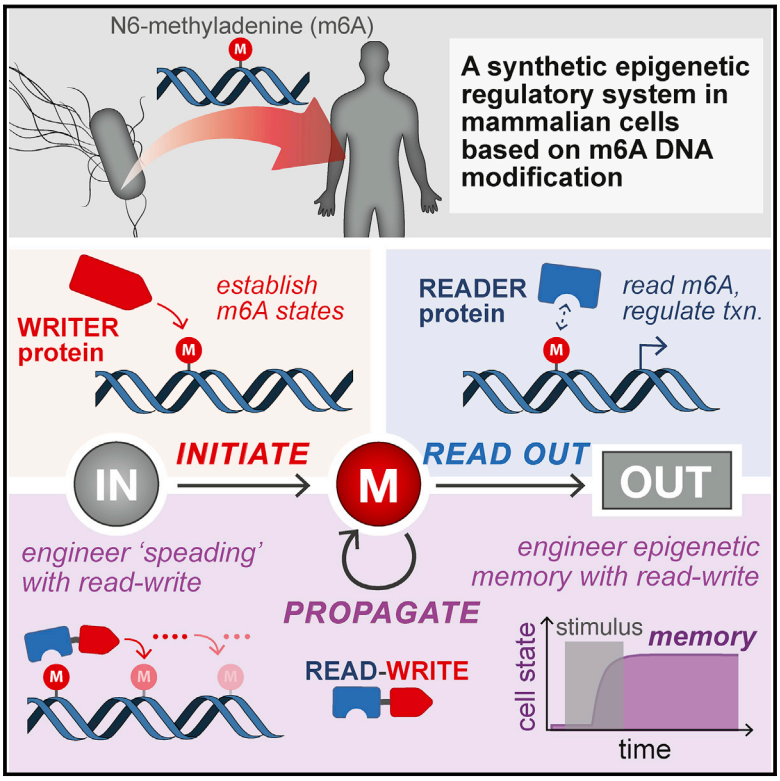
Article

Cell

Engineering Epigenetic Regulation Using Synthetic Read-Write Modules

Graphical Abstract



Authors

Minhee Park, Nikit Patel, Albert J. Keung, Ahmad S. Khalil

Correspondence

ajkeung@ncsu.edu (A.J.K.), khalil@bu.edu (A.S.K.)

In Brief

A synthetic, modular, and programmable read-write system allows isolated and orthogonal epigenetic control in mammalian cells.

HIGHLIGHTS

- A synthetic epigenetic regulatory system in human cells using m6A DNA modification
- Engineered writers and readers of m6A enable construction of regulatory circuits
- Read-write circuits drive spatial propagation and hallmarks of chromatin spreading
- Read-write circuits enable epigenetic memory of transcriptional states

Park et al., 2019, Cell *176*, 1–12
January 10, 2019
https://doi.org/10.1016/j.cell.2018.11.002

CellPress

Please cite this article in press as: Park et al., Engineering Epigenetic Regulation Using Synthetic Read-Write Modules, Cell (2019), https://doi.org/10.1016/j.cell.2018.11.002

Article

Cell

Engineering Epigenetic Regulation Using Synthetic Read-Write Modules

Minhee Park,[1,2] Nikit Patel,[1,2] Albert J. Keung,[3,*] and Ahmad S. Khalil[1,2,4,5,*]
[1]Biological Design Center, Boston University, Boston, MA 02215, USA
[2]Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA
[3]Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, NC 27606, USA
[4]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA
[5]Lead Contact
*Correspondence: ajkeung@ncsu.edu (A.J.K.), khalil@bu.edu (A.S.K.)
https://doi.org/10.1016/j.cell.2018.11.002

SUMMARY

Chemical modifications to DNA and histone proteins are involved in epigenetic programs underlying cellular differentiation and development. Regulatory networks involving molecular writers and readers of chromatin marks are thought to control these programs. Guided by this common principle, we established an orthogonal epigenetic regulatory system in mammalian cells using N6-methyladenine (m6A), a DNA modification not commonly found in metazoan epigenomes. Our system utilizes synthetic factors that write and read m6A and consequently recruit transcriptional regulators to control reporter loci. Inspired by models of chromatin spreading and epigenetic inheritance, we used our system and mathematical models to construct regulatory circuits that induce m6A-dependent transcriptional states, promote their spatial propagation, and maintain epigenetic memory of the states. These minimal circuits were able to program epigenetic functions *de novo*, conceptually validating "read-write" architectures. This work provides a toolkit for investigating models of epigenetic regulation and encoding additional layers of epigenetic information in cells.

INTRODUCTION

Genetically identical cells can produce distinct gene expression and phenotypic states that persist through cell division, a capability that is fundamental to the processes of environmental adaptation, cellular differentiation, and multicellular development. These heritable states, which do not involve changes in DNA sequence, are maintained and transmitted by self-propagating epigenetic mechanisms that persist in the absence of an initial stimulus. Chemical modifications to DNA and histone proteins have been implicated in these epigenetic programs (Berger, 2007; Bernstein et al., 2007; Feinberg, 2007; Kouzarides, 2007), and mechanisms for the propagation of certain modifications have been proposed (Bonasio et al., 2010; Moazed, 2011). These commonly invoke a core regulatory motif involving molecular species that perform basic operations on chromatin, namely, "writers" that place marks and "readers" that interpret them. To investigate this core module and obtain an understanding of the basic principles of epigenetic control, it would be useful to develop a synthetic system that could establish and drive epigenetic states *de novo*.

Studies of natural chromatin systems have identified many molecular components that regulate the placement and recognition of DNA and histone modifications, and collectively these studies have proposed a set of minimal ingredients for a bona fide epigenetic system: (1) sequence-specific placement of a modification (establish); (2) recruitment of protein effectors to the modification to mediate transcriptional changes (read & regulate); and (3) a mechanism for self-propagation that persists in the absence of an inducing signal (propagate) (Gardner et al., 2011; Moazed, 2011) (Figure 1). Combined together, these modules are thought to regulate complex epigenetic phenomena, such as the formation of silent heterochromatic domains in a variety of organisms (Beisel and Paro, 2011; Grewal and Moazed, 2003; Moazed, 2011; Ratna et al., 2009). Here, a propagation mechanism is used to spread histone modifications along the chromatin template away from a nucleation site to create an altered domain. Once established, these domains and their transcriptional states can be maintained through cell division. While molecular details of these propagation mechanisms vary across chromatin systems and organisms, a common theme is a core "read-write" motif (Figure 1). Exemplified by regulators such as Swi6/Clr4 in *S. pombe* (Ragunathan et al., 2015) and HP1α/Suv39h in mammals (Lachner et al., 2001), these are believed to function as positive feedback loops by recognizing pre-existing marks and consequently mediating the placement of the same modification on a nearby or adjacent template (e.g., to enable re-establishment after cell division) (Al-Sady et al., 2013).

The complexity of natural chromatin networks can make it difficult to decipher the principles underlying epigenetic regulation. Our approach was to design a synthetic system by placing programmable control over the basic operations of writing and reading a chemical modification in cells. The functional modules of a minimum epigenetic system could be constructed with these operations and subsequently used to engineer regulatory circuits in order to explore their capacity to generate higher-order behaviors, such as epigenetic memory. In principle, the

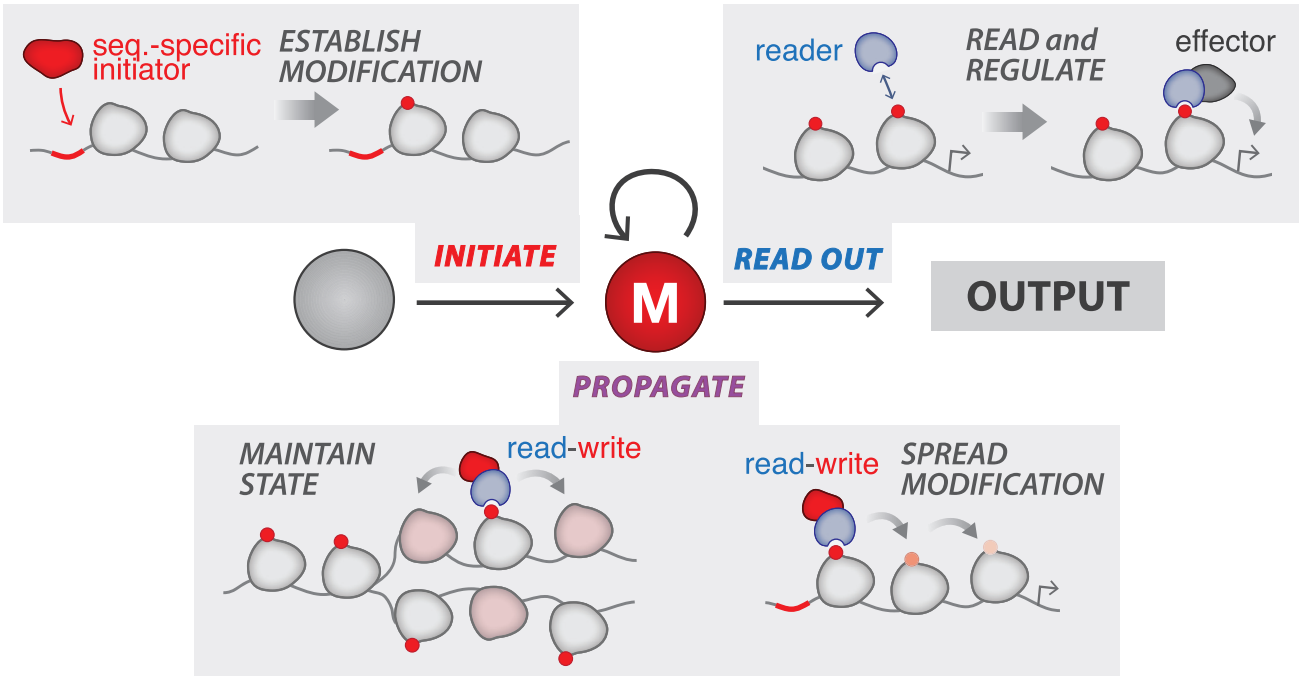


Figure 1. The Basic Functional Modules of an Epigenetic Regulatory System

(1) Initiate: "initiators" establish chromatin modifications at sequence-specific locations; (2) readout: "reader" proteins recognize modifications and mediate recruitment of regulators to establish transcriptional states; (3) Propagate: these states are propagated in the absence of the initial stimulus by read-write positive feedback mechanisms, whereby recognition of pre-existing marks is coupled to the placement of new modifications. See also Figure S1.

synthetic approach has certain advantages because natural regulatory networks are extended with still many unclear links between chromatin modifications and regulators as well as pervasive cross-talk (Lee et al., 2010). Thus, a first design challenge to developing a synthetic system is establishing well-defined, orthogonal interactions. To address this, we exploited N6-methyladenine (m6A). In contrast to cytosine methylation, which is abundant in animals and typically acts to repress genes (Bernstein et al., 2007), m6A is rarely found in metazoan genomes, and its existence and potential function remain unclear in human cells (Heyn and Esteller, 2015; O'Brown and Greer, 2016). The orthogonal properties of DNA adenine methylation were previously harnessed to develop technology for mapping chromatin-associated proteins in eukaryotic genomes (Kind et al., 2013; van Steensel and Henikoff, 2000). By transplanting this modification into human cells, we hypothesized that we could establish defined interactions for reading and writing, minimize cross-interference with pre-existing chromatin systems, and enable rapid construction of regulatory circuits that encode new and desired functions. Analogously, in natural evolution, it has been proposed that the recent emergence of the phosphotyrosine modification presented similar opportunities for rapidly evolving signal transduction systems with new functions critical to metazoan biology (Lim and Pawson, 2010).

Here, we have used m6A as the basis of encoding an additional, synthetic layer of epigenetic information in human cells. We developed synthetic factors that write and read m6A and used them to build the functional modules required of an epigenetic system. By combining these modules and identifying relevant biochemical parameter spaces using a quantitative model, we created regulatory circuits with self-perpetuating properties that can drive epigenetic behaviors, such as tunable spatial propagation of m6A marks and epigenetic memory of m6A-dependent transcriptional states. Our synthetic system thus provides a platform for programming epigenetic functions in mammalian cells and examining the core regulatory architectures underpinning epigenetic regulation.

RESULTS

Synthetic Initiator Enables Targeted m6A Enrichment at Reporter Loci

We first sought to develop a synthetic initiator module (synI) capable of establishing m6A marks in a sequence-specific manner at designer reporter loci integrated in the human genome. The general design of the module is a fusion of a Dam (*E. coli* DNA adenine methyltransferase) "writer" domain, which catalyzes methylation of adenines in GATC motifs, and an engineered zinc finger (ZF) protein, which specifically binds a 20-bp synthetic binding sequence (BS) (Figures 2A and S1). We designed two classes of reporters for this study—the Clustered Reporter and Interspersed Reporter (Figures 2A and 3A)—and generated respective reporter cell lines by singly integrating these constructs into the HEK293FT genome (see STAR Methods; Figure S1). The two reporters feature different arrangements of BS and GATC arrays placed upstream of a

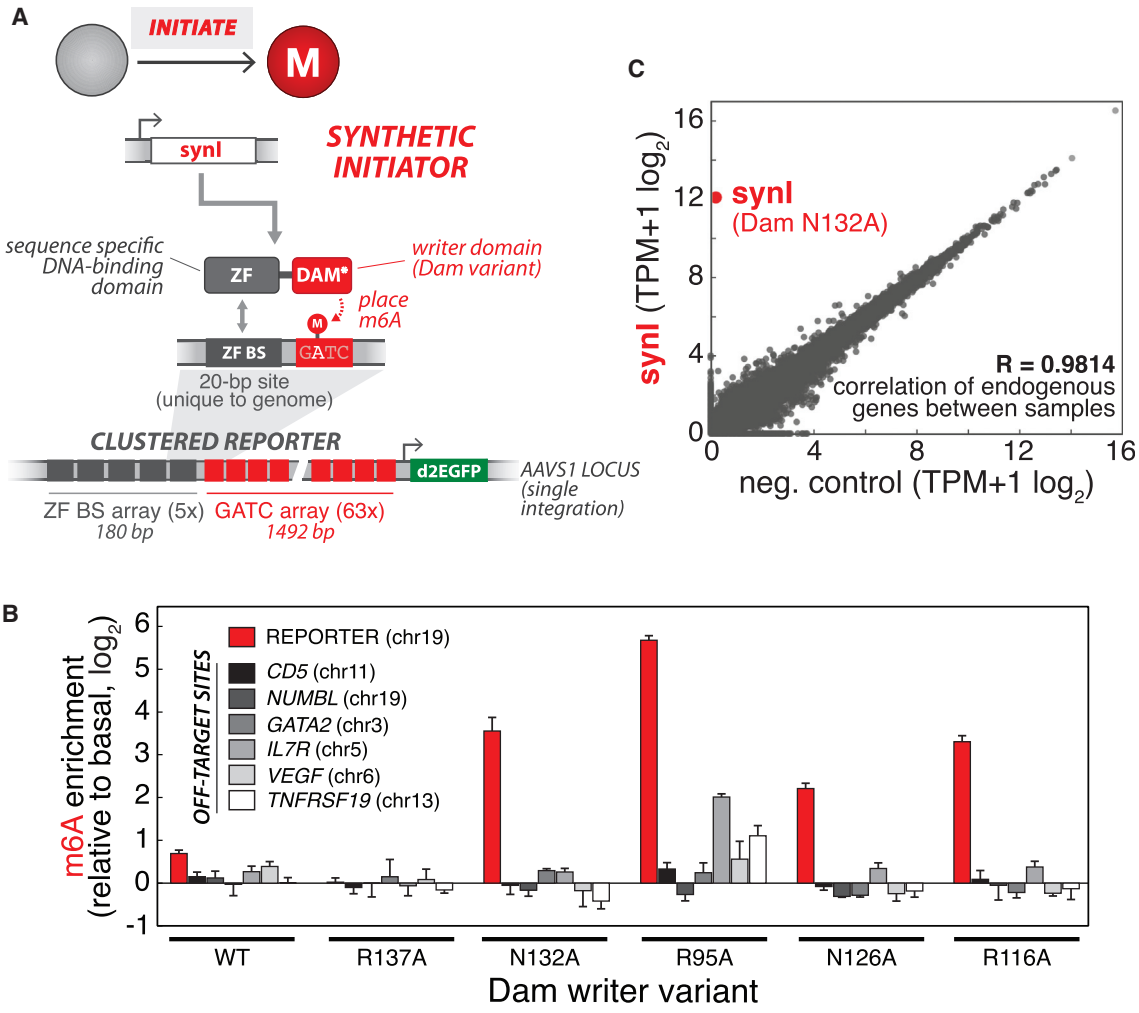


Figure 2. Engineering a Synthetic Initiator to Establish N6-Methyladenine DNA Modifications at Target Reporter Loci in Human Cells

(A) Design of synthetic initiator module (synI). synI is a fusion of a Dam (DNA adenine methyltransferase) "writer" domain and an engineered zinc finger (ZF), which specifically binds a 20-bp synthetic binding sequence (BS). synI enables *de novo* placement of m6A marks at designer reporters integrated into 293FT cells. For these experiments, we used stable cell lines harboring a singly integrated Clustered Reporter, with ZF BS and GATC arrays upstream of a *pMinCMV* driving expression of destabilized GFP (d2EGFP), as the background strain.

(B) Screening Dam variants for synI factors that induce sequence-specific enrichment of m6A at target sites. Quantification of m6A enrichment at target reporter (red) and off-target, GATC-containing endogenous loci (gray shades) by transfected synI constructs composed of different Dam mutants. Off-target loci were chosen to represent different chromosomal locations. m6A enrichment is obtained by measuring fraction methylation at a single GATC probe site in the locus of interest using m6A-qPCR, and normalizing to basal methylation induced by the Dam variant not fused to ZF (STAR Methods; Figure S2). (n = 3; error bars, SD).

(C) Expression of synI has minimal effect on the transcriptome. Correlation of transcriptome from RNA-seq measurements for reporter cells transfected with synI versus empty plasmid. Correlation coefficient of endogenous genes between samples was calculated using $\log_2$-transformed expression values. mRNA corresponding to synI is labeled. The data are representative of two biological replicates.

See also Figure S2.

promoter driving expression of a destabilized EGFP (d2EGFP). The Clustered Reporter, which features BS motifs directly upstream of a long GATC array (spanning 1.5 kb), was designed to enable methylation measurements at single GATC resolution and accordingly to facilitate studies of spatial dynamics. The Interspersed Reporter, with intermixed BS and GATC motifs, was designed to couple m6A states with transcriptional reporter outputs and accordingly to facilitate temporal studies of transcriptional dynamics and memory.

To identify a synI factor that can preferentially nucleate our reporter locus, we generated a library of Dam variants (DAM*, Figures 2A and S1). Because the wild-type Dam enzyme is known to be highly active, we chose to express the library at low levels (using a minimal CMV promoter, *pMinCMV*) to minimize global (non-specific) methylation (Figure S2G) (van Steensel and Henikoff, 2000). Additionally, we hypothesized that, by lowering intrinsic Dam activity and DNA affinity through mutations, we could identify a variant whose activity is more highly dependent

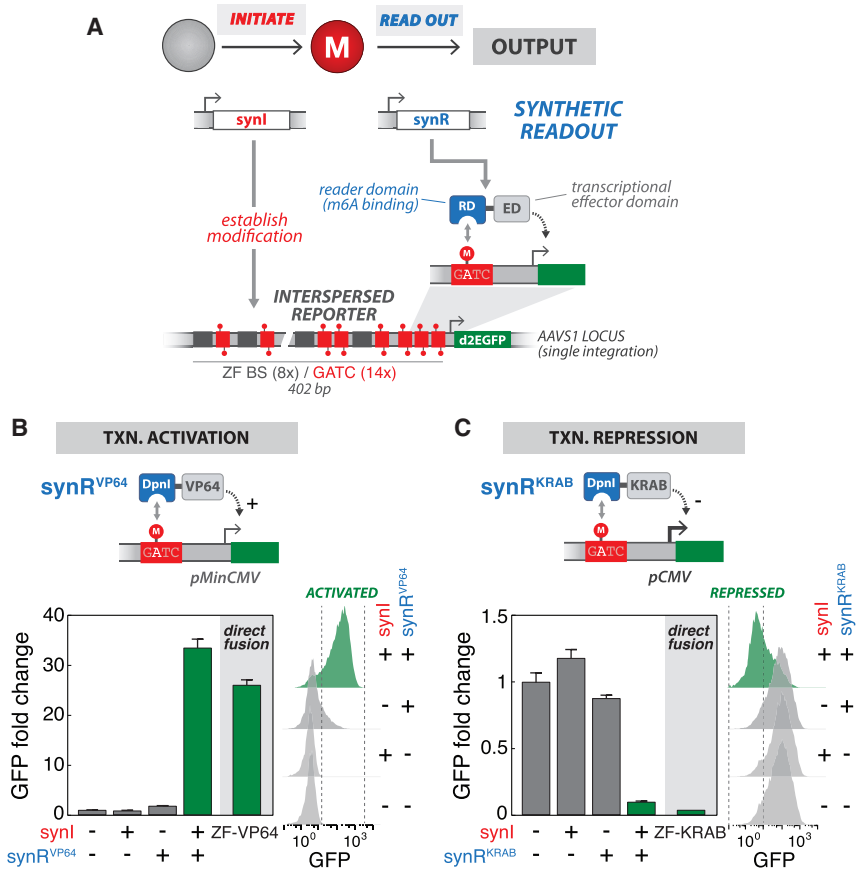


Figure 3. Programming m6A-Dependent Transcriptional States with Engineered Reader Modules

(A) Design of synthetic readout module (synR). synR is a fusion of an m6A "reader" domain (RD, binding domain of DpnI [aa 146–254]) and a transcriptional effector domain (ED). m6A marks established by synI are specifically recognized by synR, which in turn regulates transcriptional activity of a reporter gene. For these experiments, we used stable cell lines harboring a singly integrated Interspersed Reporter, with intermixed ZF BS and GATC sites upstream of a promoter (*pMinCMV* for activation or *pCMV* for repression), as the background strains.

(B) Programming m6A-mediated transcriptional activation. Top: schematic of the synR^{VP64} module, a fusion of DpnI m6A RD and VP64 transcriptional activation domain, which drives activation of a reporter gene via m6A recognition. Bottom: GFP fluorescence intensity, measured by flow cytometry, for cells transfected with indicated combinations of synI and synR^{VP64} expression constructs, or a direct ZF-VP64 fusion. Bottom left shows fold change of geometric mean GFP intensity normalized to the −/− condition (n = 3; error bars, SD); bottom right shows raw flow cytometry distributions.

(C) Programming m6A-mediated transcriptional repression. Top: Schematic of the synR^{KRAB} module, a fusion of DpnI m6A RD and KRAB transcriptional repressive domain, which drives repression of a reporter gene via m6A recognition. Bottom: GFP fluorescence intensity, measured by flow cytometry, for cells transfected with indicated combinations of synI and synR^{KRAB} expression constructs, or a direct ZF-KRAB fusion (n = 3; error bars, SD).

See also Figure S3.

on ZF binding (McNamara et al., 2002; Smith and Ford, 2007). We generated two expression constructs for each variant: fusion to ZF (synI, targeted) and mCherry (synI^{NT}, non-targeted) (Figures S2D and S2E). We then transfected each construct into the Clustered Reporter cell line and used an adapted m6A-qPCR assay to measure adenine methylation frequency at the reporter (see STAR Methods; Figures S2A–S2C). We found that single mutations to residues that mediate DNA phosphate group contact, which are known to affect the biochemical activity of Dam (Coffin and Reich, 2009; Horton et al., 2006), generally showed an enrichment in reporter m6A levels. Here, m6A enrichment is defined as targeted methylation (by synI) normalized to basal methylation, induced by the same Dam variant (synI^{NT}, non-targeted) (Figures S2D–S2G). In order to identify a factor with minimal off-target activity, we screened these synI variants and compared m6A enrichment at the reporter (red) to "off-target" endogenous loci (gray), chosen to represent different chromosomal locations and GATC frequencies (Figure 2B). We selected the ZF-Dam (N132A) fusion, which will henceforth be referred to as synI. SynI expression was found to have minimal effect on the 293FT transcriptome (Figures 2C, S2H, and S2I), cell cycle, and cell viability (Figures S2J and S2K). Together, these results establish a synthetic initiator module capable of sequence-specific placement of m6A marks at reporter loci in human cells.

Programming m6A-Dependent Transcriptional States with Engineered Reader Modules

Chromatin modifications can modulate gene transcription through several mechanisms, including through reader proteins that recognize specific, or combinations of, marks and recruit transcriptional effector functions (Berger, 2007; Gardner et al., 2011; Kouzarides, 2007). Armed with the ability to nucleate m6A marks, we next sought to engineer reader modules (Haynes and Silver, 2011) that recognize and translate these modifications into defined transcriptional outputs. We designed a synthetic readout module (synR), composed of fusions of an m6A reader domain (RD, binding domain of *S. pneumoniae* DpnI), which selectively recognizes methylated GATC (Kind et al., 2013; Siwek et al., 2012), and modular transcriptional effector domains (EDs) (Figure 3A). We generated expression constructs for synR modules harboring different EDs: synR^{VP64} (VP64 activation domain), synR^{KRAB} (KRAB repressive domain), and synR^{HP1} (HP1α chromo shadow domain). We then transfected combinations of the constructs into Interspersed Reporter cell lines (harboring either *pMinCMV* for synR activators or full-length *pCMV* for synR repressors), and measured GFP reporter output (see STAR Methods). The synR modules induced significant reporter activation or repression, only when expressed in combination with synI; moreover, these transcriptional changes were similar in levels to those induced by a direct transcriptional

regulator (direct ZF-ED fusions) (Figures 3B, 3C, and S3B–S3D). We further confirmed that reporter m6A levels were enriched only in cells expressing synI (Figure S3A), and that the presence of GATC motifs was required for transcriptional regulation by synI and synR (Figure S3E). This two-module circuit (synI, synR) was found to function on both integrated and episomal reporters (Figures 3B, 3C, and S3B–S3D) as well as to have minimal and orthogonal effects on the cellular transcriptome (Figure S3F). Taken together, we have developed a synthetic two-module regulatory system that utilizes engineered readers to establish m6A-dependent transcriptional states and logic.

To increase the versatility of this synthetic gene regulatory system, we next sought to develop a version in which synI activity could be readily directed to desired sequences without the need to redesign its DNA targeting domain. We created a CRISPR-guided version of the initiator by fusing the selected Dam (N132A) variant to the *S. pyogenes* dCas9 protein (synI$^{d\text{-}Cas9}$) (Figures S1 and S3G). When expressed in combination with single guide RNAs (gRNAs) targeting various locations in the BS array of the Clustered Reporter, we found that synI^{dCas9} was capable of preferentially enriching m6A levels at our reporter for certain gRNAs (Figure S3G). Moreover, when combined with synR, two-module circuits based on synI^{dCas9} (in place of ZF-based synI) were also able to drive transcriptional regulation of reporters (Figure S3H).

Finally, we wondered how our reporter constructs, which we have shown can be artificially modified and regulated by our synthetic m6A system, generally compare with naturally occurring GATC distributions in the human genome. This might inform future applications or improvements of our regulatory system for arbitrary genomic contexts, where GATC distributions cannot be precisely controlled. Based on a genome-wide bioinformatics analysis, we found that GATC sites are indeed naturally present in most human promoter regions, with a median of ~3–4 motifs (Figures S3I and S3J). Moreover, when we tested the performance of our synthetic system for a set of reporters harboring different numbers of ZF BS and GATC motifs, we found that it is possible to regulate reporters with GATC frequencies equivalent to those of natural promoter frequencies, albeit at more modest levels (Figure S3K).

Engineering Spatial Propagation with Read-Write Regulatory Circuits

Cells have mechanisms for propagating chromatin states in space and time, mechanisms that mediate the "spreading" of natural chromatin domains associated with silent transcription, and the faithful transmission of these altered domains across cell division. It is thought that a core regulatory feature driving these self-perpetuating mechanisms is the "read-write" system (Al-Sady et al., 2013). Inspired by this, we focused first on developing a minimal read-write (RW) module to construct regulatory circuits that drive spatial propagation of m6A. To enable these studies, we developed a small-molecule inducible initiator (synI^{IND}), which uses abscisic acid (ABA)-induced dimerization to enable temporal control over initiation of m6A states (Liang et al., 2011; Figure S4). We then designed a synthetic RW module (synRW), which is a fusion of m6A RD (binding domain of DpnI) and Dam writer domain (DAM*) (Figure 4A).

Our objective was to identify RW circuit designs capable of driving spatial propagation. Specifically, we sought to identify synRW variants that, when combined with synI^{IND} and synR modules, could spread m6A from a nucleation site. To do this, we generated a synRW variant library (varying synRW expression levels and Dam mutants) and devised a simple, phenotypic screen for spatial propagation behavior (Figure 4B). For a full description of the library, screen, and analysis, see STAR Methods. Briefly, the screen leverages the Clustered Reporter's long GATC domain (with 20-bp inter-GATC spacing) separating the nucleation site from the reporter gene. Devoid of a propagation mechanism, reporter cells stably expressing the two-module circuit (synI^{IND}, synR^{VP64}) are not activated (Figure S6A). As a result, the synRW library can be screened in these cells for candidates that produce reporter activation (+synI^{IND}, Figure 4B), as well as in cells lacking initiator to screen out spurious cases for which downstream reporter activation is induced independent of m6A nucleation (–synI^{IND}, Figure 4B). Variants emerging from this screen would, in principle, represent potential candidates for the RW module of three-module regulatory circuits (synI^{IND}, synR^{VP64}, synRW) that drive inducible spatial propagation leading to reporter activation.

Clustering analysis of our screen results based on GFP expression revealed a strong clustering by synRW Dam mutants (Figure 4B). Specifically, we observed a relationship between methylation activity of synRW and propagation, quantified by a "spreading score" metric we developed to score the phenotypic outcomes of the screen (Figure 4C; STAR Methods). This suggested to us that intermediate writer activity (i.e., intermediate levels of Dam methylation activity) may be an important design criterion for the synRW module to produce these propagation phenotypes. To explore the generality of this result, we turned to quantitative modeling, adapting a previously described model of chromatin spreading dynamics (Hodges and Crabtree, 2012) to capture the essential features of our system (see STAR Methods; Figure S5). We used our model to simulate and interrogate how synRW properties, such as methylation writer activity, affect spatial m6A profiles; this revealed a similar relationship to that observed from our experimental screen (Figure 4C, blue line).

Supported by our screen and simulation results, we selected a high-scoring, intermediate-activity synRW (*pUBC*: DpnI-Dam [R95A]) and integrated this construct to generate cells stably expressing a full RW "propagation circuit" (synI^{IND}, synR^{VP64} synRW). We then tested the circuit by triggering m6A nucleation and measuring m6A profiles across the domain over time. Cells expressing our circuit exhibited a growing m6A domain over time, in contrast to a control circuit lacking synRW (Figures 4D and S6C). Moreover, and as desired, robust propagation was dependent on m6A nucleation, as cells lacking initiator (–synI^{IND}) did not exhibit propagating m6A domains (Figure S6C). Interestingly, and consistent with recent studies on the dynamics of heterochromatin formation in mammalian cells (Hathaway et al., 2012), establishment of this m6A domain occurred over relatively slow timescales (~days). Consistent with our previous observations for synI, expression of synRW has no adverse effects on the cell cycle and viability (Figures S2J and S2K) and led to only modest increases in methylation at chosen off-target,

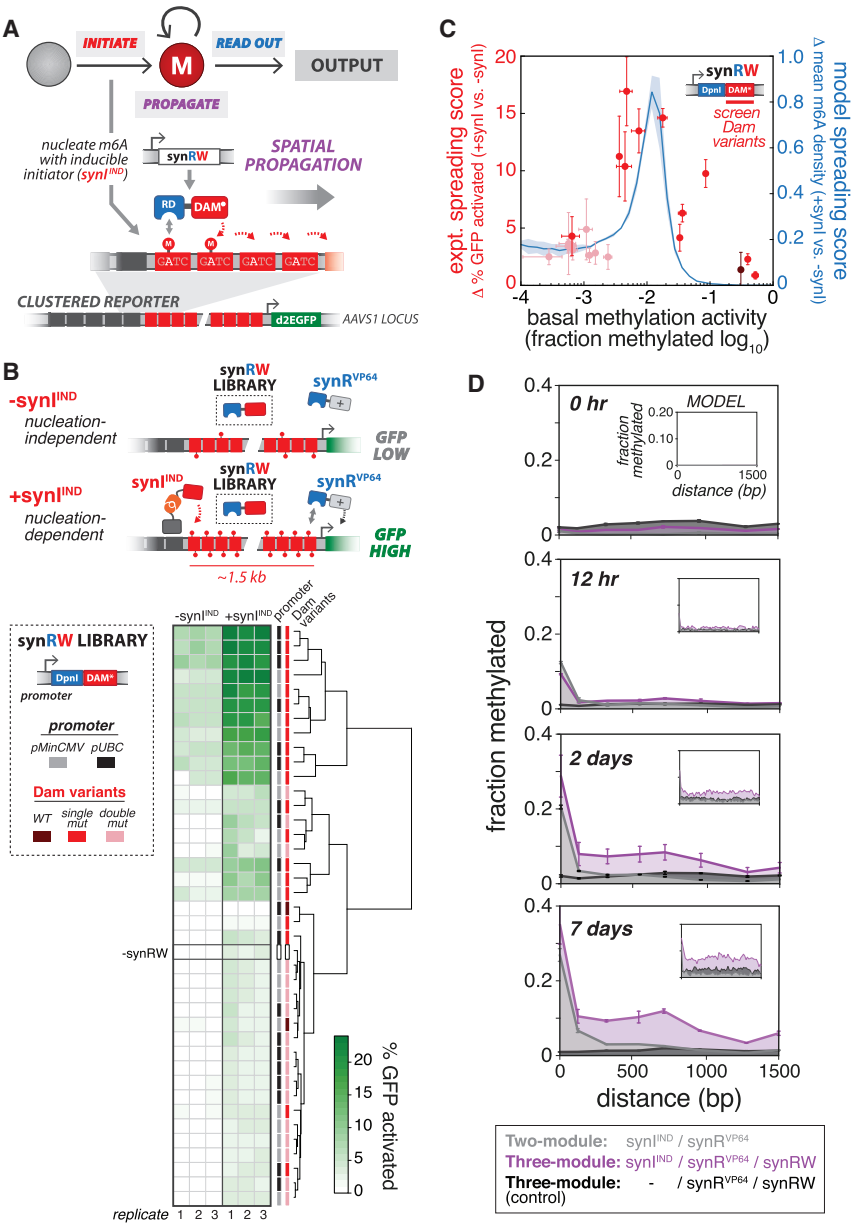


Figure 4. Engineering Spatial Propagation of m6A with Three-Module Read-Write Circuits

(A) Design of synthetic read-write module (synRW) for propagating m6A modifications over a domain. synRW is a fusion of DpnI m6A RD and a Dam writer domain (Dam variant). synRW recognizes pre-existing m6A marks and catalyzes methylation of nearby GATC motifs, creating local reinforcement and spreading of m6A to larger domains. For these experiments, we used Clustered Reporter cell lines, with *pMinCMV* driving expression of d2EGFP, as the background strain.

(B) Screening synRW variants for nucleation-dependent spreading behavior. Top: schematic of the "spatial propagation screen" design. A library of synRW variants was screened for the ability to activate a distal reporter gene in cells with synthetic initiator (+synI^{IND}, GFP High), but not in cells lacking initiator (–synI^{IND}, GFP Low). Bottom left: the synRW library featuring different promoter strengths and Dam mutant writers. Each member of the library was individually transfected into Clustered Reporter lines stably expressing synR^{VP64}, either with (+synI^{IND}) or without (–synI^{IND}) stable expression of the inducible initiator. Bottom right: screen results. Heatmap of percentage of GFP-activated cells, quantified by flow cytometry 4 days following transfection of synRW (with continuous 200 μM ABA induction of synI^{IND}). Hierarchical clustering analysis based on similarity in percentage of GFP-activated cells is shown.

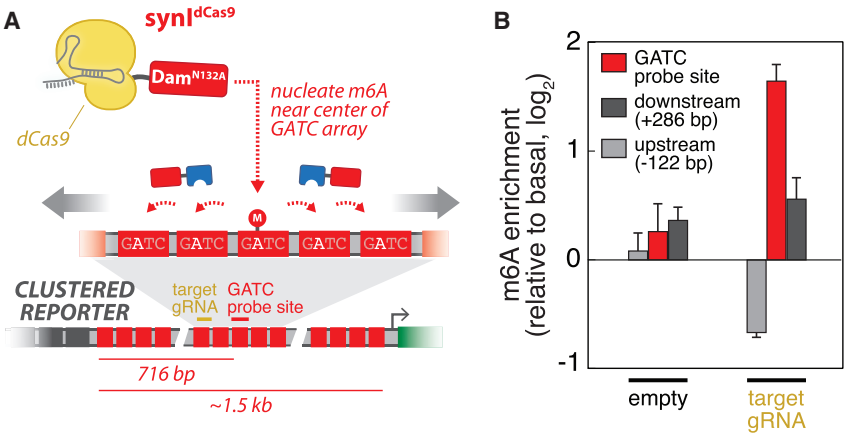
(C) Summary and modeling of spatial propagation screen. We defined quantitative metrics to score propagation propensity for synRW library members: "expt. spreading score" (red; n = 3, error bars, SD) is the difference between percentage of GFP-activated cells with and without synI^{IND} measured from the experimental screen; "model spreading score" (blue; mean ± SD) is the difference between model-computed m6A density at promoter-proximal sites with and without synI (see STAR Methods; Figure S5). Shades of red correspond to Dam mutant writers: WT, single mutants, double mutants (dark to light, respectively).

(D) m6A profiles measured across the GATC array and over time for cells stably expressing the three-module "propagation circuit" (purple) or circuits lacking either synRW (gray) or synI^{IND} (black). Cells were continuously induced with 200 μM ABA. Model simulations are shown in insets (with b_{ZF} = 10, b_{DpnI} = 100, see STAR Methods; Figure S5) (n = 3; error bars, SD).

See also Figures S4, S5, and S6.

endogenous loci (Figure S2L). These results demonstrate the development of a three-module regulatory circuit that can drive spatial m6A propagation in order to synthetically modify a domain and regionally control the expression of genes.

We next explored how different compositions and expression levels of the modular components making up the propagation circuit affect spreading dynamics. A synI "trigger" and constitutive synRW were both required for m6A propagation (Figures 4D and S6C). However, the synR module, while necessary for screening and coupling m6A activity to transcriptional output states, is presumably not essential to the circuit's ability to drive m6A spreading (acting only to compete with synRW for binding sites). To test this, we generated stable cell lines with a propagation circuit lacking synR and measured m6A profiles after triggering nucleation. Indeed, these cells also exhibited propagating m6A domains with similar profiles to those of the full three-module circuit, albeit with slightly elevated overall m6A levels across the domain by day 7; these elevated m6A profiles were also observed for control circuits lacking synI or synRW (Figure S6D). These results suggest that using competitors for methylated GATC substrates—i.e., "futile" reader modules like synR—may provide strategies for tuning methylation levels and suppressing



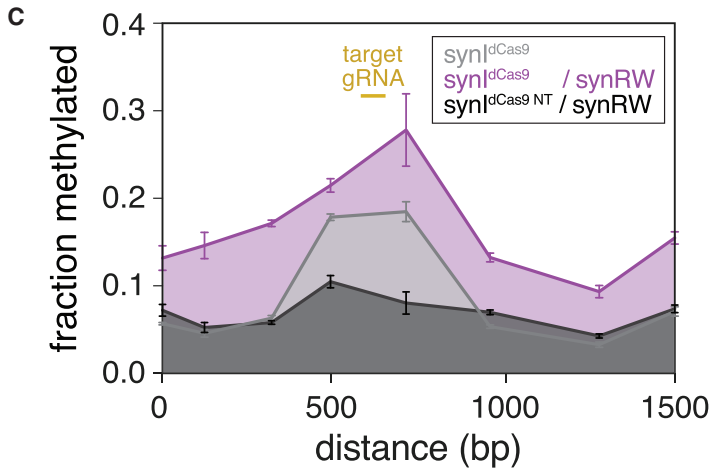


Figure 5. Synthetic Propagation Circuits Induce Bi-directional Spreading

(A) Using a CRISPR-guided synthetic initiator (synI^{dCas9}) to investigate m6A propagation symmetry. synI^{dCas9} is recruited to the center of the Clustered Reporter GATC array to establish a nucleation site and to observe synRW-induced m6A propagation away from the site (see Figures S3G and S3H).

(B) synI^{dCas9} enables targeted m6A enrichment near the center of the GATC reporter array. Quantification of m6A enrichment at a central GATC probe site (red; 716 bp downstream of ZF array) and sites located downstream (dark gray; +286 bp) and upstream (light gray; −122 bp) of the probe site following transfection of Clustered Reporter lines with synI^{dCas9} and single gRNA (or empty) cassette. m6A enrichment is obtained as previously described, normalizing to basal methylation induced by synI^{dCas9} alone (n = 3; error bars, SD).

(C) m6A profiles measured across the GATC array for Clustered Reporter cell lines transfected with: synI^{dCas9} and target gRNA (gray); synI^{dCas9}, target gRNA, and synRW (purple); synI^{dCas9}, empty gRNA, and synRW (black) (n = 3; error bars, SD).

See also Figure S6.

basal writing at desired time points, presumably through altering the kinetics of propagation.

Clustering analysis of our screen results revealed no obvious correlation or sensitivity in spreading behavior for the two different promoters used to drive synRW expression (*pMinCMV* and *pUBC*) (Figure 4B). We reasoned this was likely because the transfection conditions used in screening produced saturating concentrations of synRW for reactions on a limited number of available methylated GATC sites. To further investigate how synRW levels may affect propagation dynamics, we used our screen-selected synRW variant (*pUBC*: DpnI-Dam [R95A]) to perform a dosing experiment, in which we varied the concentration of synRW plasmid transfected into Clustered Reporter cells (with versus without synI^{IND}) and quantified difference in methylation at the promoter proximal end of the GATC array (Figure S6E). By varying plasmid concentrations, we observed a dose-dependent relationship in propagation outcomes as a function of synRW levels. The dose-response curve varied in time, shifting to lower synRW threshold concentrations by day 4 (the saturation peak was well below the concentration used in the screen (100 ng), suggesting that the conditions used in the original screen were likely producing saturating synRW for both promoters).

Next, we investigated the behavior of the propagation circuit for different configurations of the reporter. First we tested the effect of varying GATC density, which may have implications for potential applications of our synthetic epigenetic regulatory system in natural genomic contexts where GATC spatial frequencies cannot be controlled. We generated three Clustered Reporter variant lines with inter-GATC spacers of increasing length (50, 212, 800 bp) (Figure S6F); our canonical reporter uses 20-bp spacers, the shortest possible while still allowing individual GATC measurements with m6A-qPCR. m6A profiles induced by our propagation circuit (with transfected synRW) showed a monotonic decrease in the capacity for spreading from 50- to 800-bp spacers, where we observed complete extinction.

Finally, inspired by recent observations that H3K9me3 profiles grow symmetrically from a nucleation site (Hathaway et al., 2012), we wondered whether our synthetic RW circuit could also drive symmetric propagation patterns. To test this, we exploited our CRISPR-guided initiator (synI^{dCas9}), which could be directed to the center of the reporter GATC array with a single gRNA to induce m6A enrichment (Figures 5A and 5B). Upon triggering a central nucleation site, cells expressing the synRW module exhibited bi-directional enrichment in m6A, while control circuits did not (Figure 5C).

Engineering Epigenetic Memory with Read-Write Regulatory Circuits

In principle, RW positive feedback loops could also provide a mechanism to establish epigenetic memory. Thus, we wondered whether our propagation circuit could mediate the maintenance

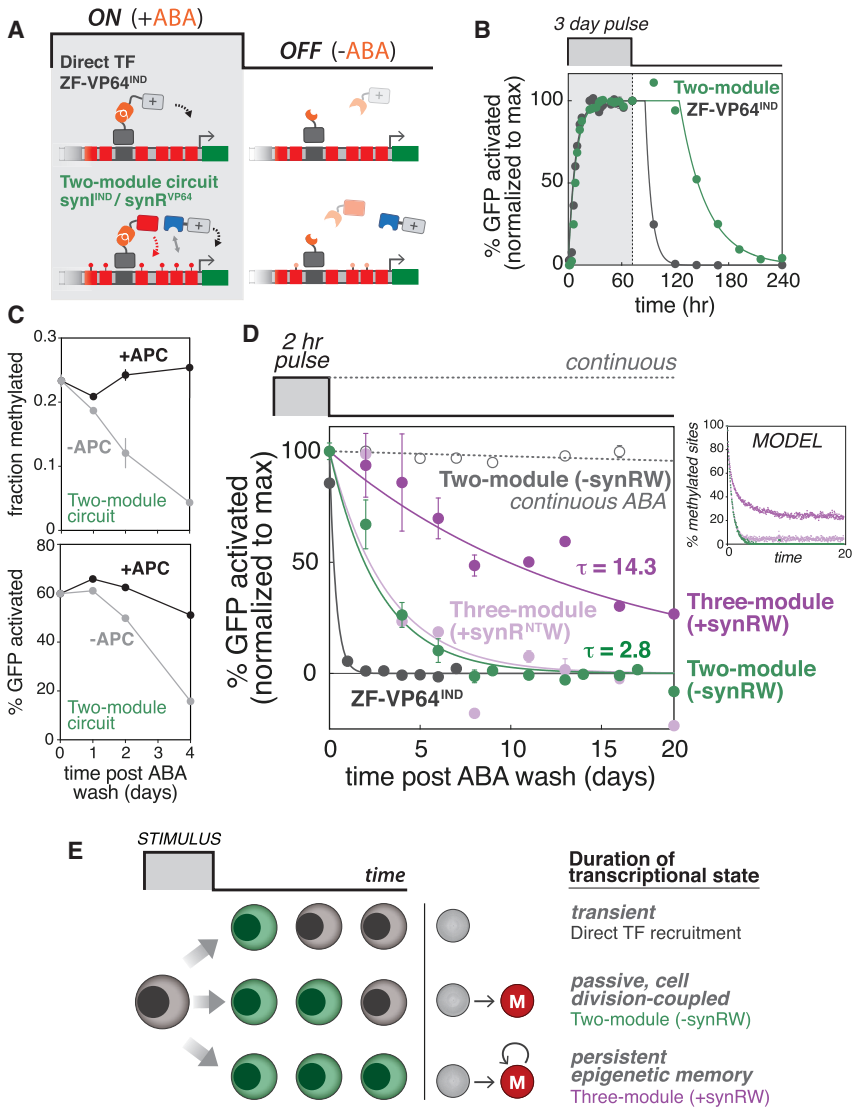


Figure 6. Engineering Epigenetic Memory with Three-Module Read-Write Circuits

(A) Schematic depicting ON (+ABA) and OFF (–ABA) states with direct TF recruitment (ZF-VP64^{IND}, top) and m6A two-module regulatory circuit (synI^{IND}, synR^{VP64}; bottom). Addition of ABA mediates dimerization to form an active ZF-VP64^{IND} (top) or synI^{IND} initiator (bottom). Removal of ABA allows analysis of deactivation dynamics and epigenetic memory. For these experiments, we used Interspersed Reporter cell lines, with *pMinCMV* driving expression of d2EGFP, as the background strain.

(B) Activation/deactivation dynamics for ZF-VP64^{IND} (gray) versus two-module m6A circuit (green) in response to a transient (3 day) ABA pulse. The percentage of GFP-activated cells was quantified by flow cytometry (see Figure S4E). Dots are data points (n = 3; error bars, SD); lines represent sigmoidal curve fits to ON phase and exponential fits with time delay to OFF phase.

(C) Maintenance of m6A (top) and transcriptional state (bottom) by inhibiting DNA replication. Cells stably expressing m6A two-module circuit were induced with a 3-day ABA pulse, and then ABA was washed out and cells were maintained in media with or without 5 μg/mL APC. Fraction methylation and percentage of GFP-activated cells were quantified at the indicated time points following ABA washout (n = 3; error bars, SD).

(D) Three-module propagation circuit induces epigenetic memory. Deactivation dynamics following transient (2 hr) ABA pulse for cells stably expressing: three-module propagation circuit (purple), the circuit lacking synRW (green), and the circuit with a reader-defective synR^{NT}W (light purple), or cells stably expressing ZF-VP64^{IND} (gray). Also plotted are cells stably expressing the two-module circuit continuously induced with 200 μM ABA (gray, dotted line). The percentage of GFP-activated cells was quantified by flow cytometry. Dots are data points (n = 3; error bars, SD); lines represent exponential fits to OFF phase. To the right are model simulation results of m6A maintenance for the different m6A circuit components (same color code). The percentage of methylated GATC sites (out of 14 available in the Interspersed Reporter) are plotted as a function of time (see STAR Methods).

(E) Proposed model for engineering transcriptional states with varying durations and epigenetic memory using synthetic m6A operations. Direct TF recruitment induces transient transcriptional states (top). Depositing m6A marks facilitates passive, cell-division coupled persistence of the induced transcriptional state (middle). Regulatory circuits utilizing read-write positive feedback can induce persistent epigenetic memory (bottom).

See also Figures S4 and S7.

and transmission of an induced transcriptional state through cell division, as compared with the transient state induced by a direct transcription factor (TF). To test this, we first followed the response of Interspersed Reporter cells stably expressing either an inducible ZF-VP64^{IND} or the m6A two-module circuit (synI^{IND}, synR^{VP64}) to a transient pulse of ABA (Figure 6A). Both exhibited rapid reporter activation followed by deactivation upon removal of ABA; however, the m6A system exhibited a delay in the timescale of deactivation (Figure 6B). In contrast to direct TF recruitment, the decay dynamics of the two-module circuit depends on the lifetime of the modified m6A state, which in the absence of an active removal mechanism is largely governed by passive dilution through semiconservative replication. We tested this by blocking DNA replication using Aphidicolin (APC). As expected, without APC, m6A levels roughly halved every 2 days with a concomitant loss of GFP-activated cells, whereas blocking DNA replication led to persistence of m6A and corresponding GFP states (Figures 6C and S7A).

We next tested whether our three-module propagation circuit could transform this "passive", cell-division-coupled state into a more durable memory. The timescale of deactivation of the two-module circuit (without RW) is approximately the same for all induction pulse lengths tested (Figure S4F), which is once again consistent with a passive dilution model governed by cell division. Thus, for our inducing signal, we selected a short (2 hr) ABA pulse because it is sufficient to induce the activated state

in cells (longer pulses led to only slight increases in fraction of activated cells) and tested whether our circuit could create a memory of this transient signal. We subjected cells to this short pulse of ABA and tracked the percentage of activated cells following removal of ABA for a total of 20 days (Figure 6D). Our propagation circuit led to higher levels of maintenance of GFP-activated cells (dark purple), extending the decay time significantly beyond ZF-VP64^{IND} (gray) and the two-module circuit lacking synRW (green), which was fully lost by 8 days. Notably, with the three-module RW circuit, approximately half of the cellular population remained GFP-positive for ~10 generations. To ensure this was not simply the result of a Dam dosage effect (by addition of synRW module), but indeed dependent on the RW mechanism, we also tested a control circuit in which the synRW RD was replaced with mCherry (synR^{NT}W). This abolished memory, returning the deactivation phase to two-module circuit levels (–synRW, light purple). Model simulations of this experiment, using the same model parameter set from the spatial dynamics simulations, effectively reproduced these dynamics (Figures 6D and S7E; STAR Methods). The memory induced by our circuit resembled GFP maintenance profiles resulting from blocking DNA replication with APC, and as before was dependent on m6A initiation by synI^{IND} (Figures S7B and S7C). Additionally, cells stably expressing the propagation circuit were actively dividing and transmitting the epigenetic state to progeny, not simply slowing cell growth (Figure S7D).

Taken together, we have demonstrated that minimal synthetic circuits utilizing RW positive feedback can establish epigenetic memory of transient signals. Additionally, by combining different modular operations, we found that the duration of a transcriptional state can be tuned from (1) a highly transient state regulated by direct TF recruitment, to (2) a passive, cell-division-coupled persistent state regulated by circuits that deposit and recognize m6A, and finally to (3) longer lasting epigenetic memory regulated by circuits that additionally utilize RW modules (Figure 6E).

DISCUSSION

Chromatin is a substrate for a complex assortment of chemical modifications made to DNA and histone proteins. These dynamic modifications can influence genome structure and orchestrate the recruitment of effector protein complexes, thereby playing essential roles in regulating gene transcription and other biological processes (Berger, 2007; Bernstein et al., 2007; Feinberg, 2007; Kouzarides, 2007). Additionally, certain modifications have been implicated as carriers of epigenetic information, and significant efforts have been focused on understanding how these marks contribute to mechanisms for transmission of heritable transcriptional states independent of DNA sequence (Bonasio et al., 2010; Moazed, 2011). Here, we used a synthetic biology approach to construct an epigenetic regulatory system in mammalian cells from first principles. Specifically, we developed a suite of synthetic factors that execute reading and writing operations with the orthogonal DNA chemical modification m6A and assembled them into regulatory architectures believed to be core engines for driving epigenetic functions. Importantly, by creating variants of these factors that capture a range of biochemical properties and then combining them into circuits both experimentally and computationally, we were able to identify operating regimes for an array of epigenetic behaviors.

The first component of our system (synI) is a factor that enables sequence-specific establishment of m6A. We found that to achieve targeted m6A enrichment at artificial reporter loci, it was necessary to reduce the intrinsic activity of WT Dam by mutating residues that interact with the DNA phosphate group; similar variants (e.g., K9A) were also found to be effective in the DamIP system (Xiao and Moore, 2011). These results suggest that DNA specificity of WT Dam, which likely has strong evolutionary preference for GATC to support critical functions such as DNA repair and immunity, can be reprogrammed by protein engineering efforts aimed at first reducing binding energy. Our selected Dam variant was shown to function with two different DNA-targeting technologies (ZFs and CRISPR-dCas9), and the CRISPR-guided version of synI in particular may open up future possibilities for artificially modifying and regulating endogenous loci. However, this will require assessing the genome-wide specificities (and off-target profiles) of synI, a critical and yet unresolved issue, especially for the CRISPR-guided version, which showed higher propensity for m6A enrichment at our limited panel of off-target loci (Figure S3G). More broadly, the efficacy of our system at endogenous loci would need to be established before it can be translated into an epigenome and cellular engineering platform, including characterizing its performance at more complex genomic contexts (nucleosome positioning, barrier elements, etc.).

Epigenome editing tools that allow modifications to be induced at specific genomic sites have recently emerged to help understand causal links between chromatin modifications and gene regulation, and to explore therapeutic strategies for a number of diseases (Braun et al., 2017; de Groote et al., 2012; Hilton et al., 2015; Maeder et al., 2013; Mendenhall et al., 2013; Park et al., 2016; Thakore et al., 2016). Among these, there has been significant previous effort dedicated to developing site-specific DNA methylation tools, typically utilizing bacterial cytosine methyltransferases to enable CpG methylation and heritable gene silencing (Nomura and Barbas, 2007; Smith and Ford, 2007; Xiong et al., 2017). Instead, our approach was to leverage N6-methyladenine (m6A). While m6A is a well-established and abundant DNA modification in bacteria, much less is known about its existence and role in eukaryotic chromosomes. With the advent of sensitive detection techniques, studies have recently emerged reporting the discovery of DNA m6A in the genomes of certain eukaryotes, including *C. reinhardtii*, *D. melanogaster*, and *C. elegans*, and putative roles in transcriptional regulation, nucleosome positioning, embryogenesis, and epigenetic inheritance have been suggested (Fu et al., 2015; Greer et al., 2015; Zhang et al., 2015). Its presence in mammalian genomes has been more elusive and potentially controversial. Though recent reports claimed to observe trace amounts of DNA m6A in mouse embryonic stem cells and human cells (Wu et al., 2016; Xiao et al., 2018), validation studies using ultrasensitive ultra-high-performance liquid chromatography-mass spectrometry (UHPLC-MS) methods find no evidence for it, potentially attributing its observation to bacterial DNA or mammalian mRNA sources (Schiffers et al., 2017).

We next combined m6A deposition using synI with recruitment of transcriptional regulators using engineered synR readers. This minimal two-module circuit was sufficient to establish m6A-dependent transcriptional states. Moreover, the modular architecture of synR provided flexibility to encode different logical outputs depending on the fusion domain. In principle, other types of regulatory and EDs could be recruited to m6A-modified domains using this mechanism, providing a platform for combinatorial or enhanced recruitment of regulatory function through both genomic and epigenomic signatures.

An epigenetic mark in the strict sense is one that can persist after removal of an initial stimulus. The "read-write" motif is believed to be a core regulatory mechanism driving this self-amplification and contributing to the epigenetic inheritance of diverse chromatin modifications, but direct investigations of the RW principle have been difficult (Moazed, 2011). Using this as a blueprint, we engineered a synthetic "read-write" module and tested whether it provides a positive feedback mechanism capable of propagating and facilitating heritable maintenance of m6A states. Critically, our synthetic RW circuits displayed self-perpetuating behaviors and were able to program epigenetic functions *de novo* that do not rely on potentially confounding, endogenous chromatin mechanisms. Moreover, the behaviors we programmed exhibited hallmarks of natural epigenetic phenomena. For example, m6A distributions in our spatial propagation studies resembled those of propagating H3K9 methylation in HP1-induced heterochromatin, which underlies gene silencing and position-effect variegation, and both were shown to propagate symmetrically away from a nucleation site (Hathaway et al., 2012). Additionally, our synthetic circuits were capable of inducing epigenetic memory, enabling transcriptional states to be heritably maintained for >10 days in the majority of cells upon removal of an inducing signal.

Our synthetic approach enables investigation into how epigenetic behaviors are regulated by different configurations of the underlying molecular modules. For example, spatial propagation and epigenetic memory were both strictly dependent on the initiation module (synI). Furthermore, we showed how spatial propagation could, in principle, be tuned by adjusting synRW levels through a pseudo dose-response curve (within a range where available methylated GATC sites are presumably not limiting). Finally, though propagation did not strictly depend on the synR module, this additional factor could be used to help set and/or lower methylation levels, including basal levels. It is interesting to speculate that "futile" RDs, such as synR, could be operating to compete with and accordingly adjust the effective concentration of synRW via its dose-response curve to shape evolving m6A profiles.

By recombining variants of a reader and writer domain, we were able to generate synthetic modules that reconstitute the functional requirements of an epigenetic regulatory system. The modular simplicity of this design may highlight why this core regulatory solution seems to appear in different cellular contexts and for diverse chemical modifications, ranging from DNA methylation to histone post-translational modifications (PTMs) and protein phosphorylation. The modularity of the system may also provide a facile means for building additional layers of regulation atop this core circuitry. Indeed, it has recently been suggested that positive feedback based exclusively on RW mechanisms may constitute weaker forms of epigenetic inheritance (Audergon et al., 2015; Ragunathan et al., 2015), and that additional heritable feedback loops may be necessary to increase stability of epigenetic states. In support of this, the epigenetic memory we engineered based on minimal synthetic RW circuits was not as stable or long-lived as some of the memory observed in a recent study in which silencers were dynamically recruited to reporters (Bintu et al., 2016). Additional mechanisms for enhancing epigenetic memory and inheritance have been proposed—including sequence-specific elements (Laprell et al., 2017; Wang and Moazed, 2017), chromatin remodeling (Kundu et al., 2007), and cooperativity between modifications (Rudner et al., 2005; Rusché et al., 2002)—and a recent intriguing example showed that positive feedback involving siRNA amplification can directly couple to and strengthen (RW) histone PTM feedback loops (Yu et al., 2018). In principle, additional regulatory elements, connections, and feedback loops could be constructed within our synthetic system to explore their functional roles.

Early pioneering work in synthetic biology illustrated the power of using minimal synthetic genetic circuits to explore cellular regulation and engineer emergent properties in bacteria, such as cellular memory and oscillations (Elowitz and Leibler, 2000; Gardner et al., 2000). As interest in eukaryotic and especially mammalian systems has grown recently, it is becoming increasingly important to develop foundational tools and frameworks for manipulating the different organizational layers of mammalian gene regulation. Our work provides an initial set of molecular building blocks and a circuit engineering framework to aid in connecting the individual regulatory components making up complex systems like chromatin to its salient, higher-order properties like epigenetic memory and spreading. Much like synthetic circuits in bacteria provided strong evidence for quantitative control mechanisms underlying switch-like and oscillatory behaviors in natural systems, our work offers insights into the long-standing "read-write" propagation hypothesis in chromatin biology. In the future, our synthetic platform may also be adapted to provide new systems for cell and epigenome engineering, such as for manipulating genome architecture, molecular recording through epigenomic changes, and new forms of dynamic gene expression control. Finally, we envision our system could be expanded to incorporate other molecular operations, such as erasing, and chemical modifications with which to synthetically encode and regulate additional layers of epigenetic information in cells.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell culture
  - Cell line generation
- METHOD DETAILS
  - Cloning and plasmid construction
  - Transfection

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and three tables and can be found with this article online at https://doi.org/10.1016/j.cell.2018.11.002.

ACKNOWLEDGMENTS

We thank M. Maeder, J. Sander, and J.K. Joung for generously providing the zinc finger (ZF) construct, H. VanBenschoten for help with cloning, C.-H. Wang and P. Mehta for modeling discussions, and members of the Khalil laboratory for insightful comments on the manuscript. We thank the Tufts University Core Facility (TUFC) for their Next-Gen sequencing services. This work was supported by a NIH-NIGMS Ruth L. Kirschstein Postdoctoral Fellowship (1F32GM105272-01A1 to A.J.K.), a NSF Expeditions in Computing grant (CCF-1522074 to A.S.K.), a NSF grant (MCB-1713855 to A.S.K.), and a DARPA grant (W911NF-11-2-0056 to A.S.K.). A.S.K. also acknowledges funding from the NIH Director's New Innovator Award (1DP2AI131083-01), NSF CAREER Award (MCB-1350949), and DARPA Young Faculty Award (D16AP00142).

AUTHOR CONTRIBUTIONS

M.P., A.J.K., and A.S.K. conceived the study. M.P. and A.J.K. developed the system with help from N.P. M.P. planned and performed all experiments with guidance from all authors. M.P. and N.P. performed and analyzed the RNA-seq experiments. N.P. performed the theoretical modeling. M.P. and A.S.K. wrote the paper with input from all authors.

DECLARATION OF INTERESTS

A provisional US patent has been filed based on this work.

Received: January 26, 2018
Revised: August 31, 2018
Accepted: October 31, 2018
Published: December 6, 2018

REFERENCES

Al-Sady, B., Madhani, H.D., and Narlikar, G.J. (2013). Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread. Mol. Cell *51*, 80–91.

Audergon, P.N., Catania, S., Kagansky, A., Tong, P., Shukla, M., Pidoux, A.L., and Allshire, R.C. (2015). Epigenetics. Restricted epigenetic inheritance of H3K9 methylation. Science *348*, 132–135.

Beisel, C., and Paro, R. (2011). Silencing chromatin: Comparing modes and mechanisms. Nat. Rev. Genet. *12*, 123–135.

Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407–412.

Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The mammalian epigenome. Cell *128*, 669–681.

Bintu, L., Yong, J., Antebi, Y.E., McCue, K., Kazuki, Y., Uno, N., Oshimura, M., and Elowitz, M.B. (2016). Dynamics of epigenetic regulation at the single-cell level. Science *351*, 720–724.

Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular signals of epigenetic states. Science *330*, 612–616.

Braun, S.M.G., Kirkland, J.G., Chory, E.J., Husmann, D., Calarco, J.P., and Crabtree, G.R. (2017). Rapid and reversible epigenome editing by endogenous chromatin regulators. Nat. Commun. *8*, 560.

Coffin, S.R., and Reich, N.O. (2009). Escherichia coli DNA adenine methyltransferase: The structural basis of processive catalysis and indirect readout. J. Biol. Chem. *284*, 18390–18400.

de Groote, M.L., Verschure, P.J., and Rots, M.G. (2012). Epigenetic Editing: Targeted rewriting of epigenetic marks to modulate expression of selected target genes. Nucleic Acids Res. *40*, 10596–10613.

Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. Nature *403*, 335–338.

Feinberg, A.P. (2007). Phenotypic plasticity and the epigenetics of human disease. Nature *447*, 433–440.

Fu, Y., Luo, G.Z., Chen, K., Deng, X., Yu, M., Han, D., Hao, Z., Liu, J., Lu, X., Dore, L.C., et al. (2015). N6-methyldeoxyadenosine marks active transcription start sites in Chlamydomonas. Cell *161*, 879–892.

Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in Escherichia coli. Nature *403*, 339–342.

Gardner, K.E., Allis, C.D., and Strahl, B.D. (2011). Operating on chromatin, a colorful language where context matters. J. Mol. Biol. *409*, 36–46.

Greer, E.L., Blanco, M.A., Gu, L., Sendinc, E., Liu, J., Aristizábal-Corrales, D., Hsu, C.H., Aravind, L., He, C., and Shi, Y. (2015). DNA Methylation on N6-Adenine in C. elegans. Cell *161*, 868–878.

Grewal, S.I., and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. Science *301*, 798–802.

Hathaway, N.A., Bell, O., Hodges, C., Miller, E.L., Neel, D.S., and Crabtree, G.R. (2012). Dynamics and memory of heterochromatin in living cells. Cell *149*, 1447–1460.

Haynes, K.A., and Silver, P.A. (2011). Synthetic reversal of epigenetic silencing. J. Biol. Chem. *286*, 27176–27182.

Heyn, H., and Esteller, M. (2015). An adenine code for DNA: A second life for N6-methyladenine. Cell *161*, 710–713.

Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol. *33*, 510–517.

Hodges, C., and Crabtree, G.R. (2012). Dynamics of inherently bounded histone modification domains. Proc. Natl. Acad. Sci. USA *109*, 13296–13301.

Horton, J.R., Liebert, K., Bekes, M., Jeltsch, A., and Cheng, X. (2006). Structure and substrate recognition of the Escherichia coli DNA adenine methyltransferase. J. Mol. Biol. *358*, 559–570.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-cell dynamics of genome-nuclear lamina interactions. Cell *153*, 178–192.

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature *529*, 490–495.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693–705.

Kundu, S., Horn, P.J., and Peterson, C.L. (2007). SWI/SNF is required for transcriptional memory at the yeast GAL gene cluster. Genes Dev. *21*, 997–1004.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116–120.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Laprell, F., Finkl, K., and Müller, J. (2017). Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. Science *356*, 85–88.

Lee, J.S., Smith, E., and Shilatifard, A. (2010). The language of histone crosstalk. Cell *142*, 682–685.

Liang, F.S., Ho, W.Q., and Crabtree, G.R. (2011). Engineering the ABA plant stress pathway for regulation of induced proximity. Sci. Signal. *4*, rs2.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Lim, W.A., and Pawson, T. (2010). Phosphotyrosine signaling: Evolving a new cellular communication system. Cell *142*, 661–667.

Maeder, M.L., Angstman, J.F., Richardson, M.E., Linder, S.J., Cascio, V.M., Tsai, S.Q., Ho, Q.H., Sander, J.D., Reyon, D., Bernstein, B.E., et al. (2013). Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nat. Biotechnol. *31*, 1137–1142.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

McNamara, A.R., Hurd, P.J., Smith, A.E., and Ford, K.G. (2002). Characterisation of site-biased DNA methyltransferases: Specificity, affinity and subsite relationships. Nucleic Acids Res. *30*, 3818–3830.

Mendenhall, E.M., Williamson, K.E., Reyon, D., Zou, J.Y., Ram, O., Joung, J.K., and Bernstein, B.E. (2013). Locus-specific editing of histone modifications at endogenous enhancers. Nat. Biotechnol. *31*, 1133–1136.

Moazed, D. (2011). Mechanisms for the inheritance of chromatin states. Cell *146*, 510–518.

Nomura, W., and Barbas, C.F., 3rd. (2007). In vivo site-specific DNA methylation with a designed sequence-enabled DNA methylase. J. Am. Chem. Soc. *129*, 8676–8677.

O'Brown, Z.K., and Greer, E.L. (2016). N6-methyladenine: A conserved and dynamic DNA mark. Adv. Exp. Med. Biol. *945*, 213–246.

Park, M., Keung, A.J., and Khalil, A.S. (2016). The epigenome: The next substrate for engineering. Genome Biol. *17*, 183.

Ragunathan, K., Jih, G., and Moazed, D. (2015). Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. Science *348*, 1258699.

Ratna, P., Scherrer, S., Fleischli, C., and Becskei, A. (2009). Synergy of repression and silencing gradients along the chromosome. J. Mol. Biol. *387*, 826–839.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet. *16*, 276–277.

Rudner, A.D., Hall, B.E., Ellenberger, T., and Moazed, D. (2005). A nonhistone protein-protein interaction required for assembly of the SIR complex and silent chromatin. Mol. Cell. Biol. *25*, 4514–4528.

Rusché, L.N., Kirchmaier, A.L., and Rine, J. (2002). Ordered nucleation and spreading of silenced chromatin in Saccharomyces cerevisiae. Mol. Biol. Cell *13*, 2207–2222.

Schiffers, S., Ebert, C., Rahimoff, R., Kosmatchev, O., Steinbacher, J., Bohne, A.V., Spada, F., Michalakis, S., Nickelsen, J., Müller, M., and Carell, T. (2017). Quantitative LC-MS Provides No Evidence for m^6 dA or m^4 dC in the Genome of Mouse Embryonic Stem Cells and Tissues. Angew. Chem. Int. Ed. Engl. *56*, 11268–11271.

Siwek, W., Czapinska, H., Bochtler, M., Bujnicki, J.M., and Skowronek, K. (2012). Crystal structure and mechanism of action of the N6-methyladenine-dependent type IIM restriction endonuclease R.DpnI. Nucleic Acids Res. *40*, 7563–7572.

Smith, A.E., and Ford, K.G. (2007). Specific targeting of cytosine methylation to DNA sequences in vivo. Nucleic Acids Res. *35*, 740–754.

Thakore, P.I., Black, J.B., Hilton, I.B., and Gersbach, C.A. (2016). Editing the epigenome: Technologies for programmable transcription and epigenetic modulation. Nat. Methods *13*, 127–137.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: Discovering splice junctions with RNA-seq. Bioinformatics *25*, 1105–1111.

van Steensel, B., and Henikoff, S. (2000). Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. Nat. Biotechnol. *18*, 424–428.

Wang, X., and Moazed, D. (2017). DNA sequence-dependent epigenetic inheritance of gene silencing and histone H3K9 methylation. Science *356*, 88–91.

Wu, T.P., Wang, T., Seetin, M.G., Lai, Y., Zhu, S., Lin, K., Liu, Y., Byrum, S.D., Mackintosh, S.G., Zhong, M., et al. (2016). DNA methylation on N(6)-adenine in mammalian embryonic stem cells. Nature *532*, 329–333.

Xiao, R., and Moore, D.D. (2011). DamIP: Using mutant DNA adenine methyltransferase to study DNA-protein interactions in vivo. Curr. Protoc. Mol. Biol. *Chapter 21*, Unit 21.

Xiao, C.L., Zhu, S., He, M., Chen, Zhang, Q., Chen, Y., Yu, G., Liu, J., Xie, S.Q., Luo, F., et al. (2018). N(6)-methyladenine DNA modification in the human genome. Mol. Cell *71*, 306–318.

Xiong, T., Meister, G.E., Workman, R.E., Kato, N.C., Spellberg, M.J., Turker, F., Timp, W., Ostermeier, M., and Novina, C.D. (2017). Targeted DNA methylation in human cells using engineered dCas9-methyltransferases. Sci. Rep. *7*, 6732.

Yu, R., Wang, X., and Moazed, D. (2018). Epigenetic inheritance mediated by coupling of RNAi and histone H3K9 methylation. Nature *558*, 615–619.

Zhang, G., Huang, H., Liu, D., Cheng, Y., Liu, X., Zhang, W., Yin, R., Zhang, D., Zhang, P., Liu, J., et al. (2015). N6-methyladenine DNA modification in Drosophila. Cell *161*, 893–906.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Abscisic acid	Sigma Aldrich	Cat# A1049-250MG
Aphidicolin from Nigrospora sphaerica	Sigma Aldrich	Cat# A0781-5MG
Polyethylenimine (PEI)	Polysciences	Cat# 23966-1
FxCycle Far Red	Thermo Fisher Scientific	Cat# F10348
RNaseA	QIAGEN	Cat# 19101
Hygromycin B	GIBCO	Cat# 10-687-010
Puromycin Dihydrochloride	GIBCO	Cat# A1113803
Blasticidin S HCl	GIBCO	Cat# A1113903
LightCycler® 480 SYBR Green I Master	Roche	Cat# 4887352001
DpnII (1,000 units; 10,000 units/ml)	NEB	Cat# R0543S
Critical Commercial Assays		
Genomic DNA extraction: DNeasy Tissue kit	QIAGEN	Cat# 69506
Click-iT Plus EdU Pacific Blue Flow Cytometry Assay Kit	Thermo Fisher Scientific	Cat# C10636
RNeasy Plus Mini Kit	QIAGEN	Cat# 74134
QIAshredder	QIAGEN	Cat# 79656
CellTrace Far Red Cell Proliferation Kit	Thermo Fisher Scientific	Cat# C34564
Deposited Data		
Raw RNA sequencing data	This study	SRA: PRJNA488081
Human genome sequence	UCSC Genome Browser	Version hg38; https://genome.ucsc.edu/
Experimental Models: Cell Lines		
HEK293FT	Thermo Fisher Scientific	Cat# R70007
Human: Interspersed Reporter (pMinCMV)	This study	MP153
Human: Clustered Reporter. (pMinCMV)	This study	MP175
Human: Intrs. pMinCMV Reporter + synR^{VP64}	This study	MP243
Human: Intrs. pMinCMV Reporter + synI^{IND} + synR^{VP64}	This study	MP244
Human: Clust. pMinCMV Reporter + synR^{VP64}	This study	MP252
Human: Clust. pMinCMV Reporter + synI^{IND} + synR^{VP64}	This study	MP253
Human: Intrs. pMinCMV Reporter + synR^{VP64} + synRW(pUBC-DpnI-Dam(R95A))	This study	MP422
Human: Intrs. pMinCMV Reporter + synI^{IND} + synR^{VP64} + synRW(pUBC-DpnI-Dam(R95A))	This study	MP423
Human: Clust. pMinCMV Reporter + synR^{VP64} + synRW(pUBC-DpnI-Dam(R95A))	This study	MP427
Human: Clust. pMinCMV Reporter + synI^{IND} + synR^{VP64} + synRW(pUBC-DpnI-Dam(R95A))	This study	MP428
Additional derived cell lines and further information	This study	Table S2
Oligonucleotides		
no GATC reference (GFP) forward qPCR primer GTGAACCGCATCGAGCTGAAG	This study	N/A
no GATC reference (GFP) reverse qPCR primer TGTTGCCGTCCTCCTTGAAGTC	This study	N/A
GATC (20bp from ZF BS) forward qPCR primer TAAAGGCTTACTGAGCACTA	This study	N/A
GATC (20bp from ZF BS) reverse qPCR primer TGTGATTCAGAGACAACTTC	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
GATC (140bp from ZF BS) forward qPCR primer AATCGTTGCGTAATCTACAA	This study	N/A
GATC (140bp from ZF BS) reverse qPCR primer TTGCGAAAGTTGGAGAAATA	This study	N/A
Additional oligonucleotide pairs and qPCR conditions	This study	Table S3
Recombinant DNA		
pUBC ZF-VP64	This study	pMP258
Inters. (8XZFBS 14XGATC)-pMinCMV-GFPd2-RbGpA pGK-PuroR-T2A-mCh-BGHpA AAVS1 donor	This study	pMP472
Clust. (5XZFBS 63XGATC)-pMinCMV-GFPd2-RbGpA pGK-PuroR-T2A-mCh-BGHpA AAVS1 donor	This study	pMP498
pMinCMV-NLS-ABI1cs-ZF-NLS-P2A-Dam(N132A)-PYL1cs-HA pGK-BlastR	This study	pMP597
pUBC-DpnI(aa146-254)-VP64-V5 pGK-ZeoR	This study	pMP650
pUBC-DpnI(aa146-254)-3XFLAG-Dam(R95A) pGK-HygroR	This study	pMP926
pUBC-mCh-3XFLAG-Dam(R95A) pGK-HygroR	This study	pMP967
gRNA_AAVS1-T2 plasmid	(Mali et al., 2013)	Addgene: #41820
VP12 humanSpCas9-Hf1 plasmid	(Kleinstiver et al., 2016)	Addgene: #72247
Additional recombinant DNA constructs and further information	This study	Table S1
Software and Algorithms		
FlowJo v8	FlowJo, LLC.	https://www.flowjo.com/solutions/flowjo/downloads
GraphPad Prism	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
Bowtie2	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
TopHat	(Trapnell et al., 2009)	http://ccb.jhu.edu/software/tophat/index.shtml
featureCounts	(Liao et al., 2014)	http://bioinf.wehi.edu.au/featureCounts/
EMBOSS-dreg	(Rice et al., 2000)	http://www.bioinformatics.nl/cgi-bin/emboss/dreg
Other		
Attune NxT Flow Cytometer	Thermo Fisher Scientific	Attune NxT
LightCycler 480 Instrument II	Roche	Cat# 05015243001

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ahmad S. Khalil (khalil@bu.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture

The background cell line for all experiments in this study was 293FT cell line (Thermo Fisher Scientific). Cells were cultured in Dulbecco's modified Eagle's medium with L-Glutamine, 4.5 g/L Glucose and Sodium Pyruvate (DMEM, Thermo Fisher Scientific) supplemented with 10% Tet-system approved fetal bovine serum (FBS, Clontech), 1% GlutaMAX supplement (Thermo Fisher Scientific), 1% MEM Non-Essential Amino Acids (NEAA, Thermo Fisher Scientific) solution and 1% penicillin-streptomycin (Thermo Fisher Scientific). Cells were split every 3 days, and maintained at 37°C and 5% CO_2 in a humidified incubator.

Cell line generation

Cell lines used in this study are listed in Table S2, and were generated by genome-integrating constructs into the 293FT cell line, a fast-growing variant of the human female embryonic kidney 293 cell line stably expressing the SV40 large T antigen (Thermo Fisher Cat# R70007). The 293FT cell line was authenticated by morphology check with microscope.

Reporter lines were generated by site-specific integration of reporter constructs (Figure S1) using CRISPR-Cas9-mediated homologous recombination into the *AAVS1* (*PPP1R2C*) locus as follows: 60,000 cells were plated in a 48-well plate and co-transfected the following day with 70 ng of gRNA_AAVS1-T2 plasmid (Addgene 41820), 70 ng of VP12 humanSpCas9-Hf1 plasmid (Addgene 72247), and 175 ng of donor reporter plasmid using PEI. Donor reporter plasmids contain flanking arms homologous to the *AAVS1* locus, a puromycin resistance cassette, and constitutive mCherry expression (Figure S1). After transfection, cells were cultured in 2 μg/mL puromycin selection for at least 2 weeks with splitting 1:10 every 3 days, then monoclonal populations for each reporter cell line were isolated by limiting dilution in 96-well plates.

All other stable lines were generated by lentiviral integration of indicated constructs (encoding synI, synR, synRW modules and/or respective controls) into specific reporter lines. Lentivirus was produced by PEI co-transfection of 293FT cells with the donor plasmid, along with packaging vectors pCMVR8.74 (Addgene 22036), pAdVAntage (Promega), and pMD2.G (Addgene 12259). Virus was harvested with centrifugation (300 g, 5 min) and was added/incubated into specific reporter lines for three days, followed by selection in appropriate selection media: blasticidin (10 μg/ml), zeocin (100 μg/ml), and/or hygromycin (200 μg/ml).

METHOD DETAILS

Cloning and plasmid construction

Plasmid constructs used in this study are listed in Table S1 and their designs described in Figure S1. All constructs were constructed using standard molecular biology techniques and Gibson isothermal assembly. Donor plasmids for CRISPR-Cas9-induced reporter knock-in were constructed by PCR and subsequent Gibson assembly of components into the pCAGEN mammalian expression vector (Addgene 1160), digested with SalI/HindIII. Donor plasmids for lentiviral integration were constructed by PCR and subsequent Gibson assembly of components into pFUGW (Addgene 14883), digested with PacI/XhoI. During cloning, plasmids were transformed and prepped in *E. coli* TOP10 (Thermo Fisher Scientific). After sequence-verification, final reporter vectors were transformed and propagated in the *dam-/dcm-* strain, *E. coli* K12 ER2925 (NEB).

Transfection

For all transient transfection experiments, plasmid constructs (Figure S1) were transfected into indicated stable cell lines using polyethylenimine (PEI, 7.5 mM linear PEI stock, nitrogen/phosphorus ratio of 20, Polysciences). 60,000 cells were plated in 48-well plates and transfected the following day with a total of 100-300 ng DNA, including a pCAG-iRFP720 (Table S1) transfection control plasmid. Cells were collected and prepared for either flow cytometry or qPCR analysis 3 days after transfection, unless otherwise noted.

m6A-qPCR assay for measuring adenine methylation

We adapted a previously described qPCR-based assay to quantitatively measure adenine methylation at specific genomic sequences/loci (van Steensel and Henikoff, 2000). To obtain fraction methylated values reported throughout this paper, we used the assay to calculate the ratio of amplified DNA, protected from DpnII digestion, for a GATC site(s) of interest relative to a non-GATC reference site, which serves as an internal control to account for variation in DNA amount in each sample (Figures S2A–S2C). First, total genomic DNA (gDNA) was isolated using a DNeasy Blood and Tissue Kit (QIAGEN) according to the manufacturer's instructions (with addition of 4 μL of 100 mg/mL RNase A), eluting in 300 μL elution buffer. A 35 μL aliquot of the resulting gDNA was incubated for 16 hr at 37°C with or without 2 units of DpnII (NEB), followed by a 20 min heat inactivation at 65°C. Next, four qPCR reactions were prepared to amplify: (1) DpnII-digested GATC site(s), (2) undigested GATC site(s), (3) DpnII-digested reference site, (4) undigested reference site (Figure S2B). Specifically, qPCR reactions using a 1:10 dilution of the digested gDNA samples were prepared using LightCycler 480 SYBR Green I Master Kit (Roche) according to the manufacturer's instruction. qPCR reactions were performed on a LightCycler 480 Instrument II (Roche) with a total reaction volume of 20 μL (5 μL of DNA, 0.5 μM of forward primer, 0.5 μM of reverse primer, 10 μL of 2X SYBR Green Master Mix), using the following cycle conditions: (i) pre-incubation: 95°C for 10 min; (ii) amplification (45 cycles): 95°C for 10 s, [annealing temperature] for 20 s, 72°C for [extension time]; (iii) melting curve: 95°C for 5 s, 65°C for 1 min, 97°C at ramp rate 0.11°C/s; (iv) cooling: 40°C for 10 s. PCR primer sequences (listed in Table S3) were designed to flank the GATC site of interest or reference site. Annealing temperatures and extension times for specific primer sets are also listed in Table S3. Fraction methylated is then computed from the resulting qPCR Ct values using the ΔΔCt method/equation shown in Figure S2B.

To obtain "m6A enrichment" (e.g., for a synI factor), fraction methylation at a GATC probe site in the locus of interest was measured using m6A-qPCR and normalized to basal methylation induced by the Dam variant not fused to ZF (synI^{NT}, mCherry-Dam fusion) (Figures S2D and S2E).

Flow cytometry

Flow cytometry measurements were performed using an Attune NxT Flow Cytometer (Thermo Fisher Scientific) equipped with a high-throughput auto-sampler. Typically, 50,000 or 70,000 events were acquired for transient transfection or stable cell line experiments, respectively. Cells were gated by forward (FSC) and side scatter (SSC) distributions, and either iRFP or mCherry expression for transfection- or integration-positive populations, respectively. For experiments with transient transfection of synI and synR modules

(Figure 3), geometric means of the GFP fluorescence distributions were calculated using FlowJo (Treestar Software). GFP fold change was then calculated by normalizing mean GFP intensity to reporter-only controls, unless otherwise noted. For spreading and memory experiments (with integrated constructs), the percentage of GFP activated cells was quantified using a GFP+ gate that contains ~0.5%–1% of negative control cells (Figure S4E). Flow cytometer laser/filter configurations used in this study were: Click-iT Plus Edu Pacific Blue (405 nm laser, 440/50 emission filter), EGFP (488 nm, 510/10), mCherry (561 nm, 615/25), FxCycle Far Red or CellTrace Far Red (638 nm, 670/14), iRFP-720 (638 nm, 720/30).

Cell cycle assay

Cells cultured for 3 days with or without 200 μM abscisic acid (ABA, Sigma Aldrich) were labeled with Click-iT EdU Pacific Blue (Click-iT Plus EdU Pacific Blue Flow Cytometry Assay Kit, Thermo Fisher Scientific) to monitor DNA replication and FxCycle Far Red Stain (Thermo Fisher Scientific) to measure DNA content. Labeling was performed according to the manufacturers' instructions, with a 1.5 hr incubation in 10 μM EdU, and 30 min additional incubation with 200 nM FxCycle Far Red Stain and 1 μL of RNaseA (100 mg/mL). Cells were analyzed using flow cytometry.

RNA Sequencing

RNA-seq measurements were performed on two biological replicates per experimental condition. Total RNA was purified from ~1 million cells using the RNeasy Plus Mini Kit (QIAGEN) and QIAshredder (QIAGEN), according to the manufacturer's instructions, three days following transfection. Sequencing libraries were prepared at the Tufts University Core Facility (TUCF Genomics) using the TruSeq Stranded mRNA Library Prep Kit (Illumina). 50-bp single-end reads were sequenced on an Illumina HiSeq 2500.

DNA sequences for our synthetic constructs (reporter, synI, synR, ZF-VP64) were appended to the human UCSC genome (version hg19), and genome indices were built using the Bowtie 2 software (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml). Sequencing reads were aligned to this indexed genome using the TopHat software (http://ccb.jhu.edu/software/tophat/index.shtml), and the mapped reads were counted for genomic features using featureCounts (http://bioinf.wehi.edu.au/featureCounts/). Differential expression analysis was performed in R using the DESeq2 analysis package. Multiple hypothesis correction was performed using the Benjamini-Hochberg procedure with a FDR of < 1%.

Screen for spatial propagation

Our objective was to devise a phenotypic screen that would allow for identification of three-module, read-write (RW) circuit designs that drive spatial propagation behaviors. Specifically, we sought to identify synRW variants that, when coupled with synI^{IND} and synR modules, could propagate m6A across a domain in a manner that is dependent on m6A nucleation. Below we provide a detailed description of how we generated a library of synRW variants, developed a phenotypic screen to identify synRW candidates, and finally analyzed the screen results.

Generation of synRW module library

We created a synRW library by varying two biochemical properties: expression level and Dam writer activity. To vary expression level, we placed synRW expression under the control of two promoters of different strength: *pMinCMV* (weak) and *pUBC* (strong) (Figure S6B). To vary writer activity, we used a series of Dam mutants that have a range of methylation activities (Figures 4B and S1). Most of these mutations target residues responsible for mediating DNA phosphate group contact, either within or flanking the GATC sequence, which are known to affect the biochemical activity of the molecule (Coffin and Reich, 2009; Horton et al., 2006). In total, this collection represented 11 single residue mutants and 9 double residue mutants, along with WT Dam. The affinity of the DpnI RD represents another potential, tunable biochemical property of the synRW module; however, we chose to keep the RD fixed across the library because (1) less has been done to identify mutants and characterize their biochemical properties (relative to Dam) and (2) we wanted to maintain a manageable library size to transfect, culture, and assay in arrayed format.

We used Gibson isothermal assembly to construct the library, and then cloned the collection into a lentiviral vector (Figure S1).

Screen design

To identify three-module RW circuits that can drive spatial propagation, we developed a phenotypic screen in Clustered Reporter cell lines. The screen leverages the long GATC domain (~1.5 kb) that separates a nucleation site (ZF BS array) from the reporter gene (Figure 4B). Devoid of a mechanism for propagation across this domain, reporter cells stably expressing the two-module circuit (synI^{IND}, synR^{VP64}) do not activate the reporter (Figure S6A), since marks established at the distal nucleation site cannot effectively mediate promoter regulation by synR. We therefore screened the synRW library in these cells for candidates that lead to reporter activation (+synI^{IND}, Figure 4B), as well as in cells lacking synI^{IND} to screen out spurious cases for which reporter activation is independent of m6A nucleation (-synI^{IND}, Figure 4B). Variants emerging from this screen would represent potential candidates for the RW module of three-module regulatory circuits (synI^{IND}, synR^{VP64}, synRW) that drive inducible spatial propagation leading to reporter activation.

To perform the screen, we transfected 100 ng of each synRW construct (and 50 ng of pCAG-iRFP720 transfection marker) into two cell lines (60,000 cells, in triplicates): (1) Clustered Reporter cells stably expressing the two-module circuit (synI^{IND}, synR^{VP64}), and (2) Clustered Reporter cells stably expressing only synR^{VP64}. Cells triggered at the same time (6 hr after transfection) and induced continuously thereafter with 200 μM of ABA were harvested 4 days after transfection, whereupon half of the cells were assayed for GFP activation by flow cytometry (Figure 4B) and the remaining half collected for m6A-qPCR analysis of methylation profiles.

Screen analysis

To examine our screen results, define thresholds, and guide circuit designs, we performed hierarchical cluster analysis on the GFP expression patterns (similarity in % GFP activated data, treating each replicate individually; heatmap.2 function in R). Satisfyingly, the unbiased analysis distinguished -synI^{IND} from +synI^{IND} cells (vertical dendrogram not shown in Figure 4B). The analysis also revealed a number of interesting features. Library members divided into two parental clusters: one with strong GFP activation in the +synI^{IND} case (top) and the other with weak or no GFP activation (bottom) (Figure 4B). We used these clusters to define a threshold of circuits exhibiting functional (top) versus non-functional (bottom) propagation phenotypes, with the -synRW control circuit occupying the latter.

Examining molecular components within these clusters, we find that the non-functional cluster possessed all the Dam double mutants (as well as WT Dam), whereas the functional cluster was composed entirely of Dam single mutants. These results suggest that methylation activity of the synRW module could be a key factor in the design of synthetic propagation circuits. On the other hand, the two promoters we tested were scattered across parental and sub clusters, meaning there was no obvious preference in phenotypic outcomes for either promoter. This suggests that the transfection conditions used for screening likely produced saturating concentrations of synRW for reactions on a limited number of available methylated GATC sites. We tested whether this was the case by performing an analogous experiment with the other engineered reader module: synR. Indeed, when we transfected Interspersed Reporter cells with a two-module circuit, placing expression of synR^{VP64} under the control of either weak *pMinCMV* or strong *pUBC*, we observed a similar insensitivity in reporter output (Figure S6B).

The unbiased analysis of our propagation screen revealed a strong clustering of circuits based on synRW Dam mutants (Figure 4B). To further examine the relationship between writer methylation activity and the results of the screen, we defined a quantitative metric to score propagation propensity: ‘‘expt. spreading score’’ is computed as the difference (Δ) in % GFP activated for cells with and without synI^{IND} (Figures 4B and 4C). Higher spreading scores correspond to circuits that drive high levels of reporter activation preferentially in cells with m6A nucleation (+synI^{IND}). In Figure 4C, the spreading score for each synRW library member is plotted as a function of its basal Dam methylation activity, as previously measured (Figures S2D and S2F). (Plotted is the average value for both promoters since screen results for each synRW were similar across both promoters – a plot of the full library can be found in Figure S5D). Based on this analysis, we find that the highest scoring circuits contain synRW variants with intermediate methylation activity, potentially an important design feature of this module for producing the propagation phenotype.

We selected three synRW variants representing low, intermediate and high Dam writer activity to further investigate (Figures S5D–S5G). We measured corresponding m6A profiles across the Clustered Reporter GATC array in cells that were transfected with or without the synRW module (see STAR Methods). As expected, a low-activity synRW module produced no enrichment in m6A across the array over control cells lacking synRW (Figure S5G, left). A high-activity synRW module, on the other hand, led to significantly higher levels of m6A over cells lacking synRW; however, this came at a cost, as m6A levels were enriched in cells with or without synthetic initiator (Figure S5G, right). We speculate that the corresponding low reporter/spreading score (Figure S5F) is likely a result of global off-target m6A enrichment induced by the high-activity synRW, similar to what we observed with high-activity synI factors, which could serve to titrate synR factors away from the reporter. Finally, profiles measured for the intermediate-activity synRW showed evidence for an enlarged m6A domain in cells expressing both synI^{IND} and synRW (Figure S5G, middle). These results further validate the ‘‘spreading score’’ as a useful phenotypic metric for identifying circuits capable of driving spatial propagation behaviors.

Model of m6A spatial dynamics

Model description

Despite the molecular complexities inherent to chromatin regulatory systems, previous studies have shown that behaviors like nucleation and propagation of histone modifications along a chromosome can be captured by simple general models (Hathaway et al., 2012; Hodges and Crabtree, 2012). We therefore hypothesized that such models could also be used to capture spatial propagation of m6A by synthetic RW circuits, providing a general guide for their design and construction.

We adapted a previously described chromatin spreading model (Hodges and Crabtree, 2012) in order to explore how the properties of synI and synRW affect m6A spatial dynamics. We modeled the array of GATC sites in the Clustered Reporter as a discrete, one-dimensional lattice with 63 sites (Figure S5). Each position (j) on the lattice (I_j) corresponds to a GATC site in the reporter, where methylated and unmethylated states are denoted by values of $I_j = 1$ and $I_j = 0$, respectively (Figure S5A). Four reactions govern this model: basal (non-targeted) methylation, sequence-specific nucleation, reader-mediated propagation, and mark turnover. We implemented these reactions with the following rules:

(1) Basal methylation by synI or synRW occurs at a rate k_{synI_act} or k_{synRW_act}, respectively, to convert any unmethylated GATC site ($I_j = 0$) to methylated ($I_j = 1$).

(2) Sequence-specific nucleation by synI occurs at a rate $b_{ZF} \cdot k_{synI\_act}$ to convert the first lattice site from unmethylated ($I_1 = 0$) to methylated ($I_1 = 1$). k_{synI_act} is the basal methylation rate of the synI Dam writer domain, while b_{ZF} serves as a specificity multiplier to increase the methylation frequency by synI at the nucleation site.

(3) Reader-mediated propagation by synRW occurs at a rate $b_{Dpn1} \cdot k_{synRW\_act}$ to convert unmarked sites ($I_j = 0$) to methylated, if any neighboring GATC sites are marked ($I_{j-1} = 1$ or $I_{j+1} = 1$). k_{synRW_act} is the basal methylation rate of the synRW Dam writer

domain, while b_{Dpn1} serves as a specificity multiplier to increase the methylation frequency of unmarked sites adjacent to marked ones.

(4) Mark turnover occurs at rate k_{turn} during which any methylated GATC site ($I_j = 1$) can be converted to unmethylated ($I_j = 0$).

Model parameterization

For all simulations, we set $k_{turn} = 0.05\ hr^{-1}$ to approximate dilution of m6A modifications by cell division (~20 hr doubling time). In the model, synI activity is described by the non-specific methylation (k_{synI_act}) of each Dam mutant as well as the specificity (b_{ZF}) conferred by the ZF domain at the nucleation site (I_1). To approximate the range of non-specific methylation activity (k_{synI_act}) values across the Dam library, we used the following relationship:

$$f_{meth\_basal} = \frac{k_{synI\_act}}{k_{turn} + k_{synI\_act}}$$

where f_{meth_basal} is the experimentally obtained values of basal methylation for each Dam mutant (Figure S2D) and $k_{turn} = 0.05\ hr^{-1}$. We found that non-specific activity values ranged from $k_{synI\_act} = 10^{-6}$ to $10^{1}\ hr^{-1}$. Subsequently, to parameterize the ZF specificity multiplier (b_{ZF}) for each synI in the library, we calculated steady-state methylation (f_{meth}) by synI across a range of k_{synI_act} and b_{ZF} values with the following relationship:

$$f_{meth\_targ} = \frac{k_{synI\_act} + b_{ZF} \cdot k_{synI\_act}}{k_{turn} + k_{synI\_act} + b_{ZF} \cdot k_{synI\_act}}$$

where f_{meth_target} is the experimentally obtained values of targeted methylation for each Dam mutant (Figure S2E) and $k_{turn} = 0.05\ hr^{-1}$. We found that our synI module library is captured by specificity multiplier values, ranging between $b_{ZF} = 1 - 100$ (Figure S5B). We chose values of $k_{synI\_act} = 5 \cdot 10^{-4}\ hr^{-1}$ and $b_{ZF} = 10$, which closely approximated the behavior of the selected synI (Dam N132A) featured in our three-module propagation circuit.

Stochastic simulations of spatial propagation

To model the synRW library screen in Figure 4B, we ran stochastic simulations of synRW activity for a range of k_{synRW_act} (10^{-6} to $10^{1}\ hr^{-1}$ as obtained above for k_{synI_act}) and b_{Dpn1} values, with and without synI. We defined a "model spreading score" to assess m6A propagation from the nucleation site ($j = 1$) for each synRW library member. This was defined as the difference in mean m6A density at the 15 most downstream GATC sites ($j = 49$ to $j = 63$), with and without synI (Figure S5C). All stochastic simulations were implemented using the Gillespie algorithm and were run to steady-state (t = 5000 hr).

We found that the model spreading score distribution closely resembled the experimental spreading score distribution for b_{Dpn1} values between 50 and 200 (Figures S5D and S5E, see also Figure 4C with $b_{Dpn1} = 100$). Moreover, the model results showed a similar relationship between methylation activity of synRW and propagation propensity, where intermediate methylation levels were predicted to yield the highest scoring circuits. Interestingly, the model also predicted that, if reader specificity could be increased (higher b_{Dpn1}), then one could generate synRW factors that drive high levels of propagation with lower writer activity (Figure S5E).

In general, model-generated m6A spatial profiles agreed well with experimentally measured profiles for the range of synRW variants, representing low, intermediate and high Dam writer activity (Figure S5G). For example, high-activity synRW was predicted to produce nucleation-independent methylation, while low-activity synRW was predicted to yield weak spreading from the nucleation site. Furthermore, our model captured the dynamics of the growing m6A domain induced by our three-module "propagation circuit" (with $b_{Dpn1} = 100$, $k_{synRW\_act} = 3.4 \cdot 10^{-4}\ hr^{-1}$) (Figure 4D). Taken together, this simple, four-parameter model can effectively capture the essential features of our m6A synthetic propagation system.

In order to increase the predictive power of this model, future iterations could include off-target methylation by synI and synRW (at other genomic loci), as well as incorporate transcriptional regulation by synR. Describing these properties may be necessary for predicting targeted reporter activity from the distribution of methylation across the genome.

Epigenetic memory experiments

~120,000 cells were initially plated in multiple wells of a 6-well plate, and incubated either with or without 200 μM ABA. ABA was washed out at indicated times by aspirating out ABA-containing media and adding back fresh media. At indicated time points following ABA washout, approximately half of the cells were re-plated and continued in culture, while the rest were harvested for downstream analysis. For Aphidicolin (APC) experiments, re-plated cells were continued in culture with or without 5 μg/mL Aphidicolin (APC).

Model of epigenetic memory

Stochastic simulations of epigenetic memory

We applied our model framework to investigate epigenetic memory induced by synthetic RW circuits. Specifically, we modeled the effect of different synRW variants on the methylation decay dynamics of an initially fully methylated Interspersed Reporter. The Interspersed Reporter was described as a one-dimensional lattice with 14 methylated GATC sites ($I_{j\ =\ 1\text{-}14} = 1$). We tracked methylation across the lattice in the presence and absence of synRW ($b_{Dpn1} = 100$, $k_{synRW\_act} = 3.4 \cdot 10^{-4}\ hr^{-1}$), as well as for a

reader-defective synR^{NT}W (b_{Dpn1} = 0, k_{synRW_act} = 3.4•10^{-4} hr^{-1}). Simulated methylation dynamics showed close agreement with experiments, wherein methylation profiles decay rapidly due to passive dilution in simulations without synRW or with a control synR^{NT}W, and show maintenance of methylation over 20 days in simulations with a functional synRW (Figure 6D). Finally, we simulated decay profiles for synRW variants across a range of b_{Dpn1} values (Figure S7E). We found that higher b_{Dpn1} values corresponded to a larger percentage of the lattice retaining its methylation state over time. These results suggest that epigenetic memory could be strengthened by improving the binding activity of the synRW RD in future designs.

Cell proliferation assay

~120,000 cells were initially plated in multiple wells of a 6-well plate, and incubated either with or without 200 μM ABA. ABA was washed out after 3 days by aspirating out ABA-containing media and adding back fresh media. Cells were harvested with trypsin and brought to suspension. Cells were then stained with 5 μM CellTrace Far Red, according to manufacturer's instructions for labeling cells in suspension (CellTrace Far Red Cell Proliferation Kit, Thermo Fisher Scientific). About half of the stained cells were analyzed with flow cytometry (Day 0), while the rest were re-plated for continued culture. Thereafter at indicated time points following ABA washout, approximately half of the cells were re-plated and continued in culture, while the rest were resuspended in media for flow cytometry analysis.

Human genome sequence analysis of GATC motifs

Promoter sequences, defined as the region 1500 bp upstream of each TSS, for all human transcripts by chromosome were retrieved from the UCSC genome browser (version hg38) using 'GENCODE v24' track. The start coordinates of the GATC motif within the retrieved promoter sequences were mapped out using the EMBOSS-dreg bioinformatics tool (http://www.bioinformatics.nl/cgi-bin/emboss/dreg). The retrieved coordinates were further analyzed and plotted for the number of GATC motif occurrences and inter-GATC distances using custom R scripts.

QUANTIFICATION AND STATISTICAL ANALYSIS

FlowJo was used to extract geometric mean fluorescence values or the percentage of GFP activated cells from flow cytometry measurements. Microsoft Excel and GraphPad Prism software were used to process data. Statistical details such as N and error calculations are provided in figure legends.

DATA AND SOFTWARE AVAILABILITY

Raw RNA-seq data for 293FT transcriptome analysis upon expression of synI and/or synR newly reported in this paper is available at SRA: PRJNA488081.

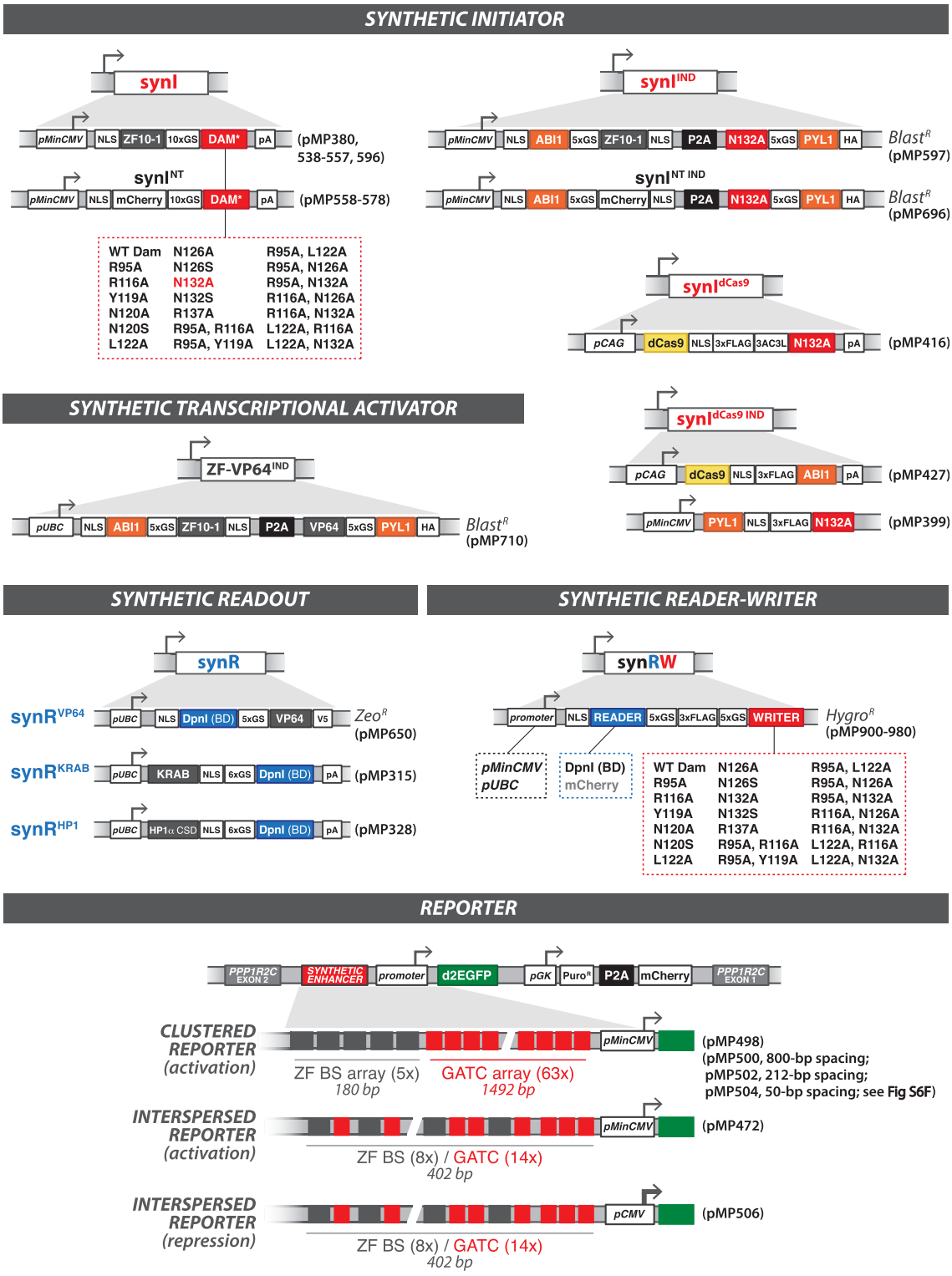


Figure S1. Overview of Genetic Components and Constructs, Related to Figure 1

Synthetic molecular components and corresponding plasmid constructs (pMPXXX, listed in Table S1) are indicated. Plasmid constructs for synI^{IND}, synR, synRW and ZF-VP64^{IND}, listed with resistance markers, indicate constructs used for lentiviral integration of cells. Reporter constructs were singly-integrated into the *AAVS1* (*PPP1R2C*) locus using CRISPR-Cas9 (see STAR Methods). Stable cell lines derived from combinations of these constructs are listed in Table S2. NLS, SV40 nuclear localization sequence; FLAG, epitope tag; GS, glycine-serine linker; ZF10-1, engineered zinc finger (ZF) designed to bind synthetic 20-bp nucleotide

(legend continued on next page)

sequence (BS = cGGCGTAGCCGATGTCGCGc); DAM*, *E. coli* DNA adenine methyltransferase (Dam) variants (mutations denoted in red dashed box); dCas9, nuclease-deactivated *S. pyogenes* Cas9; 3AC3L, flexible artificial linker (SSGNSNANSRGPSFSSGLVPLSLRGSH); DpnI (BD), DpnI binding domain (aa 146-254); VP64, tetramer of Herpes simplex VP16 transcriptional activation domain; KRAB, human Krüppel associated box; HP1α CSD, human Hp1α chromo shadow domain; d2EGFP, EGFP destabilized with degradation domain (fusion with aa 422-461 of mouse ornithine decarboxylase); *pCMV,* CMV promoter; *pMinCMV*, minimal CMV promoter; *pUBC*, human ubiquitin C promoter; *pGK*, human phosphoglycerate kinase promoter; *pCAG*, strong mammalian synthetic CAG promoter; P2A, porcine teschovirus-1 2A ribosome skipping peptide; ABI1, ABA insensitive 1 complementary surfaces (aa 126-423); PYL1, PYR1-like complementary surfaces (aa 33-209); HA, epitope tag; *Blast*^{R} Blasticidin resistance; *Hygro*^{R}, Hygromycin resistance; *Puro*^{R}, Puromycin resistance; *Zeo*^{R}, Zeocin resistance.

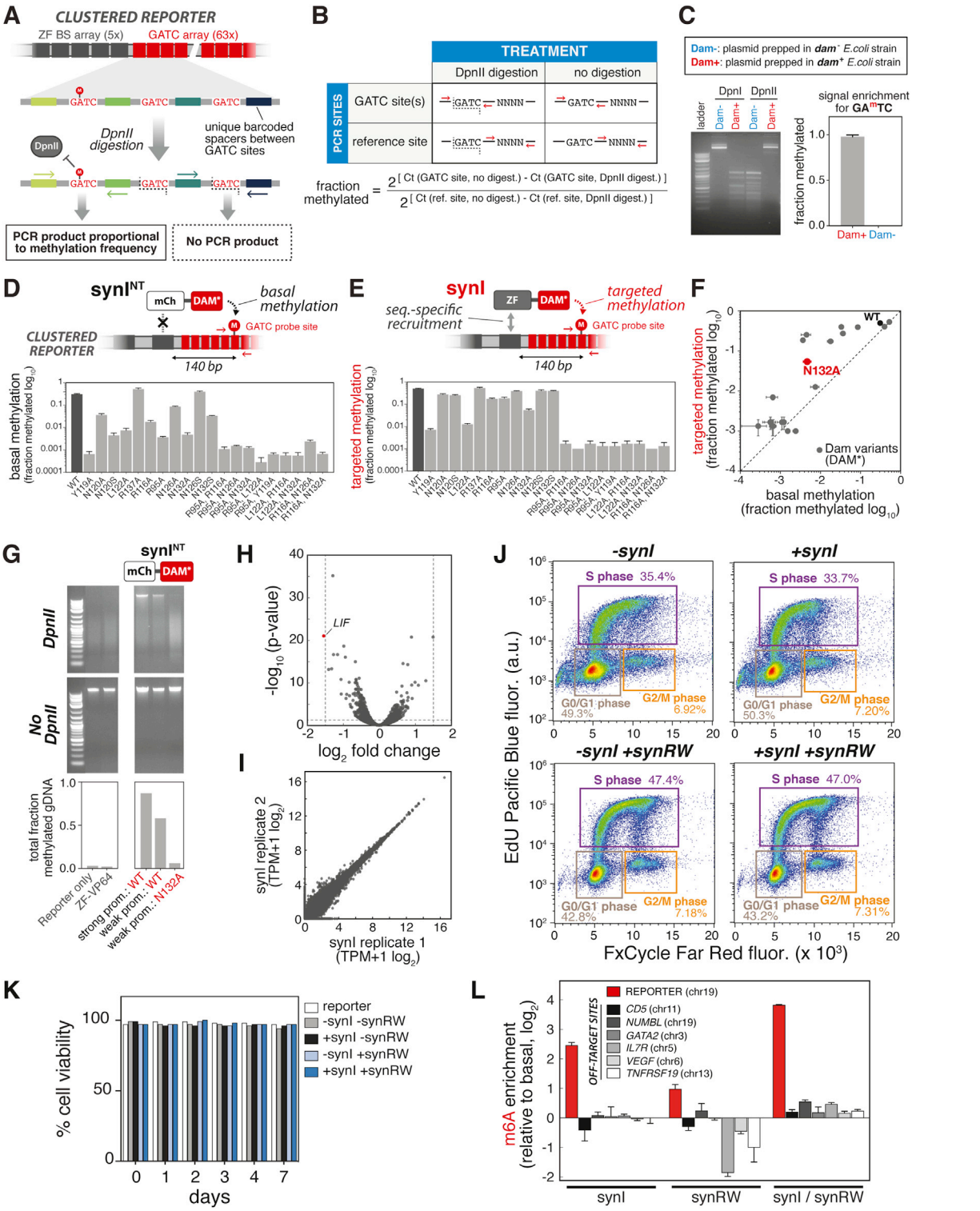


Figure S2. Characterizing *De Novo* m6A Establishment at Reporter Locus and Its Effect on the Transcriptome, Cell-Cycle, and Cell Viability, Related to Figure 2

(A) Schematic of the m6A-qPCR assay for quantifying adenine methylation frequency at the Clustered Reporter. GATC sites in the Clustered Reporter are flanked by unique 20-bp barcode spacers (colored bars), allowing single GATC site resolution in methylation levels. Genomic DNA is isolated and digested with DpnII, then intact GATC sites are qPCR amplified using primers (colored arrows) unique to the barcodes (see STAR Methods). Sequences containing unmethylated

(legend continued on next page)

GATC are cut by DpnII yielding no PCR product, while methylated sequences are protected from digestion yielding a PCR product proportional to the methylation frequency.

(B) Overview of digestion treatments and PCR sites used to quantify methylation frequency at GATC site(s) of interest with m6A-qPCR assay. Methylation frequency is calculated as the fraction of DNA that is resistant to DpnII digestion (see STAR Methods).

(C) Validation of the m6A-qPCR assay for quantifying methylated DNA. Episomal reporter plasmid was transformed and propagated in either *dam*$^-$ (Dam-, K12 ER2925) or *dam*$^+$(Dam+, TOP10) *E. coli*. Plasmids were then isolated and treated with either DpnI or DpnII. Left: Digested plasmids were run on a 1% agarose gel, confirming that plasmid derived from Dam+ is fully adenine-methylated (cut by DpnI, protected from DpnII digestion) while plasmid derived from Dam- is unmethylated. Right: Quantification of fraction methylation for the two plasmids as measured by the m6A-qPCR assay. (n = 3; error bars, SD).

(D) "Basal" adenine methylation levels for library of synI^{NT} factors, non-targeting writers composed of fusion of mCherry and Dam variant. Fraction methylation was measured using m6A-qPCR at a GATC probe site proximal to the ZF BS array (6th site, 140 bp downstream of end of ZF array). (n = 3; error bars, SD).

(E) "Targeted" adenine methylation levels for library of synI factors, targeting writers composed of fusion of sequence-specific ZF and Dam variant. (n = 3; error bars, SD).

(F) Distribution of targeted (synI) versus basal (synI^{NT}) adenine methylation levels for Dam variants. Diagonal line represents no enrichment in reporter-specific methylation over "background" methylation. Dam N132A and WT are marked in red and black, respectively. (n = 3; error bars, SD).

(G) Lowering expression level and using a mutant Dam leads to lower m6A levels genome-wide. Cells were transfected with the indicated constructs: Interspersed Reporter only, ZF-VP64 constitutively expressed from *pUBC*, synI^{NT} (Dam WT) constitutively expressed from a strong promoter (*pCLPIT* with no Dox induction), synI^{NT} (Dam WT) constitutively expressed from a weak promoter (*pMinCMV*), and synI^{NT} (Dam N132A) constitutively expressed from a weak promoter (*pMinCMV*). Genomic DNA (gDNA) was isolated 2 days following transfection, digested with or without DpnII, run out on an agarose gel (top), and the bands used to quantify total fraction methylated DNA (bottom). Undigested bands appearing at the top of each lane in DpnII-treated samples represent methylated (protected) gDNA, while unmethylated gDNA fragments that are digested by DpnII appear as a smear (from ~0.3-2 kb). Total fraction methylated gDNA is calculated as undigested gDNA in DpnII-treated samples normalized by total gDNA.

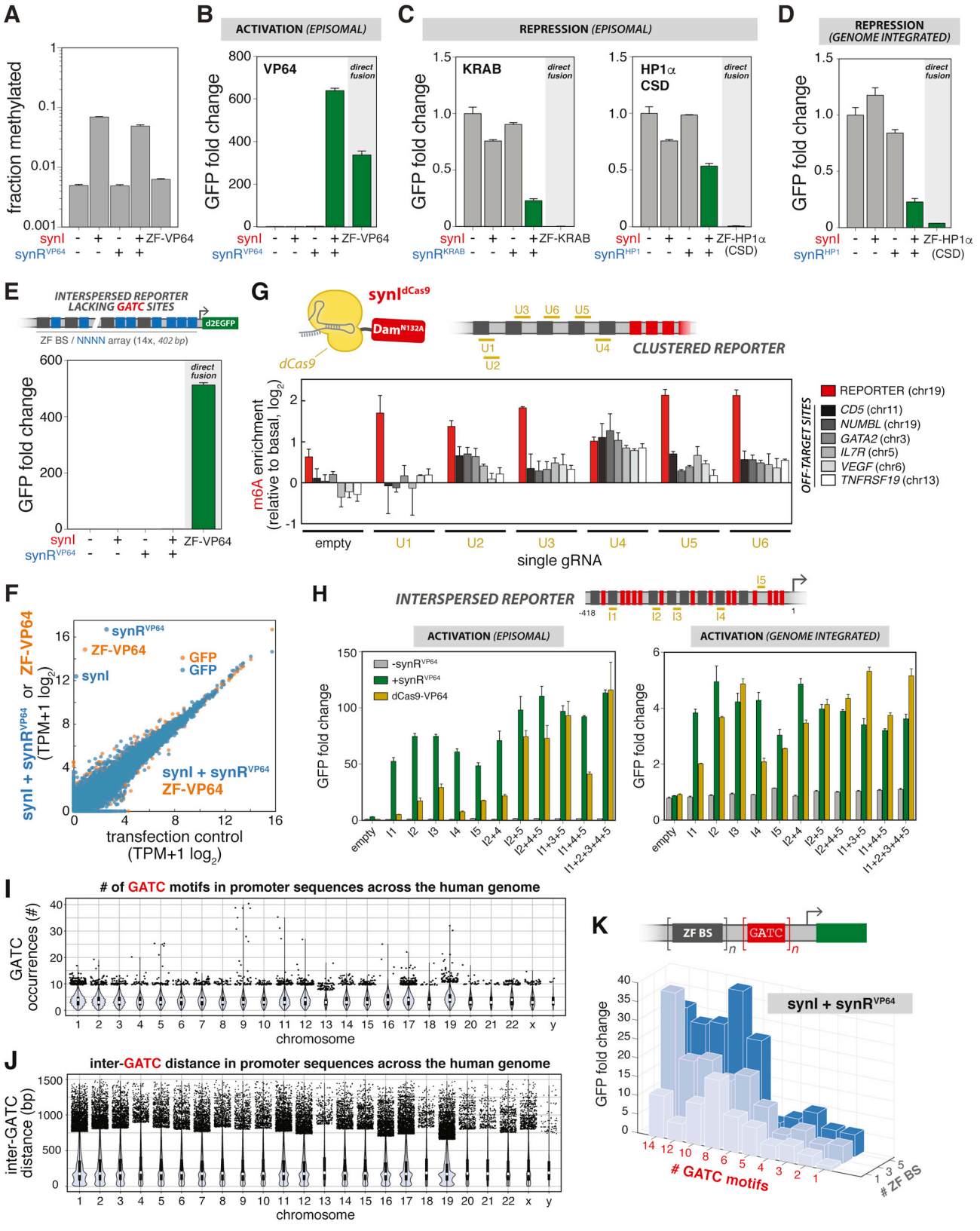
(H) Volcano plot generated from differential transcript (DEseq2) analysis of RNA-seq measurements for cells transiently transfected with synI (N132A, constitutive expression) versus empty plasmid (see also Figure 2C). Only one transcript showed significant enrichment based on differential expression > $\log_2$ 1.5 and FDR < 0.01 (*LIF*, red point). We found no evidence for the presence of a near match (having < 6-bp mismatch) to the ZF BS in genomic sequences 1-kb upstream and downstream of *LIF*, suggesting that the changes observed are due to a small level of noise in RNA-seq measurements. The data represent two independent biological replicates for each condition.

(I) Correlation of transcriptome from RNA-seq measurements of synI-expressing sample replicates. Correlation coefficient is 0.98, calculated using $\log_2$ transformed expression values.

(J) Cell cycle is not affected by expression of synthetic writers: synI and synRW (see Figures 4, 5, and 6). Flow cytometry plots of Interspersed Reporter cell lines stably expressing different combinations of synthetic writers, labeled with EdU-Pacific Blue to analyze DNA replication and FxCycle Far Red to quantify DNA content. Shown on top are cells expressing a small molecule inducible synI (uninduced = -synI, induced with 200 μM ABA = +synI; see also Figure S4), and on bottom are cells expressing both the inducible synI and a synRW. Cell percentages in different phases of the cell cycle are quantified.

(K) Cell viability is not affected by expression of synthetic writers: synI and synRW (see Figures 4, 5, and 6). Quantification of Trypan Blue dye exclusion for Interspersed Reporter cell lines stably expressing different combinations of synthetic writers (uninduced = -synI, induced with 200 μM ABA = +synI; see also Figure S4).

(L) Assessment of m6A enrichment (on- and off-target) by synthetic writers. m6A enrichment at target reporter (red bars; GATC probe site) and off-target, GATC-containing endogenous loci (gray bars) is shown for Clustered Reporter cell lines transfected with synI, synRW, or both. Off-target loci are same as Figure 2B. m6A enrichment is obtained by measuring fraction methylation at a GATC probe site in the locus of interest using m6A-qPCR, and normalizing to basal methylation induced by Dam (N132A) not fused to ZF. (n = 3; error bars, SD).



(legend on next page)

Figure S3. Characterizing m6A-Dependent Transcriptional Regulation by synI and synR Modules, Related to Figure 3

(A) Methylation levels quantified by m6A-qPCR for Clustered Reporter cell lines transfected with indicated combinations of synI and $synR^{VP64}$ expression constructs. (n = 3; error bars, SD).

(B) m6A-mediated transcriptional activation of episomal reporter. synI and $synR^{VP64}$ expression constructs (or direct ZF-VP64 fusion) were co-transfected into cells along with an Interspersed *pMinCMV* Reporter plasmid (pMP472) in the combinations shown. GFP fluorescence intensity, measured by flow cytometry, was normalized to the −/− condition to obtain GFP fold change. (n = 3; error bars, SD).

(C) m6A-mediated transcriptional repression of episomal reporter. Left: synI and $synR^{KRAB}$ expression constructs (or direct ZF-KRAB fusion) were co-transfected into cells along with an Interspersed *pCMV* Reporter plasmid (pMP506) in the combinations shown, and GFP fluorescence assayed by flow cytometry. Right: Identical experiment but replacing KRAB with the Hp1α chromo shadow domain (CSD). (n = 3; error bars, SD).

(D) m6A-mediated transcriptional repression by synI and $synR^{HP1}$ modules. GFP fluorescence intensity, measured by flow cytometry, for Interspersed Reporter cell lines (with *pCMV*) transfected with indicated combinations of synI and $synR^{HP1}$ expression constructs (or a direct ZF-HP1α (CSD) fusion). (n = 3; error bars, SD).

(E) GATC sites are required for transcriptional regulation by synI and synR modules. An Interspersed *pMinCMV* Reporter lacking GATC sites was generated by replacing the 14 GATC motifs with 4-bp random sequences (NNNN, blue boxes). synI and $synR^{VP64}$ expression constructs (or direct ZF-VP64 fusion) were co-transfected into cells along with the GATC-lacking reporter plasmid in the combinations shown. (n = 3; error bars, SD).

(F) Transcriptome analysis of m6A-mediated transcriptional activation by synI and $synR^{VP64}$ modules. Correlation of transcriptome from RNA-seq measurements of Interspersed Reporter cell lines co-transfected with either a two-module system (synI and $synR^{VP64}$, blue) or ZF-VP64 (orange) versus reporter cells transfected with empty plasmid. mRNAs corresponding to synI, $synR^{VP64}$, ZF-VP64, and GFP are labeled.

(G) m6A placement and enrichment at reporter locus by dCas9-based synthetic initiator ($synI^{dCas9}$). $synI^{dCas9}$ is a fusion of Dam N132A and dCas9. m6A enrichment at target reporter (red bars; GATC probe site) and off-target, GATC-containing endogenous loci (gray bars) is shown for Clustered Reporter Cell lines co-transfected with $synI^{dCas9}$ and either single gRNAs (gold, U1-U5) or empty gRNA cassette. Off-target loci are same as Figure 2B. m6A enrichment is obtained by measuring fraction methylation at a GATC probe site in the locus of interest using m6A-qPCR, and normalizing to basal methylation induced by $synI^{dCas9}$ alone. (n = 3; error bars, SD).

(H) m6A-mediated transcriptional activation by $synI^{dCas9}$ and $synR^{VP64}$ modules. Left: Cells were co-transfected with Interspersed *pMinCMV* Reporter plasmid (pMP472), $synI^{dCas9}$ and $synR^{VP64}$ expression constructs (or direct dCas9-VP64 fusion), and gRNAs (gold, I1-I5) in the combinations shown, and GFP fluorescence assayed by flow cytometry. gRNA target locations in the Interspersed Reporter are indicated above. Right: Identical experiment but performed with Interspersed Reporter cell lines. GFP fluorescence intensity, measured by flow cytometry, was normalized to the reporter only control to obtain GFP fold change. (n = 3; error bars, SD).

(I) Violin plot showing the number of GATC occurrences in all human promoters for each chromosome. We defined a promoter as the region 1.5 kb upstream of a transcription start site (TSS).

(J) Violin plot showing the inter-GATC distance in all human promoters for each chromosome. (I-J) Boxplots depicting median counts (white circles) and interquartile ranges (black boxes) are overlaid onto violin plots. Outliers are shown in black dots.

(K) Reporter activation as a function of different numbers of ZF BS and GATC motifs. synI and $synR^{VP64}$ expression constructs were co-transfected into cells along with reporter constructs designed with indicated numbers of ZF BS and GATC repeats upstream of a *pMinCMV* promoter. GFP fluorescence intensity, measured by flow cytometry, was normalized to the respective reporter only control to obtain GFP fold change. (n = 3; mean values).

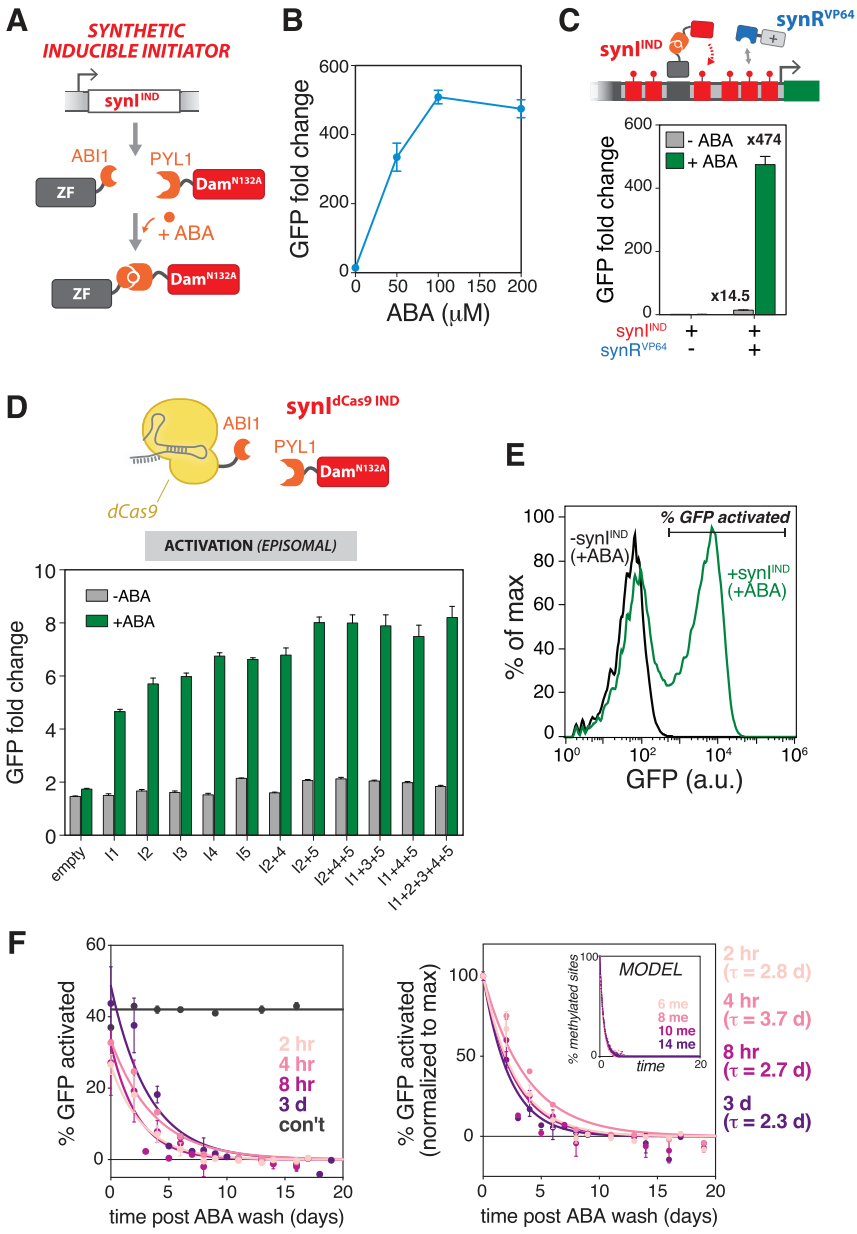


Figure S4. Design and Characterization of a Small-Molecule Inducible Initiator: synI^{IND}, Related to Figures 4 and 6

(A) Schematic of the synthetic inducible initiator, synI^{IND}, which uses abscisic acid (ABA)-induced dimerization to trigger and temporally control m6A placement. ZF and Dam (N132A) domains are fused to ABI1 and PYL1 domains, respectively, and co-expressed from a bicistronic cassette (see Figure S1). Upon addition of ABA, the two halves are reconstituted and localized to the ZF BS array.

(B) Dose-dependent activation of episomal reporter by an inducible two-module circuit. Cells were co-transfected with Interspersed *pMinCMV* Reporter, synI^{IND}, and synR^{VP64} expression constructs. GFP fluorescence intensity was measured by flow cytometry 2 days after ABA induction, and normalized to reporter only-transfected cells to obtain GFP fold change. (n = 3; error bars, SD).

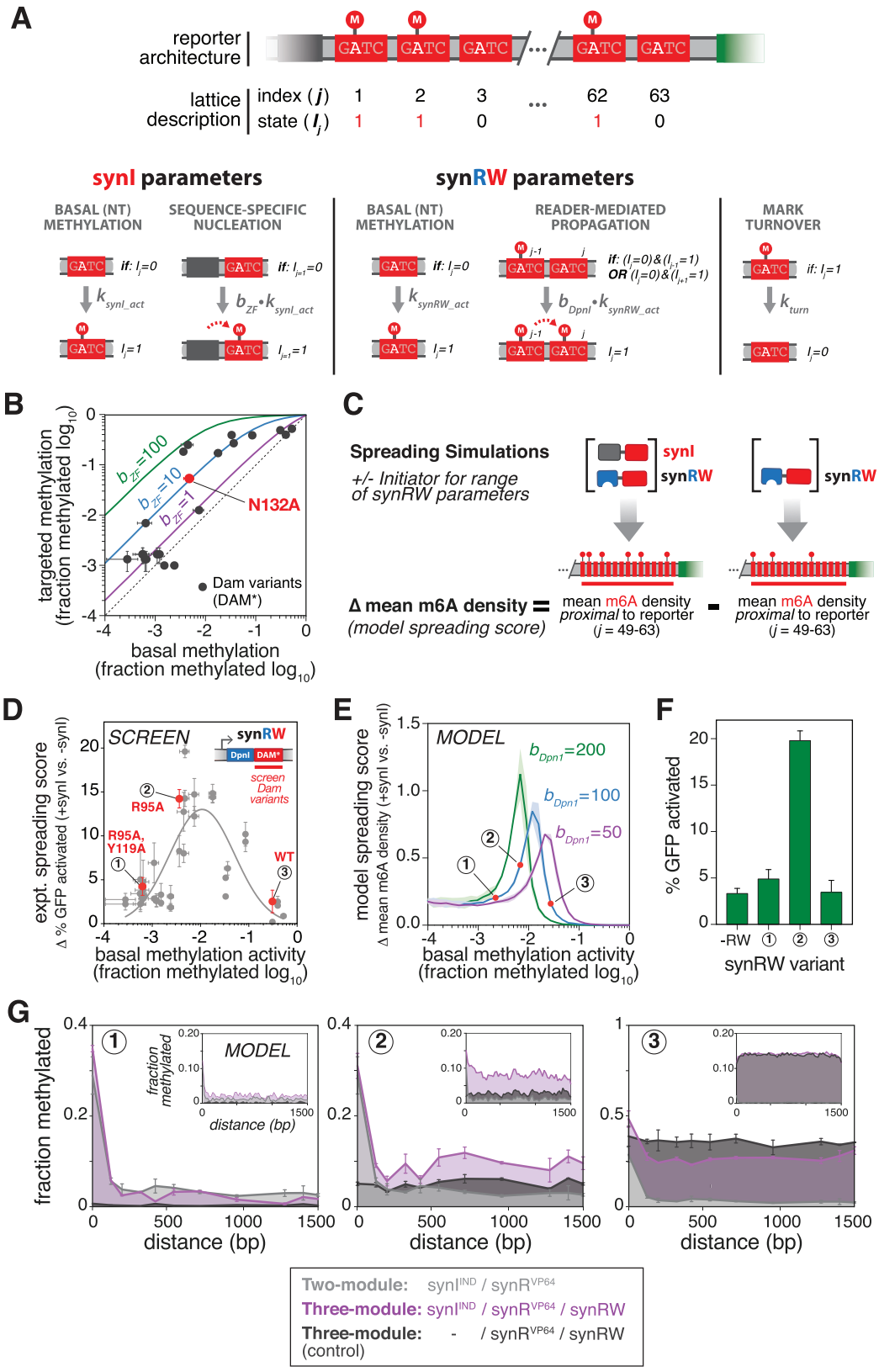
(C) Dynamic range of activation of episomal reporter by an inducible two-module circuit. Cells were co-transfected with Interspersed *pMinCMV* Reporter plasmid and indicated combinations of synI^{IND} and synR^{VP64} expression constructs. Values indicate GFP fold change over reporter only control. +ABA = 200 μM ABA. (n = 3; error bars, SD).

(D) Inducible CRISPR-guided initiator, synI$^{dCas9\ IND}$. Top: Schematic of synI$^{dCas9\ IND}$. Below: Dynamic range of activation of episomal reporter by a two-module circuit using synI$^{dCas9\ IND}$. Cells were co-transfected with Interspersed *pMinCMV* Reporter, synI$^{dCas9\ IND}$ and synR^{VP64} expression constructs, and gRNAs (I1-I5) in the combinations shown, and GFP fluorescence assayed by flow cytometry. gRNA target locations are shown in Figure S3H. Values indicate GFP fold change over reporter only control. +ABA = 200 μM ABA. (n = 3; error bars, SD).

(E) Quantifying percentage of GFP activated cells induced by the two-module circuit (synI^{IND}, synR^{VP64}). Flow cytometry GFP fluorescence distributions for Interspersed Reporter cell lines stably expressing two-module system (+synI^{IND}) or synR^{VP64} only (-synI^{IND}) with 200 μM ABA induction (+ABA). The threshold for GFP activated cells is indicated.

(legend continued on next page)

(F) Deactivation dynamics following transient induction with synI^{IND}. Interspersed Reporter cell lines stably expressing inducible two-module circuit (synI^{IND}, synR^{VP64}) were induced with 200 μM ABA pulses of different duration, and the percentage of GFP activated cells quantified by flow cytometry over 20 days after ABA washout. Left: Absolute percentage of activated cells. Right: Normalized to maximum value. Dots are data points (n = 3; error bars, SD); lines represent exponential fits to the OFF phase. Model simulations of % methylated sites as a function of time for different initial methylation states are shown in inset (see STAR Methods; Figure 6D).



(legend on next page)

Figure S5. Model of m6A Spatial Dynamics, Related to Figure 4

(A) Lattice model of chromatin spreading dynamics. In this model, our synthetic promoter is represented by a one-dimensional array of 63 GATC sites (index j), each of which can be in one of two states: unmethylated ($l_j = 0$), and methylated ($l_j = 1$). The dynamics are controlled by four reactions: (1) Basal (non-targeted) methylation governed by rates k_{synI_act} (for synI) and k_{synRW_act} (for synRW); (2) Sequence-specific nucleation proximal to the ZF BS array at a rate $b_{ZF} \cdot k_{synI\_act}$; (3) If nearest neighbor site is methylated, reader-mediated writing occurs at a rate $b_{DpnI} \cdot k_{synRW\_act}$; (4) Turnover from any methylated GATC site at a constant rate k_{turn}.

(B) Model parameterization of m6A nucleation by synI. Model profiles of steady-state targeted versus basal methylation by synI for different values of b_{ZF} (colored lines), overlaid onto the experimentally determined values for the Dam mutant library (see Figure S2F). synI (Dam N132A) is described well with $b_{ZF} = 10$ (see STAR Methods).

(C) Computational workflow for simulating and quantifying spatial spreading. Mean m6A density at promoter-proximal sites (15 downstream sites, j = 49-63) is computed after simulating synRW reactions for a range of k_{synRW_act} and b_{DpnI} values (see STAR Methods). To obtain a "model spreading score" (analogous to expt. spreading score used in screen (Figure 4B)), we computed the difference (Δ) in mean m6A density for simulations with and without synI.

(D) Summary of spatial propagation screen results. The "expt. spreading score" is a quantitative metric used to score propagation propensity for synRW library members. This is plotted for all library members as a function of Dam methylation activity. Three synRW variants representing different levels of writer (methylation) activity are highlighted: 1) Low = Dam (R95A, Y119A); 2) Intermediate = Dam (R95A); 3) High = Dam (WT). (n = 3; error bars, SD). Data are fitted with a log-normal distribution.

(E) Summary of spatial propagation simulation results. The "model spreading score" is a quantitative metric used to score propagation propensity for the simulated synRW library members. Model-generated profiles of "model spreading score" (Δ mean m6A density) are plotted as a function of Dam methylation activity. Simulation results are shown for three different parameter values of b_{DpnI} (reader-mediated writing specificity multiplier), and the three representative synRW variants (low, intermediate, high Dam writer activity) are highlighted.

(F) Spreading-dependent reporter activation in Clustered Reporter cells (stably expressing $synI^{IND}$ and $synR^{VP64}$), transfected with one of three synRW variants: low, intermediate, high Dam writer activity. Percentage of GFP activated cells was quantified as previously described. Cells were continuously induced with 200 μM ABA. (n = 3; error bars, SD).

(G) m6A profiles measured across the GATC array in Clustered Reporter cells (stably expressing $synI^{IND}$ and $synR^{VP64}$), transfected with one of three synRW variants: low, intermediate, high Dam writer activity. Cells were continuously induced with 200 μM ABA. (n = 3; error bars, SD). Model simulation results are shown in insets with $b_{ZF} = 10$, $b_{DpnI} = 100$, $k_{synRW\_act} = 1.1 \cdot 10^{-4}$ (left), $3.4 \cdot 10^{-4}$ (middle), and $1.4 \cdot 10^{-3}$ hr^{-1} (right).

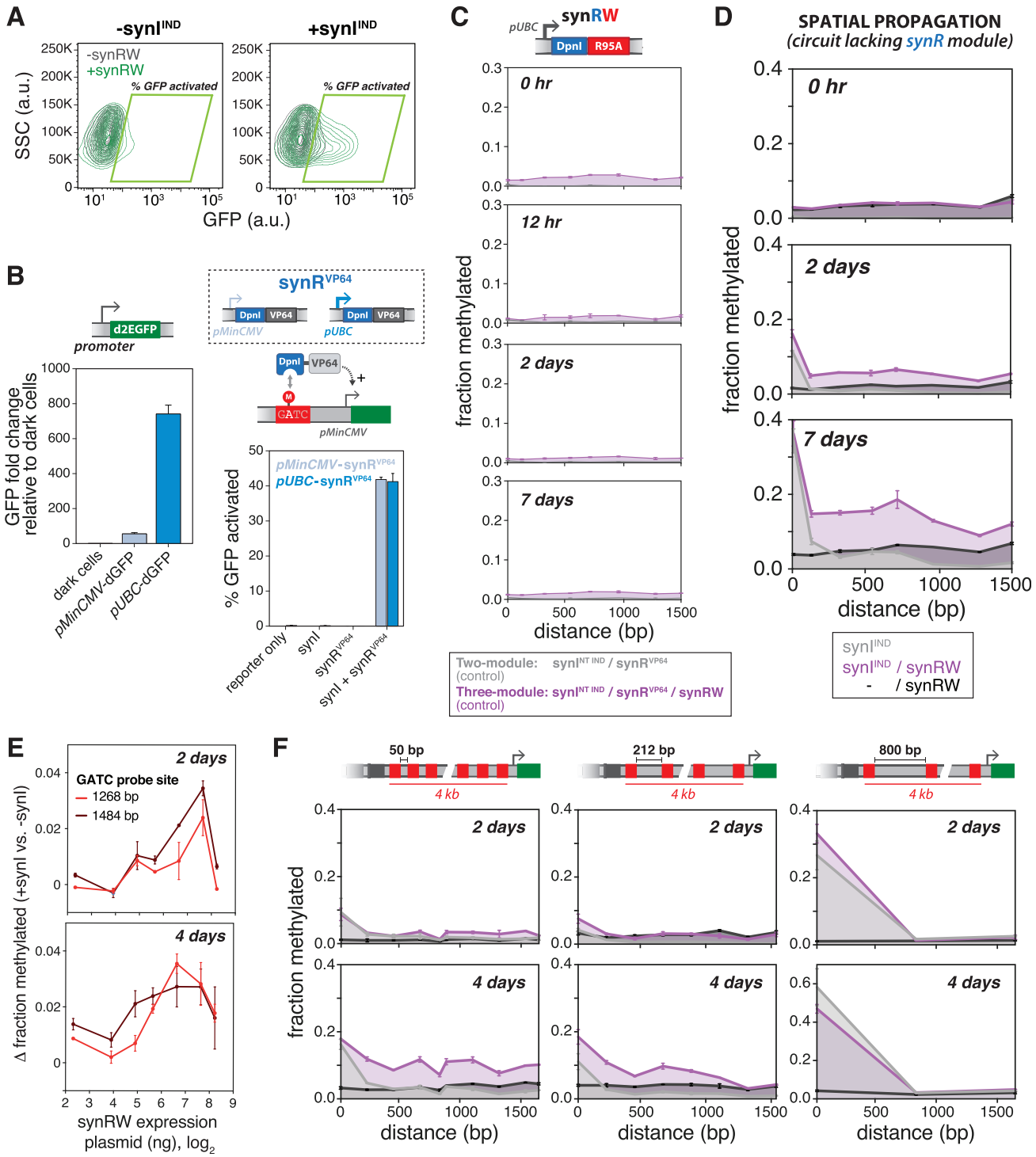


Figure S6. Characterization and Controls for Synthetic Spatial Propagation, Related to Figures 4 and 5

(A) Flow cytometry plots of representative circuit candidate chosen from the spatial propagation screen (see Figure 4B). The percentage of GFP activated cells for a given synRW candidate is quantified in cells with (+synI^{IND}) and without (-synI^{IND}) inducible initiator.

(B) m6A-mediated transcriptional regulation by the two-module circuit is insensitive to synR expression level. Left: Relative expression levels for a weak (*pMinCMV*) and strong (*pUBC*) promoter, assayed by measuring GFP expression from plasmid-transfected cells. GFP fluorescence intensity, measured by flow cytometry, was normalized to dark cells to obtain GFP fold change. (n = 3; error bars, SD). Right: m6A-mediated activation by synthetic two-module system, with different expression levels of synR^{VP64}. Interspersed Reporter lines were co-transfected with synI and synR^{VP64}, driven by either weak *pMinCMV* or strong *pUBC*, and the percentage of GFP activated cells was quantified by flow cytometry. (n = 3; error bars, SD).

(legend continued on next page)

(C) m6A profiles measured across the GATC array and over time for cells stably expressing a control propagation circuit, containing a non-targeting synI ($synI^{NT\ IND}$). Cells were continuously induced with 200 μM ABA. (n = 3; error bars, SD).

(D) m6A profiles measured across the GATC array and over time for cells stably expressing the propagation circuit without the synR module. Cells were continuously induced with 200 μM ABA. (n = 3; error bars, SD).

(E) Spatial propagation as a function of synRW dosage. The difference in fraction methylated (measured by m6A-qPCR at promoter proximal sites) between Clustered Reporter cells with and without $synI^{IND}$ is plotted for varying concentrations of transfected synRW-expressing plasmid. Cells were also stably expressing synR and were continuously induced with 200 μM ABA. (n = 3; error bars, SD).

(F) Synthetic spatial propagation on Clustered Reporters of different inter-GATC spacings. Reporter cell lines were generated by singly integrating respective Clustered Reporter variant constructs. Reporter cell lines stably expressing $synI^{IND}$ and $synR^{VP64}$ were transfected with synRW and continuously induced with 200 μM ABA. (n = 3; error bars, SD).

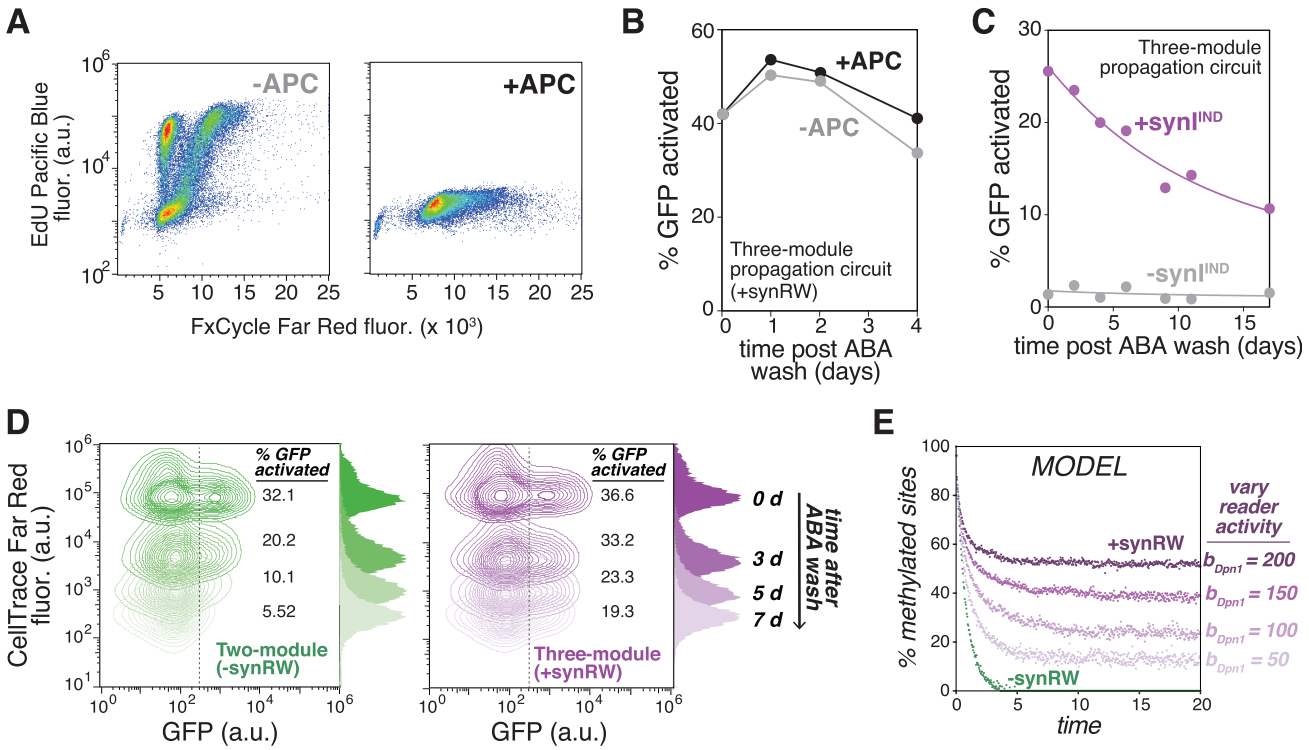


Figure S7. Characterization of Epigenetic Memory Induced by Three-Module, Read-Write Circuits, Related to Figure 6

(A) Using Aphidicolin treatment to inhibit DNA replication in cells. Flow cytometry plots of Interspersed Reporter cells (stably expressing synI^{IND} and synR^{VP64}) that were either control-treated (-APC, top) or treated with the DNA replication inhibitor Aphidicolin (+APC, bottom) (see STAR Methods). DNA replication was monitored with 5-ethynyl-2′-deoxyuridine (EdU) Pacific Blue, which is incorporated into newly synthesized DNA. Cells were also stained with FxCycle Far Red to measure DNA content.

(B) Maintenance of the m6A-dependent transcriptional state by synRW. Cells stably expressing the propagation circuit were induced with a 3 day ABA pulse, then ABA was washed out and cells were maintained in media with or without 5 μg/mL APC. (n = 3; error bars, SD).

(C) m6A establishment by synI^{IND} is required for GFP activation and epigenetic memory. Cells stably expressing either the three-module propagation circuit (+synI^{IND}) or the circuit lacking initiator (-synI^{IND}) were induced with a 2 hr ABA pulse, and the percentage of GFP activated cells was quantified at the indicated time points following ABA washout. Curves represent exponential fits. (n = 3; error bars, SD).

(D) Cells expressing the three-module propagation circuit are actively dividing and transmitting the epigenetic state. Flow cytometry plots of cells stably expressing either the propagation circuit (+synRW) or the circuit lacking synRW (-synRW), labeled with CellTrace Far Red to trace cell divisions across time following a 3 day ABA pulse. Dashed line represents cutoff for percentage of GFP activated cells, showing that GFP+ state is transmitted to progeny in cellular populations expressing the propagation circuit.

(E) Simulation results of m6A maintenance for varying synRW reader activity. The percentage of methylated GATC sites (out of 14 available in the Interspersed Reporter) are plotted as a function of time for simulations in which the reader activity parameter, b_{DpnI}, is varied (see STAR Methods).