Rosa26 docking sites for investigating genetic circuit silencing in stem cells

Michael Fitzgerald[1], Chelsea Gibbs[1] and Tara L. Deans[1]*

[1] Department of Biomedical Engineering, University of Utah, Salt Lake City, UT, 84112, USA
* To whom correspondence should be addressed. Tel: 801-585-7859; Email: tara.deans@utah.edu

ABSTRACT

Approaches in mammalian synthetic biology have transformed how cells can be programmed to have reliable and predictable behaviour. The majority of mammalian synthetic biology has been accomplished using immortalized cell lines that are easy to grow, and easy to transfect. Additionally, genetic circuits that integrate into the genome of these immortalized lines remain functional for many generations, often for the lifetime of the cells. However, when genetic circuits are integrated into the genome of stem cells, gene silencing is observed within a few generations. To investigate the stability of genetic circuits and their propensity to silence in stem cells, the Rosa26 locus of mouse pluripotent stem cells was modified to contain docking sites for site-specific integration of genetic circuits. We show that the silencing of genetic circuits can be reversed with the addition of sodium butyrate, a histone deacetylase inhibitor. These findings demonstrate an approach to stabilize the function of genetic circuits in pluripotent stem cells to ensure robust function over many generations. Altogether, this work introduces an approach to overcome silencing of genetic circuits in pluripotent stem cells that enable the possibility of leveraging synthetic biology in pluripotent stem cells to meet therapeutic needs.

INTRODUCTION

Pluripotent stem cells have the potential to augment tissue regeneration, in addition to creating cell-specific in vitro diagnostics and drug screens because they are capable of self-renewing and differentiating into any cell type (1). Induced pluripotent stem (iPS) cells can be derived from mature tissue cells from individuals by expressing key transcription factors, which has the potential to facilitate personalized medicine (2). However, despite advances in stem cell culture techniques, differentiation can be inefficient, laborious, expensive, or otherwise intractable (3,4). It has been proposed that programming reliable cell behaviour using approaches in synthetic biology can be used for directing cell fate decisions to enhance their therapeutic potential. For example, using genetic circuits to reprogram cells to facilitate precise gene regulation can be used to enhance cell differentiation outcomes for tissue engineering and regenerative medicine applications. Moreover, in order to capitalize on the tight gene control of genetic circuits, their stability in the genome, or ability to maintain function, for many generations is critical.

Novel genetic tools built by synthetic biologists have transformed how cells can be reprogrammed that include genetic programs to introduce switching (5-12), oscillations (13-18), logic gates (19-21), and biosensing (22-28) behaviours into cells. Assembling simple genetic parts into more complex gene circuits can reliably and predictably control cell behaviours (29,30). To date, the majority of mammalian synthetic biology has taken place in easy-to-grow and easy-to-transfect cells, derived

from immortalized cell lines. These model cell lines are useful for enhancing our understanding of synthetic gene circuits that underscore the potential of synthetic biology tools, however, these cell lines may not be good predictors of the challenges that arise in stem cells (31). Plasmid and viral gene delivery systems have shown to lose expression over weeks of cell culture, which is thought to be a consequence of epigenetic modifications of the inserted DNA (32-34). The current understanding of transgene silencing suggests that silencing of genes can occur through the methylation of the expression cassette and/or the formation of heterochromatin, both of which facilitate changes in gene expression; however, the circumstances that trigger these mechanisms are still being elucidated. To overcome these challenges, several strategies to mitigate the effects of transgene silencing have been used, such as the inclusion of universal chromatin opening elements (35,36), scaffold/matrix attachment regions (37), CG free plasmid backbones (38), minicircle DNA (39), genomic insulators (40), and targeted integration in open chromatin regions (41). These approaches can significantly lessen transgene silencing, however the effects are not universal (42,43). Currently, the impact of genetic circuit silencing is not known, and strategies to alleviate silencing have yet to be explored. Genetic circuits are distinct from plasmid and viral expression cassettes because they contain multiple genetic modules that makeup the circuits, resulting in larger DNA constructs that need to be inserted into cells. Since genetic circuits endow cells with tight gene control, they have been a focus for directing stem cell fate to enable both timed and tuned gene expression that match the evolving requirements of the differentiating cells as they undergo cell fate decisions (44). Altogether, reliable stem cell programming with genetic circuits will require tools that integrate into the genome, have predictable expression patterns in differentiated cells, and avoid disruption of endogenous genes and pathways.

One potential approach to avoid epigenetic silencing of integrated genetic circuits is to engineer genomic safe harbour sites within the genome that support the integration of genetic material to those specific genomic locations, or loci. Additionally, safe harbour loci within the genome limit transgene silencing caused by unpredictable genome interactions associated with random insertions (32,45). The Rosa26 locus in mice has been observed to facilitate the ubiquitous expression of transgenes in developing and adult tissue when inserted at this site, suggesting that transgenes are active in germ cells, in addition to the differentiated progeny of those cells (46,47). Therefore, the Rosa26 locus is an ideal location to target for inserting genetic circuits to study their stability and function, in stem cells.

Here, genetic circuit silencing was studied by observing circuit function in pluripotent stem cells by engineering a mouse embryonic stem (ES) cell line with three ϕC31 docking sites on one allele of the Rosa26 locus to function as a landing pad for large genetic circuits that have a matching recombinase. The genetic circuit, LTRi, was chosen because it is relatively complex in architecture and the genetic circuit permits tuneable control of gene expression (7,48,49).

MATERIAL AND METHODS

Design of plasmids

The homology repair template for CRISPR was Gibson assembled to include 1kb homology arms, an FRT-flanked neomycin resistance gene, 3x attP sites, and a blue fluorescent protein (BFP) expression cassette (Addgene #36086). The 1kb upstream and 1kb downstream arms were amplified from purified mouse genome from AB2.2 ES cells (ATCC #SCRC-1023) and verified by sequencing. The Cas9/gRNA plasmid was obtained from the University of Utah Mutation Generation and Detection Core (gRNA homology: TGGGCGGGAGTCTTCTGGGC). Modifications were made to the LTRi genetic switch to exchange viral promoters with non-viral promoters, namely mEF1 and rEF1 promoters from the pVItro1-msc plasmid (InvivoGen #pvitro1-mcs) and the addition of an attP docking site recognition sequence. The modified LTRi genetic switch (mLTRi) was constructed by cloning the transgene module into the vector module with DraIII and XhoI cut sites, and the individual modules were put together by Gibson assembly.

Cell culture

Mouse embryonic stem (ES) cells (ATCC #SCRC-1023) were maintained in high glucose knockout DMEM (Life Technologies #10829−018) supplemented with 15% ES certified FBS (LifeTechnologies #10439024), 1% nonessential amino acids (Life Technologies #11140050), 1mM L-glutamine (Life Technologies #25030−081), 0.1mM 2-mercaptoethanol (Life Technologies #21985023), and 200 units/mL penicillin and streptomycin. The ES cells were plated on mitomycin C treated SNLP cells (ATCC SCRC-1049) that are G418 resistant and secrete leukemic inhibitory factor (LIF) at an abundant level to maintain ES cell pluripotency. All cell lines were grown in a humidified 5% CO_2, 37°C incubator. The SNLP feeder cells were grown in high glucose DMEM medium supplemented with 10% FBS, and 200 units/mL penicillin and streptomycin until seeded with mouse ES cells, at which time the media conditions were as stated for the ES cells.

CRISPR modified pluripotent stem cells

Plasmid DNA containing the repair template with three tandem attP sites was co-transfected (1:1, Jetprime VWR#89129) with a plasmid containing Rosa26-Cas9/gRNA into mouse ES cells when confluency reached 70%. The cells were selected by adding 200μg/ml neomycin (G418 Life Technologies #10131035) to the growth medium. Resistant clones were expanded and screened for the on-site genomic edit by genomic PCR. Copy number qPCR (Power Sybr, Thermofisher #4368577) was used to determine off-site integration. The neomycin resistance gene was removed by transfecting and flow sorting (Beckon Dickenson FACSAria) the candidate cell line with a plasmid harboring GFP-flip recombinase and assessing the reacquisition of neomycin sensitivity.

Docking plasmids

mLTRi-GFP cell lines were established by co-transfection with □C31 integrase in the 3X-attP AB2.2 ES cells. Cells that had the mLTRi-GFP plasmid contained a neomycin resistant cassette so these transfected ES cells were selected by G418, and resistant lines were clonally expanded and further screened by inducibility with 250□M IPTG. One of the positive clones was chosen and used for the entire study reported here.

Quantitative PCR primer design

All primers were design with NCBI's primer blast to have a PCR product size between 70bp and 200bp and a melting temperature between 58°C and 62°C (Table S1). The primers for known copy number (GAPDH and ZFY1) were design to not span and exon junction because the template was genomic DNA. The primers were tested to ensure that there was no alignment with non-specific sequences. Each primer was tested using a 2x dilution of genomic DNA to ensure a single melting curve peak to ensure specific binding to only the desired location.

Quantitative PCR

Genomic DNA was isolated from the ES cells using lysis buffer (100mMTris-Cl, 5mM EDTA, 200mM NaCl, 0.2% SDS) and proteinase K as previously reported (50). The DNA was then precipitated in $C_2H_3NaO_2$ and isopropanol and washed in EtOH. After genomic DNA elution and quantification using a NanoDrop (ThermoFisher) and a 1ng/µl stock was created. From the 1ng/ul stock a 2x dilution was prepared for each qPCR experiment. Each well of the qPCR reaction contained: 10ul of Power SYBR Green PCR Master Mix (ThermoFisher #4367659), 2ul of a 3□M forward primer that anneals to, 2ul of a 3□M reverse primer, genomic DNA, and up to 6ul of H_2O. The experiments were performed in triplicates on 96 well plates and used the StepOne™ Real-Time PCR System (ThermoFisher) where a standard curve was generated for each gene. The Ct value vs. the log of amount of DNA was plotted. The slope was then calculated for each gene and the unknown 3x attP was compared to the known genes. Given that two copies of GAPDH exist in the genome, and only one copy of ZFY1, which is a gene on the Y chromosome (we used male ES cells) exist in the genome, CT values and slopes were compared to the attP sample.

Epigenetic modification and flow cytometry

Pluripotent stem cells were induced with 250□M IPTG in the presence or absence of epigenetic modifying drugs at the concentrations noted in text and within the figure legends. Modifying drugs were purchased from the following: sodium butryrate (VWR #89147), 5-azacytodine (Sigma Aldrich #A2385). Cells were treated with the drug for 24 hours with and without IPTG and GFP expression was assessed by flow cytometry.

RESULTS

Generation of ϕC31 docking sites in the Rosa26 locus of pluripotent stem cells

Clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 technology enable the targeting of specific locations within the genome. Utilizing CRISPR/Cas9 with customized guide RNAs (gRNAs) to target specific locations within the genome offers a useful method for engineering landing pads for genetic circuits at desired locations within the genome. Indeed, CRISPR/Cas9 can be used for inserting new sequences of DNA, however, the efficiency is significantly decreased when inserting DNA sequences larger than 5kb (51). Because most genetic circuits that function in mammalian cells are larger than 5kb, inserting a landing pad that has a site-specific recombinase can increase

efficiency of integration of these larger DNA constructs into the genome when a matching recombinase is present in the plasmid containing the DNA to be inserted (52).

To allow for targeted and efficient integration of genetic circuits into the genome of pluripotent stem cells, ϕC31 docking sites were inserted into the Rosa26 locus (Figure 1A) of mouse ES cells using CRISPR/Cas9 technology. Three tandem ϕC31 integrase *attP* sites were inserted that serve as a landing pad for larger DNA additions containing an *attB* sequence (47). To create an efficient screening method, a neomycin resistance gene flanked by *FRT* sites was added upstream of the 3x *attP* sites inside the homology arms of the repair DNA plasmid, and outside of the arms, a blue fluorescent protein (BFP) cassette was included (Figure 1B). This plasmid was co-transfected with Cas9 endonuclease and a gRNA that targets the Rosa26 locus. Adding neomycin (G418) to the media of transfected cells enables for the selection of ES colonies that have the landing pad integrated into the genome at the Rosa26 locus (Figure 1C). CRISPR/Cas9 technology has shown to have off target effects, namely whole or parts of the repair template integrate into the genome at random locations, rather than just the repair DNA flanked by the homology arms in the targeted location (53-56). To screen for off-target insertions, a BFP cassette was added outside of the homology arms in the repair DNA template (Figure 1B). Therefore, any ES colony expressing BFP would indicate that the repair plasmid was incorrectly integrated into the genome and discarded. The ES cells resistant to G418 that did not express BFP were selected and verified by PCR for the on-site insertion of *attP* sites in the Rosa26 locus with primers designed to span part of the neomycin resistance gene and the genomic DNA beyond the homology arm (Figure 2A). ES clones with an on-site edit produced a 1.6kb amplicon (Figure 2B) that contained an XbaI site and when isolated and cut with XbaI endonuclease produced the predicted 1.1kb and 500bp bands (Figure 2C). DNA sequencing was also performed and further confirmed the insertion of the repair DNA in the Rosa26 locus.

Screen of the mouse genome for correct genomic edits

To assess the number of docking sites integrated into the mouse genome, quantitative PCR (qPCR) was performed to verify whether additional integrations occurred elsewhere within the genome (Figure 2D). By referencing two genes of known copy number, ZFY1 on the Y chromosome (one copy), and GAPDH on chromosome 6 (two copies), it was possible to determine how many times the docking site was inserted into the genome. Primers were designed to anneal at the end of the 3x *attP* sequence and included part of the genomic Rosa26 sequence to give a ~100mer amplicon (Figure 2A). The mean CT values of each primer set vs. the log of DNA (ng) were plotted as previously described (50). The slope of CT values vs. log of DNA of each sample was calculated and compared to the slope of ZFY1 and GAPDH. Since the slope of the attP amplicon matched that of the ZFY1, it shows that the docking site was added once (Figure 2D). Once correct docking was confirmed, the neomycin resistance gene was removed by transfecting the confirmed heterozygous ES line with a plasmid harboring flip recombinase, leaving the 3x *attP* sites (Figure 2E). To validate that the resistant cassette was removed, the *FLP* transfected ES cells were clonally expanded, assessed for neomycin

sensitivity, and verified by PCR. The amplification of the wild type ES Rosa26 locus is expected to be 1.1kb, while the addition of the *attP* sites without neomycin is expected to increase in size to 1.4kb (Figure 2F). Results indicate that the ES cell line has the Rosa26 site-specific addition of 3x *attP* sites at one allele and is nowhere else in the genome.

Docking genetic circuits

To assess the stability LTRi_EGFP in pluripotent stem cells, the genetic circuit was integrated into the Rosa26 locus at the docking sites. First, to rule out the possibility of CMV and RSV being silenced (57), LTRi was modified to replace the original CMV and RSV promoters with the non-viral promoters mouse Elongation Factor-1 (mEF1), and rat Elongation Factor-1 (rEF1) (Figure 3A-B). Site-specific docking of the modified LTRi (mLTRi)_EGFP genetic circuit was accomplished by adding an *attB* sequence to the plasmid. Along with the genetic circuit, the plasmid also contained a neomycin resistance cassette for G418 selection of transfected ES cells. The mLTRi genetic circuit containing an *attB* sequence was co-transfected with ϕC31 integrase, resulting in the mLTRi genetic circuit stably integrated into the genome. ES cells resistant to G418 were clonally expanded and screened for their response to the chemical inducer, isopropyl β-D-1-thiogalactopyranoside (IPTG), a small molecule that activates gene expression of the genetic circuit (7,48,49). To confirm docking, the IPTG responsive cell lines were screened for integration into Rosa26 by conventional PCR with primers flanking the Rosa locus and the genetic circuit, which is expected to produce a 1.3kb amplicon (Figure S1A and B).

Assessing the stability of genetic circuits in Rosa26 locus

To investigate the stability of the mLTRi genetic circuit in pluripotent stem cells, mLTRi was docked into the Rosa26 locus of mouse ES cells and the cells expressing EGFP in the presence of IPTG were single cell sorted and cultured for more than three weeks in the presence or absence of IPTG. EGFP was assessed using microscopy and results show that ES colonies that are cultured beyond three weeks significantly lose their ability to express EGFP in the presence of IPTG compared to cells in earlier time points (Figure 3C-D'). Altogether, these data suggest that the decrease in gene expression strength is likely due to the genetic circuit being susceptible to epigenetic modifications, specifically to methylation and/or histone acetylation (33,34,58).

Reactivation of silenced mLTRi in pluripotent stem cells

To investigate the epigenetic silencing of mLTRi, two common mechanisms of transgene silencing that are frequently cited as barriers when introducing exogenous DNA were studied. The first, promoter methylation, occurs with the methylation of cytosine residues in CpG sequences by the cytosine DNA methyltransferase (DNMT1) enzyme that can be reversed with methyltransferase inhibitor 5-aza-deoxycytidine (AzaC) (59-64). The second, histone acetylation, is a process where an acetyl group is added to a lysine residue on the tail of a histone by histone deacetylase, which can be prevented by sodium butyrate (NaB), a known histone deacetylase inhibitor (HDACi) (65-67).

To test whether the silenced genetic circuit could be reactivated, ES colonies that lost their ability to respond to IPTG were grown in the presence of the inhibitors AzaC or NaB for 48 hours, and EGFP expression was assessed using flow cytometry (Figure 4A). In line with histone acetylation dependent silencing, EGFP expression was reactivated in the presence of NaB, suggesting that the genetic circuit was silenced by histone acetylation. Next, to determine the optimal concentration of NaB for recovering genetic circuit function, various concentrations of NaB were added to the media of silenced ES cells in the presence of 250μM IPTG (Figure 4B). While 5M NaB showed the most recovery of EGFP expression, cells grown in this condition stop growing after 24 hours and were dead after 48 hours. We found that the 500μM concentration of NaB also slowed the growth of the ES cells, and while they did not die after 48 hours, their morphology changed and the cells did not appear as healthy as the no NaB control. The 250μM concentration of NaB did not appear to alter the growth rate or the morphology of the ES cells over time. Therefore, we conducted our experiments at 250μM NaB. To determine whether EGFP expression dynamics could be recovered, ES cells that recovered their EGFP expression in the presence of 250μM NaB were sorted by flow cytometry and grown in various concentrations of IPTG (Figure 4C) over various time points. These data show that exposure to 250μM NaB for 48 hours reverses the silencing of mLTRi_EGFP and allows for the recovery of the genetic circuit.

DISCUSSION

To date, genetic circuits in mammalian cells have primarily been reported in easy to grow and easy to transfect immortalized cell lines. Pluripotent stem cells have the potential to give rise to all cell types in the body and can propagate indefinitely under the right culturing conditions. Because pluripotent stem cells represent a single cell source that can make large contributions toward currently unmet clinical needs for regenerating damaged and diseased tissue, tightly controlling specific genes is critical for effectively driving stem cell differentiation into desired lineages. Novel genetic tools built by synthetic biologists allow for such control in a variety of mammalian cell lines. However, gene expression from transgene expression systems has been shown to have heterogenic expression patterns that are often silenced by epigenetic modifications over time (59,60). To overcome this limitation, we have engineered mouse embryonic stem cells with ϕC31 docking sites using CRISPR/Cas9 technology to allow for the targeted insertion of genetic circuits into the Rosa26 locus. This docking site functions as a landing pad for genetic circuits that have a matching recombinase to be targeted for insertion at this genome location. Docked mLTRi in this location demonstrated robust circuit function for up to three weeks of culture, however, after three weeks EGFP expression significantly decreased and mLTRi was no longer responsive to IPTG induction. To overcome this silencing in pluripotent stem cells, we show that adding the HDAC inhibitor, NaB, to the media reverses the silencing of the mLTRi genetic circuit and enables its original function to be rescued.

Taken together, this study demonstrates that genetic circuits can be inserted into the genome of pluripotent stem cells and if circuit function diminishes over time, NaB can be added to the growth

media to re-establish circuit function. These data raise the exciting possibility of using synthetic biology in pluripotent stem cells for many therapeutic applications.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENT

We thank the members of the Deans Lab for helpful discussions on the manuscript, and the University of Utah Flow Cytometry Facility that is funded by the National Cancer Institute [5P30CA042014-24], and by the National Centre for Research of the National Institutes of Health [1S10RR026802-01].

FUNDING

This work was supported by the University of Utah start-up funds, the National Science Foundation CAREER Program [CBET-1554017 to TLD], the Office of Naval Research Young Investigator Program [N00014-16-1-3012 to TLD], and the National Institute of Health Trailblazer Award [1R21EB025413-01 to TLD]. This work was also supported by the University of Utah Undergraduate Research Opportunities Program (UROP) to CG.

CONFLICT OF INTEREST

None.

REFERENCES

1. Singh, V.K., Kalsan, M., Kumar, N., Saini, A. and Chandra, R. (2015) Induced pluripotent stem cells: applications in regenerative medicine, disease modeling, and drug discovery. *Front Cell Dev Biol*, **3**, 2.
2. Takahashi, K. and Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, **126**, 663-676.
3. Abdelalim, E.M. and Emara, M.M. (2015) Advances and challenges in the differentiation of pluripotent stem cells into pancreatic beta cells. *World J Stem Cells*, **7**, 174-181.
4. Youssef, A.A., Ross, E.G., Bolli, R., Pepine, C.J., Leeper, N.J. and Yang, P.C. (2016) The Promise and Challenge of Induced Pluripotent Stem Cells for Cardiovascular Applications. *JACC Basic Transl Sci*, **1**, 510-523.
5. Callura, J.M., Cantor, C.R. and Collins, J.J. (2012) Genetic switchboard for synthetic biology applications. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 5850-5855.
6. Chang, A.L., Wolf, J.J. and Smolke, C.D. (2012) Synthetic RNA switches as a tool for temporal and spatial control over gene expression. *Current opinion in biotechnology*, **23**, 679-688.
7. Deans, T.L., Cantor, C.R. and Collins, J.J. (2007) A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell*, **130**, 363-372.
8. Fitzgerald, M., Gibbs, C., Shimpi, A.A. and Deans, T.L. (2017) Adoption of the Q Transcriptional System for Regulating Gene Expression in Stem Cells. *ACS synthetic biology*, **6**, 2014-2020.
9. Gardner, T.S., Cantor, C.R. and Collins, J.J. (2000) Construction of a genetic toggle switch in Escherichia coli. *Nature*, **403**, 339-342.

10. Green, A.A., Silver, P.A., Collins, J.J. and Yin, P. (2014) Toehold switches: de-novo-designed regulators of gene expression. *Cell*, **159**, 925-939.
11. Kramer, B.P., Viretta, A.U., Daoud-El-Baba, M., Aubel, D., Weber, W. and Fussenegger, M. (2004) An engineered epigenetic transgene switch in mammalian cells. *Nature biotechnology*, **22**, 867-870.
12. Weber, W. and Fussenegger, M. (2011) Molecular diversity--the toolbox for synthetic gene switches and networks. *Curr Opin Chem Biol*, **15**, 414-420.
13. Danino, T., Mondragon-Palomino, O., Tsimring, L. and Hasty, J. (2010) A synchronized quorum of genetic clocks. *Nature*, **463**, 326-330.
14. Elowitz, M.B. and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**, 335-338.
15. Fung, E., Wong, W.W., Suen, J.K., Bulter, T., Lee, S.G. and Liao, J.C. (2005) A synthetic gene-metabolic oscillator. *Nature*, **435**, 118-122.
16. Judd, E.M., Laub, M.T. and McAdams, H.H. (2000) Toggles and oscillators: new genetic circuit designs. *BioEssays : news and reviews in molecular, cellular and developmental biology*, **22**, 507-509.
17. Stricker, J., Cookson, S., Bennett, M.R., Mather, W.H., Tsimring, L.S. and Hasty, J. (2008) A fast, robust and tunable synthetic gene oscillator. *Nature*, **456**, 516-519.
18. Tigges, M., Marquez-Lago, T.T., Stelling, J. and Fussenegger, M. (2009) A tunable synthetic mammalian oscillator. *Nature*, **457**, 309-312.
19. Guet, C.C., Elowitz, M.B., Hsing, W. and Leibler, S. (2002) Combinatorial synthesis of genetic networks. *Science*, **296**, 1466-1470.
20. Schukur, L. and Fussenegger, M. (2016) Engineering of synthetic gene circuits for (re-)balancing physiological processes in chronic diseases. *Wiley Interdiscip Rev Syst Biol Med*, **8**, 402-422.
21. Siuti, P., Yazbek, J. and Lu, T.K. (2013) Synthetic circuits integrating logic and memory in living cells. *Nature biotechnology*, **31**, 448-452.
22. Bowsher, C.G. and Swain, P.S. (2014) Environmental sensing, information transfer, and cellular decision-making. *Current opinion in biotechnology*, **28C**, 149-155.
23. Brenner, M., Cho, J.H. and Wong, W.W. (2017) Synthetic biology: Sensing with modular receptors. *Nature chemical biology*, **13**, 131-132.
24. Daringer, N.M., Dudek, R.M., Schwarz, K.A. and Leonard, J.N. (2014) Modular extracellular sensor architecture for engineering mammalian cell-based devices. *ACS synthetic biology*, **3**, 892-902.
25. Feng, J., Jester, B.W., Tinberg, C.E., Mandell, D.J., Antunes, M.S., Chari, R., Morey, K.J., Rios, X., Medford, J.I., Church, G.M. *et al.* (2015) A general strategy to construct small molecule biosensors in eukaryotes. *Elife*, **4**, 1-23.
26. Slomovic, S. and Collins, J.J. (2015) DNA sense-and-respond protein modules for mammalian cells. *Nature methods*, **12**, 1085-1090.
27. Smole, A., Lainscek, D., Bezeljak, U., Horvat, S. and Jerala, R. (2017) A Synthetic Mammalian Therapeutic Gene Circuit for Sensing and Suppressing Inflammation. *Molecular therapy : the journal of the American Society of Gene Therapy*, **25**, 102-119.
28. Youk, H. and Lim, W.A. (2014) Secreting and sensing the same molecule allows cells to achieve versatile social behaviors. *Science*, **343**, 1242782.
29. MacDonald, I.C. and Deans, T.L. (2016) Tools and applications in synthetic biology. *Adv Drug Deliv Rev*, **105**, 20-34.
30. Weisenberger, M.S. and Deans, T.L. (2018) Bottom-up approaches in synthetic biology and biomaterials for tissue engineering applications. *J Ind Microbiol Biotechnol*, **45**, 599-614.
31. Ceroni, F. and Ellis, T. (2018) The challenges facing synthetic biology in eukaryotes. *Nature reviews. Molecular cell biology*.
32. Papapetrou, E.P. and Schambach, A. (2016) Gene Insertion Into Genomic Safe Harbors for Human Gene Therapy. *Molecular therapy : the journal of the American Society of Gene Therapy*, **24**, 678-684.
33. Godecke, N., Zha, L., Spencer, S., Behme, S., Riemer, P., Rehli, M., Hauser, H. and Wirth, D. (2017) Controlled re-activation of epigenetically silenced Tet promoter-driven transgene expression by targeted demethylation. *Nucleic acids research*, **45**, e147.
34. Bencsik, R., Boto, P., Szabo, R.N., Toth, B.M., Simo, E., Balint, B.L. and Szatmari, I. (2016) Improved transgene expression in doxycycline-inducible embryonic stem cells by repeated chemical selection or cell sorting. *Stem Cell Res*, **17**, 228-234.

35. Muller-Kuller, U., Ackermann, M., Kolodziej, S., Brendel, C., Fritsch, J., Lachmann, N., Kunkel, H., Lausen, J., Schambach, A., Moritz, T. *et al.* (2015) A minimal ubiquitous chromatin opening element (UCOE) effectively prevents silencing of juxtaposed heterologous promoters by epigenetic remodeling in multipotent and pluripotent stem cells. *Nucleic acids research*, **43**, 1577-1592.
36. Zhang, F., Frost, A.R., Blundell, M.P., Bales, O., Antoniou, M.N. and Thrasher, A.J. (2010) A ubiquitous chromatin opening element (UCOE) confers resistance to DNA methylation-mediated silencing of lentiviral vectors. *Molecular therapy : the journal of the American Society of Gene Therapy*, **18**, 1640-1649.
37. Zhao, C.P., Guo, X., Chen, S.J., Li, C.Z., Yang, Y., Zhang, J.H., Chen, S.N., Jia, Y.L. and Wang, T.Y. (2017) Matrix attachment region combinations increase transgene expression in transfected Chinese hamster ovary cells. *Sci Rep*, **7**, 42805.
38. Bazzani, R.P., Pringle, I.A., Connolly, M.M., Davies, L.A., Sumner-Jones, S.G., Schleef, M., Hyde, S.C. and Gill, D.R. (2016) Transgene sequences free of CG dinucleotides lead to high level, long-term expression in the lung independent of plasmid backbone design. *Biomaterials*, **93**, 20-26.
39. Chen, Z.Y., Riu, E., He, C.Y., Xu, H. and Kay, M.A. (2008) Silencing of episomal transgene expression in liver by plasmid bacterial backbone DNA is independent of CpG methylation. *Molecular therapy : the journal of the American Society of Gene Therapy*, **16**, 548-556.
40. Macarthur, C.C., Xue, H., Van Hoof, D., Lieu, P.T., Dudas, M., Fontes, A., Swistowski, A., Touboul, T., Seerke, R., Laurent, L.C. *et al.* (2012) Chromatin insulator elements block transgene silencing in engineered human embryonic stem cell lines at a defined chromosome 13 locus. *Stem Cells Dev*, **21**, 191-205.
41. Strathdee, D., Ibbotson, H. and Grant, S.G. (2006) Expression of transgenes targeted to the Gt(ROSA)26Sor locus is orientation dependent. *PloS one*, **1**, e4.
42. Hoffmann, D., Schott, J.W., Geis, F.K., Lange, L., Muller, F.J., Lenz, D., Zychlinski, D., Steinemann, D., Morgan, M., Moritz, T. *et al.* (2017) Detailed comparison of retroviral vectors and promoter configurations for stable and high transgene expression in human induced pluripotent stem cells. *Gene Ther*, **24**, 298-307.
43. Saunders, F., Sweeney, B., Antoniou, M.N., Stephens, P. and Cain, K. (2015) Chromatin function modifying elements in an industrial antibody production platform--comparison of UCOE, MAR, STAR and cHS4 elements. *PloS one*, **10**, e0120096.
44. Tsankov, A.M., Gu, H., Akopian, V., Ziller, M.J., Donaghey, J., Amit, I., Gnirke, A. and Meissner, A. (2015) Transcription factor binding dynamics during human ES cell differentiation. *Nature*, **518**, 344-349.
45. Soriano, P. (1999) Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature genetics*, **21**, 70-71.
46. Casola, S. (2010) Mouse models for miRNA expression: the ROSA26 locus. *Methods in molecular biology*, **667**, 145-163.
47. Tasic, B., Hippenmeyer, S., Wang, C., Gamboa, M., Zong, H., Chen-Tsai, Y. and Luo, L. (2011) Site-specific integrase-mediated transgenesis in mice via pronuclear injection. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 7902-7907.
48. Deans, T.L., Singh, A., Gibson, M. and Elisseeff, J.H. (2012) Regulating synthetic gene networks in 3D materials. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 15217-15222.
49. Singh, A., Deans, T.L. and Elisseeff, J.H. (2013) Photomodulation of Cellular Gene Expression in Hydrogels. *Acs Macro Letters*, **2**, 269-272.
50. Yun, J.J., Heisler, L.E., Hwang, II, Wilkins, O., Lau, S.K., Hyrcza, M., Jayabalasingham, B., Jin, J., McLaurin, J., Tsao, M.S. *et al.* (2006) Genomic DNA functions as a universal external standard in quantitative real-time PCR. *Nucleic acids research*, **34**, e85.
51. Zhang, L., Jia, R., Palange, N.J., Satheka, A.C., Togo, J., An, Y., Humphrey, M., Ban, L., Ji, Y., Jin, H. *et al.* (2015) Large genomic fragment deletions and insertions in mouse using CRISPR/Cas9. *PloS one*, **10**, e0120396.
52. Gaidukov, L., Wroblewska, L., Teague, B., Nelson, T., Zhang, X., Liu, Y., Jagtap, K., Mamo, S., Tseng, W.A., Lowe, A. *et al.* (2018) A multi-landing pad DNA integration platform for mammalian cell engineering. *Nucleic acids research*, **46**, 4072-4086.
53. Sawatsubashi, S., Joko, Y., Fukumoto, S., Matsumoto, T. and Sugano, S.S. (2018) Development of versatile non-homologous end joining-based knock-in module for genome editing. *Sci Rep*, **8**, 593.

54. Zhang, X.H., Tee, L.Y., Wang, X.G., Huang, Q.S. and Yang, S.H. (2015) Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol Ther Nucleic Acids*, **4**, e264.
55. Kim, S., Kim, D., Cho, S.W., Kim, J. and Kim, J.S. (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res*, **24**, 1012-1019.
56. Peng, R., Lin, G. and Li, J. (2016) Potential pitfalls of CRISPR/Cas9-mediated genome editing. *FEBS J*, **283**, 1218-1231.
57. Brooks, A.R., Harkins, R.N., Wang, P., Qian, H.S., Liu, P. and Rubanyi, G.M. (2004) Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle. *J Gene Med*, **6**, 395-404.
58. Bestor, T.H. (2000) Gene silencing as a threat to the success of gene therapy. *The Journal of clinical investigation*, **105**, 409-411.
59. Tycko, B. (2000) Epigenetic gene silencing in cancer. *The Journal of clinical investigation*, **105**, 401-407.
60. Wolffe, A.P. and Matzke, M.A. (1999) Epigenetics: regulation through repression. *Science*, **286**, 481-486.
61. Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes & development*, **16**, 6-21.
62. Shalginskikh, N., Poleshko, A., Skalka, A.M. and Katz, R.A. (2013) Retroviral DNA methylation and epigenetic repression are mediated by the antiviral host protein Daxx. *J Virol*, **87**, 2137-2150.
63. Lubbert, M. (2000) DNA methylation inhibitors in the treatment of leukemias, myelodysplastic syndromes and hemoglobinopathies: clinical results and possible mechanisms of action. *Curr Top Microbiol Immunol*, **249**, 135-164.
64. Stresemann, C. and Lyko, F. (2008) Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. *Int J Cancer*, **123**, 8-13.
65. Mariani, M.R., Carpaneto, E.M., Ulivi, M., Allfrey, V.G. and Boffa, L.C. (2003) Correlation between butyrate-induced histone hyperacetylation turn-over and c-myc expression. *J Steroid Biochem Mol Biol*, **86**, 167-171.
66. Davie, J.R. (2003) Inhibition of histone deacetylase activity by butyrate. *J Nutr*, **133**, 2485S-2493S.
67. Eberharter, A. and Becker, P.B. (2002) Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO reports*, **3**, 224-229.

TABLE AND FIGURES LEGENDS

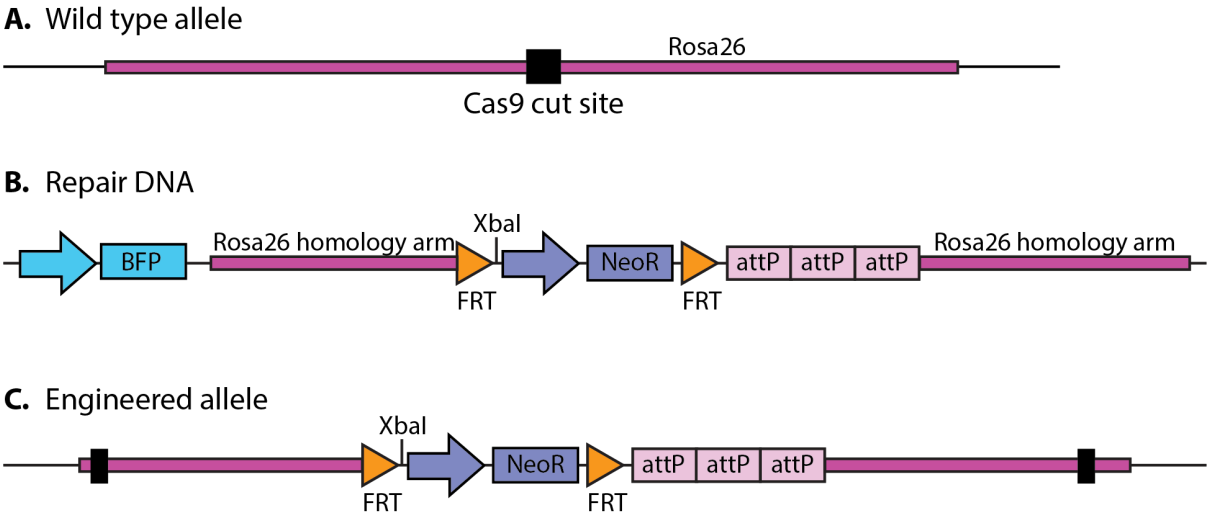

Figure 1: Approach for inserting docking sites into mouse pluripotent stem cells. **A.** Wild type Rosa26 allele with the Cas9 cut site (black square). **B**. Engineered repair DNA with 3x attP docking sites (pink squares). Neomycin resistance flanked by FRT sites (orange triangles) was added to enable G418 selection of the genomic insertion of the docking sites. Blue fluorescence protein (BFP) was added outside of the homology arms to screen for off-target, random integration of the repair DNA. **C**. The engineered Rosa26 allele with the repair DNA successfully inserted into the genome.

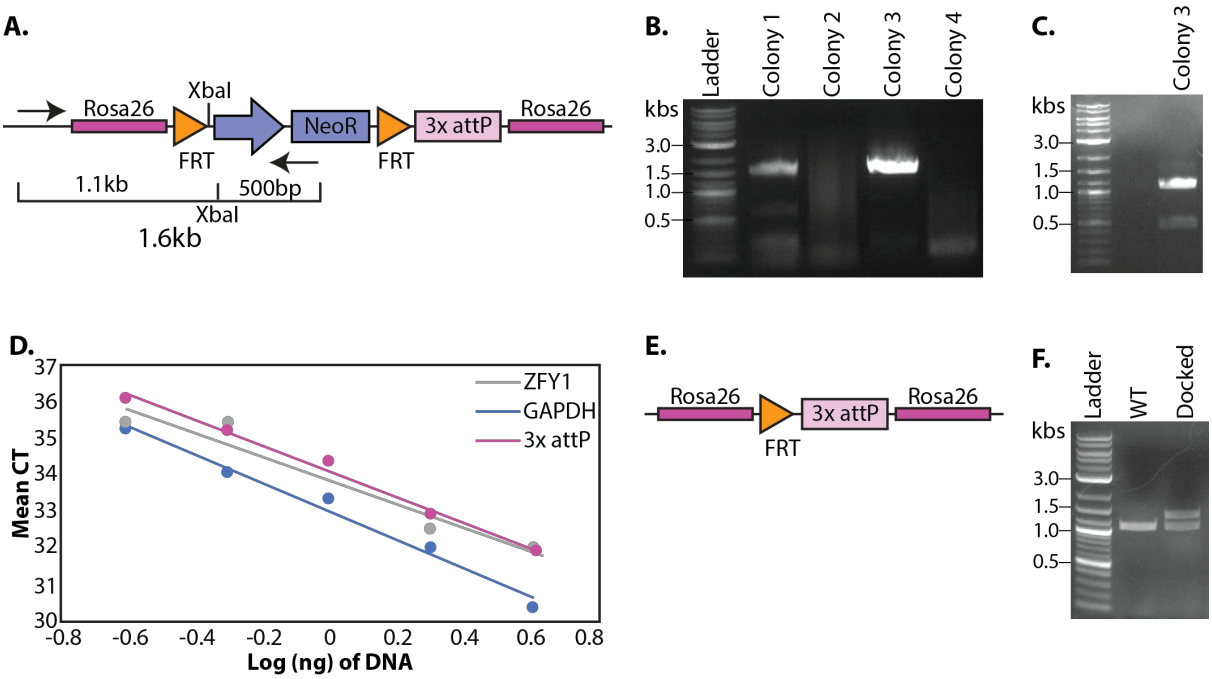

Figure 2: Screening for docking site integration. A. Schematic of PCR screen to confirm integration into the Rosa26 locus. Correct integration yields an expected amplicon size of 1.6kb. **B**. PCR on the genomic DNA of four different ES colonies. **C**. Colony 3 amplicon was gel isolated and cut with XbaI to confirm integration of the repair DNA. **D**. Quantitative PCR on ZFY1 (grey), GAPDH (blue), and the targeted Rosa26 allele with 3x attP sites (pink). **E.** Schematic of the genomic edited site in ES clones with confirmed attP docking sites in the Rosa26 locus after transfection with flp recombinase to remove the neomycin resistance gene. **F.** PCR comparing the wild time (WT) ES cells without the docking site and with a confirmed ES colony harboring the 3x attP docking site.

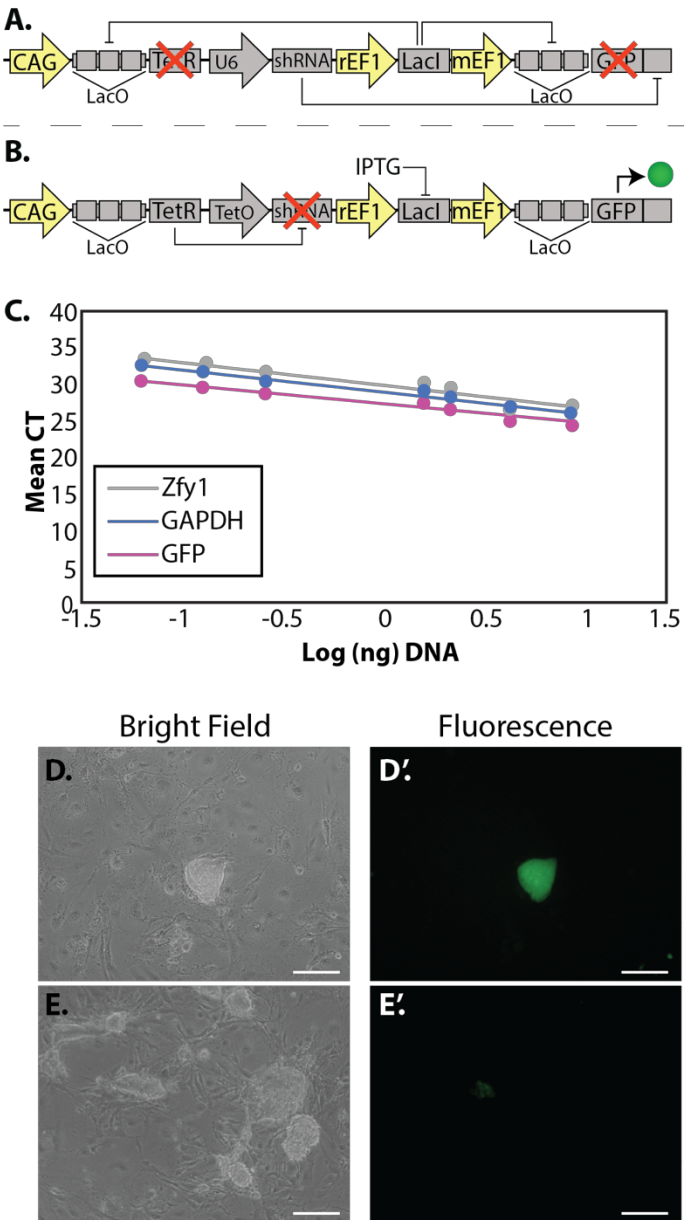

Figure 3: Integrating mLTRi_EGFP into the Rosa26 locus. A. Schematic of mLTRi in the off state: LacI repressor proteins are constitutively expressed and bind to the lac operator sites upstream of TetR and GFP. This causes transcriptional repression of TetR and GFP. With the repression of TetR, shRNA is transcribed by the U6 promoter and complementarily binds to the synthetic target sequence located in the 3'UTR of the GFP mRNA. **B.** Schematic of mLTRi in the on state: in the induced state, IPTG binds to the LacI proteins, producing a conformational change in the repressor proteins. This causes them to no longer bind to the lac operator sites, which allows for the transcription of GFP and TetR. The Tet repressor proteins bind to the tet operator site located in the U6 promoter, repressing the transcription of the shRNA. The resulting effect is a robust expression of GFP. **C.** Bright field image of stably integrated mLTRi into the Rosa26 locus of mouse ES cells 7 days after selection with G418. **C'**. The presence of 250μM IPTG in the culture media induces robust expression of EGFP 7 days after G418 selection. **D**. Bright field image of stably integrated mLTRi 30 days after selection with G418. **D'.** The presence of 250μM IPTG in the culture no longer induces EGFP expression 30 days after G418 selection. Scale bars, 200μm.

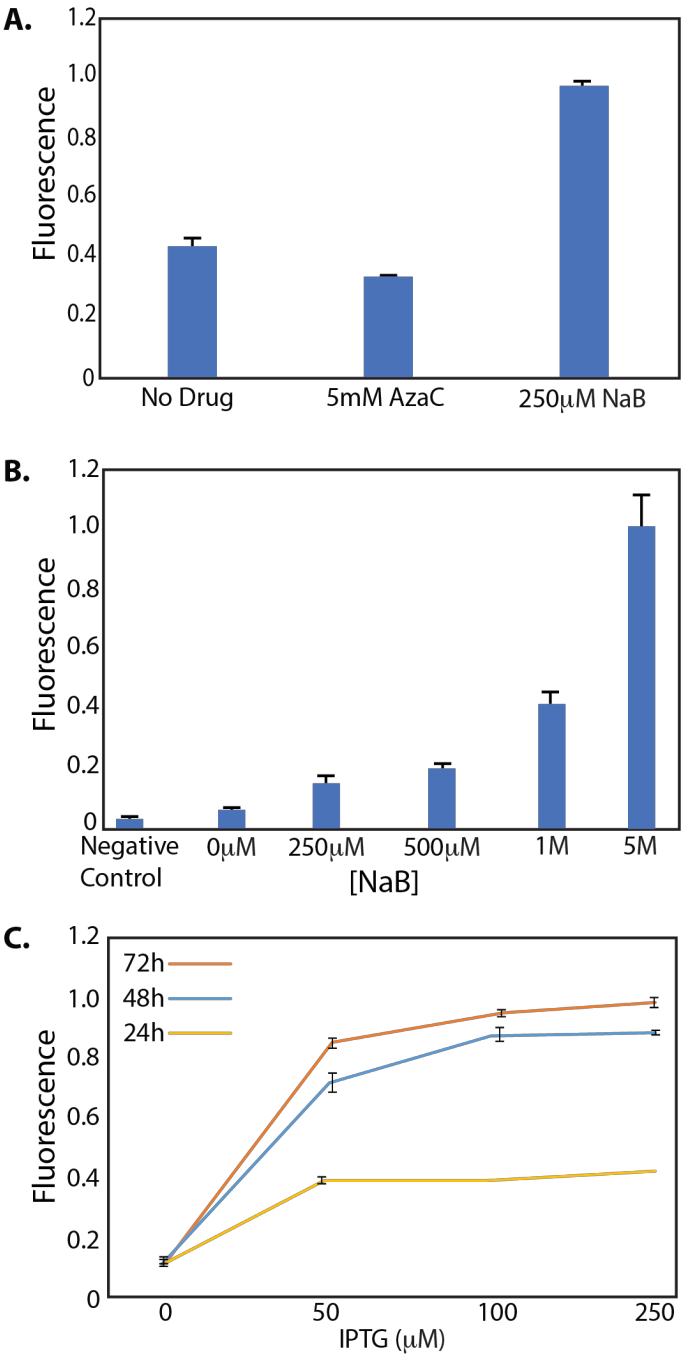

Figure 4: Quantification of EGFP expression using flow cytometry. A. ES colonies with mLTRi_EGFP integrated into the Rosa26 locus that stopped expressing EGFP in the presence of IPTG after weeks of culturing were exposed to either 5-aza-deoxycytidine (AzaC) or sodium butyrate (NaB) and 250μM IPTG in the media. EGFP expression was assessed after 48 hours of exposure to AzaC or NaB. **B.** Mean EGFP fluorescence of cells after 48 hours of growing with 250μM IPTG in the media and increasing the concentration of NaB. **C.** Mean EFGP of recovered switching dynamics Cultures were maintained with 250μM NaB and varying concentrations of IPTG over 24 hours (yellow), 48 hours (blue), and 72 hours (orange). Each data point represents the mean EGFP expression in at least three independent experiments, and the error bars represent the standard deviation between these experiments.

Supplemental Information

Area to amplify	Forward primer	Reverse primer	Size of amplicon	Type of PCR
GAPDH	GCCCAATACGGCCAAATCT	CACCTATGGTGCAACAGTATTCC	125	qPCR
ZFY1	AGCTTGACCTGCAAAGGAAGA	GCCCAATACGGCCAAATCT	137	qPCR
3x attP	CCGAAAACGATATCCTTACGACC	CTTCCCTCGTGATCTGCAACT	70	qPCR
Screening Primers	CGTGTAGAATGCCATGAGTCAAGC	CGCCTTCTTGACGAGTTCTTCTGAG	1646	conventional
TetR	CACCGCTAATTCAAAGCAACCGGTGATATCTCCTCTAGAGTCGACCTGCAGG	GTATCTTATCATGTCTGGCCAGCTCAAGTTCTGCTTTAATAAGATCTGAATTC	1440	Gibson
REF1	CAGATCTTATTAAAGCAGAACTTGAGCTGGCCAGACATGATAAGATAC	CGTATAACGTTACTGGTTTCATGGTGGCGGTTGCTTTGAATTAGCGGTGGCTTTCACAAC	2801	Gibson
MEF	CGTCACGTGGTGCGTTTTGCCTCGAGAACCATGGACCTGCAGGGCCTGAAATAAC	CTGCAGGTCGACTCTAGAGGAGATATCACCGGTTGCTTTGAATTAG	1566	Gibson
LacI	GTTGTGAAAGCCACCGCTAATTCAAAGCAACCGCCACCATGAAACCAGTAACGTTATACG	GCACGTTTTGTGTCATTGGGGAAACCTGCTCTCAAACCTTCCTCTTCTTCTTAGGAGGCC	1188	Gibson
IRES	GGCCTCCTAAGAAGAAGAGGAAGGTTTGAGAGCAGGTTTCCCCAATGACACAAAACG	CTGTGAGGACTGAGGGGCCTGAAATGAGCCCAGCTTTCTATGCAACCCAAGGAC	2018	Gibson
RNAi	CTGAGTCCTTGGGTTGCATAGAAAGCTGGGCTCATTTCAGGCCCCTCAG	GTTATTTCAGGCCCTGCAGGTCCATGGTTCTCGAGGCAAAACGCACCACGTG	1056	Gibson
TRZ	CGGACCGGTAGCTAGCATCCGGAAAGATCTCTGTACAAGTAA	CGCCCAATACTCATGATTCTCGAGCGCGTTGGCC	404	Gibson

	AGCGGCCGC	GATTCATTAATG		
Amp Origin	AACGCGCTCGAGAATCATGAGTATTGGGCGCTCTTCCGC	CAATAATCAATGTCGACACCACGTGACGAAAGGGCCTCGTGATAC	2005	Gibson
Intron w/LacO	CTCTGCTAACCATGCCTACAGAGATTTAAAGCTCTAAGG	GAGATCTTTCCGGATGCTAGCTACCGGTCCGCAAGCTTGGTTGGAATC	447	Gibson
Bactin Promoter	CTTTCGTCACGTGGTGTCGACATTGATTATTGACTAG	CTTTAAATCTCTGTAGGCATGGTTAGCAGAGGCTC	1630	Gibson
Genome 1	TTTGGAACAAGAGTCCGTCTCCACCGGACGCGGCCATGG	CCTGGTTTATGCCTTCTAGAAAGACTGGAGTTGCAGATCACGAGGG	943	Gibson
Azurite	GGTGGTTCCTTTTGGAATGCCTCGAGCTCATGACCAAAATCCCTTAAC	TGGCCGCGTCCGGTGGAGACGGACTCTTGTTCCAAACTGG	2450	Gibson
RF	TCTGCAACTCCAGTCTTTCTAGAAGGCATAAACCAGGTCGTAAGGATATC	CCAGGTTAGCCTTTAAGCCGTCGAGGCCGCGAAGTTCCTATAC	2161	Gibson
Genome 2	GCCTCGACGGCTTAAAGGCTAACCTGGTGTGTG	AGCTCGAGGCATTCCAAAAGGAACCACCTTTTAC	980	Gibson

Supplementary Table: Table of primers used in the study.

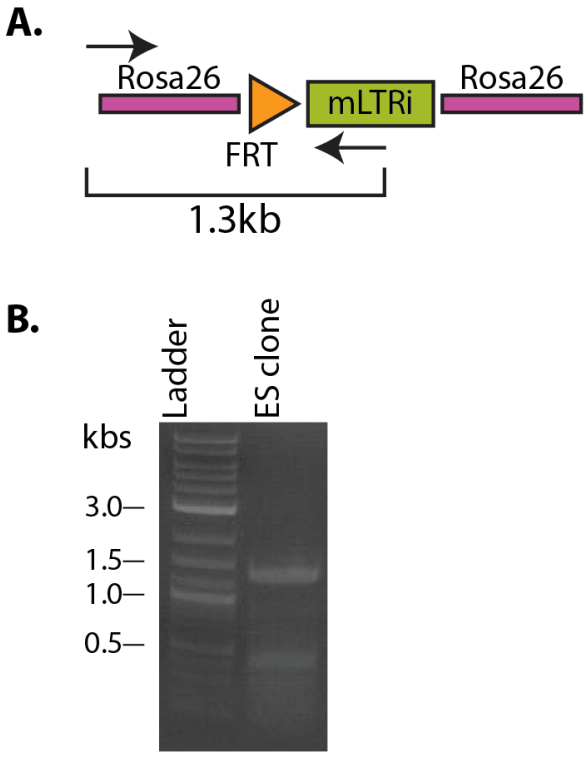

Figure S1: Screening docked mLTRi genetic circuit. A. Schematic of primer design for PCR of genomic DNA with mLTRi integrated into the Rosa26 docking site. **B.** PCR results confirming on-site integration with primers that span the genome and part of the inserted genetic circuit.