTCTCaGGTCTCaCAGC	MTK0_027
4a	5' gcatCGTCTCatcGGTCTCaATCC	5' gcatCGTCTCaGGTCTCaGCCA	MTK0_027
4b	5' gcatCGTCTCatcGGTCTCaTCGG	5' gcatCGTCTCaGGTCTCaCAGC	MTK0_027
5	5' gcatCGTCTCatcGGTCTCaGCTG	5' gcatCGTCTCaGGTCTCaTGTACGTCTCtNNNN	MTK0_027
5 (reverse connectors)	5' gcatCGTCTCatcGGTCTCaGCTGCGTCTCaNNNN	5' gcatCGTCTCaGGTCTCaTGTA	MTK0_027
6	5' gcatCGTCTCatcGGTCTCaTACA	5' gcatCGTCTCaGGTCTCaACTC	MTK0_027
7	5' gcatCGTCTCatcGGTCTCaGAGT	5' gcatCGTCTCaGGTCTCaTCGG	MTK0_027
8	5' gcatCGTCTCatcGGTCTCaCCGA	5' gcatCGTCTCaGGTCTCaAGGG	MTK0_027
8a	5' gcatCGTCTCatcGGTCTCaCCGA	5' gcatCGTCTCaGGTCTCaATTG	MTK0_027
8b	5' gcatCGTCTCatcGGTCTCaCAAT	5' gcatCGTCTCaGGTCTCaAGGG	MTK0_027
234	5' gcatCGTCTCatcGGTCTCaAACG	5' gcatCGTCTCaGGTCTCaCAGC	MTK0_027
234 (SpCas9 sgRNA)	5' TGTTTGNNNNNNNNNNNNNNNNNNNG	5' TAAACN'N'N'N'N'N'N'N'N'N'N'N'N'N'N'N'N'N'CA	MTK234_002
234 (SaCas9 sgRNA)	5 TGTTTGNNNNNNNNNNNNNNNNNNG	5' ATAACN'N'N'N'N'N'N'N'N'N'N'N'N'N'N'N'N'N'CA	MTK234_050

Table 3.4: Primers for MTK part domestication (continued)

Part	Forward Primer	Notes
1	5' gcatCGTCTCatcGGTCTCaCCCTCGTCTCaNNNN	NNNN is barcode for connector, see supplemental table 4
1 (reverse connectors)	5' gcatCGTCTCatcGGTCTCaCCCT	NNNN is barcode for connector, see supplemental table 4
2	5' gcatCGTCTCatcGGTCTCaAACG	
3	5' gcatCGTCTCatcGGTCTCaTATG	
3a	5' gcatCGTCTCatcGGTCTCaTATG	
3b	5' gcatCGTCTCatcGGTCTCaTTCT	
4	5' gcatCGTCTCatcGGTCTCaATCCtaa	
4a	5' gcatCGTCTCatcGGTCTCaATCC	remove first 3 bases (taa) for multicistronic use
4b	5' gcatCGTCTCatcGGTCTCaTCGG	NNNN is barcode for connector, see supplemental table 4
5 (reverse connectors)	5' gcatCGTCTCatcGGTCTCaGCTGCGTCTCaNNNN	NNNN is barcode for connector, see supplemental table 4
6	5' gcatCGTCTCatcGGTCTCaTACA	
7	5' gcatCGTCTCatcGGTCTCaGAGT	
8	5' gcatCGTCTCatcGGTCTCaCCGA	
8a	5' gcatCGTCTCatcGGTCTCaCCGA	
8b	5' gcatCGTCTCatcGGTCTCaCAAT	
234	5' gcatCGTCTCatcGGTCTCaAACG	
234 (SpCas9 sgRNA)	5' TGTTTGNNNNNNNNNNNNNNNNNNNG	N(20) are guide specific sequence, N' is reverse complement
234 (SaCas9 sgRNA)	5 TGTTTGNNNNNNNNNNNNNNNNNNG	N(20) are guide specific sequence, N' is reverse complement

Table 3.5: Connector part overhangs

Connector	Overhang
LS, LS'	5' CTGA
L1, R1	5' CCAA
L2, R2	5' GATG
L3, R3	5' GTTC
L4, R4	5' GGTA
L5, R5	5' AAGT
L6, R6	5' CCCT
L7, R7	5' GCGG
L8, R8	5' TTTA
RE, RE'	5' AGCA

Bibliography

[1] D. B. Berry and A. P. Gasch. Stress-activated genomic expression changes serve a preparative role for impending stress in yeast. *Molecular Biology of the Cell*, 19, 2008.

[2] A. P. Gasch et al. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell*, 11, 2000.

[3] S. Hohmann and W. H. Yeast Stress Responses Mager. *(Springer Science Business Media*. 2007.

[4] N. Hao and E. K. O'Shea. Signal-dependent dynamics of transcription factor translocation controls gene expression. *Nat. Struct. Mol. Biol*, 19, 2011.

[5] Y. Lin, C. H. Sohn, C. K. Dalal, L. Cai, and M. B. Elowitz. *Combinatorial gene regulation by modulation of relative pulse timing.* 2015.

[6] D. J. Lew and S. I. Reed. A cell cycle checkpoint monitors cell morphogenesis in budding yeast. *J. Cell Biol*, 129, 1995.

[7] M. Moreno-Torres, M. Jaquenoud, De Virgilio, and C. Torc1 controls G1–s cell cycle transition in yeast via Mpk1 and the greatwall kinase pathway. *Nature Communications*, 6, 2015.

[8] A. Rowley, G. C. Johnston, B. Butler, M. Werner-Washburne, and R. A. Singer. Heat

shock-mediated cell cycle blockage and G1 cyclin expression in the yeast Saccharomyces cerevisiae. *Mol. Cell. Biol*, 13, 1993.

[9] G. Bellí, E. Garí, M. Aldea, and E. Herrero. Osmotic stress causes a G1 cell cycle delay and downregulation of Cln3/cdc28 activity in Saccharomyces cerevisiae. *Mol. Microbiol*, 39, 2001.

[10] K. Yano et al. *Mih1/Cdc25 is negatively regulated by Pkc1 in Saccharomyces cerevisiae.* 2013.

[11] X. Escoté, M. Zapater, J. Clotet, and F. Posas. Hog1 mediates cell-cycle arrest in G1 phase by the dual targeting of Sic1. *Nat. Cell Biol*, 6, 2004.

[12] J. Clotet et al. Phosphorylation of Hsl1 by Hog1 leads to a G2 arrest essential for cell survival at high osmolarity. *EMBO J*, 25, 2006.

[13] Y. H. Ho, E. Shishkova, J. Hose, J. J. Coon, and A. P. Gasch. Decoupling Yeast Cell Division and Stress Defense Implicates mRNA Repression in Translational Reallocation during Stress. *Curr. Biol*, 28, 2018.

[14] Y. L. Chang. et al. *Yeast Cip1 is activated by environmental stress to inhibit Cdk1-G1 cyclins via Mcm1 and Msn2/4. Nat. Commun. 8*, 8(56), 2017.

[15] M. J. Brauer et al. Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. *Mol. Biol. Cell*, 19, 2008.

[16] C. Lu, M. J. Brauer, and D. Botstein. Slow growth induces heat-shock resistance in normal and respiratory-deficient yeast. *Mol. Biol. Cell*, 20, 2009.

[17] K. Kochanowski, L. Morinishi, S. Altschuler, and L. Wu. *Drug persistence - from antibiotics to cancer therapies.* 2018.

[18] J. Brewster, T. de Valoir, N. Dwyer, E. Winter, and M. Gustin. An osmosensing signal transduction pathway in yeast. *Science*, 259, 1993.

[19] Y. Goulev et al. *Nonlinear feedback drives homeostatic plasticity in HO stress response*. Elife 6, 2017.

[20] S. M. O'Rourke and I. Herskowitz. Unique and redundant roles for HOG MAPK pathway components as revealed by whole-genome expression analysis. *Mol. Biol. Cell*, 15, 2004.

[21] H. Saito and F. Posas. *Response to hyperosmotic stress*. 2012.

[22] S. Hohmann. Osmotic stress signaling and osmoadaptation in yeasts. *Microbiol. Mol. Biol. Rev*, 66, 2002.

[23] D. Muzzey, C. A. Gómez-Uribe, J. T. Mettetal, and A. van Oudenaarden. *A systems-level analysis of perfect adaptation in yeast osmoregulation*. 2009.

[24] J. T. Mettetal, D. Muzzey, C. Gómez-Uribe, and A. van Oudenaarden. *The frequency dependence of osmo-adaptation in Saccharomyces cerevisiae*. 2008.

[25] A. A. Granados et al. *Distributing tasks via multiple input pathways increases cellular survival in stress*. Elife 6, 2017.

[26] D. E. Quain and C. A. Boulton. Growth and metabolism of mannitol by strains of Saccharomyces cerevisiae. *J. Gen. Microbiol*, 133, 1987.

[27] S. P. Banavar et al. Mechanical feedback coordinates cell wall expansion and assembly in yeast mating morphogenesis. *PLoS Comput. Biol*, 14, 2018.

[28] L. M. Blank and U. Sauer. *TCA cycle activity in Saccharomyces cerevisiae is a function of the environmentally determined specific growth and glucose uptake rates*. 2004.

[29] C. A. Suarez-Mendez et al. *Interaction of storage carbohydrates and other cyclic fluxes with central metabolism: A quantitative approach by non-stationary C metabolic flux analysis*. 2016.

[30] A. Blomberg. Metabolic surprises in Saccharomyces cerevisiae during adaptation to saline conditions: questions, some answers and a model. *FEMS Microbiology Letters*, 182, 2000.

[31] S. Özsezen et al. *Inference of the High-Level Interaction Topology between the Metabolic and Cell-Cycle Oscillators from Single-Cell Dynamics.* Cell Syst 9, 354–365, 2019.

[32] B. P. Tu, A. Kudlicki, M. Rowicka, and S. L. McKnight. *Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes.* 2005.

[33] J. C. Ewald, A. Kuehne, N. Zamboni, and J. M. Skotheim. The yeast cyclin-dependent kinase routes carbon fluxes to fuel cell cycle progression. *Mol. Cell*, 62, 2016.

[34] G. Zhao, Y. Chen, L. Carey, and B. Futcher. Cyclin-dependent kinase co-ordinates carbohydrate metabolism and cell cycle in s. *cerevisiae*, 62, 2016.

[35] S. Nwaka and H. Holzer. Molecular biology of trehalose and the trehalases in the yeast Saccharomyces cerevisiae. *Prog. Nucleic Acid Res. Mol. Biol*, 58, 1998.

[36] A. Mitchell, P. Wei, and W. A. Lim. Oscillatory stress stimulation uncovers an Achilles heel of the yeast MAPK signaling network. *Science*, 350, 2015.

[37] L. J. García-Rodríguez, R. Valle, A. Durán, and C. Roncero. Cell integrity signaling activation in response to hyperosmotic shock in yeast. *FEBS Lett*, 579, 2005.

[38] R. G. Jones and C. B. Thompson. Tumor suppressors and cell metabolism: a recipe for cancer growth. *Genes Dev*, 23, 2009.

[39] M. Yuan, S. B. Breitkopf, X. Yang, and J. M. Asara. A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nat. Protoc*, 7, 2012.

[40] C. Engler et al. A golden gate modular cloning toolbox for plants. *ACS Synth. Biol*, 3, 2014.

[41] M. Hernanz-Koers et al. Fungalbraid: A GoldenBraid-based modular cloning platform for the assembly and exchange of DNA elements tailored to fungal synthetic biology. *Fungal Genet. Biol*, 116, 2018.

[42] M. E. Lee, W. C. DeLoache, B. Cervantes, and J. E. Dueber. A highly characterized yeast toolkit for modular. *Multipart Assembly. ACS Synth. Biol*, 4, 2015.

[43] A. Pérez-González et al. Adaptation of the GoldenBraid modular cloning system and creation of a toolkit for the expression of heterologous proteins in yeast mitochondria. *BMC Biotechnol. 17*, 17(80), 2017.

[44] B. Pollak et al. Loop assembly: a simple and open system for recursive fabrication of DNA circuits. *New Phytol*, 222, 2019.

[45] A. D. Halleran, A. Swaminathan, and R. M. Murray. Single day construction of multigene circuits with 3g assembly. *ACS Synth. Biol*, 7, 2018.

[46] E. H. Akama-Garren et al. A modular assembly platform for rapid generation of DNA constructs. *Sci. Rep. 6*, 6:16836, 2016.

[47] X. Duportet et al. A platform for rapid prototyping of synthetic gene networks in mammalian cells. *Nucleic Acids Res*, 42, 2014.

[48] A. Martella, M. Matjusaitis, J. Auxillos, S. M. Pollard, and Y. Cai. Emma: An extensible mammalian modular assembly toolkit for the rapid design and production of diverse expression vectors. *ACS Synth. Biol*, 6, 2017.

[49] E. Weber, C. Engler, R. Gruetzner, S. Werner, and S. Marillonnet. *A modular cloning system for standardized assembly of multigene constructs*. 2011.

[50] M. Liu et al. Genomic discovery of potent chromatin insulators for human gene therapy. *Nat. Biotechnol*, 33, 2015.

[51] A. L. Szymczak and D. A. A. Vignali. Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin. Biol. Ther*, 5, 2005.

[52] Z. Liu et al. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci. Rep*, 7, 2017.

[53] M. Hermann et al. Binary recombinase systems for high-resolution conditional mutagenesis. *Nucleic Acids Res*, 42, 2014.

[54] K. Yusa, L. Zhou, M. A. Li, A. Bradley, and N. L. Craig. A hyperactive piggyBac transposase for mammalian applications. *Proc. Natl. Acad. Sci. U. S. A*, 108, 2011.

[55] M. M. Nalaskowski, P. Ehm, S. Giehler, and G. W. Mayr. A toolkit for graded expression of green fluorescent protein fusion proteins in mammalian cells. *Anal. Biochem*, 428, 2012.

[56] S. Regot, J. J. Hughey, B. T. Bajar, S. Carrasco, and M. W. Covert. *High-sensitivity measurements of multiple kinase activities in live single cells*. 2014.

[57] J. R. Smith et al. Robust. *Persistent Transgene Expression in Human Embryonic Stem Cells Is Achieved with AAVS1-Targeted Integration. Stem Cells*, 26, 2008.

[58] M. Sadelain, E. P. Papapetrou, and F. D. Bushman. Safe harbours for the integration of new DNA in the human genome. *Nat. Rev. Cancer*, 12, 2011.

[59] J. G. Doench et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol*, 34, 2016.

[60] M. A. Moreno-Mateos et al. Crisprscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods*, 12, 2015.

[61] N. C. Yeo et al. An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat. Methods*, 15, 2018.

[62] S. Kiani et al. Cas9 gRNA engineering for genome editing, activation and repression. *Nat. Methods*, 12, 2015.

[63] J. F. Margolin et al. Krüppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. U. S. A*, 91, 1994.

[64] A. Chavez et al. Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods*, 12, 2015.

[65] L. S. Qi et al. *Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.* 2013.

[66] J. Emanuel, A. Marzi, H. Filoviruses: Ecology Feldmann, Molecular Biology, and Evolution. Adv. *Virus Res*, 100, 2018.

[67] E. Mühlberger, M. Weik, V. E. Volchkov, H. D. Klenk, and S. Becker. Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J. Virol*, 73, 1999.

[68] T. Hoenen and H. Feldmann. Reverse genetics systems as tools for the development of novel therapies against filoviruses. *Expert Rev. Anti. Infect. Ther*, 12, 2014.

[69] W. Tao, T. Gan, M. Guo, Y. Xu, and J. Zhong. *Novel Stable Ebola Virus Minigenome Replicon Reveals Remarkable Stability of the Viral Genome.* J. Virol, 2017.

[70] K. Brauburger et al. Analysis of the highly diverse gene borders in Ebola virus reveals a distinct mechanism of transcriptional regulation. *J. Virol*, 88, 2014.

[71] M. Trunschke et al. *The L-VP35 and L-L interaction domains reside in the amino terminus of the Ebola virus L protein and are potential targets for antivirals.* 2013.

[72] A. Groseth et al. The Ebola virus ribonucleoprotein complex: a novel VP30-L interaction identified. *Virus Res*, 140, 2009.

[73] T. Hoenen et al. Inclusion bodies are a site of ebolavirus replication. *J. Virol*, 86, 2012.

[74] M. R. Edwards et al. *High-Throughput Minigenome System for Identifying Small-Molecule Inhibitors of Ebola Virus Replication.* 2015.

[75] L. S. Uebelhoer et al. High-throughput, luciferase-based reverse genetics systems for identifying inhibitors of Marburg and Ebola viruses. *Antiviral Res*, 106, 2014.

[76] V. Lohmann and R. Bartenschlager. On the history of hepatitis C virus cell culture systems. *Chem*, 57, 2014.

[77] M. E. J. Woolhouse et al. Temporal trends in the discovery of human viruses. *Proc. Biol. Sci*, 275, 2008.

[78] M. Woolhouse, F. Scott, Z. Hudson, R. Howey, and M. Chase-Topping. Human viruses: discovery and emergence. *Philos. Trans. R. Soc. Lond. B Biol. Sci*, 367, 2012.

[79] B. Adamson et al. A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell 167*, 1867.

[80] P. Datlinger et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods*, 14, 2017.

[81] A. Dixit et al. Perturb-seq: Dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell 167*, 1853.

[82] E. Yeung et al. *Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks.* Cell Syst 5, 11–24, 2017.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Alain Bonny
032BF516E5844C0...
Author Signature

3/3/2020
Date