Enhanced CRISPR-based DNA demethylation by Casilio-ME-mediated RNA-guided coupling of methylcytosine oxidation and DNA repair pathways

Aziz Taghbalout, Menghan Du, Nathaniel Jillette, Wojciech Rosikiewicz, Abhijit Rath, Christopher D. Heinen, Sheng Li & Albert W. Cheng

Here we develop a methylation editing toolbox, Casilio-ME, that enables not only RNA-guided methylcytosine editing by targeting TET1 to genomic sites, but also by co-delivering TET1 and protein factors that couple methylcytosine oxidation to DNA repair activities, and/or promote TET1 to achieve enhanced activation of methylation-silenced genes. Delivery of TET1 activity by Casilio-ME1 robustly alters the CpG methylation landscape of promoter regions and activates methylation-silenced genes. We augment Casilio-ME1 to simultaneously deliver the TET1-catalytic domain and GADD45A (Casio-ME2) or NEIL2 (Casio-ME3) to streamline removal of oxidized cytosine intermediates to enhance activation of targeted genes. Using two-in-one effectors or modular effectors, Casilio-ME2 and Casilio-ME3 remarkably boost gene activation and methylcytosine demethylation of targeted loci. We expand the toolbox to enable a stable and expression-inducible system for broader application of the Casilio-ME platforms. This work establishes a platform for editing DNA methylation to enable research investigations interrogating DNA methylomes.
DNA methylation is part of the multifaceted epigenetic modifications of chromatin that shape cellular differentiation, gene expression, and maintenance of cellular homeostasis. Aberrant DNA methylation is implicated in various diseases including cancer, imprinting disorders, and neurological diseases. Developing tools to directly edit the methylation state of a specific genomic locus is of significant importance both for studying the biology of DNA methylation as well as for development of therapies to treat DNA methylation-associated diseases.

In mammalian cells, the 5-methylcytosine (5mC) epigenetic mark generated by covalent linkage of a methyl group to the 5th position of the cytosine ring of CpG sequences is catalyzed by one of the three canonical DNA methyltransferases DMNT1, DNMT3A, and DNMT3B. DNA methylation is dynamic and involves demethylation pathways which erase 5mC to restore unmethylated DNA. Active demethylation involves the ten–eleven translocation (TET) family of methylcytosine dioxygenases that iteratively oxidize 5mC into 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) intermediates. Subsequently, 5fC and 5caC are processed by the base-excision repair (BER) machinery to restore unmethylated cytosines. Restoration of an intact DNA base is initiated by DNA glycosylases that excise damaged bases to generate an apurinic/apyrimidinic site (AP site) for processing by the rest of the BER machinery. Thymine DNA glycosylase (TDG)-based BER has been functionally linked to TET1-mediated demethylation, suggesting an interplay between TET1 and enzymes of the BER machinery to actively erase 5mC oxidation to DNA repair.

Advances in artificial transcription factor (ATF) technologies have enabled direct control of gene expression and epigenetic states. CRISPR/Cas9-based technologies allow much flexibility and scalability because the specificity is programmable by a single guide RNA (sgRNA). Tethering of TET1 or DNMT3a to genomic targets by use of ATFs has been shown to allow targeted removal or deposition of DNA methylation. However, these ATF systems have inherent limitations in enabling multiplexed targeting, effector multimerization, or formation of protein complexes at the targeted sequence. We recently developed the Casilio system which uses an extended sgRNA scaffold to assemble protein factors at target sites, enabling multimerization, differential multiplexing, and potentially stoichiometric complex formation. Here we develop an advanced DNA methylation editing technology which allows targeted bridging of TET1 activity to BER machinery to efficiently alter the epigenetic state of CpG targets and activate methylation-silenced genes. Casilio-DNA Methylation Editing (ME) platforms enable targeted delivery of the TET1 effector (Casio-ME1) or in association with GADD45A (Casio-ME2) or NEIL2 (Casio-ME3) to achieve enhanced 5mC demethylation and gene activation. We show that Casilio-ME- mediated delivery of TET1 activity to gene promoters induces robust cytosine demethylation within the targeted CpG island (CGI) and activation of gene expression. When systematically compared to other reported methylation editing systems, Casilio-ME shows superior activities in mediating transcriptional activation of methylation-silenced gene and 5mC demethylation. The ability of Casilio-ME to mediate co-delivery of TET1 activity along with other protein factors, which enhance turnover of oxidized cytosine intermediates, paves the way for new areas of research to efficiently address the cause–effect relationships of DNA methylation in normal and pathological processes.

**Results**

**Casilio-ME1 delivers TET1 activity to a genomic target site.** Casilio-ME1 is a three-component DNA Methylation Editing platform built on Casilio which uses nuclease-deficient Cas9 (dCas9), an effector module made of Pumilio/FBF (PUF) domain linked to an effector protein, and a modified sgRNA containing PUF-binding sites (PBS) (Fig. 1a)30. The dCas9/sgRNA ribonucleoprotein complex binds DNA targets without cutting to serve as a RNA-guided DNA-binding vehicle whose specificity is dictated by the spacer sequence of the sgRNA and a short protospacer adjacent motif located on the target DNA. PUF-tethered effectors are recruited to the ribonucleoprotein complex via binding to their cognate PBS present on the sgRNA scaffold. PUF domains are found in members of an evolutionarily conserved family of eukaryotic RNA-binding proteins whose specificity is encoded within their structural tandem repeats, each of which recognizes a single ribonucleobase. PUF domains can be programmed to bind to any 8-mer RNA sequence, e.g., PUFa and PUFc used in this study were designed to bind PBSa (UUGAUGUA) and PBSc (UUGAUGUA), respectively. Multiple PBS added in tandem to the 3’ end of the sgRNA allow concurrent recruitment of multiple PUF-effectors to targeted DNA sequences without interfering with dCas9 targeting, and therefore allow amplification of the response to associated effector modules.

To enable targeted cytosine demethylation and subsequent activation of methylation-silenced genes, we built a DNA methyl editor TET1-effector Casilio-ME1 as a protein fusion of hTET1 catalytic domain (TET1(CD)) to the carboxyl end of PUFa (Fig. 1a). We chose as a target the MLH1 promoter region that is part of a large CGI whose aberrant hypermethylation induces MLH1-silencing in 10–30% of colorectal and other cancers. MLH1 is silenced in HEK293T cells and therefore represents a clinically relevant model for developing Casilio-ME.

To test the system, cells were transiently transfected with plasmids encoding Casilio-ME1 components PUFa-TET1(CD) effector, dCas9 and six MLH1-promoter-targeting sgRNAs each containing five copies of PBSa (Fig. 1a). This resulted in robust MLH1 activation as indicated by the obtained fold changes in MLH1 mRNA in cells collected on day 3 post-transfection (Fig. 1b, c upper panel). In contrast, MLH1 activation was not obtained with a non-targeting sgRNA (NT-sgRNA) (Fig. 1b), indicating that Casilio-ME1-mediated MLH1 activation requires specific targeting of the PUFa-TET1(CD) module directed by the programmable sgRNAs.

Evidence that the Casilio-ME1-induced activation of MLH1 results from TET1-mediated 5mC demethylation came from high throughput bisulfite sequencing (BSeq) of MLH1 amplions derived from the cells analyzed in Fig. 1b. BSeq showed that targeted delivery of the TET1(CD) effector induces a profound decrease in CpG methylation frequency within the MLH1 promoter region (Fig. 1c lower panel). Demethylation activity was prominently higher within CpGs neighboring MLH1-sgRNA sites (Fig. 1c lower panel (arrows)), and seemed to spread away, albeit with relatively reduced activities. These data indicate that Casilio-ME1 mediates delivery of TET1
activity to promoter regions to induce 5mC demethylation within the targeted CGI and subsequent activation of the methylation-silenced gene.

Comparison of Casilio-ME1 with other TET1 delivery systems. Although other technologies enabling targeted delivery of TET1 activity to genomic loci have been reported to induce activation of methylation-silenced genes, a direct comparison of their efficiency is lacking. Here we compared Casilio-ME1 efficiency to alter expression of methylation-regulated genes to alternative technologies for 5mC demethylation that are based on TALEs (transcription activator like effector), dCas9/MS2 or dCas9/SunTag systems (Fig. 1d). We therefore assembled four TALE-TET1(CD) fusions each designed to bind to one of the four MLH1-sgRNA target sequences used for dCas9-based delivery.
systems. Relative quantitation of MLHI mRNA indicated that the SunTag, TALEs or MS2 based systems only achieved 63%, 7% or 1%, respectively, of the Casilio-ME1-mediated activation level (Fig. 1e). BSeq analysis of MLHI promoter comparing Casilio-ME1 and SunTag systems showed that Casilio-ME1 induced stronger demethylation at most of the Cpg sites examined (Supplementary Fig. 1a, b). The efficient MLHI activation obtained with Casilio-ME1 delivery of TET1 activity as compared to SunTag is not driven by Cpg DNA methylation as similar trends were obtained when one sgRNA was used for targeting MLHI Cpg (Supplementary Fig. 1c). These results suggest that Casilio-ME1-mediated delivery of TET1 activity to Cpg target enables stronger 5mC demethylation and gene activation compared to published systems.

Co-delivery of TET1(CD) and GADD45A enhances gene activation. Active 5mC erasure is a two-step process initiated by TET1-mediated iterative 5mC oxidations followed by base-excision (BER) or nucleotide-excision (NER) repair conversions of oxidized intermediates to cytosines. Thus, coupling these two steps could streamline Cpg active erasure to efficiently activate methylation-silenced genes. Because GADD45A promotes TET1 activity and/or could recruit key player(s) of DNA repair, we sought to augment Casilio-ME by constructing an upgraded version Casilio-ME2 which simultaneously recruits TET1(CD) and GADD45A to bridge 5mC oxidation with DNA repair at a specific genomic locus.

To test Casilio-ME2 in inducing MLHI activation in a comparison with Casilio-ME1 and our previously reported Casilio-p65HSF1 activator, we introduced plasmids encoding the protein fusions of PUFa-p65HSF1 (activator), PUFa-TET1(CD) (Casilio-ME1), PUFa-GADD45A-TET1(CD) (Casilio-ME2.1), or GADD45A-PUFa-TET1(CD) (Casilio-ME2.2) (Fig. 2a), along with dCas9 and sgRNA plasmids into HEK293T cells. Casilio-ME1-mediated MLHI activation was 46% higher than that obtained with the PUFa-p65HSF1 activator module (Fig. 2b). Interestingly, when GADD45A was added as part of Casilio-ME2.1 or Casilio-ME2.2 TET1-effectors, MLHI mRNA expression was augmented by 3 and 6-fold, respectively, compared to Casilio-ME1 (Fig. 2b). This enhanced MLHI activation, obtained with GADD45A as part of the TET1 effector modules, does not result from higher expression of Casilio-ME2 effectors (Supplementary Fig. 2). Thus, coupling of GADD45A and TET1(CD) as a two-in-one effector enhances TET1-mediated activation of methylation-silenced genes.

To obtain further evidence that co-delivery of TET1(CD) and GADD45A effectors to target sites enhances gene activation compared to delivery of TET1(CD) alone, we fused each effector to a separate PUF protein, i.e., TET1(CD) to PUFa-Casilio-ME1 delivered with GADD45A to PUFa-Casilio-ME2 and used sgRNA containing both PUFs and PBSs to tether the respective PUF-fusion to the sgRNA scaffold (Fig. 2c). When Casilio-ME2.3 (PUFc-GADD45A) or Casilio-ME2.4 (GADD45A-PUFc) components were introduced to cells, 3- and 6-fold increase in TET1-mediated MLHI activation was obtained, respectively (Fig. 2d). However, no MLHI expression was detected using Casilio-ME2.3 and ME2.4 systems when a catalytically dead TET1(CD) (dTET1(CD)) replaced wild-type TET1(CD), indicating that the observed GADD45A-mediated stimulation of gene activation requires the oxidative activity of TET1 (Fig. 2d). Similarly, no MLHI activation was obtained when the GADD45A module of Casilio-ME2.3 and ME2.4 systems were introduced into cells without the PUFa-TET1(CD) component (Fig. 2c, d), indicating that expression of GADD45A alone, in the absence of the TET1 module, does not mediate gene activation. In addition, when the sgRNAs contained PBSa but lacked PBSd required for tethering PUFc-associated modules, GADD45A modules failed to stimulate MLHI activation (Supplementary Fig. 3). Thus, enhancement of gene activation in Casilio-ME2 requires co-delivery of GADD45A and TET1-effector modules to the target site.

Co-delivery of TET1(CD) and BER enzymes. TDG, NEIL1, and NEIL2 have been functionally linked to active DNA demethylation as they are involved in the initial step of removing oxidized cytosines 5C and 5caC produced by TET1 activities. Because initiating repair of oxidized cytosines by the BER machinery might be a rate limiting step to TET1-mediated activation of methylation-silenced genes, we reasoned that coupling TET1 activities with DNA glycosylases could facilitate 5mC active erasure and enhance subsequent gene activation.

We therefore linked NEIL1, NEIL2, NEIL3, or TDG to the PUFa-TET1(CD) effector as single-chain protein fusions and looked for potential gains in Casilio-ME1-mediated MLHI activation. Among these, only NEIL2 fusions showed enhanced activation of MLHI expression (Supplementary Fig. 4). Casilio-ME3.1 and Casilio-ME3.2, in which NEIL2 is fused N-terminally to PUFa or between PUFa and TET1(CD) of the PUFa-TET1(CD) effector, respectively, increased MLHI activation 4-fold in the presence of MLHI-sgRNAs compared to Casilio-ME1 (Fig. 3a, b). Thus, co-delivering TET1(CD) and NEIL2 as a two-in-one effector to target sites improves activation of 5mC-silenced genes.

To further show that NEIL2 promotes demethylation-mediated gene activation, TET1(CD) and NEIL2 were co-delivered as separate effectors to MLHI promoter regions by fusing TET1 (CD) and NEIL2 to PUFa and PUFc, respectively, and using
Fig. 2 Casilio-ME2 mediates dual delivery of TET1(CD) and GADD45A to targeted genomic sites. a Schematic representation of the indicated Casilio-ME and Casilio platforms showing effector modules of PUFa fusion proteins used to transfect cells analyzed in (b). TET1(CD) (black), GADD45A (blue), p65HSF1 (green), PUFa (light gray), amino (N), and carboxyl (C) termini of protein fusions and occupancy of PBSa are arbitrarily shown. b Plot showing MLH1 mRNA relative levels (mean fold change ± S.E.M.; n = 3) in cells transfected with components of Casilio-ME1, Casilio-ME2.1, Casilio-ME2.2, or Casilio/p65HSF1 in the presence of MLH1-sgRNAs or NT-sgRNA as indicated. Cells were collected 3 days after transfection. Drawing of promoter regions with the MLH1-sgRNAs used (A-F), CpGs (lollipops), and TSS (arrow) is shown above the plot. ***p < 0.0005, one-way ANOVA. c Schematic representation of the indicated Casilio-ME2 platforms showing effector modules of PUFa and PUFc fusion proteins used to transfect cells analyzed in panel (d). TET1(CD) (black), GADD45A (blue), PUFa (light gray), PUFc (orange), sgRNA containing both PBSa and PBSc, amino (N) and carboxyl (C) termini of protein fusions are arbitrarily shown. d MLH1 mRNA relative levels (mean fold change ± S.E.M.; n = 3) in cells transfected with components of Casilio-ME1, Casilio-ME2.3, or Casilio-ME2.4 in the presence of MLH1-sgRNAs or NT-sgRNA is shown. When indicated PUFa-TET1(CD), effector component of Casilio-ME2.3 and Casilio-ME2.4, was replaced by a catalytically dead PUFa-TET1(CD) effector containing TET1-inactivating mutations H1671Y, D1673A57, or omitted. MLH1 promoter with the sgRNAs used (A-F), CpGs (lollipops), and TSS (arrow) is depicted above the plot. ***p < 0.0005, one-way ANOVA.
sgRNAs comprising both PBSa and PBSc (Fig. 3c). Binding of these effectors to a sgRNA scaffold brings TET1(CD) and NEIL2 into close proximity, and potentially enables coupling of DNA demethylation with BER. When Casilio-ME3.3 (PUFc-NEIL2) or Casilio-ME3.4 (NEIL2-PUFc) components were used, MLH1 activation was increased by 7-fold as compared to Casilio-ME1 (Fig. 3d). Taken together, these results show that co-delivery of TET1(CD) and NEIL2 DNA glycosylase/AP-lyase stimulates activation of a methylation-silenced gene.

To determine whether the enhanced gene activation obtained with Casilio-ME3.3 and ME3.4 systems requires co-targeting of TET1(CD) and NEIL2 to a genomic site and does not result from NEIL2 over expression, we disabled targeting of NEIL2 effector modules by using sgRNAs comprising PBSa but lacking the PBSc required for targeting PUFc-based NEIL2 effectors (Supplementary Fig. 5a). Cells transfected with Casilio-ME3.3 or ME3.4 components comprising sgRNAs that lacked PBSc tethering sites showed no significant gains in TET1-mediated MLH1 activation, indicating that enhanced TET1-mediated gene activation requires co-targeting of NEIL2 and TET1(CD) modules via an RNA scaffold (Supplementary Fig. 5b). Thus, these data show that co-delivery of NEIL2 and TET1(CD) to genomic loci synergistically promotes TET1-mediated gene activation, likely via facilitated coupling of 5mC demethylation and BER activities to efficiently restore unmethylated cytosine to targeted sites.

**Comparison of Casilio-ME platforms.** Casilio-ME2 and Casilio-ME3 platforms showed an enhanced activation of a methylation-silenced gene compared to Casilio-ME1. Here we sought to compare these platforms to one another in their efficiencies to activate MLH1 and alter methylation landscape of targeted CGI. The comparison included the previously reported dCas9-TET1 as an alternative system for reference. Normalized MLH1 levels, to those obtained with dCas9-TET1 system, showed that Casilio-ME3 platforms showed an enhanced activation of a methylation-silenced gene compared to Casilio-ME1. Here we sought to compare these platforms to one another in their efficiencies to activate MLH1 and alter methylation landscape of targeted CGI. The comparison included the previously reported dCas9-TET1 as an alternative system for reference. Normalized MLH1 levels, to those obtained with dCas9-TET1 system, showed that Casilio-ME3 platforms showed an enhanced activation of a methylation-silenced gene compared to Casilio-ME1.
ME2.2 gave augmented MLH1 activation higher than Casilio-ME1, ME2.1, ME3.1, and ME3.2 platforms (Supplementary Fig. 6a). This augmented MLH1 activation does not necessarily require targeting with multiple sgRNAs as a similar trend in MLH1 activation was obtained with one sgRNA targeting MLH1 promoter region (Fig. 6b). When modular Casilio-ME2.3, ME2.4, and ME3.4 platforms were compared, Casilio-ME2.4 showed the most MLH1 activation (Supplementary Fig. 6c). Only background MLH1 mRNA levels could be detected with TET1 (CD)-dead mutants of the Casilio-ME platforms (Casilio-dME) (Supplementary Fig. 6a, c), indicating that TET1 activity is required for gene activation and that delivery of GADD45A or NEIL2 without TET1 oxidative activity is not sufficient for activating methylation-silenced genes.

To ask whether the augmented MLH1 activation of Casilio-ME2.2 and Casilio-ME3.1 came from an increased efficiency in 5mC erasure, we performed BSeq and oxidative BSeq (oxBSeq) by NATURE COMMUNICATIONS| (2019) 10:4296 | https://doi.org/10.1038/s41467-019-12339-7 | www.nature.com/naturecommunications ME2.2 gave augmented MLH1 activation higher than Casilio-ME1, ME2.1, ME3.1, and ME3.2 platforms (Supplementary Fig. 6a). This augmented MLH1 activation does not necessarily require targeting with multiple sgRNAs as a similar trend in MLH1 activation was obtained with one sgRNA targeting MLH1 promoter region (Fig. 6b). When modular Casilio-ME2.3, ME2.4, and ME3.4 platforms were compared, Casilio-ME2.4 showed the most MLH1 activation (Supplementary Fig. 6c). Only background MLH1 mRNA levels could be detected with TET1 (CD)-dead mutants of the Casilio-ME platforms (Casilio-dME) (Supplementary Fig. 6a, c), indicating that TET1 activity is required for gene activation and that delivery of GADD45A or NEIL2 without TET1 oxidative activity is not sufficient for activating methylation-silenced genes.

To ask whether the augmented MLH1 activation of Casilio-ME2.2 and Casilio-ME3.1 came from an increased efficiency in 5mC erasure, we performed BSeq and oxidative BSeq (oxBSeq) by NATURE COMMUNICATIONS| (2019) 10:4296 | https://doi.org/10.1038/s41467-019-12339-7 | www.nature.com/naturecommunications ME2.2 gave augmented MLH1 activation higher than Casilio-ME1, ME2.1, ME3.1, and ME3.2 platforms (Supplementary Fig. 6a). This augmented MLH1 activation does not necessarily require targeting with multiple sgRNAs as a similar trend in MLH1 activation was obtained with one sgRNA targeting MLH1 promoter region (Fig. 6b). When modular Casilio-ME2.3, ME2.4, and ME3.4 platforms were compared, Casilio-ME2.4 showed the most MLH1 activation (Supplementary Fig. 6c). Only background MLH1 mRNA levels could be detected with TET1 (CD)-dead mutants of the Casilio-ME platforms (Casilio-dME) (Supplementary Fig. 6a, c), indicating that TET1 activity is required for gene activation and that delivery of GADD45A or NEIL2 without TET1 oxidative activity is not sufficient for activating methylation-silenced genes.

To ask whether the augmented MLH1 activation of Casilio-ME2.2 and Casilio-ME3.1 came from an increased efficiency in 5mC erasure, we performed BSeq and oxidative BSeq (oxBSeq) by NATURE COMMUNICATIONS| (2019) 10:4296 | https://doi.org/10.1038/s41467-019-12339-7 | www.nature.com/naturecommunications ME2.2 gave augmented MLH1 activation higher than Casilio-ME1, ME2.1, ME3.1, and ME3.2 platforms (Supplementary Fig. 6a). This augmented MLH1 activation does not necessarily require targeting with multiple sgRNAs as a similar trend in MLH1 activation was obtained with one sgRNA targeting MLH1 promoter region (Fig. 6b). When modular Casilio-ME2.3, ME2.4, and ME3.4 platforms were compared, Casilio-ME2.4 showed the most MLH1 activation (Supplementary Fig. 6c). Only background MLH1 mRNA levels could be detected with TET1 (CD)-dead mutants of the Casilio-ME platforms (Casilio-dME) (Supplementary Fig. 6a, c), indicating that TET1 activity is required for gene activation and that delivery of GADD45A or NEIL2 without TET1 oxidative activity is not sufficient for activating methylation-silenced genes.

To ask whether the augmented MLH1 activation of Casilio-ME2.2 and Casilio-ME3.1 came from an increased efficiency in 5mC erasure, we performed BSeq and oxidative BSeq (oxBSeq) by NATURE COMMUNICATIONS| (2019) 10:4296 | https://doi.org/10.1038/s41467-019-12339-7 | www.nature.com/naturecommunications ME2.2 gave augmented MLH1 activation higher than Casilio-ME1, ME2.1, ME3.1, and ME3.2 platforms (Supplementary Fig. 6a). This augmented MLH1 activation does not necessarily require targeting with multiple sgRNAs as a similar trend in MLH1 activation was obtained with one sgRNA targeting MLH1 promoter region (Fig. 6b). When modular Casilio-ME2.3, ME2.4, and ME3.4 platforms were compared, Casilio-ME2.4 showed the most MLH1 activation (Supplementary Fig. 6c). Only background MLH1 mRNA levels could be detected with TET1 (CD)-dead mutants of the Casilio-ME platforms (Casilio-dME) (Supplementary Fig. 6a, c), indicating that TET1 activity is required for gene activation and that delivery of GADD45A or NEIL2 without TET1 oxidative activity is not sufficient for activating methylation-silenced genes.

To ask whether the augmented MLH1 activation of Casilio-ME2.2 and Casilio-ME3.1 came from an increased efficiency in 5mC erasure, we performed BSeq and oxidative BSeq (oxBSeq) by NATURE COMMUNICATIONS| (2019) 10:4296 | https://doi.org/10.1038/s41467-019-12339-7 | www.nature.com/naturecommunications ME2.2 gave augmented MLH1 activation higher than Casilio-ME1, ME2.1, ME3.1, and ME3.2 platforms (Supplementary Fig. 6a). This augmented MLH1 activation does not necessarily require targeting with multiple sgRNAs as a similar trend in MLH1 activation was obtained with one sgRNA targeting MLH1 promoter region (Fig. 6b). When modular Casilio-ME2.3, ME2.4, and ME3.4 platforms were compared, Casilio-ME2.4 showed the most MLH1 activation (Supplementary Fig. 6c). Only background MLH1 mRNA levels could be detected with TET1 (CD)-dead mutants of the Casilio-ME platforms (Casilio-dME) (Supplementary Fig. 6a, c), indicating that TET1 activity is required for gene activation and that delivery of GADD45A or NEIL2 without TET1 oxidative activity is not sufficient for activating methylation-silenced genes.

To ask whether the augmented MLH1 activation of Casilio-ME2.2 and Casilio-ME3.1 came from an increased efficiency in 5mC erasure, we performed BSeq and oxidative BSeq (oxBSeq) by NATURE COMMUNICATIONS| (2019) 10:4296 | https://doi.org/10.1038/s41467-019-12339-7 | www.nature.com/naturecommunications ME2.2 gave augmented MLH1 activation higher than Casilio-ME1, ME2.1, ME3.1, and ME3.2 platforms (Supplementary Fig. 6a). This augmented MLH1 activation does not necessarily require targeting with multiple sgRNAs as a similar trend in MLH1 activation was obtained with one sgRNA targeting MLH1 promoter region (Fig. 6b). When modular Casilio-ME2.3, ME2.4, and ME3.4 platforms were compared, Casilio-ME2.4 showed the most MLH1 activation (Supplementary Fig. 6c). Only background MLH1 mRNA levels could be detected with TET1 (CD)-dead mutants of the Casilio-ME platforms (Casilio-dME) (Supplementary Fig. 6a, c), indicating that TET1 activity is required for gene activation and that delivery of GADD45A or NEIL2 without TET1 oxidative activity is not sufficient for activating methylation-silenced genes.

To ask whether the augmented MLH1 activation of Casilio-ME2.2 and Casilio-ME3.1 came from an increased efficiency in 5mC erasure, we performed BSeq and oxidative BSeq (oxBSeq) by NATURE COMMUNICATIONS| (2019) 10:4296 | https://doi.org/10.1038/s41467-019-12339-7 | www.nature.com/naturecommunications ME2.2 gave augmented MLH1 activation higher than Casilio-ME1, ME2.1, ME3.1, and ME3.2 platforms (Supplementary Fig. 6a). This augmented MLH1 activation does not necessarily require targeting with multiple sgRNAs as a similar trend in MLH1 activation was obtained with one sgRNA targeting MLH1 promoter region (Fig. 6b). When modular Casilio-ME2.3, ME2.4, and ME3.4 platforms were compared, Casilio-ME2.4 showed the most MLH1 activation (Supplementary Fig. 6c). Only background MLH1 mRNA levels could be detected with TET1 (CD)-dead mutants of the Casilio-ME platforms (Casilio-dME) (Supplementary Fig. 6a, c), indicating that TET1 activity is required for gene activation and that delivery of GADD45A or NEIL2 without TET1 oxidative activity is not sufficient for activating methylation-silenced genes.
target genes in Casilio-ME2.2 transfected cells (Supplementary Fig. 10c). Nonetheless, the activated expression of the four genes required TET1 activity and fully functional GADD45A, indicating that expression of GADD45A as part of PUFa-TET1 (CD) protein fusion is not sufficient for inducing the observed changes in expression of untargeted genes (Supplementary Fig. 10c). The CRISPR associated off-target activity could likely be reduced by using dCas9 variants with higher fidelity as a delivery vehicle\(^{40-43}\).

**Portability of the Casilio-ME platforms.** To show that Casilio-ME platforms enable efficient activation of other 5mC-silenced genes and in different cell types, we measured changes in
expression of MGMT, SOX17, RHOXF2 (HEK293T), CDH1 (U2OS), and GSTP1 (LNCap) in cells transfected with the components of Casilio-ME platforms. This showed that Casilio-ME platforms also enabled activation of these methylation-silenced genes. Casilio-ME2.2, Casilio-ME3.1, or ME3.2 produced significantly enhanced gene activations compared to ME1 targeting for the genes tested (Supplementary Fig. 11a-e). The improved gene activation is consistent with the increased demethylation efficiency obtained by Casilio-ME3.2 targeting GSTP1 CGI as compared to Casilio-ME1 (Supplementary Fig. 11g, h). Expression of PUFa protein fusions of Casilio-dME1, dME2.2, and dME3.1, containing catalytically inactive TET1(CD), in the absence of sgRNAs failed to activate MGMT, SOX17, and RHOXF2 (Supplementary Fig. 11f), indicating that expression of GADD45A or NEIL2 within close proximity of their substrates 5mC demethylation system is shown. For each box plot the thick line inside the box represents the median value and the surrounding bottom and top lines represent the 25th and 75th percentiles. The whiskers represent min and max values, the x represents the mean value. *p < 0.05, and **p < 0.001, two-way ANOVA.

**Box plot of frequencies of different CpG variants measured by BSeq and oxBSseq across the MLH1 promoter regions in cells transfected with the indicated 5mC demethylation system. Statistical significance of differences in methylation patterns were tested.**

**Casio-ME2 platforms and dCas9-TET1 system.** This study establishes a modular RNA-guided DNA methylation editing platform that not only recruits the TET1 effector to initiate DNA demethylation by 5mC oxidations, but also delivers protein factors to facilitate coupling 5mC oxidation to DNA repair pathways to effectively restore intact DNA to targeted sites. Such dual delivery enhanced 5mC demethylation at CGI target and augmented gene activations when compared to TET1(CD) delivered alone. In addition to the robustness of the platform, the modular design of Casilio-ME allows a high degree of tunability and flexibility in editing 5mC epigenetic marks.

Turnover of 5fC and 5caC by DNA repair machinery lags behind TET1-mediated 5mC oxidations as these intermediates accumulate before getting converted to unmethylated cytosine. Coupling TET1 activity with BER or NER pathways could accelerate 5fC and 5caC turnover, thereby enhancing activation of methylation-silenced genes. Consistent with this idea, Casilio-ME2 and Casilio-ME3 platforms designed to facilitate coupling TET1 and DNA repair activities gave an enhanced gene activation and CpG demethylation of targeted sites. This enhanced gene activation requires TET1 catalytic activity, fully functional GADD45A or NEIL2 proteins and co-targeting relevant effectors in close proximity to genomic target sites.

Previous studies revealed interesting functional and physical interactions among proteins involved in oxidizing 5mC and removal of oxidized cytosine intermediates via BER or NER. NEIL2 promotes substrate turnover by TET1 during DNA demethylation. GADD45A physically interacts with TET1 or TDG and seems to promote TET1 activity, and enhances removal of 5fC and 5caC by TDG. GADD45A also recruits repair enzymes such as the 3′-NER endonuclease XPG to genomic sites DNA. As GADD45A is devoid of any enzymatic activity, it was proposed to function as a liaison protein to physically couple 5mC oxidation with DNA repair. Consistent with these observations, Casilio-mediated co-targeting of TET1(CD) with GADD45A or NEIL2 within close proximity of their substrates enhanced 5mC demethylation and activation of methylation-silenced genes. However, it was reported that GADD45A and TDG failed to enhance demethylation of methylated plasmid by TET1(CD) in vitro, and the addition of TDG to Casilio-ME modules failed to augment gene activation. TDG protein fusions tested here might not be functional or other factor(s) might be required for TDG to produce enhanced gene activation. The enhanced activation of methylation-silenced genes observed could alternatively be achieved by TDG-independent pathways, by perhaps recruiting yet to be found players that could act by enhancing TET1/BER activities in the presence of GADD45A or NEIL2, or inhibiting DNA methyltransferases leading to replication-dependent demethylation at the targeted sites. The mechanisms by which these potential partners enhance 5mC erasure are in need of further studies. The Casilio-ME2 and Casilio-ME3 have the potential to be used in such mechanistic studies as different combinations of protein assemblies that include mutants or other proteins could be tested. Future characterization of protein domains enabling an enhanced TET1-mediated gene activation and of protein interactions taking place

**Inducible Casilio-ME platform.** To enable tunable and “on-demand” targeted DNA demethylation and gene activation, we constructed piggyBac (PB) transposon vectors hosting doxycycline-inducible PB Casilio-ME1 cassettes (DIP_Casilio- ME1) where the expression of dCas9 and PUFa-TET1(CD) is under the control of Tet-On promoters (Supplementary Fig. 12a). A DIP_Casilio-ME1 stable cell line is established by piggyBac transposition followed by antibiotic selection. When the DIP_Casilio-ME1 cell line was transiently transfected with targeting sgRNAs, we obtained robust MLH1 activation in the presence of doxycycline. MLH1 mRNA level also increased in response to increasing amounts of doxycycline. Without doxycycline added, only background levels of MLH1 were detected and no detectable amounts of Casilio-ME1 protein components were observed. Western blot analysis of protein extracts from transfected cells (Supplementary Fig. 12b, c). This DIP_Casilio-ME1 will enable establishment of isogenic cell lines that can be used to study different target CGIs in a tunable manner by supplying different target-specific sgRNAs and adjusting doxycycline dosage.

**Discussion.** This study establishes a modular RNA-guided DNA methylation editing platform that not only recruits the TET1 effector to initiate DNA demethylation by 5mC oxidations, but also delivers
at TET1-targeted genomic sites could lead to further improvement of the Castilio-ME platform and shed light on 5mC editing in mammalian cells.

Different levels of activation of methylation-silenced genes could be obtained by using one of the three flavors of Castilio-ME and by varying doxycycline concentrations with the DIP Castilio-ME platform described here. This equips Castilio-ME platforms with the capability to fine-tune gene activation. These Castilio-ME platforms significantly expand 5mC editing capability to efficiently address the causal-effect relationships of methylcytosine epigenic marks in numerous biological and pathological systems.

Methods

Cell culture and transfection. HEK293T and U2OS cells (both from ATCC) were cultivated in Dulbecco’s modified Eagle’s medium (DMEM) with 10% fetal bovine serum (FBS)(Lonza), 4% Glutamax (Gibco), 1% Sodium Pyruvate (Gibco), and penicillin-streptomycin (Gibco) in an incubator set to 37 °C and 5% CO2. LNCaP cells were obtained from ATCC and cultivated in RPMI-1640 supplemented with 10% FBS. Doxycycline (Dox) (Sigma) (1 μg/ml) or as otherwise indicated was added on the day of transfections with a daily change of media supplemented with Dox. Cells were seeded into 12-well plates at 150,000 cells per well the day before being transfected with plasmids each encoding dCas9 (100 ng), sgRNAs (100 ng) or PUF-fusion (200 ng) in the presence of Attractene or Lipofectamine 3000 transfection reagents according to manufacturers’ instructions (Qiagen). Western blot signals from the same plasmid ratio and total amount of DNA were used in cell transfections with components of dCas9/MS2 and SunTag systems. In the two-component systems, where TET1(CD) was fused to dCas9 or TALEs, 200 ng effector, 100 ng sgRNAs, and 100 ng empty vector of plasmid DNA were used. The combinations of MLH1 sgRNAs used were based on the obtained Castilio-ME-mediated MLH1 activation efficiency in preliminary experiments. However, no sgRNAs optimizations were performed for the other methylation-silenced genes. Nucleofections of LNCaP cells were performed by using 4D-nucleofector according to manufacturer’s instructions (Lonza) using 400 ng plasmid DNA. Cells were harvested 3 days after transfection or as otherwise indicated with media changes at 24 h post-transfection, and cell pellets were used for extraction of RNA, genomic DNA and protein using AllPrep DNA/RNA/ Protein Mini Kit according to the manufacturer’s instructions (Qiagen). Stable and Dox-inducible expression cell line was generated by transfecting three plasmids including the hyperactive transposase plasmid (byPBase) and the indicated PiggyBac vectors hosting cassette enabling a Dox-inducible expression of dCas9 or PUFa-TET1(CD) effector. The transfected cells were then subjected to double selection in the presence of blasticidin and hygromycin.

Plasmid constructions. A list of plasmids with links to their Addgene entries are provided in Supplementary Table 1. Detailed descriptions and sequences of oligonucleotides and proteins are given in the Supplementary Tables 2-5. The plasmids pCas9-dCas9-3αFat2A/IIID and pCasG-scVscG≈4GFPETT1(CD) (Addgene #82560 and 82561, respectively), pdCas9-Tet1-CD, and pcDNA3.1-MS2-Tet1-CD (Addgene #83340 and 83341, respectively) were gifts from Izuho Hatada and Ronggui Hu, respectively.

Quantitative RT-PCR analysis. Harvested cells were washed with Dulbecco’s phosphate-buffered saline (DPBS), centrifuged at 125 × g for 5 min and then the flash-frozen pellets were stored at −80 °C. Extracted RNA (500 ng–2 μg) were used as templates to make cDNA libraries using a High Capacity RNA-to-cDNA kit (Applied Biosystems), TaqMan gene expression assays were designed using Applied Biosystems, Primer fast DNA Bisulfite Kit, True Methyl oxBS Module and genomic DNA according to manufacturers’ instructions (Qiagen and NuGen, respectively). For oxBSeg, oxidation reactions were carried out in parallel and all treated samples developed same orange color expected for successful oxidative reactions. The average bisulfite conversion rates of cytosines in oxidized and non-oxidized sample sets were 0.996 and 0.995, respectively. Bisulfite treated DNA served as templates to PCR-amply three DNA fragments of 350–400 bp or a single fragment that cover MLHI or GSTP1 promoter region. Using the TaqTag PrepMix according to manufacturer instructions (Zymo Research). The MLHI PCR fragments were then cloned by SLIC into EcoRI-linearized pUC19 plasmid using T4 DNA polymerase. Ten independent positive clones for each sample were then subjected to Sanger sequencing to determine methylation profiles based on time-course mediated cytosine to thymine conversion frequencies. Individual Cytosine methylations obtained from bisulfite converted DNA templates and from unconverted DNA were subjected to high throughput sequencing (2 × 250 paired-end reads) conducted at Genewiz (South Plainfield, NJ, USA). Fifty to 120 thousand reads were obtained per sample. Sequences analysis of the placental samples were performed to determine methylation frequencies was performed by using BiQ Analyzer 3 with minimal bisulfite conversion rate and sequence identity set to 97 and 95%, respectively54. Reads from high throughput amplicon sequencing, on the other hand, were analyzed for 5mC and 5hmC by using BiQ Analyzer HiMod with minimal read quality score, alignment score, sequence identity and bisulfite conversion rate set to 30, 100, 0.9, and 0.9, respectively52. Obtained levels of 5mC and 5hmC and 5mC from two experiments were concordant across the CpGps. Bisulite Analyzer HiMod was used without sequence identity filter to analyze sequence integrity of MLHI amplicons derived from genomic DNA without bisulfite treatment.

Reduced representation bisulfite sequencing (RRBS). Library preparation for RRBS was performed according to manufacturer’s instructions (Diagenode). Briefly, 100 ng of genomic DNA for each sample was enzymatically digested, end-repaired and ligated with an adapter. Samples with different adaptors were then pooled together and subjected to bisulfite treatment followed by purification steps. The pooled DNA was PCR-amplified and then cleaned up with Ampure XP beads. Libraries were quantified with real time qPCR and sequenced using Illumina NextSeq (1 × 75 single end reads). Forty to 60 million reads per sample were obtained. PCR amplification levels, RRBS quality, and genome coverage were determined on the hg38 human reference genome using Bismark (version 0.16.0)53. CGp sites with coverage lower than 10 or higher than 400 reads were filtered out and Pearson’s correlation of methylation profile across indicated samples were computed by using methylKit (version 1.0.0)55.

RNA sequencing. RNA sequencing libraries were prepared for independent replicates of mRNA samples at Genewiz (South Plainfield, NJ, USA) by using NEBNext Ultra RNA Library Prep Kit for Illumina according to manufacturer’s instructions (New England Biolabs). Briefly, mRNA was enriched with Oligo(dT) beads, fragmented for 15 min at 94 °C and then reverse transcribed followed by second strand cDNA synthesis. cDNA fragments were end repaired and adenylated at 3′ ends, and universal adapters were ligated to cDNA fragments, followed by index addition and library enrichment by limited cycle PCR. The libraries were validated on the Agilent TapeStation (Agilent Technologies), and quantified by using Qubit 2.0 Fluorometer (Invitrogen) and quantitative PCR (KAPA Biosystems). The libraries were clustered on two lanes of a flowcell then loaded on the Illumina HiSeq instrument according to manufacturer’s instructions. The samples were sequenced using a 2 × 150 paired-end configuration and image analysis and base calling were conducted by the HiSeq Control Software. Generated raw sequence data (bcl files) from Illumina HiSeq was converted into fastq files and de-multiplexed using Illumina’s bcl2fastq 2.17 software. Thirty to 40 million read pairs were obtained per sample. Reads were then quantitated by Salmon into transcript estimates56, then subjected to DESeq2 for differential gene expression analysis57. Genes with low read counts in all samples were filtered out to eliminate noise in the analysis. Differentially expressed genes were called with adjusted p-values less than 0.05. MA-plots and FPKM scatter plots were generated using DESeq2.

Western blot analysis. Protein cell extracts (30 μg) were separated by electrophoresis on 10% SDS-polyacrylamide gels and then transferred to nitrocellulose membranes at 100 V for 1 h using Bjerum Schafer-Nielsen buffer with SDS. Blocked membrane in 5% Blotting-Grade Blocker (BioRad) in TBS-T (50 mM Tris pH 7.6, 200 mM NaCl, 0.1% Tween 20) were incubated overnight at 4 °C with the indicated antibodies, and then protein bands were detected using Horseradish peroxidase-conjugated secondary antibodies (Sigma) and Clarity Western ECL Substrate according to manufacturer’s instructions (BioRad). Monoclonal anti-Flag (dilution 1:1000, Sigma cat #F1804), monoclonal CRISPR/Cas9 (dilution 1:5000, Epitope cat #A-9000-100) and monoclonal anti-β-actin (dilution 1:5000, Sigma cat #MAB1501) antibodies were used according to manufacturers’ instructions. Blots were imaged using a Bio-Rad ChemiDoc. Uncorrected and unprocessed images of the blots are supplied in Source Data file.

Cell proliferation assay. Cells were split 48 h after transfection in serial dilutions as indicated, incubated for 48 h and then MTT assay was performed according to manufacturer’s instructions (Thermo Fisher Scientific), and absorbance at 540 nm was recorded using a SpectraMax M5 plate reader (Molecular devices).
Statistical analyses. Information on replication, statistical tests and presentation are given in the respective figure legends. GraphPad Prism 8 and Microsoft Excel were used to perform the indicated tests. Differences in all comparisons were considered significant if the obtained p values were less than 0.05.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The authors declare that the data that support the findings of this study are included in the published article and in the Supplemental Information, and are available from corresponding author upon reasonable request. Data containing RNAseq and RBBR raw sequencing read files were deposited onto sequencing read archive (SRA) with accession number PRJNA515359. The source data underlying Figs. 1b, 1c, 1e, 2b, 2d, 3b, 3d, and 4a-c and Supplementary Figs. 2b, 7b-c and 12c are provided as a Source Data file. New plasmids have been deposited to the Addgene repository.

Received: 20 May 2019; Accepted: 4 September 2019; Published online: 20 September 2019

References

1. Robertson, K. D. DNA methylation and human disease. Nat. Rev. Genet 6, 597–610 (2005).
2. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. Nat. Rev. Genet 14, 204–220 (2013).
3. Jones, P. A. & Liang, G. Rethinking how DNA methylation patterns are maintained. Nat. Rev. Genet 10, 805–811 (2009).
4. Lyko, F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. Nat. Rev. Genet 19, 81–92 (2018).
5. Liu, X. et al. CRISPR-based approach for targeted DNA demethylation. Cell Discov. 2, 1609 (2016).
6. Nishino, H. et al. TDG-mediated active DNA demethylation: mechanism, function and beyond. Nat. Rev. Genet 18, 517–534 (2017).
7. Kohl, R. M. & Zhang, Y. Tet enzymes, TDG, and the dynamics of DNA demethylation. Nature 502, 472–479 (2013).
8. Hashimoto, H., Hong, S., Bhagwat, A. S., Zhang, X. & Cheng, X. Excision of 5-hydroxymethylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. Nucleic Acids Res 40, 10203–10214 (2012).
9. Zhang, L. et al. Thymine DNA glycosylase specifically recognizes 5-carboxy-5-methylcytosine-modified DNA. Nat. Chem. Biol. 8, 328–330 (2012).
10. He, Y. F. et al. Tet-mediated formation of 5-carboxy-5-methylcytosine and its excision by TDG in mammalian DNA. Science 333, 1303–1307 (2011).
11. Muller, U., Bauer, C., Siegl, M., Rottach, A. & Leonhardt, H. Tet-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase-dependent gene demethylation. Nucleic Acids Res 42, 8592–8604 (2014).
12. Schomacher, L. et al. NEIL DNA glycosylases promote substrate turnover by TDg during DNA demethylation. Nat. Struct. Mol. Biol. 23, 116–124 (2016).
13. Niehrs, C. & Schafer, A. Active DNA demethylation by Gadd45 and DNA repair. Trends Cell Biol. 22, 220–227 (2012).
14. Barroso, G. et al. Gadd45α promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature 445, 671–675 (2007).
15. Schuermann, D., Weber, A. R. & Scharf, P. Active DNA demethylation by DNA repair: facts and uncertainties. DNA Repair (Amst) 44, 92–102 (2016).
16. Kienhoffner, S. et al. GADD45α physiologically and functionally interacts with TET1. Differentiation 90, 59–68 (2015).
17. Li, Z. et al. Gadd45α promotes DNA demethylation through TDG. Nucleic Acids Res 43, 3986–3997 (2015).
18. van Tol, N. & van der Zaal, B. J. Artificial transcription factor-mediated regulation of gene expression. Plant Sci. 225, 58–67 (2014).
19. Polecio, I., Verma, N., Mejia-Ramirez, E., Huangfu, D. & Raya, A. CRISPR/Cas9-Based Engineering of the Epigenome. Cell Stem Cell 21, 431–447 (2017).
20. Gaj, T., Gersbach, C. A. & Barbas, C. F. 3rd ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 31, 397–405 (2013).
21. Cheng, A. W. et al. Multiplexed activation of endogenous genes by CRISPR-coronavirus, an RNA-guided transcriptional activator system. Cell 123, 1163–1171 (2013).
22. Gilbert, L. A. et al. Genome-scale CRISPR-mediated control of gene repression and activation. Cell 159, 647–661 (2014).
23. Maeder, M. L. et al. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nat. Biotechnol. 31, 1137–1141 (2013).
24. Morita, S. Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. Nat. Biotechnol. 34, 1060–1065 (2016).
25. Liu, X. et al. Editing DNA Methylation in the Mammalian Genome. Cell 163, 233–247 e217 (2015).
26. Xu, X. et al. A CRISPR-based approach for targeted DNA demethylation. Cell Discov. 2, 1609 (2016).
27. Nishino, H. & Zhang, Y. Mammalian Tet Parathion as a hydrogen donor for the TDG-mediated active DNA demethylation mechanism. Chem. Sci. 7, 6690–6693 (2016).
28. Zhang, L. et al. Thymine DNA glycosylase specificity for 5-methylcytosine oxidation dynamics. Cell 167, 1641–1656 (2016).
29. Sanchez, R. et al. Solution structure of human growth arrest and DNA damage 5alpha (Gadd45alpha) and its interactions with proliferating cell nuclear antigen (PCNA) and Aurora A kinase. J. Biol. Chem. 285, 22196–22201 (2010).
30. Aas, D. et al. Identification of a zinc finger domain in the human NEIL2 (Nei-like-2) protein. J. Biol. Chem. 279, 47132–47138 (2004).
31. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas9 nucleases in human cells. Nat. Biotechnol. 31, 822–826 (2013).
32. Hsu, P. D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. 31, 827–832 (2013).
33. Harris, R. A. et al. Comparison of sequencing-based methods to profile DNA methylation and identification of monosaccharide epigenetic modifications. Nat. Biotechnol. 28, 1097–1105 (2010).
34. Albi, Z., Denard, C. A. & Zhao, H. Modular assembly of designer PUF proteins for specific post-transcriptional regulation of endogenous RNA. J. Biol. Eng. 8, 7 (2014).
35. Li, J. et al. Biology of CRISPR-Cas9–PAM hybrids for gene regulation and genomic labeling. Cell 163, 254–257 (2016).
36. Cheng, A. W. et al. Casilaso: a versatile CRISPR-Cas9-Pumilio hybrid for gene regulation and genomic labeling. Cell 163, 254–257 (2016).
37. Vales, C. et al. 5-carboxy-5-methylcytosine oxidation dynamics. Cell 153, 692–706 (2013).
38. Le May, N. et al. NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. Mol. Cell 38, 54–60 (2010).
39. Schmitz, K. M. et al. TAF12 recruits Gadd45α and the nucleotide excision repair complex to the promoter of RNA genes leading to active DNA demethylation. Mol. Cell 33, 344–353 (2009).
40. Weber, A. R. et al. Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. Nat. Commun. 7, 10806 (2016).
41. Yusa, K., Zhou, L., Li, M. A., Bradley, A. & Craig, N. L. A hyperactive piggyBac transposase for mammalian applications. Proc. Natl Acad. Sci. USA 108, 1531–1536 (2011).
42. Wang, T., Wei, J. I., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. Science 343, 80–84 (2014).
43. Jeong, J. Y. et al. One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics studies. Appl. Environ. Microbiol. 78, 5440–5443 (2012).
44. Bock, C. et al. BiQ Analyzer: visualization and quality control for DNA methylation data from bisulphite sequencing. Bioinformatics 21, 4067–4068 (2005).
45. Becker, D. et al. BiQ Analyzer HiMod: an interactive software tool for high-throughput locus-specific analysis of 5-methylcytosine and its oxidized derivatives. Nucleic Acids Res 42, W201–W207 (2014).
Acknowledgements
This work has been supported by the Jackson Laboratory internal grants (to A.C.), National Human Genome Research Institute 1R01HG009900 (to A.C) and National Cancer Institute P30CA034196 (to A.C), Leukemia Research Foundation New Investigator Grant (to S.L.), The Jackson Laboratory Director’s Innovation fund 19000-17-31 (to S.L.), The Jackson Laboratory Cancer Center New Investigator Award (to S.L.), and National Cancer Institute of the National Institutes of Health P30CA034196 (to S.L.), and the National Institutes of Health grant CA115783 (to C.D.H.).

Author contributions
A.T., C.H. and A.C. conceived and designed the study. A.T., N.J., M.D., A.R. performed the experiments. W.R. and S.L. analyzed RRBS data. A.T., C.H., and A.C. analyzed data and wrote the manuscripts.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-019-12339-7.

Correspondence and requests for materials should be addressed to A.W.C.

Peer review information Nature Communications thanks Rahul Kohli and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019