Izrada modularnog sustava CRISPR/dCas9 upotrebom ortolognih proteina Cas9 i njegova primjena u epigenetičkim manipulacijama genskih lokusa

Josipović, Goran

Doctoral thesis / Disertacija

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:756265

Sveučilište u Zagrebu

PRIRODOSLOVNO-MATEMATIČKI FAKULTET

BIOLOŠKI ODSJEK

Goran Josipović

IZRADA MODULARNOG SUSTAVA CRISPR/dCas9 UPOTREBOM ORTOLOGNIH PROTEINA Cas9 I NJEGOVA PRIMJENA U EPIGENETIČKIM MANIPULACIJAMA GENSKIH LOKUSA

DOKTORSKI RAD

Zagreb, 2021

University of Zagreb

FACULTY OF SCIENCE

DEPARTMENT OF BIOLOGY

Goran Josipović

DESIGN OF CRISPR/dCas9 MODULAR SYSTEM USING ORTHOGONAL Cas9 PROTEINS AND ITS APPLICATION IN EPIGENETIC MANIPULATION OF GENE LOCI

DOCTORAL THESIS

Zagreb, 2021

Ovaj je doktorski rad izrađen na Zavodu za molekularnu biologiju Biološkog odsjeka Prirodoslovno-matematičkog fakulteta, pod vodstvom prof. dr. sc. Vlatke Zoldoš, u sklopu Sveučilišnog poslijediplomskog doktorskog studija Biologije pri Biološkom odsjeku Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu.

ZAHVALE

Hvala mentorici prof. dr. sc. Vlatki Zoldoš na pruženoj prilici, povjerenju, potpori i svim savjetima danima prilikom izrade ovog doktorata.

Zahvaljujem kolegama iz grupe, kao i ostalim članovima ZMB-a koji su bili spremni priskočiti u pomoć kad god je trebalo.

Posebno hvala Samiri i Mariji na svim riješenim dilemama, problemima, na svakodnevnoj nesebičnoj pomoći i prijateljstvu.

Petra, hvala na savjetima, društvu i svemu zbog čega si mi uljepšala rad na fakultetu.

Veliko hvala mojoj obitelji bez kojih ne bih bio osoba kakva sam danas.

Ovaj doktorski rad posvećen je mojim roditeljima.

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek

Doktorski rad

IZRADA MODULARNOG SUSTAVA CRISPR/dCas9 UPOTREBOM ORTOLOGNIH PROTEINA Cas9 I NJEGOVA PRIMJENA U EPIGENETIČKIM MANIPULACIJAMA GENSKIH LOKUSA

GORAN JOSIPOVIĆ

Prirodoslovno-matematički fakultet, Sveučilište u Zagrebu

Otkrivanje uzročno posljedičnih veza između epigenetičkih oznaka u kromatinu i regulacije ekspresije gena omogućeno je primjenom molekularnih alata koji mogu ciljano mijenjati epigenom. U ovom istraživanju izradio sam modularni sustav temeljen na CRISPR/dCas9 metodologiji za ciljanu epigenetičku manipulaciju i direktnu regulaciju ekspresije gena, sastavljen od glavnih modula: promotora RNA polimeraze II, ortologa dCas9, efektorske domene, selekcijskog biljega, te modula za primanje više molekula sgRNA potrebnih za navođenje konstrukta na ciljno mjesto u genomu. Ciljanjem parova genskih lokusa *BACH2-HNF1A* i *IL6ST-MGAT3* s fuzijama antagonističkog djelovanja (metilacija *vs*. demetilacija DNA), uz korištenje dva ortologa dSpCas9 i dSaCas9 u stanicama HEK293, pokazao sam da je moguće promijeniti stupanj metilacije DNA kojeg prati promjena u ekspresiji tih gena. Također sam pokazao da epigenetičkom manipulacijom para gena *MGAT3-HNF1A*, u stanicama BG1, mogu promijeniti glikanski fenotip. Istovremenim korištenjem fuzijskih konstrukata TET1-dSaCas9 i VPR-dSpCas9 postigao sam sinergistički učinak na ekspresiju gena *HNF1A* koji je postojan kroz duži vremenski period. Dodatno sam unaprijedio modularni sustav u svrhu smanjenja nespecifičnog učinka.

(135 stranica, 33 slike, 9 tablica, 267 literaturnih navoda, jezik izvornika hrvatski)

Ključne riječi: sustav CRISPR/dCas9, ortologni proteini Cas9, ciljana manipulacija epigenoma, regulacija ekspresije gena, nespecifični učinak sustava CRISPR/dCas9

Mentor: prof. dr. sc. Vlatka Zoldoš

Ocjenjivači: izv. prof. dr. sc. Aleksandar Vojta
prof. dr. sc. Mirjana Pavlica
prof. dr. sc. Srećko Gajović

University of Zagreb
Faculty of Science
Department of Biology

Doctoral thesis

DESIGN OF CRISPR/dCas9 MODULAR SYSTEM USING ORTHOGONAL Cas9 PROTEINS AND ITS APPLICATION IN EPIGENETIC MANIPULATION OF GENE LOCI

GORAN JOSIPOVIĆ
Faculty of Science, University of Zagreb

Establishing a causal relationship between epigenetic marks in chromatin and the regulation of gene expression is made possible by the use of molecular tools that can specifically alter the epigenome. In this study, I developed a modular system based on the CRISPR/dCas9 methodology for targeted epigenetic manipulation and direct regulation of gene expression, consisting of major modules: RNA polymerase II promoter, dCas9 ortholog, effector domain, selection marker and module that can receive multiple sgRNA molecules to guide the construct in the genome. By targeting *BACH2-HNF1A* and *IL6ST-MGAT3* gene pairs with fusions showing antagonistic functionalities (DNA methylation *vs*. demethylation), using two orthologs dSpCas9 and dSaCas9 in HEK293 cells, I showed that it is possible to change the DNA methylation accompanied by a change in gene expression. I also showed that by epigenetic manipulation of *MGAT3-HNF1A* gene pair, in BG1 cells, the glycan phenotype can be altered. By simultaneously using the fusion constructs TET1-dSaCas9 and VPR-dSpCas9, I achieved a synergistic effect on *HNF1A* gene expression that is stable over a longer period of time. I further improved the modular system in order to reduce the off-target effect.

(135 pages, 33 figures, 9 tables, 267 references, original in croatian)

Keywords: CRISPR/dCas9 system, Cas9 orthologous proteins, targeted epigenome editing, direct regulation of gene expression, CRISPR/dCas9 off-target effect

Supervisor: Vlatka Zoldoš, PhD, Professor

Reviewers: Aleksandar Vojta, PhD, Associate Professor
Mirjana Pavlica, PhD, Professor
Srećko Gajović, PhD, Professor

SADRŽAJ

U DOKTORSKOJ DISERTACIJI OBJEDINJENI SU SLIJEDEĆI ZNANSTVENI RADOVI:

1. **Josipović G**, Zoldoš V, Vojta A. Active fusions of Cas9 orthologs. *J Biotechnol*. 2019;301:18-23.
2. **Josipović G**, Tadić V, Klasić M, Zanki V, Bečeheli I, Chung F, Ghantous A, Keser T, Madunić J, Bošković M, Lauc G, Herceg Z, Vojta A, Zoldoš V. Antagonistic and synergistic epigenetic modulation using orthologous CRISPR/dCas9-based modular system. *Nucleic Acids Res*. 2019;47(18):9637-9657.
3. Tadić V, **Josipović G**, Zoldoš V, Vojta A. CRISPR/Cas9-based epigenome editing: An overview of dCas9-based tools with special emphasis on off-target activity. *Methods*. 2019;164-165:109-119.

1 UVOD

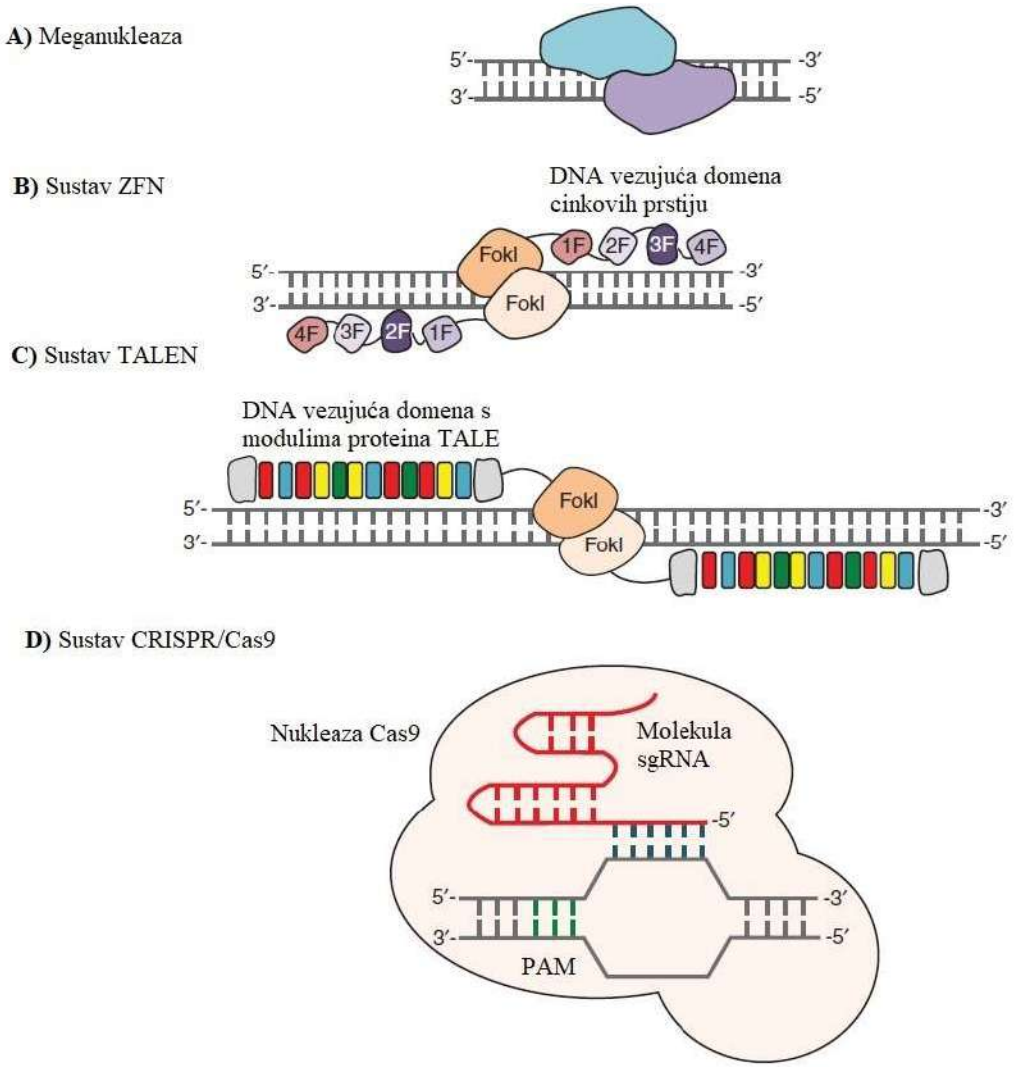
1.1 Ciljana manipulacija genoma

Otkrićem restrikcijskih enzima 1970. godine, kao jedne od linija obrana bakterija protiv strane virusne DNA, počelo se razvijati razdoblje genetičkog inženjerstva. Induciranjem dvolančanog loma u molekulu DNA otvorena je mogućnost za uvođenje stranih fragmenata molekule DNA u originalni genom. Takva manipulacija genoma omogućava razumijevanje funkcije određenog gena, popravak istoga, te promjenu ili uvođenje novih funkcija u genomu. Prvi pokušaji ciljanog manipuliranja genoma na temelju homologne rekombinacije[1,2] pokazali su se izrazito neefikasnim te potvrdili važnost induciranog dvolančanog loma u molekulu DNA[3]. Primjenom meganukleaza, endonuklaza koje prepoznaju dugačke odsječke molekule DNA, bilo je moguće uvesti dvolančani lom (Slika 1.A). Glavni nedostatak ove metode je kompleksnost primjene meganukleaza u pogledu odabira prave za ciljani gen, kao i modifikacija proteinskih regija zaslužnih za stvaranje kontakata s molekulom DNA kako bi se povećala mogućnost ciljanja drugih sljedova DNA[4–6].

Otkriće cinkovih prstiju, koji specifično vežu molekulu DNA na temelju prepoznavanja određenog slijeda nukleotida (nt), omogućen je razvoj ciljanih i programiranih nukleaza. Sklapanjem više cinkovih prstena (najčešće tri do četiri) pomoću kratkih peptidnih lanaca, od kojih svaki prepoznaje 3 para baza (pb), moguće je povećati specifičnost na temelju prepoznavanja dužeg slijeda u genomu[7]. Programirane nukleaze s domenom cinkovih prstiju (engl. *zinc-finger nucleases*, ZFNs) možemo dobiti vezanjem DNA cijepajuće domene endonukleaze FokI iz vrste *Flavobacterium okeanokoites* na sklopljene module cinkovih prstiju čime je omogućeno uvođenje loma u molekuli DNA na točno određenom mjestu u genomu (Slika 1.B)[8]. Endonukleazu FokI moguće je ciljati na različita mjesta u genomu promjenom cinkovih prstiju. Kako je za aktivnost endonukleaze FokI bitna homodimerizacija potrebno je koristiti dva odvojena modula cinkovih prstiju koji se vežu na susjednim pozicijama na ciljnom mjestu u genomu[9,10]. Primjena programiranih nukleaza s domenom cinkovih prstiju ostala je limitirana zbog problema kompleksnosti sklapanja cinkovih prstiju, te zbog nedostatka fleksibilnosti prepoznavanja bilo koje sekvence u genomu.

Otkriće proteina TALE (engl. *transcription activator-like effector (TALE) proteins*) iz biljne patogene bakterije roda *Xanthomonas* dodatno je pojednostavilo fleksibilnost ciljanja bilo kojeg slijeda u genomu. Dvije aminokiseline na pozicijama 12 i 13 nazvane RVD (engl.

repeat variable diresidues) u N-terminalnoj DNA vezujućoj domeni proteina TALE zaslužne su za prepoznavanje jedne baze u DNA[11,12]. Programirana nukleaza s modularnom domenom vezanja na DNA, odnosno sustav TALEN (engl. *TALE nuclease*), također sadrži endonukleazu FokI koja se specifično navodi na određeno mjesto u genomu sklapanjem više monomera proteina TALE (Slika 1.C)[13,14]. Za razliku od cinkovih prstena proteini TALE pokazuju veću specifičnost[15], međutim za efikasno vezanje sustava potreban je timin (T) ispred 5' kraja vezanja monomera proteina TALE[16]. Direktnom evolucijom monomera proteina TALE danas su dostupne mutante koje ne zahtijevaju timin neposredno ispred 5' kraja za efikasno vezanje[17]. Glavni nedostatci metode TALEN su osjetljivost vezanja na metilirani citozin (C), što zahtjeva poznavanje metilacijskog statusa ciljane genomske regije, te novi dizajn i sklapanje DNA slijeda za proteinsku regiju koja će ciljati željeno mjesto u genomu[18]. Otkrićem sustava CRISPR (engl. *Clustered Regulatory Interspaced Short Palindromic Repeats*) pojednostavljena je cijela tehnologija manipulacije genoma navođenjem nukleaze na željeno mjesto u genomu pomoću malene molekule RNA (Slika 1.D) čime je ova metoda postala superiorna u odnosu na prethodne.



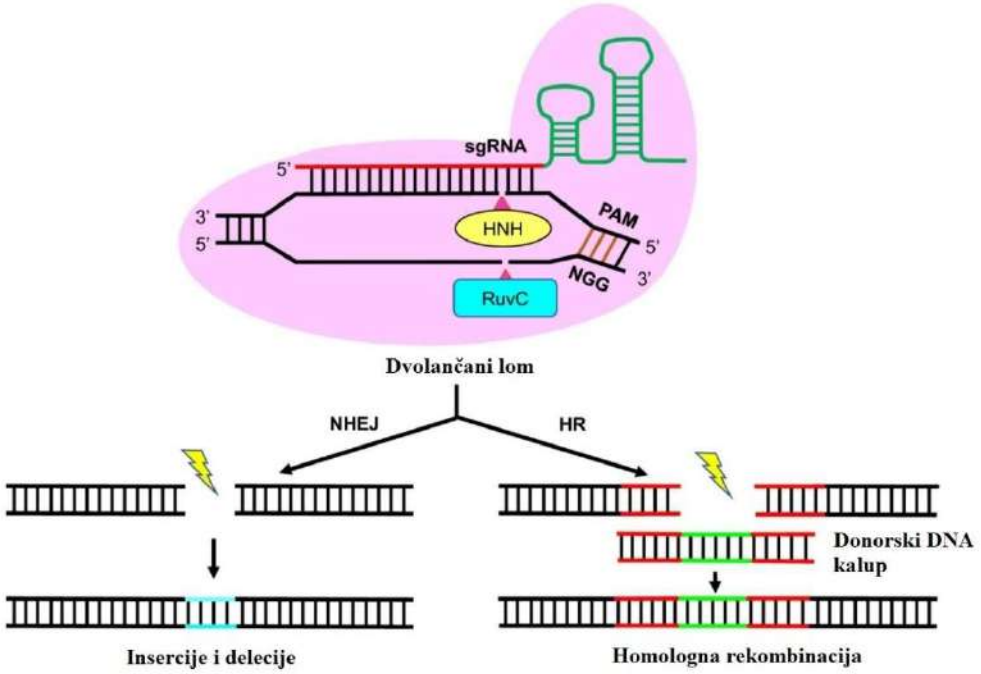
Slika 1. Shematski prikaz metoda za ciljanu manipulaciju genoma. (A) Meganukleaze ciljano prepoznaju i režu duže sljedove molekule DNA. **(B)** Sustav ZFN sadrži DNA vezujuću domenu sačinjenu od više modula cinkovih prstiju koji specifično prepoznaju sekvencu dužine od 3 pb te dovode nukleazu FokI na specifično mjesto u genomu. **(C)** Sustav TALEN koristi proteine TALE iz patogene bakterije *Xanthomonas* od kojih svaki monomer proteina TALE prepoznaje jednu bazu u molekuli DNA, te se na taj način vezanjem više monomera može specifično navoditi nukleazu FokI. **(D)** Sustav CRISPR/Cas9 temelji se na navođenju nukleaze Cas9 pomoću malene molekule sgRNA. Prvih 20 nt na 5' kraju molekule sgRNA zaslužno je za navođenje i specifično vezanje molekule DNA. Za vezanje nukleaze Cas9 na određeno mjesto u genomu potrebna je i sekvenca PAM, 5'-NGG-3' (gdje N predstavlja bilo koju od četiri dušične baze A, C, T ili G), koja se mora nalaziti neposredno nizvodno od mjesta vezanja molekule sgRNA. Preuzeto i prilagođeno iz Gaj i sur.[19].

1.2 Sustav CRISPR/Cas9

Proučavanjem gena *iap* u vrsti *Escherichia coli* otkrivene su ponavljajuće regije dužine 29 pb. Posebnost ovih regija bila je u njihovoj organizaciji – naime, isprekidane su jedinstvenim neponavljajućim sljedovima dužine 32 pb[20]. Ubrzo je termin CRISPR dodijeljen ovakvoj organizaciji ponavljajućih regija, te su razvojem metoda sekvenciranja otkrivene CRISPR regije u oko 40% bakterija i 90% arheja[21]. Saznanje da neponavljajući sljedovi potječu iz virusne DNA i drugih mobilnih elemenata[22–24], te otkriće skupine gena Cas[25] (engl. *CRISPR-associated genes*) u neposrednoj blizini lokusa CRISPR, ubrzo je razjasnilo ulogu ovoga sustava u zaštiti bakterija i arheja od stranih nukleinskih kiselina[23,24]. Sustavi CRISPR/Cas djeluju različitim mehanizmima ovisno o efektorskim proteinima, a podijeljeni su u: 1) sustav klase I (tip I, III i IV) gdje više proteina čini jedan kompleks koji se navodi na stranu nukleinsku kiselinu i 2) sustav klase II (tip II, V i VI) koji sadrži jedan efektorski protein Cas (Cas9, Cas12 ili Cas13)[26,27].

Sustav CRISPR/Cas9 iz vrste *Streptococcus pyogenes* najbolje je proučen i najčešće korišten. Osnovne komponente tog sustava su: endonukleaza Cas9 (SpCas9), molekula CRISPR-RNA (crRNA) nužna za specifično prepoznavanje komplementarne molekule DNA i molekula tracrRNA (engl. *trans-activating CRISPR RNA*) važna za sparivanje s molekulom crRNA te stvaranje aktivnog kompleksa s nukleazom Cas9[28–30]. Za vezanje i aktivaciju nukleaze Cas9 potrebna je i sekvenca PAM (engl. *protospacer adjacent motif*) koja se mora nalaziti neposredno nizvodno od mjesta vezanja molekule crRNA, a ona predstavlja slijed 5′-NGG-3[28,31]. Značajno pojednostavljenje sustava CRISPR/Cas9 za primjenu u stanicama eukariota dobiveno je spajanjem molekula crRNA i tracrRNA u kimernu molekulu sgRNA (engl. *single guide RNA*) čime se smanjio broj potrebnih komponenti[28]. Prvih 20 nt koji se nalaze na 5' kraju molekule sgRNA zaslužno je za specifično vezanje na komplementarni slijed u molekuli DNA. Promjenom te kratke sekvence sustav se može navoditi na bilo koje mjesto u genomu nizvodno kojeg se nalazi sekvenca PAM što sam sustav čini jednostavnim za primjenu[28,32,33]. Mehanizam prepoznavanja ciljanog mjesta u genomu uključuje traženje i prepoznavanje sekvence PAM od strane nukleaze Cas9, nakon čega slijedi lokalno razdvajanje lanaca molekule DNA u blizini sekvence PAM, te pokušaj hibridizacije prvih 20 nt molekule sgRNA s ciljnim lancem molekule DNA[31,34]. Za aktivaciju nukleaze Cas9 bitno je pravilno komplementarno sparivanje proksimalne regije PAM, koja se nalazi 10 do 12 nt neposredno uzvodno od sekvence PAM, tzv. regija „seed“ molekule sgRNA[28,30,31,35]. Nukleaza Cas9

tolerira pogreške u distalnoj regiji PAM što uzrokuje nespecifičnu aktivnost sustava CRISPR/Cas9[36,37].



Slika 2. Shematski prikaz uvođenja i razrješenja dvolančanog loma u molekuli DNA pomoću sustava CRISPR/Cas9. Nukleaza Cas9 (označena ružičastom bojom) navodi se na određeno mjesto u genomu pomoću molekule sgRNA čijih je prvih 20 nt na 5' kraju zaslužno za komplementarno sparivanje s molekulom DNA. Sekvenca PAM, koja se mora nalaziti neposredno nizvodno od mjesta vezanja molekule sgRNA, nužna je za vezanje i aktivaciju nukleaze Cas9. Nukleazne domene HNH i RuvC uvode dvolančani lom u molekulu DNA. Inducirani dvolančani lom u molekuli DNA može se popraviti na dva načina: popravak DNA nehomolognim sparivanjem krajeva (NHEJ) ili homolognom rekombinacijom (HR). Preuzeto i prilagođeno iz Cai i sur.[38].

Dovođenjem sustava CRISPR/Cas9 na željeno mjesto u genomu nukleaza Cas9 uvodi dvolančani lom u molekulu DNA na točno određenom mjestu, 3 pb uzvodno od sekvence PAM, pomoću nukleaznih domena HNH i RuvC. Nukleazna domena HNH uvodi lom u lanac molekule DNA komplementaran sekvenci molekule sgRNA, dok nukleazna domena RuvC reže suprotan lanac[28,39,40]. Inducirani dvolančani lom u molekuli DNA pomoću sustava CRISPR/Cas9 može se popraviti na dva načina (Slika 2.). Najčešći mehanizam popravka prisutan u stanicama tijekom cijelog staničnog ciklusa je popravak nehomolognog sparivanja

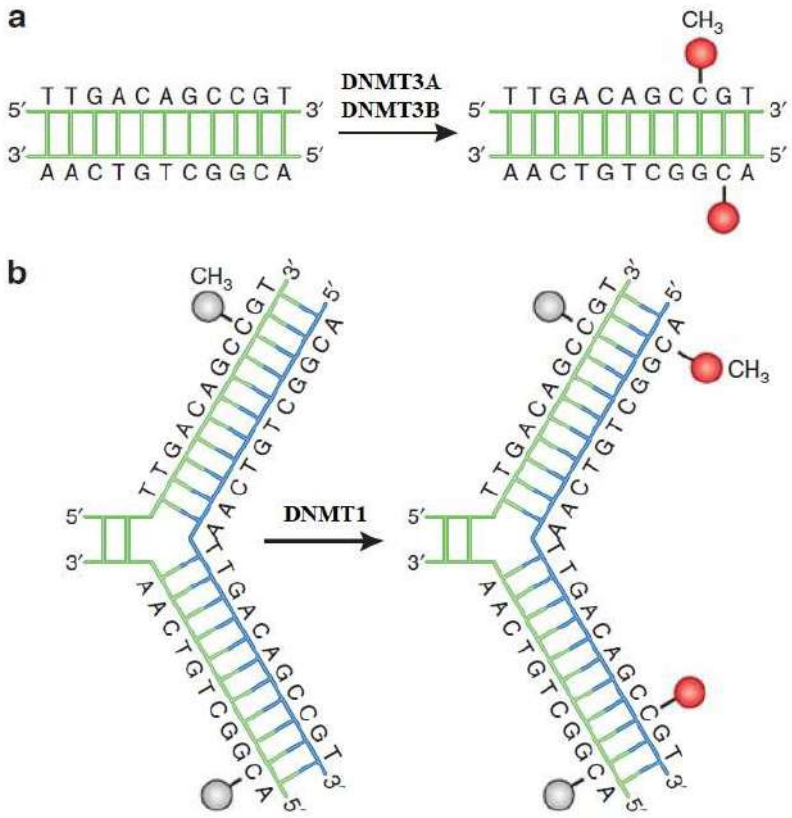
krajeva (engl. *non-homologous end joining*, NHEJ) koji uključuje direktno spajanje krajeva zahvaćenih dvolančanim lomom uz manje insercije i delecije. Rezultat takvog popravka su pogreške koje mogu dovesti do narušavanja funkcije gena[41]. Alternativno, popravak vođen homolognom rekombinacijom zahtjeva homologni DNA kalup te omogućava precizne promjene u genomu[42].

1.3 Metilacija i demetilacija molekule DNA

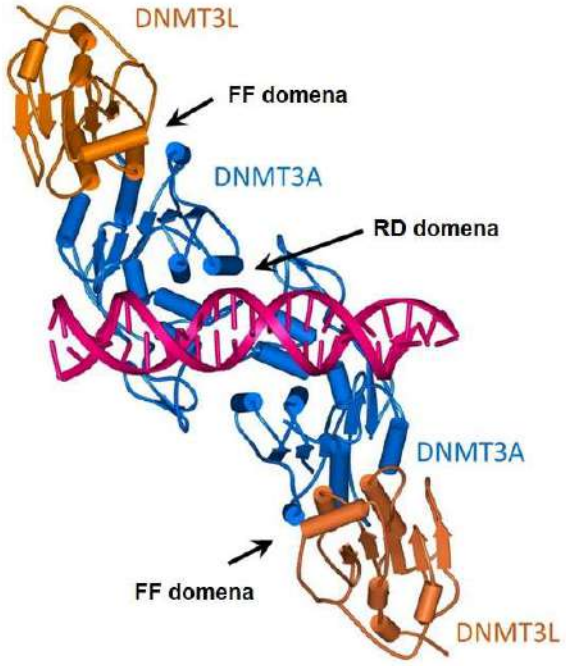
Metilacija DNA jedan je od važnih epigenetičkih mehanizama uključenih u regulaciju ekspresije gena te posljedično u razvoj i diferencijaciju[43,44]. U somatskim stanicama sisavaca preko 98% metilacije DNA uglavnom zahvaća citozine unutar CpG dinukleotida i simetrično je raspoređena na oba lanca molekule DNA[45]. Zastupljenost CpG mjesta u genomu iznosi oko 1% te je za oko 5 do 10 puta manja u odnosu na ostale dinukleotide. Razlog tome je spontana deaminacija 5-metil-citozina (5mC) u T što rezultira tranzicijom C:G u T:A koja je sama po sebi mutagena zbog neefikasnog popravka[46]. CpG dinukleotidi su stoga najčešće grupirani unutar CpG otoka, duljine oko 200 pb ili više, a koje najčešće nalazimo unutar promotora protein kodirajućih gena blizu mjesta inicijacije transkripcije (engl. *transcription start site*, TSS)[47,48]. Razvojem novih visokoprotočnih metoda analize metilacije DNA moguće je odrediti metilacijski status genoma[49] čime je pokazano da je većina CpG otoka nemetilirana u somatskim stanicama za razliku od ostatka genoma[48]. U biljnim genomima, zbog velike količine DNA i velikog udjela puno klasa mobilnih elemenata, metiliraju se citozini i unutar konteksta CpHpG i CpHpH (H predstavlja dušične baze A, C ili T)[50]. Također, u genomima sisavaca otkrivena je metilacija citozina izvan konteksta CpG dinukleotida, najčešće unutar dinukleotida CpA čija je funkcionalna važnost opisana za pluripotentne matične stanice i stanice mozga[51–53].

Grupa enzima DNA metiltransferaza kataliziraju kemijsku modifikaciju citozina dodatkom metilne skupine (CH3) na peti atom citozina molekule DNA s donora S-adenozil-L-metionina (SAM). DNA metiltransferaze *de novo* 3A i 3B (DNMT3A i DNMT3B) kataliziraju *de novo* metilaciju DNA (Slika 3.A) prilikom embrionalnog razvoja, dok je za održavanje obrasca već uspostavljenje metilacije (Slika 3.B) prilikom replikacije molekule DNA zaslužna DNA metiltransferaza 1 (DNMT1)[54]. Istraživanja provedena na matičnim stanicama miša pokazuju da nema jasne razlike između funkcija različitih DNA metiltransferaza, te da DNMT3A i DNMT3B u određenoj mjeri sudjeluju i u održavanju metilacije DNA[55,56]. U

obitelji DNA metiltransferaza nalazi se i DNMT2, uključena u metilaciju citozina molekule tRNA[57] i metilaciju hibridnih molekula DNA-RNA[58], te DNMT3L koja ne posjeduje katalitičku aktivnost, međutim sudjeluje u uspostavljanju metilacije kroz direktnu interakciju s DNMT3A i DNMT3B[59,60]. Metiltransferaza DNMT3L se veže za N-terminalni kraj histona H3 na poziciji nemetiliranog Lys4, što je preduvjet za regrutaciju metiltransferaze DNMT3A ili DNMT3B. Ovo je jedan od primjera međusobne ovisnosti različitih epigenetičkih mehanizama, u ovom slučaju metilacije DNA i modifikacija histona[61]. DNMT3L stvara tetramerni kompleks s DNMT3A/DNMT3B čime stabilizira konformaciju aktivnog mjesta katalitičke domene DNMT3A/DNMT3B (Slika 4.)[62,63]. Za stvaranje homodimera DNMT3A-DNMT3A važna je domena RD koja sadrži polarne aminokiseline arginin i aspartat, dok je za stvaranje heterodimera DNMT3A-DNMT3L ključna domena FF s dva hidrofobna fenilalanina (Slika 4.)[63,64].

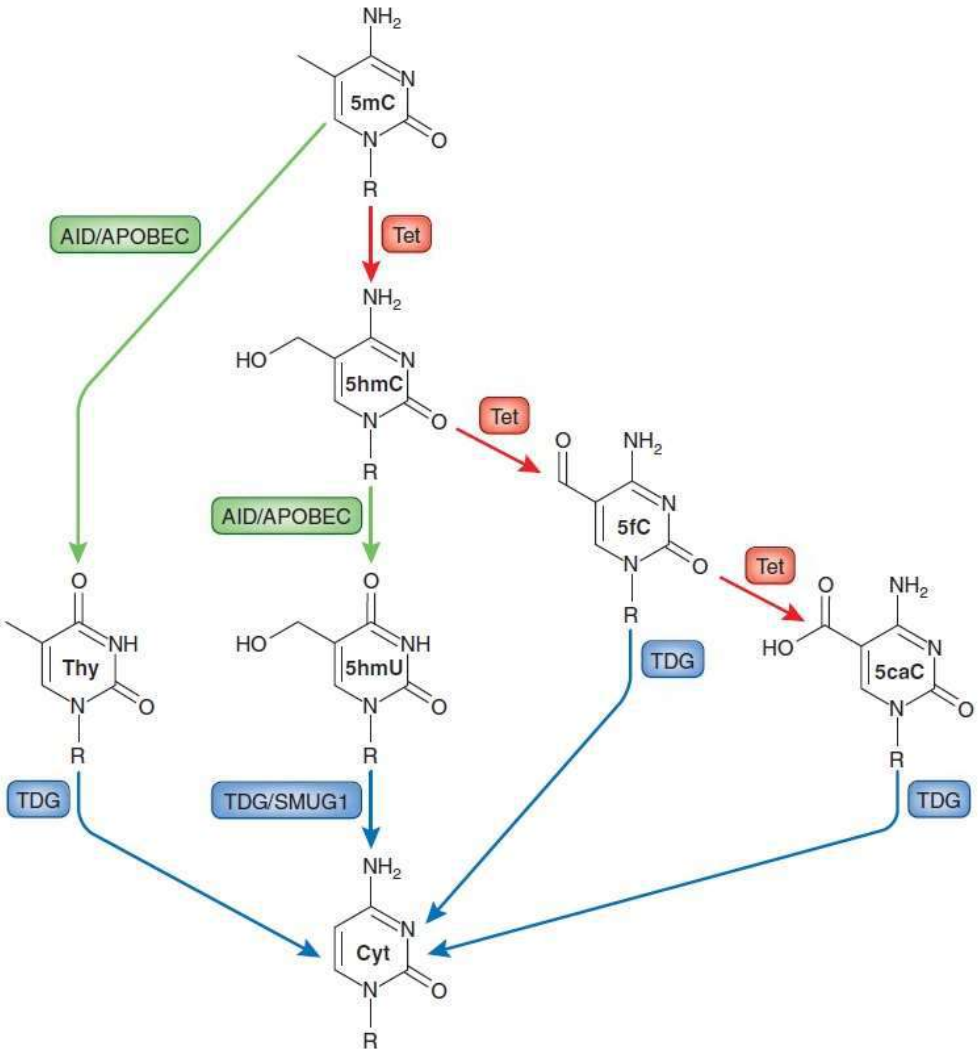


Slika 3. Mehanizam uspostavljanja i održavanja metilacije molekule DNA. (A) Metilacija DNA *de novo* uspostavljena je pomoću DNA metiltransferaza DNMT3A i DNMT3B. **(B)** DNA metiltransferaza 1 (DNMT1) održava metilaciju molekule DNA prilikom replikacije, te kao kalup koristi stari metilirani lanac. Preuzeto iz Moore i sur.[65].



Slika 4. Tetramerni kompleks metiltransferaza DNMT3L-DNMT3A u interakciji s molekulom DNA. Za aktivnost DNA metiltranferaze DNMT3A/DNMT3B važno je stvaranje tetramernog kompleksa čime se stabilizira aktivno mjesto metiltransferaze DNMT3A i omogućava njezina aktivnost. Domena FF, koja sadrži dva hidrofobna fenilalanina, bitna je u stvaranju heterodimera DNMT3A-DNMT3L. Za stvaranje homodimera DNMT3A ključna je domena RD koja sadrži polarne aminokiseline arginin i aspartat. Preuzeto i prilagođeno iz Jeltsch, Jurkowska[66].

Demetilacija molekule DNA je kompleksan mehanizam te razlikujemo pasivni i aktivni oblik. Pasivni oblik demetilacije molekule DNA ograničen je na stanice koje se dijele te je temeljen na smanjenoj razini metilacije kroz inhibiciju DNMT1[67]. Aktivni proces demetilacije, prisutan u stanicama koje se dijele i onima koje se ne dijele, temelji se na deaminaciji i/ili oksidaciji 5-metil-citozina (5mC) do različitih produkata koji aktiviraju popravak molekule DNA izrezivanjem baza (engl. *base excision repair*, BER)[67,68]. Mehanizam oksidacije posredovan je enzimima iz porodice Tet (engl. *ten-eleven translocation*) koji oksidacijom 5mC konvertiraju u 5-hidroksi-metil citozin (5hmC)[69,70]. Daljnjom oksidacijom stvara se 5-formil-citozin (5fC) do krajnjeg produkta 5-karboksi-citozina (5caC)[71]. Posljedično, aktivacijom popravka DNA izrezivanjem baza posredovano enzimom timin-DNA glikozilaza (engl. *thymine DNA glycosylase*, TDG) uklanja se među produkt oksidacije 5mC i zamjenjuje citozinom (Slika 5.)[72].



Slika 5. Mehanizam aktivne demetilacije molekule DNA. Enzimi iz porodice Tet oksidacijom 5-metil-citozina (5mC) stvaraju 5-hidroksi-metil-citozin (5hmC). Daljnjom oksidacijom stvara se 5-formil-citozin (5fC) te krajnji produkt 5-karboksi-citozin (5caC). Alternativno, 5hmC može biti konvertiran u 5-hidroksi-metil-uracil (5hmU) pomoću enzima AID/APOBEC. Drugi mehanizam aktivne demetilacije uključuje deaminaciju 5mC u timin pomoću enzima AID/APOBEC. U konačnici produkti oksidacije i deaminacije 5mC aktiviraju popravak DNA izrezivanjem baza posredovano enzimom timin-DNA glikozilaza (TDG) te zamjenu s citozinom. Preuzeto iz Moore i sur.[65].

Metilacija DNA uključena je u različite procese genoma - od utišavanja retrovirusnih i transpozonskih elemenata, čime direktno utječe na stabilnost genoma, održavanja obrasca monoalelne ekspresije utisnutih gena, održavanja inaktiviranog stanja jednog kromosoma X u

sisavaca, pa sve do same regulacije ekspresije gena[73–75]. Hipermetilirane regije najčešće su transkripcijski inaktivne zbog toga što metilacija citozina sudjeluje u stvaranju kompaktnije strukture kromatina, te pokazuju smanjenu razinu acetilacije histona, što sve direktno utječe na nesposobnost vezanja transkripcijskih faktora[76]. Oko 70% promotora gena čovjeka sadrži CpG otoke koji tijekom embrionalnog razvoja budu diferencijalno metilirani[48]. Pokazano je kako metilacija unutar CpG otoka smanjuje ekspresiju gena[77]. Također, metilacija regija koje su udaljene oko 2 kilo baznih parova (kb) od CpG otoka, a nazivaju se CpG obale (engl. *CpG shores*), korelira sa smanjenjem ekspresije gena[78]. Metilacija DNA utječe na smanjenu ekspresiju gena na različite načine. Jedan od mehanizama je direktno blokiranje vezanja transkripcijskih faktora na promotorsku sekvencu bogatu metiliranim citozinima[79,80]. Drugi mehanizam uključuje vezanje specifičnih proteina na metilirane CpG dinukleotide, koji zatim regrutiraju ostale korepresore i epigenetičke pisače i brisače, koji stavljaju i uklanjaju modifikacije na histonskim repovima. Uklanjanjem acetilne oznake ili pak dodatkom metilne skupine specifično na Lys9 u histonu H3 u konačnici rezultira takozvanom „zatvorenom“ ili represivnom strukturom kromatina[81–83]. Za razliku od CpG otoka, metilacija DNA u tijelu gena povezana je s povećanom ekspresijom nekih gena[84–86]. Studija na mišjim embrionalnim stanicama pokazuje kako stanice s dvostrukom mutacijom gena *Dnmt3b*$^{-/-}$ pokazuju veću razinu transkripata s alternativnih promotora unutar tijela gena[87]. Blokiranje alternativnih promotora metilacijom citozina jedan je od mehanizama kako metilacija tijela gena pojačava ekspresiju s glavnog promotora[88].

1.4 Regulacija ekspresije gena

Precizna regulacija ekspresije eukariotskih gena ključna je u razvoju viših eukariota te definiranju aktivnosti pojedinih gena u različitim stanicama i tkivima. Ključni regulatorni proteini, transkripcijski faktori, djeluju *in trans* te se vežu na *cis*-regulatorne elemente i regije pojačivača kako bi potaknuli ili inhibirali ekspresiju određenog genskog lokusa[89]. Kroz evoluciju razvijeni su brojni transkripcijski faktori koji se razlikuju u strukturi i domeni koja se veže za molekulu DNA. Točni mehanizmi regulacije vezanja i selektivnost transkripcijskih faktora za određene genske lokuse u različitim stanicama nisu u potpunosti razjašnjeni. Evolucijski srodni transkripcijski faktori nerijetko pokazuju sličan način vezanja za molekulu DNA iako znaju pokazivati različite funkcije[90].

Dodatnoj kompleksnosti procesa regulacije ekspresije gena pridonosi organizacija molekule DNA u kromatinsku strukturu vezanjem 147 pb oko proteinskog kompleksa sačinjenog od po dvije kopije četiri različita histona - H2A, H2B, H3 i H4[91,92]. Struktura kromatina nije homogena te je u jezgri organizirana u kompleksnije lokalne strukture i organizacije višeg reda[92,93]. Posljedično, sama struktura kromatina utječe na vezanje transkirpcijskih faktora i na aktivnost genskih lokusa. Eukromatin je dio kromatina koji je više „otvoren" i stoga je dostupan transkripcijskim faktorima i ostalim regulatornim proteinima, a njegovo manje kondenzirano stanje omogućuje post-translacijsko dodavanje kovalentnih skupina na histone, poput fosfatne i acetilne skupine, čime se lokalno olabavi veza histona s molekulom DNA. Za razliku od eukromatina, heterokromatin pokazuje inaktivne oznake kao što su metilne skupine na određenim pozicijama lizina i arginina u histonima H3 i H4 (primjerice, H3K9me3 i H3K27me3), ili pak sumoilacija koja podrazumijeva kovalentno vezanje SUMO proteina (engl. *small ubiquitin-like modifier*). Ove oznake učvršćuju vezu između histona i molekule DNA što dovodi do kompaktnije strukture kromatina, ili se pak fizički onemogućava vezanje transkripcijskih faktora i drugih regulatornih proteina u slučaju sumoilacije[94–98].

Metilacija molekule DNA također je jedan od ključnih regulatornih mehanizama u kontroli ekspresije gena (Slika 6.A). Pokazano je kako metilacija DNA u promotorskim regijama korelira sa smanjenom ekspresijom gena[77]. Postoje razni mehanizmi kojima se objašnjava utjecaj metilacije DNA na strukturu kromatina i regulaciju ekspresije gena. Jedan od mehanizama objašnjava da dodatak metilne skupine na citozine u molekuli DNA blokira vezanje transkripcijskih faktora čime se onemogućava ekspresija gena[79,80]. Drugi mehanizmi povezuju metilaciju molekule DNA s modifikacijama histonskih proteina. Glavna poveznica u tom slučaju je grupa proteina kao što su MeCP2 (engl. *Methyl-CpG-binding protein 2*) i MBD1 (engl. *Methyl-CpG-binding domain protein 1*) koji prepoznaju i vežu se za metiliranu DNA. Njihovim vezanjem dolazi do regrutacije korepresora histon deacetilaza i histon metiltransferaza, koje uklanjaju aktivne i uvode inaktivne histonske oznake, te posljedično dolazi do stvaranja kompaktnije strukture kromatina čime se opet onemogućava vezanje transkripcijskih faktora i drugih regulatornih proteina potrebnih za inicijaciju transkripcije gena[81,99,100]. Također, inaktivna histonska oznaka H3K9me3 povezena je s metilacijom DNA *de novo*. Naime, za regrutaciju metiltransferaza DNMT3A i DNMT3B na heterokromatinske regije bogate oznakom H3K9me3 bitan je protein HP1 (engl. *heterochromatin protein 1*) koji prepoznaje i veže se sa svojom kromodomenom za tri metilne skupine na Lys9 histona H3[101].

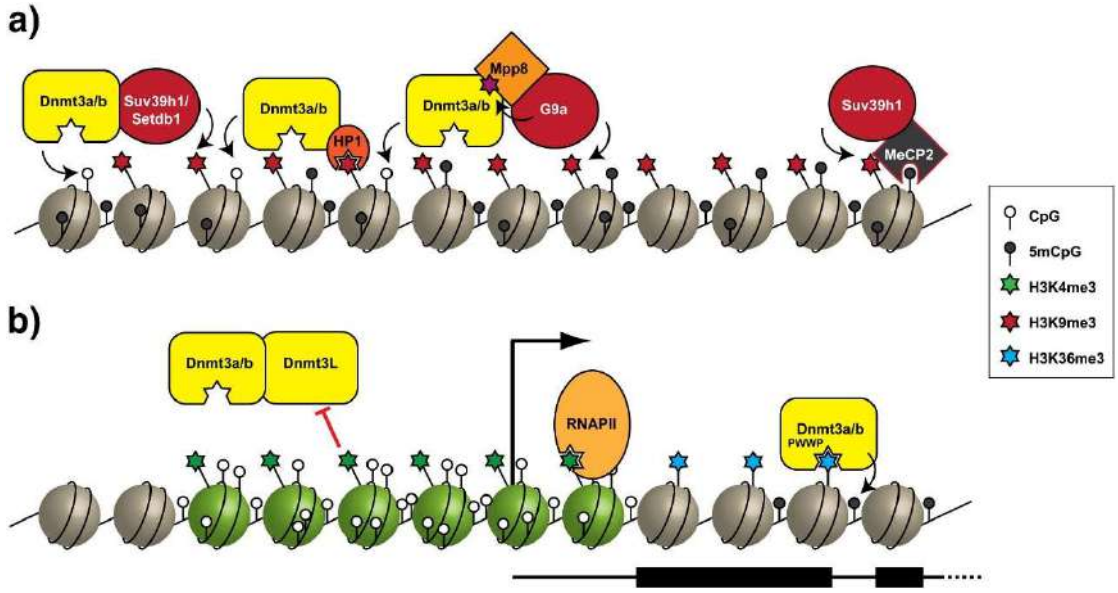
Također, i drugi proteini s kromodomenom mogu posredovati u interakcijama s DNMT3A i DNMT3B. Protein MMP8 (engl. *M-phase phosphoprotein 8*) sudjeluje u interakciji DNMT3A ili DNMT3B s histon metiltransferazom G9a čime dovodi ta dva proteina u kompleks[102]. Osim indirektnim dovođenjem DNA metiltransferaza *de novo* u područje heterokromatina pokazano je kako i same metiltransferaze DNMT3A i DNMT3B mogu direktno interagirati s histon metiltransferazama SUV39H1 i SETDB1 (Slika 6.A)[103,104].

Demetilacija molekule DNA, posredovana proteinima iz obitelji TET, dovodi do otvaranja kromatina i do povećanja ekspresije gena[67,105,106]. Pokazano je kako TET1 direktno veže transkripcijske faktore te tako utječe na ekspresiju gena[107,108]. Također, TET1 može direktno vezati i histon acetiltransferazu MOF, zaslužnu za uvođenje aktivne oznake H4K16ac, što dovodi do relaksiranije strukture kromatina i omogućava vezanje transkripcijskim faktorima[109,110]. Proteini TET mogu interagirati i s proteinom OGT (engl. *O-linked N-acetylglucosamine transferase*) koji pojačava uvođenje aktivne oznake H3K4me3 vezanjem na histon metiltransferazu SETD1[111]. Metilacija *de novo* molekule DNA ovisi o odsustvu metilne skupine na poziciji Lys4 u histonu H3 jer se za nemetilirane Lys4 specifično veže metiltransferaza DNMT3L koja pak regrutira DNMT3A/DNMT3B (Slika 6.B)[61,112]. Metilacija H3K4 stoga blokira metilaciju DNA te otvara kromatin i omogućava stvaranje transkripcijskog inicijacijskog kompleksa sačinjenog od općih transkripcijskih faktora TFIIA, TFIIB, TFIIE, TFIIF, TFIIH i TBP (engl. *TATA-box binding protein*) i RNA polimeraze tipa II[113–116].

Trimetilacija Lys36 u histonu H3 (H3K36me3) jedna je od aktivnih oznaka karakteristična za tijelo gena, a ovu oznaku u kromatin uvodi histon metiltransferaza SETD2[117]. Interakcija SETD2 s C-terminalnom domenom RNA polimeraze tipa II omogućava uvođenje aktivne oznake H3K36me3 u tijelo gena[117,118]. Metiltransferaze DNMT3A i DNMT3B specifično prepoznaju i vežu oznaku H3K36me3, te metiliraju molekulu DNA na toj genomskoj regiji što djeluje sinergistički na ekspresiju gena (Slika 6.B)[119,120].

Pokazano je kako je pozicija nukleosoma u regulatornim regijama gena točno organizirana čime je omogućeno stvaranje regija bez nukleosoma (engl. *nucleosome-depleted region, NDR*)[121–123]. Veličina takvih regija odgovara duljini molekule DNA omotane oko jednog nukleosoma, međutim može varirati u različitim tipovima stanica. Pokazano je kako transkripcijski aktivni geni sadrže u svojim regulatornim regijama područja oslobođena nukleosoma dok to nije slučaj kod transkripcijski inaktivnih gena[124,125]. Područja bez nukleosoma omogućavaju da je sekvenca DNA, koja odgovara regulatornim regijama aktivnih

gena, dostupna transkripcijskim faktorima i drugim regulatornim proteinima[126]. Na stvaranje genomskih područja oslobođenih nukleosoma utječe sama sekvenca molekule DNA, kromatinski remodelirajući kompleksi kao što je kompleks Swi/Snf, međutim i samo vezanje nekih transkripcijskog faktora, kao što su FOXA, GATA, PU.1 i AP1, može potaknuti lokalno remodeliranje kromatina[127–131].



Slika 6. Mehanizam regulacije ekspresije gena posredovan metilacijom DNA i post-translacijskim oznakama histona. (A) DNA metiltransferaze DNMT3A i DNMT3B mogu biti regrutirane na određeno mjesto u genomu pomoću proteina HP1, koji sa svojom kromodomenom prepoznaje i veže se na inaktivnu oznaku H3K9me3. Osim toga, protein MMP8 sudjeluje u interakciji metiltransferaza DNMT3A i DNMT3B s histon metiltransferazom G9a čime dovodi ta dva proteina u kompleks na molekuli DNA. Također, metiltransferaze DNMT3A i DNMT3B mogu i same direktno vezati histon metiltransferaze SUV39H1 i SETDB1. **(B)** Unutar regulatornih regija aktivnih gena, koje ne sadrže metilirane CpG dinukleotide, zastupljena je aktivna oznaka H3K4me3 koja blokira vezanje DNMT3L te na taj način onemogućava vezanje metiltransferaza *de novo* i uvođenje metilne skupine na citozine. U tijelu aktivnih gena zastupljena je aktivna oznaka H3K36me3 koju prepoznaju i vežu DNA metiltransferaze DNMT3A ili DNMT3B te metiliraju molekulu DNA što djeluje sinergistički na ekspresiju gena. Preuzeto iz Rose, Klose[132].

1.5 Primjena sustava CRISPR/dCas9 u ciljanoj manipulaciji metilacije DNA

Razumijevanje uloge određenih epigenetičkih oznaka u regulaciji ekspresije gena prvobitno se temeljilo na rezultatima korelacijskih studija. Razvojem sustava CRISPR/dCas9 omogućena je direktna manipulacija epigenetičkih oznaka te razjašnjavanje njihove uloge u kompleksnoj mreži regulacije ekspresije gena. Osnovni preduvjet za primjenu sustava CRISPR/Cas9 izvan ciljane manipulacije genoma (uvođenje mutacija) bio je inaktivirati nukleazne domene RuvC i HNH uvođenjem mutacija D10A i H840A, čime je dobivena katalitički inaktivirana nukleaza dCas9 (engl. *dead Cas9*)[133]. Pošto je inaktivna nukleaza dCas9 zadržala mogućnost vezanja na ciljno mjesto u genomu, njenom fuzijom s različitim efektorskim domenama moguće je iste dovesti na željeno mjesto u genomu bez uvođenja loma u molekulu DNA. Ovako prenamijenjen sustav CRISPR/dCas9 našao je primjenu u različitim područjima - od ciljane manipulacije epigenetičkih oznaka na molekuli DNA[134–137] ili histonskim proteinima[138–140], direktne manipulacije genske ekspresije[141–144], promjene arhitekture jezgre[145,146], vizualizacije kromosoma i lokusa[147,148] pa sve do ciljanih promjena dušičnih baza u molekuli DNA[149,150].

Direktnim vezanjem katalitičke domene DNA metiltransferaze DNMT3A na inaktiviranu nukleazu dCas9 iz bakterije *Streptococcus pyogenes*, dSpCas9 (Slika 7.), preko fleksibilnog peptidnog lanca sastava Gly_4Ser omogućeno je po prvi puta ciljano uvođenje metilacijske oznake na citozine u promotorima gena *BACH2* i *IL6ST* u modelnim stanicama HEK293. Povećanje stupnja metilacije citozina do 60% zabilježeno je u lokusu *BACH2*, dok je maksimalno povećanje stupnja metilacije citozina u lokusu *IL6ST* doseglo 35% na određenim CpG mjestima. Također, u istoj studiji je pokazano da inducirana metilacija DNA uz pomoć alata dCas9-DNMT3A navođenog specifičnim molekulama sgRNA na promotorske regije genskih lokusa *BACH2* i *IL6ST* mijenja razinu njihove transkripcijske aktivnosti, čime je pokazana uzročno-posljedična veza metilacije specifičnih CpG mjesta i ekspresije gena[136]. U drugoj studiji, fuzijom katalitičke domene DNA metiltransferaze DNMT3A na dCas9 ciljani su genski lokusi *CDKN2A* i *ARF* te je pokazano povećanje metilacije DNA do 50%[151]. Također, na inaktiviranu nukleazu dCas9 vezana je i kompletna DNMT3A te je i takav pristup pokazan uspješnim u *in vitro* i *in vivo* eksperimentima[137]. Pokazano je kako DNMT3L pojačava *de novo* DNA metilaciju kroz stvaranje tetramernog kompleksa s DNMT3A[63,152]. Vezanjem kimernog DNMT3A-DNMT3L kompleksa na protein dCas9 dobiven je fuzijski konstrukt koji pokazuje 4 do 5 puta veći efekt na povećanje stupnja metilaciju DNA nego što je postignuto korištenjem

fuzijskog konstrukta koji sadrži dCas9 i vezanu katalitičku domenu DNA metiltransferaze DNMT3A[153].

Ciljanje fuzijskog konstrukta, koji sadrži demetilazu TET1 i protein dCas9, uz pomoć specifičnih sgRNA za određeni lokus također se pokazalo vrlo uspješnim u demetilaciji citozina, te posljedičnoj reaktivaciji ciljanih gena (Slika 7.). Primjer je ciljano dovođenje fuzijskog konstrukta dCas9-TET1 unutar promotora gena *BRCA1*, što je rezultiralo demetilacijom ciljanih citozina te posljedično dovelo do reaktivacije ovog gena[154]. Također, direktna fuzija dCas9-TET1 pokazala se uspješnom u demetilaciji nekih drugih genomskih regija osim promotora. Reprogramiranje fibroblasta u mioblaste potaknuto je nakon ciljane demetilacije pojačivača (engl. *enhancer*) gena *MyoD*[137]. Ciljana demetilacija regije koja sadrži ponavljanja trinukleotida CGG dovela je do uspješne reaktivacije gena *FMR1*, odgovornog za sindrom fragilnog kromosoma X[155]. Nakon ciljane demetilacije trinukleotidnih CGG ponavljanja povećala se razina aktivnih histonskih oznaka H3K27ac i H3K4me3, dok se smanjila razina represivne oznake H3K9me3. Posljedično, reaktivacija gena *FMR1* dovela je do popravka abnormalnog fenotipa u modelnim induciranim pluripotentnim matičnim stanicama FX52 kao i u stanicama neurona dobivenih diferencijacijom. Transplantacija takvih modificiranih stanica neurona u mozak miševa pokazala je održivost normalne ekspresije gena *FMR1 in vivo* te potencijal ove metode u terapijske svrhe u slučaju sindroma fragilnog X kromosoma[155].

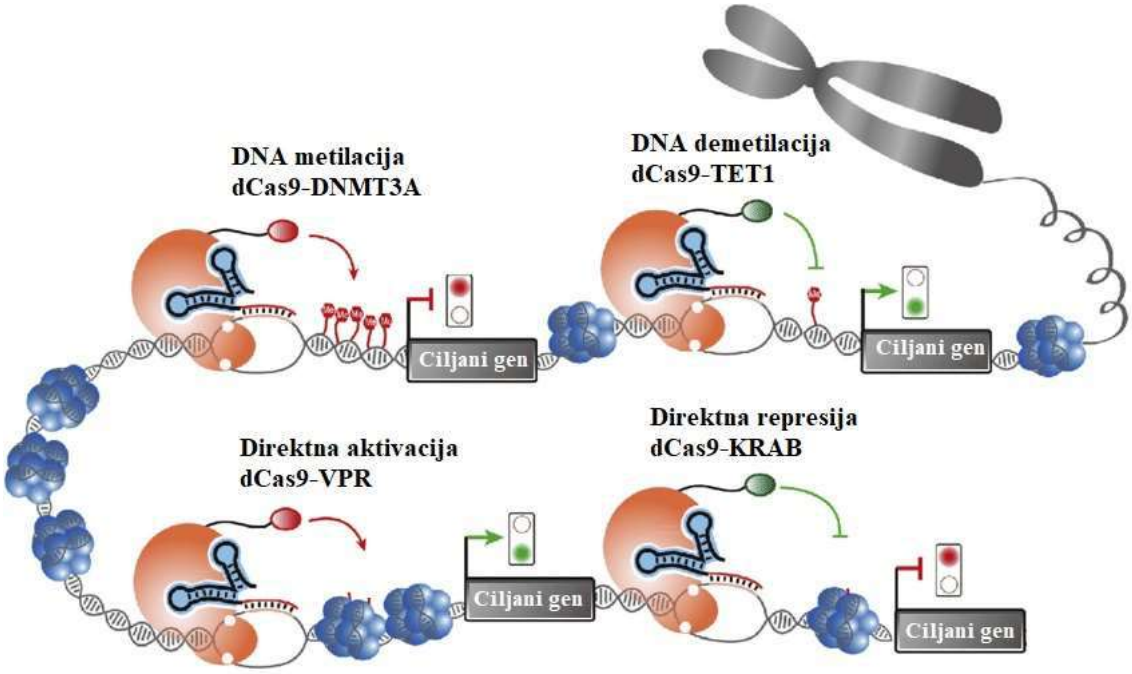
1.6 Primjena sustava CRISPR/dCas9 u direktnoj regulaciji ekspresije gena

Vezanje same inaktivirane nukleaze dCas9 na ciljno mjesto u genomu može interferirati s vezanjem drugih proteina, kao što su transkripcijski faktori ili RNA polimeraza tipa II, te na taj način direktno utjecati na transkripciju. Takav pristup, nazvan CRISPR interferencija (CRISPRi), pokazao se učinkovitim u utišavanju ekspresije ciljanih gena[133,144,156]. Gilbert i suradnici testirali su fuzije različitih represorskih domena s dCas9 u svrhu utišavanja ekspresije ciljanih gena. Domena KRAB (engl. *Kruppel associated Box*) pokazala je veći efekt na utišavanje ekspresije gena (oko 5 puta) od kromodomene proteina HP1α i domene WRPW (Trp-Arg-Pro-Trp) proteina Hes1 (engl. *Hes Family BHLH Transcription Factor 1*) koje su smanjile ekspresiju ciljanih gena svega 2 puta[144]. Domena KRAB prisutna je na N-terminalnoj regiji proteina cinkovih prstiju koji u sisavaca djeluju kao transkripcijski represori[157,158]. Dovođenjem domene KRAB na ciljno mjesto u genomu regrutira se korepresor KAP1 (engl. *KRAB-*

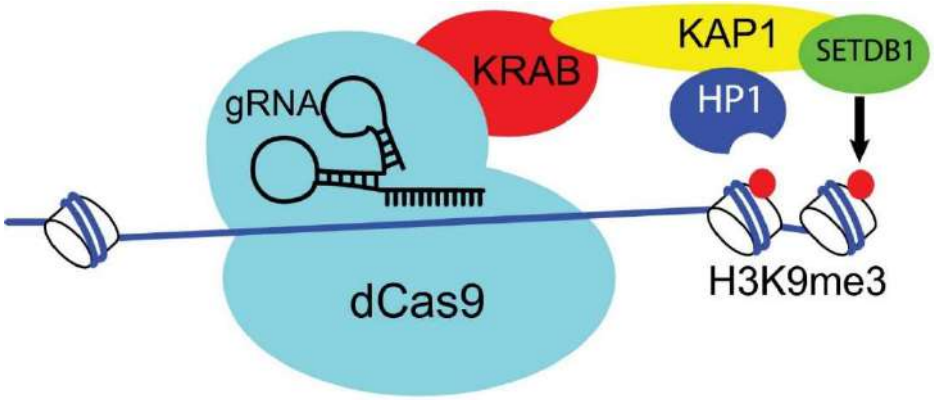
associated protein-1) koji potom regrutira protein HP1 i histon metiltransferaze, koje uvode inaktivnu oznaku H3K9me3 i tako sudjeluju u kondenzaciji kromatina (Slika 8.)[159,160]. Fuzija dCas9-KRAB pokazala se uspješnom u utišavanju različitih ciljanih gena[143,161–163]. Međutim, utišavanje ekspresije gena uzrokovano vezanjem domene KRAB na određeno mjesto u genomu nije trajno već se ekspresija gena vrati na normalnu razinu nakon određenog vremena[161,164]. Studija Gilbert i suradnika pokazala je kako je za uspješno utišavanje genske ekspresije ključno ciljanje fuzijskog konstrukta dCas9-KRAB na regiju od -50 do +300 pb u odnosu na TSS[163].

Različite aktivacijske domene također se mogu vezati na dCas9 u svrhu direktne aktivacije ekspresije gena. Vezanje domene VP64, sastavljene od četiri uzastopne kopije aktivacijske domene VP16[165] iz virusa *Herpes simplex*, na protein dCas9 i navođenje takvog fuzijskog konstrukta specifičnim molekulama sgRNA na ciljne gene uspješno je povećalo njihovu transkripcijsku aktivnost[141,166,167] Isto je pokazalo korištenje aktivacijske domene p65 iz transkripcijskog faktora NF-κB u fuzijskom konstruktu s dCas9[144]. Također, različit broj ponavljanja aktivacijske domene VP16 (VP48 i VP160) testiran je u fuziji s dCas9 te je pokazano kako je za efikasno povećanje ekspresije gena potrebno korištenje više molekula sgRNA istovremeno[168]. Najveći efekt na povećanje ekspresije gena pokazala je fuzija dCas9 s tripartitnim aktivatorom VPR sastavljenim od aktivacijskih domena VP64, p65 i Rta iz virusa *Epstein-Bar*. Fuzija dCas9-VPR pokazala je puno veću efikasnost (22 do 320 puta) od fuzije dCas9-VP64 u svrhu povećanja razine transkripcijske aktivnosti ciljanih gena[142].

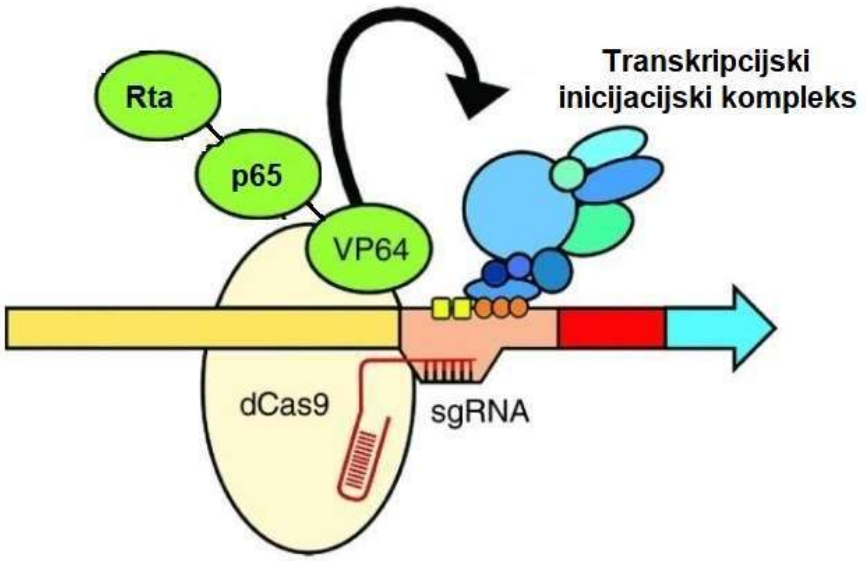
Mehanizam aktivacije ekspresije gena, nakon ciljanja s direktnim aktivatorima vezanima na protein dCas9, temelji se na regrutaciji ostalih aktivatora transkripcije. Pokazano je kako direktni aktivatori regrutiraju brojne proteine uključene u regulaciju transkripcije gena, kao što su histon acetiltrasferaza p300, opći transkripcijski faktori TFIIB, TFIID i TFIIH i kromatin remodelirajući kompleksi, a sve u svrhu otvaranja kromatina i sklapanja transkripcijskog inicijacijskog kompleksa (Slika 9.)[169–173]. U svrhu uspješne aktivacije ekspresije gena fuzijskim konstruktom koji se osniva na vezanju aktivatora na dCas9, prilikom dizajna molekula sgRNA, potrebno je voditi računa o ciljanju fuzijskog konstrukta na regiju lociranu -400 do -50 pb u odnosu na na TSS[163].



Slika 7. Primjena sustava CRISPR/dCas9 u ciljanoj manipulaciji epigenoma i direktnoj regulaciji ekspresije gena. Vezanjem katalitičke domene DNA metiltransferaze DNMT3A na inaktiviranu nukleazu dCas9 (dCas9-DNMT3A) moguće je ciljano uvoditi metilnu oznaku na citozine u molekuli DNA. Ciljanjem promotorske regije fuzijskim konstruktom dCas9-DNMT3A dolazi do kompakcije kromatina te posljedično do smanjenja ekspresije gena. Također, moguće je postići uklanjanje metilne skupine s citozina u molekuli DNA, u svrhu reaktivacije gena, navođenjem fuzijskog konstrukta dCas9-TET1 na ciljno mjesto u genomu. Vezanjem direktnog aktivatora (VPR) ili represora (KRAB) na dCas9 moguće je direktno regulirati ekspresiju gena. Preuzeto i prilagođeno iz Xu i sur.[174].



Slika 8. Utišavanje ekspresije gena pomoću fuzijskog konstrukta dCas9-KRAB. Dovođenjem domene KRAB, vezane na dCas9, na određeno mjesto u genomu regrutira se korepresor KAP1 koji potom regrutira protein HP1 i histon metiltransferazu SETDB1 te posljedično dolazi do uvođenja inaktivne oznake H3K9me3 i kondenzacije kromatina. Preuzeto i prilagođeno iz Lavender i sur.[175].



Slika 9. Aktivacija ekspresije gena pomoću fuzijskog konstrukta dCas9-VPR. Dovođenjem tripartitnog aktivatora VPR, vezanog na dCas9, pomoću specifične sgRNA na određeno mjesto u genomu dolazi do regrutacije endogenih transkripcijskih faktora, kromatin remodelirajućih kompleksa, te do slaganja transkripcijskog inicijacijskog kompleksa u području inicijacije transkripcije i početka same transkripcije specifičnog gena. Preuzeto i prilagođeno iz Limsirichai i sur.[176].

1.7 Nespecifična aktivnost sustava CRISPR/Cas9

Danas se sustav CRISPR/Cas9 uspješno koristi u svrhu manipulacija eukariotskih genoma. Međutim, mnogobrojne studije pokazuju njegov glavni nedostatak, a to je nespecifična aktivnost (engl. *off-target activity*)[177–180]. Problematiku nespecifične aktivnosti možemo podijeliti u dvije glavne kategorije. Prva kategorija obuhvaća nespecifično uvođenje lomova u genom. Za pravilnu aktivaciju i vezanje nukleaze Cas9 potrebno je točno sparivanje proksimalne regije PAM, tzv. „seed" regije, dužine od 8 do 13 nt koja se nalazi neposredno uzvodno od sekvence PAM[28,30,31,35]. Dokazano je kako se pogreške u sparivanju distalne regije PAM puno bolje toleriraju. Studijama *in vitro* potvrđeno je da nukleaza Cas9 uvodi lomove na nespecifična mjesta u genomu koja sadrže do pet krivo sparenih nukleotida[177,178,181]. Novije studije pokazale su da se može tolerirati čak i šest krivo sparenih nukleotida[182]. Također, prepoznavanje alternativnih sekvenci PAM od strane nukleaze Cas9 može rezultirati uvođenjem lomova na neželjenim mjestima u genomu[177,181,183,184]. Novija istraživanja pokazuju da molekula sgRNA može stvarati omče veličine do 4 nt (engl. *RNA bulges*) kako bi vezala nespecifično mjesto, te da omče u samoj molekuli DNA također pridonose nespecifičnoj aktivnosti sustava CRISPR/Cas9[185].

Razvijeni su različiti pristupi kojima se smanjuje razina nespecifične aktivnosti sustava CRISPR/Cas9, a oni obuhvaćaju: (1) pravilan dizajn, kemijske modifikacije i skraćivanje duljine molekule sgRNA; (2) variranje efektivne koncentracije kompleksa sgRNA-Cas9 upotrebom slabijih promotora ili direktnom dostavom pročišćenih ribonukleoproteinskih kompleksa; (3) korištenje mutiranih varijanti proteina Cas9; (4) primjenu različitih ortologa proteina Cas9; (5) korištenje proteina anti-CRISPR (engl. *Anti CRISPR*, Acr)[37,186–190].

Kromatinskom imunoprecipitacijom potvrđeno je nespecifično vezanje proteina Cas9 duž cijelog genoma[36,191,192]. Pokazano je kako svaka korištena molekula sgRNA dovodi protein Cas9 na nekoliko tisuća različitih mjesta u genomu. Također, različite molekule sgRNA dizajnirane specifično za isti gen pokazuju za nekoliko stotina puta manji ili veći broj nespecifičnih veznih mjesta u genomu[36]. Problematika nespecifičnog vezanja nukleaze Cas9 duž cijelog genoma najviše je izražena prilikom primjene sustava CRISPR/Cas9 u ciljanom manipuliranju epigenoma. Dokazano je kako fuzijski konstrukti gdje je vezana katalitička domena DNMT3A ili DNMT3B s inaktiviranom nukleazom dCas9 uzrokuju nespecifičnu metilaciju DNA duž cijelog genoma. Nespecifična metilacija DNA prvenstveno zahvaća otvorene kromatinske regije kao što su CpG otoci, promotori i 5' UTR regije (engl. *5′*

untranslated region), dok u ostalim područjima genoma, koja su inače visoko metilirana, nije zabilježeno dodatno povećanje stupnja metilacije[193]. Različite studije su pokazale kako je nespecifično i globalno povećanje u metilaciji DNA neovisno o molekuli sgRNA koja navodi određeni fuzijski konstrukt[193,194]. Korištenje inaktivirane katalitičke domene DNMT3A u fuziji s dCas9 potvrdilo je da je sama katalitička domena odgovorna za nespecifičnu metilaciju DNA[193]. Studija Galonska i suradnika pokazuje kako je nespecifična globalna metilacija DNA nasumičnog karaktera, te da je zapravo posljedica slučajnih kontakata katalitičke domene DNMT3A i kromatina u jezgri[194]. Različite studije pokazuju da je smanjenje razine ekspresije određenog fuzijskog konstrukta ili pak same katalitičke domene jedan od načina kako umanjiti nespecifičan učinak sustava CRISPR/dCas9 u ciljanom manipuliranju epigenoma[151,193,195,196].

1.8 Primjena različitih ortologa proteina Cas9 izvan ciljanog manipuliranja genoma

Najčešće korišten ortolog proteina Cas9 je onaj izoliran iz vrste *Streptococcus pyogenes* (SpCas9), a njegova primjena pokazana je uspješnom u ciljanom manipuliranju epigenoma kao i u direktnoj regulaciji ekspresije gena. Uspješno su konstruirane fuzije proteina dSpCas9 s brojnim „pisačima“ (proteini koji uvode epigenetičke oznake) i „brisačima“ (proteini koji uklanjaju epigenetičke oznake na histonskim proteinima i molekuli DNA) kao što su primjerice histon deacetilaza (HDAC)[140], histon acetiltransferaza p300[139], lizin specifična demetilaza (LSD1)[138], DNA metiltransferaza DNMT3A[136,137] i DNA demetilaza TET1[134,137]. Osim epigenetičkih efektorskih domena, na dSpCas9 uspješno su vezane transkripcijska aktivacijska domena VPR[197] i represorska domena KRAB[143] čime je omogućena direktna manipulacija transkripcijske aktivnosti ciljanih gena.

Primjena drugih ortologa proteina Cas9 ostala je dosta limitirana uglavnom na manipuliranje genoma eukariotskih stanica[190]. Protein Cas9 iz vrste *Staphylococcus aureus* (SaCas9) jedan je od najmanjih ortologa proteina Cas9 koji prepoznaje dužu sekvencu PAM slijeda 5'-NNGRRT-3' (R predstavlja purinske dušične baze). Upotreba ortologa SaCas9 pokazala se robusnom u eksperimentima *in vitro* ali i *in vivo*, uz postignutu veću specifičnost od najčešće korištenog ortologa SpCas9[198–200]. Osim u svrhu manipuliranja genoma, na inaktiviranu nukleazu dSaCas9 (sadrži mutacije D10A i N580A) uspješno je vezan direktan transkripcijski aktivator VPR (dSaCas9-VPR) te je testirana njegova učinkovitost u svrhu direktne aktivacije genske ekspresije korištenjem molekula sgRNA različite duljine za navođenje fuzijskog konstrukta[197]. Također, konstruirana je i direktna fuzija dSaCas9 s

transkripcijskom represorskom domenom KRAB (dSaCas9-KRAB). Ciljanjem gena *Pcsk9* uspješno je utišana njegova ekspresija *in vivo* te je posljedično smanjena razina kolesterola LDL u miševa[201]. Gao i suradnici uspješno su upotrijebili dva ortologna proteina SaCas9 i SpCas9 u svrhu ortogonalne regulacije genske ekspresije, aktivacije jednog i represije drugog reporter gena u reporterskoj staničnoj liniji HEK293T pSV40-EGFP pTRE3G-mCherry. Također, u stanicama HEK293T uspješno su simultano aktivirali ekspresiju gena *CD95* pomoću fuzije VPR-dSpCas9, dok su ekspresiju gena *CXCR4* povećali pomoću fuzije VPR-dSaCas9[202].

2 ZNANSTVENI RADOVI

2.1 Active fusions of Cas9 orthologs

Journal of Biotechnology 301 (2019) 18–23

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Active fusions of Cas9 orthologs

Goran Josipović, Vlatka Zoldoš, Aleksandar Vojta*

University of Zagreb, Faculty of Science, Department of Biology, Division of Molecular Biology, Horvatovac 102a, 10000 Zagreb, Croatia

ARTICLE INFO

Keywords:
Cas9 ortholog
Cas9 fusion
sgRNA
DNMT3A
Epigenome editing

ABSTRACT

Many recent epigenetic studies utilize the advantages of CRISPR/dCas9 based tools in linking certain epigenetic modification with gene expression regulation. Various multifactorial diseases often contain changed epigenetic signatures at many loci, so tools for simultaneously targeting different loci would significantly facilitate the understanding of disease pathogenesis. We tested different dCas9 orthologs (dCjCas9, dNmCas9, dSt1Cas9, dFnCas9, dSaCas9 and dSpCas9) in C-terminal fusion with DNMT3A effector domain to find candidates that potentiate effector domain to perform its function at the target site. We demonstrated that nuclear localization signals (NLS) at both termini of fusion constructs is crucial for both proper nuclear import of such large constructs as well as for maximization of targeted DNA methylation activity. We identified SpCas9, SaCas9 and CjCas9 as potential candidates for the fusion constructs. With further optimization of the SaCas9 ortholog, due to less complex PAM requirements in contrast to CjCas9, we showed that N-terminal fusion with DNMT3A (dSaCas9-DNMT3A) is optimal to exert targeted DNA methylation activity comparable to the dSpCas9-DNMT3A construct. N-terminal fusions showed better results for both Cas9 orthologs, SaCas9 and SpCas9, so it can be used as universal approach for linking different effector domains in order to obtain highly active fusions.

1. Introduction

The advent of CRISPR/Cas9 technology revolutionized the field of genome editing by providing a molecular tool that could target almost any sequence in the genome (Wang et al., 2016). The paradigm of targeting a sequence by Cas9 relies on two components besides the Cas9 protein and the target sequence: (1) the PAM sequence ("protospacer adjacent motif") and (2) a small guidance RNA molecule (sgRNA). The PAM is a short sequence (such as 5'-NGG-3' in case of SpCas9) which must occur immediately after the targeted sequence. The single guide RNA (sgRNA) consists of a "scaffold", which is the constant part responsible for Cas9 binding, and a variable part of about 20 nucleotides, which "programs" the target sequence by base complementarity. The native Cas9 is a nuclease; it can be rendered catalytically inactive (nuclease-null or "dead Cas9", dCas9) by mutation of catalytic residues, in which case it only binds its target without cutting.

CRISPR/Cas9 tools were initially used for genome editing and gene knockout, but their application has been expanded to programmable regulation of gene expression, as well as to use in epigenetic editing. Different epigenetic effector domains have been fused to catalytically inactive dCas9 (Hilton et al., 2015; Kearns et al., 2015; Kwon et al., 2017; Lei et al., 2017; Liu et al., 2016; Vojta et al., 2016), which functions as a programmable targeting moiety, thus providing the opportunity to change epigenetic marks at a specific locus. In addition to epigenetic effectors, various domains for direct regulation of gene expression can also mediate fast changes in gene expression pattern (Chavez et al., 2015; Gilbert et al., 2013; Konermann et al., 2015; Yeo et al., 2018), providing opportunities for experimental design that links gene expression changes with specific phenotypes. However, the best characterized and the most widely used Cas9 protein for such purposes comes from *Streptococcus pyogenes* CRISPR/Cas9 system (SpCas9).

Soon after the SpCas9 came into widespread use as a molecular tool for genome editing, orthologs from other bacteria were described and characterized: from *Streptococcus thermophilus* (St1Cas9) and *Neisseria meningitidis* (NmCas9) (Esvelt et al., 2013), *Staphylococcus aureus* (SaCas9) (Nishimasu et al., 2015), *Francisella novicida* (FnCas9) (Hirano et al., 2016) and, more recently, *Campylobacter jejuni* (CjCas9) (Yamada et al., 2017). The orthologs differ in their PAM requirement, size and specificity. Importantly, each ortholog binds its cognate sgRNA, recognizing it by its constant part ("scaffold"). It was soon realized that different orthologs and their sgRNAs co-expressed in the same cell do not interfere with each other, which was initially used for multicolor visualization of genomic regions (Ma et al., 2015).

Molecular tools for genome editing in a eukaryotic cell require co-expression of the Cas9 nuclease and a targeting sgRNA; Cas9 used in that way needs to have a nuclear localization signal (NLS) in order to be

* Corresponding author.
E-mail address: vojta@biol.pmf.hr (A. Vojta).

https://doi.org/10.1016/j.jbiotec.2019.05.306
Received 4 February 2019; Received in revised form 7 May 2019; Accepted 29 May 2019
Available online 31 May 2019

imported from the cytosol into the nucleus. This is an additional concern when dCas9 guides a fused effector domain to the targeted sequence in the genome. Further, the effector domain is particularly sensitive to the way it is connected with Cas9 if it needs to retain a catalytic function, change conformation or bind interaction partners. For this reason, active fusions with DNA-modifying enzymes, such as the *de novo* DNA methyltransferase DNMT3A, were described later than fusions with fluorescent proteins or simple transcriptional activators like VP64 (Stepper et al., 2017; Vojta et al., 2016).

In this work, we screened several Cas9 orthologs in fusion with the catalytic domain of DNMT3A, in a search for the best candidates for active fusions with epigenetic modifiers. We used the successfully constructed and well characterized dSpCas9-DNMT3A (Vojta et al., 2016) as a positive control and a benchmark for other Cas9 fusions.

2. Materials and methods

2.1. Reagents and oligonucleotide fragments

All enzymes used for cloning and ATP were from New England Biolabs (Ipswich, MA, USA) except Esp3I, BpiI and DTT which were from Thermo Fisher Scientific (Waltham, MA, USA). T4 DNA ligase used for ligation was from TaKaRa (Kusatsu, Shiga, Japan). QIAquick PCR Purification Kit for DNA clean-up and QIAquick Gel Extraction Kit for DNA extraction from gel were obtained from Qiagen (Hilden, Germany). ZymoPURE Plasmid Miniprep Kit used for plasmid isolations was obtained from Zymo Research (Irvine, CA, USA). Custom-synthesized oligonucleotides and sequencing services were ordered from Macrogen (Seoul, South Korea). QuikChange Lightning Site-Directed Mutagenesis Kit used for single site-directed mutagenesis, QuickChange Lightning Multi Site-Directed Mutagenesis Kit for multi-site-directed mutagenesis and Herculase II Fusion DNA Polymerase for amplification of DNA fragments used in cloning were obtained from Agilent Technologies (Santa Clara, CA, USA). Source plasmids for cloning were obtained from the Addgene repository (www.addgene.org). All reagents used for cell culture: Dulbecco's Modified Eagle Medium, heat inactivated Fetal Bovine Serum, L-Glutamine and Penicillin Streptomycin Solution were obtained from Sigma-Aldrich (St. Louis, MO, SAD). Lipofectamine 3000 Reagent used for cell transfections was obtained from Invitrogen (Carlsbad, CA, SAD). EZ DNA Methylation-Gold Kit was used for bisulfite conversion of DNA and was obtained from Zymo Research. PCR amplification of bisulfite converted DNA was done with PyroMark PCR Kit (Qiagen).

2.2. Construction of sgRNA and Cas9 modules

To test and compare different Cas9 orthologs (CjCas9, St1Cas9, NmCas9, SaCas9 and FnCas9) with the well-characterized SpCas9, a streamlined strategy was developed to enable rapid cloning dCas9 orthologs and scaffolds of their cognate sgRNA molecules. For sgRNA scaffold cloning, U6 promoter was cut out from the pdCas9-DNMT3A-PuroR_v2 plasmid (Addgene plasmid # 74407) with XbaI and BbsI and purified from gel using a QIAquick Gel Extraction Kit. Custom synthesized oligonucleotides for sg-Scaffold module, containing two BpiI restriction sites needed subsequent cloning of the variable gRNA part, two Esp3I restriction sites for cloning of sgRNA scaffold portion for each ortholog and BsaI restriction site leaving nonpalindromic 4 bp 5' overhangs needed for BsaI golden-gate assembly of final functional fusion construct, were oligo annealed (for sequences of oligonucleotides, see Supplementary Table 4), phosphorylated and cloned between XbaI and XhoI restriction sites in the modified pUK21 vector (with BsaI, Esp3I and BbsI restriction sites removed) along with the U6 promoter, using the T4 DNA Ligase. Scaffold part of the sgRNA for CjCas9, St1Cas9, NmCas9 and FnCas9 orthologs was then custom synthesized and cloned between two Esp3I sites by oligo annealing, yielding pSg-U6-Scaffold plasmids containing two BpiI restriction sites to be used for cloning of variable sgRNA parts.

For cloning of some Cas9 orthologs without removing BsaI sites, we custom synthesized a “module” vector based on pUK21, containing two Esp3I restriction sites for final assembly and NheI and KasI restriction sites needed for cloning of Cas9 protein, Gly_4Ser linker for making different Cterminal Cas9 fusions and SV40 nuclear localization signal. Each Cas9 ortholog was amplified with Herculase II Fusion DNA Polymerase according to the manufacturer’s protocol using primers containing Esp3I restriction site leaving sticky ends compatible with NheI and KasI sites in the pUK21 “module”. Nuclease active CjCas9 and FnCas9 were amplified from plasmid pRGEN-CMV-CjCas9 (Addgene plasmid #89752) and P X 408 Francisella tularensis subsp. novicida Cas9 (Addgene plasmid #68705), respectively, while nuclease inactive dSt1Cas9, dNmCas9 and dSaCas9 were amplified from M-ST1n-VP64 (Addgene plasmid #48675), M-NMn-VP64 (Addgene plasmid #48676) and p X 603-AAV-CMV::NLS-dSaCas9(D10A,N580A)-NLS-3xHA-bGHpA (Addgene plasmid #61594), respectively. Nuclease inactive dSpCas9 was amplified from plasmids used in our earlier work (Vojta et al., 2016). To obtain nuclease inactive dCjCas9, D8A mutation in RuvC-like nuclease domain was introduced during PCR amplification of CjCas9 from origin vector and H559A in HNH nuclease domain mutations was then introduced using a QuikChange Lightning Site-Directed Mutagenesis Kit according to the manufacturer’s protocol. Five codons were changed (D11A, N995A, S1473A, R1474A and R1585A) in FnCas9 to obtain the nuclease inactive dFnCas9.

2.3. Golden Gate assembly of DNMT3A fusion constructs

To test the activity of different Cas9 orthologs in C-terminal fusion with epigenetic effector domain DNMT3A, we used Golden Gate cloning with BsaI and BsmBI for assembly of different “module vectors” encoding different Cas9 orthologs and their respective sgRNA (this work, except SpCas9 and SaCas9); Cbh promoter, DNMT3A catalytic domain including a Gly4Ser linker, puromycin resistance marker and Bgh terminator were developed as a part of our comprehensive molecular toolbox (submitted). This strategy was used for assembly of final expression vectors for testing activity of fusions between Cas9 orthologs and DNMT3A catalytic domain in cell cultures. Assembled constructs were transformed into chemically competent *Escherichia coli* XL10Gold cells (Agilent Technologies). Bacteria were plated on agar with ampicillin (100 µg/ml), IPTG (1 mM) and X-Gal (120 µg/ml) for blue-white selection. Blue colonies represented uncut or re-ligated backbone vector while white colonies representing possible assembly were further analyzed for complete assembly. To test the effect of an additional C-terminal NLS on the activity of Cas9/DNMT3A fusion construct, additional assemblies were done for each Cas9 ortholog according to the described protocol, but using epigenetic effector domain DNMT3A with an additional C-terminal nucleoplasmin NLS. For N-terminal fusions, we used analogous module vectors facilitating assembly of DNMT3A N-terminally to Cas9 orthologs. Such constructs had an N-terminal SV40 NLS (at the beginning of the DNMT3A domain), and a nucleoplasmin NLS at the C-terminus of Cas9 orthologs. Puromycin resistance marker was linked via a T2A self-cleaving peptide in all constructs.

2.4. Cell culture and transfection

Human embryonic kidney cell line HEK293 was maintained in Dulbecco's Modified Eagle Medium supplemented with 10% heat inactivated fetal bovine serum, 4 mM Lglutamine, 100 U/ml penicillin and 100 µg/ml streptomycin. Cells were incubated at 37 °C in a humidified atmosphere containing 5% CO_2.

One day before transfection, HEK293 cells were seeded in 24-well plates (50 000 cells per well) and transfected the next day with Lipofectamine 3000 according to the manufacturer’s protocol. Transfection was done with 500 ng of plasmid DNA co-expressing dCas9-DNMT3A or dCas9-DNMT3A-C-NLS fusion construct and gRNA

in technical duplicates targeting *BACH2* or *MGAT5* locus (for sgRNA sequences, see Supplementary Table 3). One day after transfection, cells were re-plated to 6well plates and selected with puromycin for 48 h. The puromycin selection removed all untransfected cells and ensured that only the transfected population is analyzed. On the 6th day after transfection, cells reached almost 90–100% confluence and were harvested by incubation at 37 °C overnight in lysis buffer (50 mM Tris pH 8.5, 1 mM EDTA, 0.5% Tween 20) with proteinase K (400 µg/ml) for total DNA isolation.

For validation of DNMT3A linked to the Nterminus of dSpCas9 and dSaCas9, transfection was done as described, except that 100 ng of plasmid DNA co-expressing dCas9-DNMT3A fusion construct and gRNA in technical duplicates targeting different regions of *BACH2* locus was used. Cells were harvested on the 8th day after transfection. As negative controls, we used catalytically inactive DNMT3A effector domain with mutation E756A (Chen et al., 2005), as well as fusion constructs with active DNMT3A catalytic domain but with a non-targeting sgRNA molecule (no known target sequence in the human genome, see Supplementary Table 3).

2.5. Bisulfite conversion and pyrosequencing

Bisulfite conversion of DNA was done with EZ DNA Methylation-Gold Kit according to the manufacturer's protocol, after which the genomic regions of interest were amplified using a PyroMark PCR kit. Amplified genomic regions were then sequenced using the PyroMark Q24 Advanced pyrosequencing system (Qiagen) to quantify the methylation levels at individual CpG dinucleotides. Methylation levels were analyzed for *BACH2* and *MGAT5* gene using several developed pyrosequencing assays. Assay sequences and primers used for PCR are listed in Supplementary Tables 1 and 2.

3. Results

3.1. Modular assembly of DNMT3A fusion construct with different dCas9 orthologs

Assembly of various dCas9-DNMT3A fusions was successful, despite BsaI restriction sites present in some dCas9 orthologs. We used a modified version of our modular platform for Golden Gate cloning (submitted for peer review), with modularly switching between different dCas9 orthologs, between DNMT3A versions with or without NLS at the C-terminus, as well as between C- and Nterminal configurations of DNMT3A with respect to dCas9 (Fig. 1). The final constructs for activity testing consisted of a sgRNA unit (for specific Cas9 ortholog with pre-cloned guide RNA), the eukaryotic Cbh promoter, a dCas9 ortholog of choice (dCjCas9, dNmCas9, dSt1Cas9, dFnCas9, dSaCas9 or dSpCas9) with C-terminally linked DNMT3A catalytic domain (or a variant with additional nucleoplasmin NLS on C-terminus) through two tandem Gly_4Ser linkers, the Puromycin resistance as selection marker linked using a self-cleaving 2A peptide and a eukaryotic transcriptional terminator (Fig. 1 A and B). Alternatively, the dSaCas9 and dSpCas9 were tested in a configuration where DNMT3A was fused to N-terminus of dCas9 (Fig. 1C).

3.2. Activity of DNMT3A in C-terminal fusion with different dCas9 orthologs

Each ortholog (CjCas9, NmCas9, St1Cas9, FnCas9, SaCas9 and SpCas9) was tested in fusion with the DNMT3A catalytic domain. Based on our successfully developed C-terminal dSpCas9-DNMT3A construct which showed robust activity at the tested loci (Vojta et al., 2016), we aimed to validate other dCas9 orthologs for their ability to bring C-terminally linked DNMT3A effector domain to target locus and to induce *de novo* CpG methylation. Due to different, more complex PAM requirements than for SpCas9 and FnCas9, other Cas9 orthologs were tested on different genomic loci depending on the availability of PAM sequences. For each Cas9 ortholog, two sgRNA molecules (Supplementary Table 3) were tested to determine if DNMT3A activity profile is present downstream or upstream from PAM sequence.

The orthologs NmCas9, St1Cas9 and FnCas9 in fusion with DNMT3A catalytic domain did not induce any targeted DNA methylation, neither downstream nor upstream from PAM sequence (data not shown). Orthologs CjCas9 and SaCas9 showed minimal increase in DNA methylation upstream from the PAM sequence, each with a single sgRNA (Fig. 2A). The increase in DNA methylation for SpCas9 was up to 35% with average increase for 20% at 40 bp wide region and the peak of activity was around 20 bp downstream from PAM sequence (Fig. 2A), which was expected, since SpCas9 had been well characterized and essentially served as a positive control. Mock transfected cells (without plasmid DNA, crosses) did not show any changes in DNA methylation, just like the other negative control which consisted of non-transfected cells (not shown).

3.3. Nuclear localization signal is needed for activity

Since fusions of DNMT3A to the C-terminus of dCas9 had the potential to interfere with its NLS and were thus liable to poor nuclear import, we tested all orthologs with an additional NLS at the C-terminus of the DNMT3A catalytic domain (Fig. 1B). While NmCas9, St1Cas9 and FnCas9 remained essentially fully inactive (data not shown), we found an increase in methylation effect for SpCas9, SaCas9 and CjCas9 fusions (Fig. 2B). The increase was on the order of 15% compared to the constructs without the additional NLS. We concluded that, in addition to already characterized SpCas9, orthologs SaCas9 and CjCas9 are capable of producing active fusions with DNMT3A catalytic domain, although they are more dependent on proper nuclear targeting.

3.4. Fusions of DNMT3A to the N-terminus of dSaCas9 and dSpCas9 are active

In order to develop a streamlined approach to creating active fusions between dCas9 orthologs and catalytic domains like DNMT3A, we tested a configuration where the catalytic domain is linked at the N-terminus of dCas9 (Fig. 1C), thus avoiding interference with the NLS at the C-terminus of dCas9. We tested dSaCas9 as the most promising

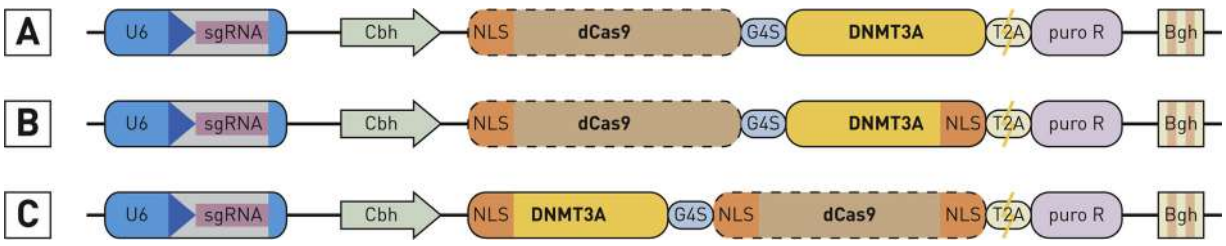


Fig. 1. Schematic representation of tested constructs. Different dCas9 orthologs were tested for DNMT3A activity at the targeted sites. The basic configuration included DNMT3A fused to dCas9 orthologs at their Ctermini, either without (A) or with (B) an additional nuleoplasmin NLS at the C-terminus of the DNMT3A catalytic domain. Additionally, fusion of DNMT3A at the N-terminus of dSpCas9 and dSaCas9 was tested (C).

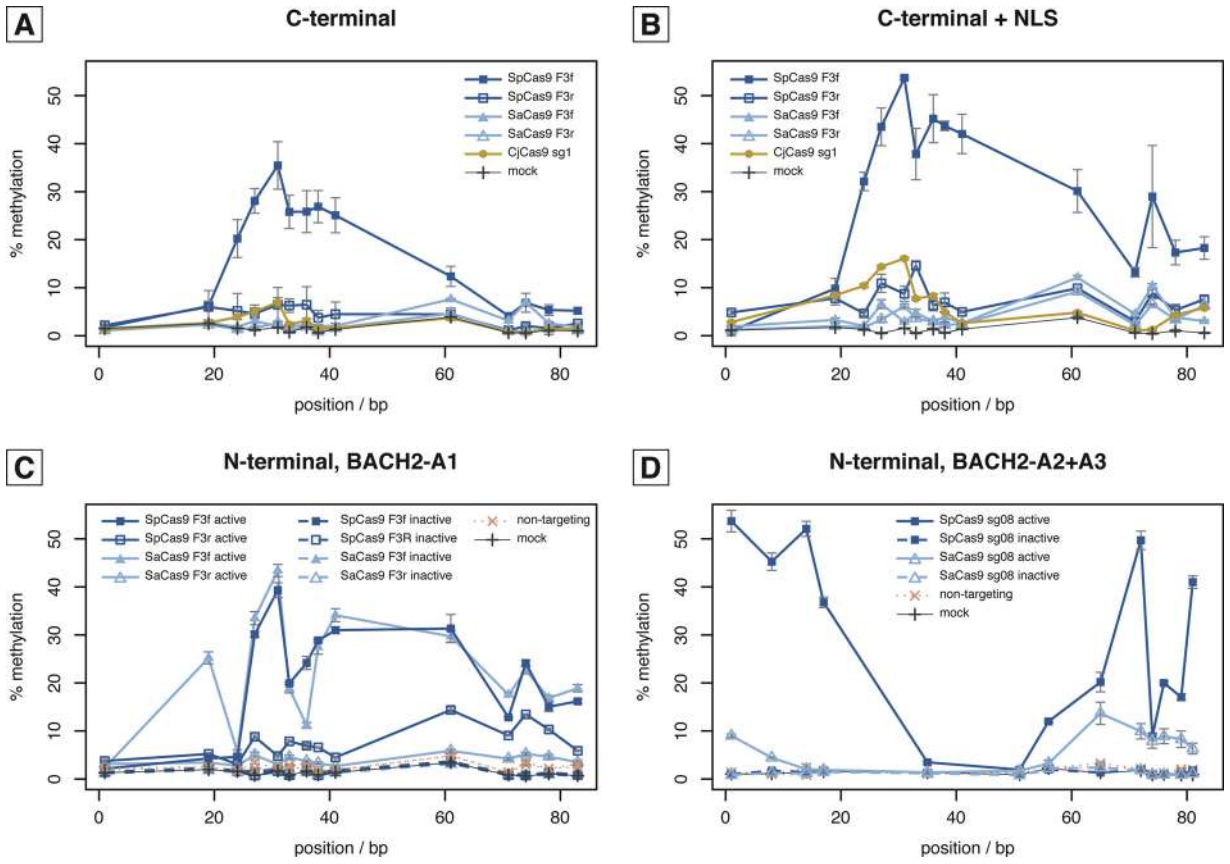


Fig. 2. DNA methylation activity of DNMT3A in fusion with dCas9 orthologs. Promoter of the *BACH2* gene was targeted with DNMT3A linked to dCas9 in various configurations. (A) When DNMT3A was linked to the C-terminus of dCas9, only dSpCas9 (squares), dSaCas9 (triangles) and dCjCas9 (circles) showed methyltransferase activity at the targeted sites. (B) Additional nucleoplasmin NLS at the C-terminus of DNMT3A increased its activity for fusions with dSpCas9, dSaCas9 and dCjCas9. In (A) and (B), BACH2-A1 pyrosequencing assay was used. (C, D) Further experiments with dSpCas9 and dSaCas9 showed their activity in fusion with DNMT3A when it was added to the N-terminus of dCas9. Activity was evident with both pyrosequencing assays BACH2A-A1 and BACH2A-A2 + A3 (C and D, respectively). Mock – untransfected cells (negative control). Non-targeting – scrambled sgRNA without a target in the genome, a negative control to demonstrate guidance-specific activity. Error bars are from three technical replicates.

ortholog against dSpCas9, which served as a positive control. In this configuration, dCas9 had NLS at both ends.

When we targeted the *BACH2* locus with sgRNA "BACH2-sg8", dSpCas9-DNMT3A construct showed high activity of up to 53% methylation increase at some targeted CpGs (Fig. 2D). Activity can be seen both upstream and downstream from PAM sequence. DNMT3A linked to N-terminus of dSaCas9 showed lower activity, up to 14% at some CpGs.

When we targeted another region of the *BACH2* promoter with sgRNA "F3f", dSpCas9- DNMT3A construct was again highly active and increased DNA methylation for up to 43%. In contrast to the other sgRNA, dSaCas9-DNMT3A construct was also highly active with sgRNA "F3f" (Fig. 2C). Activity profiles for both orthologs were almost identical. The methylation was increased for up to 40% with DNMT3A linked to Nterminus of dSaCas9. We included additional negative controls in this experiment – fusions with catalytically inactive DNMT3A, which showed no methylation activity (Fig. 2C and D).

4. Discussion

In this work, we systematically investigated strategies to expand the repertoire of Cas9 orthologs forming active fusions with catalytic domains, such as the DNA methyltransferase DNMT3A. While the SaCas9 ortholog has been used successfully for targeted gene regulation in fusion with the VPR (Fang et al., 2018; Kiani et al., 2015; Nishimasu et al., 2015) or KRAB domain (Thakore et al., 2018), effectors that carry out enzymatic reactions on the DNA molecule itself, such as the DNMT3A, can be more demanding in terms of designing fusions that retain both their catalytic activity and the targeting activity of dCas9. When we fused DNMT3A effector domain to the Cterminus of dCas9 proteins, we achieved robust activity only with SpCas9. The SaCas9 fusion construct with unmodified DNMT3A showed minimal increase in DNA methylation compared to SpCas9, mainly positioned upstream from binding site. For all other orthologs, DNA methylation remained essentially unchanged compared to non-transfected cells. Although a single NLS is sufficient for nuclear import of SpCas9 (Mali et al., 2013), some researchers reported that the addition of a linker between Cas9 and NLS can additionally improve genome editing in eukaryotes, probably because of more successful nuclear import due to the better accessibility of NLS (Shen et al., 2013). Other study reported that NLS located at both termini of a Cas9 showed the best results in human 293FT cells (Cong et al., 2013). Therefore, we speculated that the poor performance of other orthologs might at least in part be caused by inefficient nuclear import, possibly caused by steric hindrances between the NLS and the catalytic domain at the Cterminus. Thus, we added a nucleoplasmin NLS at Cterminus of DNMT3A effector domain, which resulted in detectable increase in activity of fusions with SpCas9, SaCas9 and CjCas9, while other orthologs (dNmCas9, dSt1Cas9 and dFnCas9) remained catalytically inactive. Numerous studies demonstrated that various Cas9 orthologs, like St1Cas9 (Cong et al., 2013; Esvelt et al., 2013; Müller et al., 2016), NmCas9 (Esvelt et al., 2013; Hou et al., 2013; Lee et al., 2016), CjCas9 (Kim et al., 2017) and SaCas9 (Nishimasu et al., 2015; Ran et al., 2015; Ye et al., 2016), were active as nucleases and successfully used for genome editing in human cells,

while the largest Cas9 ortholog, FnCas9, showed no activity in human 293FT cells, but was functional in mouse zygotes after microinjection (Hirano et al., 2016). A recent study demonstrated that binding of FnCas9 is largely affected by chromatin structure and that it can be restored by a novel proxy-CRISPR approach (Chen et al., 2017), so it is no surprising that this ortholog did not function in our research. Among all orthologs, only SaCas9 and CjCas9 showed activity comparable to the widely used SpCas9 (Kim et al., 2017; Ran et al., 2015).

Specificity of the dCas9-DNMT3A fusions was tested for SpCas9 in our earlier work (Vojta et al., 2016) and found to be acceptable. Our recent experiments strongly confirm this (under revision). Further, unspecific activity seems to come from overexpression of the DNMT3A catalytic domain and not from unspecific dCas9 binding. Therefore, we used the previously tested plasmid amount for transfection together with a tested promoter, thus ensuring optimal on-target to off-target activity ratio consistent with our previous successful experiments.

Certain technical limitations of this study need to be discussed. We did not test the dCas9-DNMT3A expression directly; rather, we used CpG methylation analysis to confirm both the expected activity and, indirectly, expression of the fusion protein. Further, due to the large number of constructs to be assembled and tested, we limited the scope of the experimental work to technical replicates only, within a single experiment. Our idea was to quickly screen for promising orthologs for active fusions, which yielded valuable results that served to direct our future research. We relied on our earlier experience of excellent reproducibility of the dCas9-DNMT3A methylation pattern when designing the study.

Taken together, it seems that some Cas9 orthologs (dNmCas9, dSt1Cas9 and dFnCas9) poorly tolerate fusions with catalytic domains and therefore resist repurposing from their genome editing role, while others (SaCas9 and CjCas9) critically depend on proper nuclear targeting. Of the two orthologs, we selected the SaCas9 for further optimization by testing a streamlined approach with a catalytic domain fused to the N-terminus of dCas9, which obviates the need for the catalytic domain to carry an additional NLS. We found that SpCas9 and SaCas9 represent a pair of Cas9 orthologs capable of forming active fusions with epigenome editing enzymes in the same Nterminal configuration. Finding two such orthologs means that different functionalities can be independently targeted to different loci within the same cell, in a way conceptually similar to independent targeting of three fluorescent labels by SpCas9, NmCas9 and St1Cas9 (Ma et al., 2015). However, it should be emphasized that generating catalytically active fusions is much more challenging than retaining the targeting capability of dCas9 in fusion with a fluorescent protein tag, which itself is a rigid and stable structure that does not need conformational changes or substrate binding for activity (Chudakov et al., 2010). The remaining ortholog with a potential for active fusions, the CjCas9, has the disadvantage of a complex PAM requirement that limits its targeting ability. However, it is one of the smallest Cas9 orthologs and might be reconsidered in use cases when small size is critical, such as when constructs need to be packaged into AAV particles (Jo et al., 2019).

Declaration of interest

All the listed authors hereby explicitly declare that they have no competing interests related to the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jbiotec.2019.05.306.

References

Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., Iyer E, P.R., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., Ter-Ovanesyan, D., Braff, J.L., Davidsohn, N., Housden, B.E., Perrimon, N., Weiss, R., Aach, J., Collins, J.J., Church, G.M., 2015. Highly efficient Cas9-mediated transcriptional programming. Nat. Methods.

Chen, F., Ding, X., Feng, Y., Seebeck, T., Jiang, Y., Davis, G.D., 2017. Targeted activation of diverse CRISPR-Cas systems for mammalian genome editing via proximal CRISPR targeting. Nat. Commun. 8, 14958.

Chen, Z.X., Mann, J.R., Hsieh, C.L., Riggs, A.D., Chedin, F., 2005. Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family. J. Cell. Biochem. 95, 902–917.

Chudakov, D.M., Matz, M.V., Lukyanov, S., Lukyanov, K.A., 2010. Fluorescent proteins and their applications in imaging living cells and tissues. Physiol. Rev. 90, 1103–1163.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., 2013. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339, 819–823.

Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., Church, G.M., 2013. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat. Methods 10 (11), 1116–1121.

Fang, L., Hung, S.S.C., Yek, J., El Wazan, L., Nguyen, T., Khan, S., Lim, S.Y., Hewitt, A.W., Wong, R.C.B., 2018. A simple cloning-free method to efficiently induce gene expression using CRISPR/Cas9. Mol. Ther. Nucleic Acids 14, 184–191.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR-mediated modular RNA-Guided regulation of transcription in eukaryotes. Cell 154, 442–451.

Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., Gersbach, C.A., 2015. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol. 33 (5), 510–517.

Hirano, H., Gootenberg, J.S., Horii, T., Abudayyeh, O.O., Kimura, M., Hsu, P.D., Nakane, T., Ishitani, R., Hatada, I., Zhang, F., Nishimasu, H., Nureki, O., 2016. Structure and Engineering of Francisella novicida Cas9. Cell 164 (5), 950–961.

Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.F., Sontheimer, E.J., Thomson, J.A., 2013. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proc. Natl. Acad. Sci. 110 (39), 15644–15649.

Jo, D.H., Koo, T., Cho, C.S., Kim, J.H., Kim, J.S., Kim, J.H., 2019. Long-term effects of in vivo genome editing in the mouse retina using Campylobacter jejuni Cas9 expressed via adeno-associated virus. Mol. Ther. 27, 130–136.

Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., Maehr, R., 2015. Functional annotation of native enhancers with a Cas9-histone demethylase fusion. Nat. Methods 12 (5), 401–403.

Kiani, S., Chavez, A., Tuttle, M., Hall, R.N., Chari, R., Ter-Ovanesyan, D., Qian, J., Pruitt, B.W., Beal, J., Vora, S., Buchthal, J., Kowal, E.J.K., Ebrahimkhani, M.R., Collins, J.J., Weiss, R., Church, G., 2015. Cas9 gRNA engineering for genome editing, activation and repression. Nat. Methods 12 (11), 1051–1054.

Kim, E., Koo, T., Park, S.W., Kim, D., Kim, K., Cho, H.Y., Song, D.W., Lee, K.J., Jung, M.H., Kim, S., Kim, J.H., Kim, J.H., Kim, J.S., 2017. In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. Nat. Commun. 8, 14500.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., Nureki, O., Zhang, F., 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517 (7536), 583–588.

Kwon, D.Y., Zhao, Y.T., Lamonica, J.M., Zhou, Z., 2017. Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC. Nat. Commun. 8, 15315.

Lee, C.M., Cradick, T.J., Bao, G., 2016. The neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells. Mol. Ther. 24 (3), 645–654.

Lei, Y., Zhang, X., Su, J., Jeong, M., Gundry, M.C., Huang, Y.H., Zhou, Y., Li, W., Goodell, M.A., 2017. Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. Nat. Commun.

Liu, X.S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., Shu, J., Dadon, D., Young, R.A., Jaenisch, R., 2016. Editing DNA methylation in the mammalian genome. Cell 167 (1), 233–247 e17.

Ma, H., Naseri, A., Reyes-Gutierrez, P., Wolfe, S.A., Zhang, S., Pederson, T., 2015. Multicolor CRISPR labeling of chromosomal loci in human cells. Proc. Natl. Acad. Sci. U. S. A. 112, 3002–3007.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., 2013. RNA-guided human genome engineering via Cas9. Science 339 (6121), 823–826.

Müller, M., Lee, C.M., Gasiunas, G., Davis, T.H., Cradick, T.J., Siksnys, V., Bao, G., Cathomen, T., Mussolino, C., 2016. Streptococcus thermophilus CRISPR-Cas9 systems enable specific editing of the human genome. Mol. Ther. 24, 636–644.

Nishimasu, H., Cong, L., Yan, W.X., Ran, F.A., Zetsche, B., Li, Y., Kurabayashi, A., Ishitani, R., Zhang, F., Nureki, O., 2015. Crystal structure of Staphylococcus aureus Cas9. Cell 162 (5), 1113–1126.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., Koonin, V.E., Sharp, P.A., Zhang, F., 2015. In vivo genome editing using Staphylococcus aureus Cas9. Nature 520 (7546), 186–191.

Shen, B., Zhang, J., Wu, H., Wang, J., Ma, K., Li, Z., Zhang, X., Zhang, P., Huang, X., 2013. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. Cell Res. 23, 720–723.

Stepper, P., Kungulovski, G., Jurkowska, R.Z., Chandra, T., Krueger, F., Reinhardt, R., Reik, W., Jeltsch, A., Jurkowski, T.P., 2017. Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. Nucleic Acids Res. 45, 1703–1713.

Thakore, P.I., Kwon, J.B., Nelson, C.E., Rouse, D.C., Gemberling, M.P., Oliver, M.L., Gersbach, C.A., 2018. RNA-guided transcriptional silencing in vivo with S. Aureus CRISPR-Cas9 repressors. Nat. Commun. 9 (1), 1674.

Vojta, A., Dobrinic, P., Tadic, V., Bockor, L., Korac, P., Julg, B., Klasic, M., Zoldos, V.,

2016. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Nucleic Acids Res. 44 (12), 5615–5628.
Wang, H., La Russa, M., Qi, L.S., 2016. CRISPR/Cas9 in genome editing and beyond. Ann. Rev. Biochem. 85, 227–264.
Yamada, M., Watanabe, Y., Gootenberg, J.S., Hirano, H., Ran, F.A., Nakane, T., Ishitani, R., Zhang, F., Nishimasu, H., Nureki, O., 2017. Crystal structure of the minimal Cas9 from Campylobacter jejuni reveals the molecular diversity in the CRISPR-Cas9 systems. Mol. Cell 65, 1109–1121 e1103.
Ye, L., Wang, J., Tan, Y., Beyer, A.I., Xie, F., Muench, M.O., Kan, Y.W., 2016. Genome editing using CRISPR-Cas9 to create the HPFH genotype in HSPCs: an approach for treating sickle cell disease and β-thalassemia. Proc. Natl. Acad. Sci. 113 (38), 10661–10665.
Yeo, N.C., Chavez, A., Lance-Byrne, A., Chan, Y., Menn, D., Milanova, D., Kuo, C.C., Guo, X., Sharma, S., Tung, A., Cecchi, R.J., Tuttle, M., Pradhan, S., Lim, E.T., Davidsohn, N., Ebrahimkhani, M.R., Collins, J.J., Lewis, N.E., Kiani, S., Church, G.M., 2018. An enhanced CRISPR repressor for targeted mammalian gene regulation. Nat. Methods 15, 611–616.

2.2 Antagonistic and synergistic epigenetic modulation using orthologous CRISPR/dCas9-based modular system

Published online 14 August 2019

Nucleic Acids Research, 2019, Vol. 47, No. 18 ***9637–9657***
doi: 10.1093/nar/gkz709

Antagonistic and synergistic epigenetic modulation using orthologous CRISPR/dCas9-based modular system

Goran Josipović[1,†], Vanja Tadić[1,†], Marija Klasić[1], Vladimir Zanki[2], Ivona Bečeheli[1], Felicia Chung[3], Akram Ghantous[3], Toma Keser[4], Josip Madunić[1], Maria Bošković[1], Gordan Lauc[4,5], Zdenko Herceg[3], Aleksandar Vojta[1,*] and Vlatka Zoldoš[1,*]

[1]Department of Biology, Division of Molecular Biology, Faculty of Science, University of Zagreb, Horvatovac 102a, 10000 Zagreb, Croatia, [2]Department of Chemistry, Division of Biochemistry, Faculty of Science, University of Zagreb, Horvatovac 102a, 10000 Zagreb, Croatia, [3]Epigenetics group, International Agency for Research on Cancer (IARC), 150 Cours Albert Thomas, Lyon, France, [4]Faculty of Pharmacy and Biochemistry, University of Zagreb, A. Kovačića 1, 10000 Zagreb, Croatia and [5]Genos Glycoscience Research Laboratory, Borogajska cesta 83 h, 10000 Zagreb, Croatia

Received February 06, 2019; Revised July 15, 2019; Editorial Decision August 01, 2019; Accepted August 04, 2019

ABSTRACT

Establishing causal relationship between epigenetic marks and gene transcription requires molecular tools, which can precisely modify specific genomic regions. Here, we present a modular and extensible CRISPR/dCas9-based toolbox for epigenetic editing and direct gene regulation. It features a system for expression of orthogonal dCas9 proteins fused to various effector domains and includes a multi-gRNA system for simultaneous targeting dCas9 orthologs to up to six loci. The C- and N-terminal dCas9 fusions with DNMT3A and TET1 catalytic domains were thoroughly characterized. We demonstrated simultaneous use of the DNMT3A-dSpCas9 and TET1-dSaCas9 fusions within the same cells and showed that imposed cytosine hyper- and hypo-methylation altered level of gene transcription if targeted CpG sites were functionally relevant. Dual epigenetic manipulation of the *HNF1A* and *MGAT3* genes, involved in protein *N*-glycosylation, resulted in change of the glycan phenotype in BG1 cells. Furthermore, simultaneous targeting of the TET1-dSaCas9 and VPR-dSpCas9 fusions to the *HNF1A* regulatory region revealed strong and persistent synergistic effect on gene transcription, up to 30 days following cell transfection, suggesting involvement of epigenetic mechanisms in maintenance of the reactivated state. Also, modulation of dCas9 expression effectively reduced off-target effects while maintaining the desired effects on target regions.

INTRODUCTION

The exceptional versatility of RNA-guided CRISPR/Cas9 system has led to a shift from its primary use as molecular scissors for precise genome editing to a new role of a moiety for easily programmable targeting and delivery of various functional domains for various purposes: manipulation of selected DNA loci/sequences, multicolor imaging of chromosomal loci in living cells, tracking endogenous mRNAs and/or study of chromatin conformation and dynamics (1–5). Various domains for locus-specific transcriptional activation (e.g. VP64, VPR, p300) or repression (e.g. KRAB) can been fused to nuclease-dead Cas9 (dCas9) and such CRISPR/dCas9 systems, guided to the target sequences by single guide RNA molecules (gRNA), have been used in order to manipulate gene transcription (6–11). Recent studies demonstrated that CRISPR/dCas9-based molecular tools could be repurposed for simultaneous activation and repression of different genes. The Krüppel-associated box (KRAB) and VPR domains were fused independently to two different orthologous dCas9 proteins, isolated from *Streptococcus pyogenes* (SpCas9) and *Staphylococcus aureus* (SaCas9), and such molecular tools were then used for inducible orthogonal gene regulation within the same cell (12). This and similar CRISPR activation systems were successfully used for simultaneous activation of multiple loci *in vivo* and high-throughput screening of driver and suppressor genes involved in cancer cell growth (13,14). Such

*To whom correspondence should be addressed. Tel: +385 1 4606 275; Email: vojta@biol.pmf.hr
Correspondence may also be addressed to Vlatka Zoldoš. Tel: +385 1 4606 266; Email: vzoldos@biol.pmf.hr
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

technology is extremely useful for manipulation of multiple candidate genes and has a huge potential for applications in modeling of developmental processes and disease.

A growing amount of publications demonstrates how environment and lifestyle factors are associated with coordinated changes in gene transcription due to epigenetic aberrations, and how epigenetic changes are key drivers of pathology in many cancers (15) but also in other complex diseases, including immune-mediated and inflammatory diseases (16–19). Therefore, identification of epigenetic aberrations and targeting them with new drugs, which are more precise compared to epigenetic inhibitors, have been the focus of the latest research in the fields of epigenetics, biomedicine and regenerative medicine (20). The methodology to engineer the epigenome using CRISPR/dCas9 could greatly contribute to recent advancements (21). So far, it has been demonstrated for fragile X syndrome, the most common genetic disorder, that restoration of aberrant *FMR1* gene expression can be achieved by reversing aberrant methylation at microsatellite repeats in proximity of the promoter of this gene using CRISPR/dCas9-TET1 molecular tool (22).

While simple fusions of DNMT3A and TET1 with dCas9 have already been described and used in several works (23–28), there is no universal and robust approach for creating active fusions of DNMT3A and TET1 with different dCas9 orthologs, nor a study demonstrating a simultaneous action of writers or erasers of epigenetic marks with opposite significance (activation versus silencing) within the same cells. Also, CRISPR/dCas9-based tools for simultaneous epigenetic editing have a potential to revolutionize the field of epigenetics by enabling researchers to determine causal relationship between directly manipulated individual epigenetic marks and gene transcription activity, as well as to help understand the complex chromatin layer involved in transcriptional regulation. To date, there is no clear picture of the causative relationship between a gene promoter methylation and transcriptional status of a certain gene, nor of the number and positions of CpGs in individual promoters that regulate transcription of the corresponding gene.

Based on our successfully constructed C-terminal dSpCas9-DNMT3A fusion (27), we aimed to design a backbone with exchangeable modules, thus creating a molecular toolbox that can be reused, extended with new functional modules and easily reconfigured for desired experimental setup (e.g. backbones for alternative delivery using lentiviruses etc.). Here, we describe the development and validation of a versatile platform for easy and convenient assembly of different promoters, dCas9 orthologs, effector domains (epigenetic modifiers or direct activators/silencers of gene transcription) and selection markers, along with specific gRNAs to target multiple candidate loci. We added the capability for assembling up to six gRNA-expression cassettes for both dSpCas9 and dSaCas9. Furthermore, we thoroughly characterized our new CRISPR/dCas9-based tools for epigenetic editing by examining activity profiles and time-course of their activity and demonstrating the capability of two different dCas9 fusions to perform antagonistic and synergistic activities within same cells at the same time. The candidate loci chosen for manipulation were those we study in the context of IgG glycosylation and its involvement in chronic inflammatory diseases (CID) (29). We demonstrated that dual epigenetic manipulation (i.e. DNA methylation and demethylation) of different pairs of individual loci within same cells resulted in changes in gene transcription level and consequently changes in the glycan phenotype, exemplifying the power of our molecular toolbox in testing the functional importance of modulating epigenetic states.

MATERIALS AND METHODS

Construction of the modular molecular toolbox by Golden Gate cloning

Details of the cloning steps, as well as the source plasmids, primers and oligonucleotides, plasmids for Cas9 fusions and the multi-guide system are given in the Supplementary Data. If applicable, plasmids are referenced by the number at the Addgene repository, where full sequence information is available. Assembly was done with 50 ng of the backbone plasmid pBackBone-BZ and each selected module vector, using molar ratio of 3:1 (module:backbone). Reaction was done in 1× CutSmart Buffer (New England Biolabs, Ipswich, MA, USA, NEB) with additional 1 mM ATP (NEB), 10 U of BsaI (NEB) and 350 U of T4 DNA Ligase (TaKaRa, Kusatsu, Shiga, Japan) with 30 cycles of 5 min at 37°C followed by 10 min at 16°C, after which ligase was inactivated at 50°C for 10 min. Next, BsaI was inactivated at 80°C for 10 min. Finally, 10 U of exonuclease V (RecBCD) (NEB) was added directly to the reaction along with additional 0.5 mM ATP and incubated at 37°C for 30 min to remove any remaining linear DNA. Following bacterial transformation and blue–white selection, white colonies were selected for further verification.

Construction of multi-guide system

We used two strategies for assembly of up to six gRNA modules, each carrying a different gRNA molecule into the final construct for epigenetic modulation. One is to first assemble all functional modules with BsaI (described above) and to use an empty pSg-x1-6 plasmid (depending on the number of gRNA modules to be assembled). Next, individual gRNA molecules are cloned into plasmids pSgMxA or pSgMxG (pSgM$\underline{x}$A represent plasmids for SaCas9 gRNA molecules while pSgM$\underline{x}$G represent plasmids for SpCas9 gRNA molecules, $\underline{x} = 1, 2, 3, 4, 5, 6$, depending on the order of gRNA modules to be assembled) as described (Supplementary Figure S3). Note that the pSgMxA or pSgMxG plasmids carry spectinomycin resistance, allowing counterselection with both the backbone (ampicillin) and module (kanamycin) vectors. Finally, Esp3I assembly of plasmids for cloning of individual gRNA molecules into the finished BsaI assembled construct is done in 1 × Tango Buffer (Thermo Fisher Scientific, Waltham, MA, USA) with additional 1 mM ATP, 1 mM dithiothreitol (DTT)

(Thermo Fisher Scientific), 10 U of restriction enzyme Esp3I (Thermo Fisher Scientific) and 350 U of T4 DNA Ligase with 10 cycles of 5 min at 37°C and 10 min at 16°C, followed with 15 min at 37°C and enzymes inactivation on 80°C for 10 min. The second strategy is to first assemble up to six gRNA 'individual' modules, each with inserted variable parts of gRNA molecule (as described, Supplementary Figure S4) into the empty pSg-x1-6 plasmid using the Esp3I restriction enzyme (described above). The resulting plasmid is a module vector containing multiple gRNA expression cassettes, which can be used as a standard gRNA module in the BsaI assembly.

Cell culture and transfections

Human embryonic kidney cell line HEK293 (ATCC, CRL-1573) was maintained in Dulbecco's Modified Eagle Medium (Sigma Aldrich, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (Sigma Aldrich), 4 mM L-glutamine (Sigma Aldrich), 100 U/ml penicillin and 100 μg/ml streptomycin (Sigma Aldrich). BG1, the ovarian cancer cell line, were cultivated in RPMI 1640 (Sigma-Aldrich, Missouri, USA) supplemented with 10% heat-inactivated fetal bovine serum (Sigma-Aldrich), 4 mM L-glutamine (Sigma-Aldrich), 100 U/ml penicillin and 100 μg/ml streptomycin (Sigma-Aldrich). All cells were incubated at 37°C in a humidified 5% CO_2-containing atmosphere.

Transfections of HEK293 cells were done using Lipofectamine 3000 Reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol, while PEI MAX 40K (Polysciences, Warrington, PA, USA) was used for transfection of BG1 cells. Briefly, HEK293 cells were seeded in 24- or 6-well plates day before transfection and transfected next day at around 80% of confluency with up to 1 μg of plasmid DNA. BG1 cells were seeded in 100-mm culture dishes and were transfected next day with 5 μg of individual plasmid DNA or total 10 μg of plasmid mixtures in antibiotic- and serum-free RPMI medium. The used ratio of PEI to DNA was 3:1. Cells were screened 24h post transfection for expression of fluorescent proteins (mClover3 and/or mRuby3), translated in the same reading frame as the dCas9 fusions and were selected with puromycin (Gibco Life Technologies, Grand Island, NY, USA) for 48 h. Cells were then collected at each time point depending on experiment (described below) for subsequent DNA, RNA and/or protein isolation.

Total DNA, RNA and protein isolation

In all experiments, DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) was used for total DNA isolation, while RNA was isolated using RNeasy Mini Kit (Qiagen). In experiments of simultaneous cytosine methylation and demethylation of two different gene loci in BG1 cells, total proteins were precipitated using methanol and chloroform standard protocol after cell disruption in lysis buffer (50 mM Tris pH 7.4, 0.1% Triton X, 1 mM EDTA, 135 mM sodium chloride), supplemented with Protease Inhibitor Cocktail (cOmplete™ ULTRA Tablets, EDTA-free, glass vials Protease Inhibitor Cocktail, Roche, Basel, Switzerland).

Bisulfite conversion and pyrosequencing

Bisulfite conversion of 500 ng of DNA was done with EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's protocol, after which the genomic regions of the *BACH2*, *MGAT3*, *IL6ST, LAMB1* and *HNF1A* genes were amplified using a PyroMark PCR kit (Qiagen). Amplified genomic regions were then sequenced using the PyroMark Q24 Advanced pyrosequencing system (Qiagen) and PyroMark Q24 Advanced CpG Reagents (Qiagen) to quantify methylation level at individual CpG dinucleotides. Assay sequences are listed in Supplementary Table S1, and assay maps are shown in Figures 4, 5 and 6. Sequences of PCR and pyrosequencing primers are listed in Supplementary Table S2.

Quantitative real-time PCR

Reverse transcription was done on 500 ng of total isolated RNA (prior treated with TURBO DNase (Invitrogen)) using the PrimeScript RTase (TaKaRa) and random hexamer primers (Invitrogen). RT-qPCR was performed according to the manufacturer's protocol using the 7500 Fast Real-Time PCR System, using TaqMan Gene Expression Master Mix and the following TaqMan Gene Expression Assays (Applied Biosystems, Foster City, CA, USA): Hs00167041_m1 (*HNF1A*), Hs02800695_m1 (*HPRT1*), Hs02379589_s1 (*MGAT3*), Hs00174360_m1 (*IL6ST1*) and Hs00222364_m1 (*BACH2*), according to manufacturer's protocol. For detection of the longer *MGAT3* transcript (targeting of TIS region), the Hs013692440_m1 TaqMan probe was used. Gene expression was normalized to the *HPRT1* gene and analyzed using the comparative C_t method (30). Expression was shown in all experiments as fold change (FC) relative to mock-transfected cells.

Time-course evaluation of C-terminal and N-terminal fusions of TET1 and DNMT3A catalytic domains to dSpCas9 and/or dSaCas9

Time-course evaluation of C-terminal fusion dSpCas9-TET1 was done using dSpCas9-TET1-PuroR plasmids encoding MGAT3-sg03 or LAMB1-sg01 (Supplementary Table S3). Transfections were done in biological triplicates, where control groups were cells transfected with either dSpCas9-TET1-PuroR_NCTRL (inactive fusion) co-expressing targeting gRNA or mock-transfected cells. Transfected cells were collected daily until the 8th day and later on 10th, 15th, 20th, 24th and 30th day for DNA isolation. In addition, time-course evaluation of N-terminal fusions DNMT3A-dSpCas9 and TET1-dSaCas9 was performed. DNMT3A-dSpCas9 construct was targeted to *IL6ST* using multi-guide system with four gRNA molecules, while TET1-dSaCas9 construct was targeted to *HNF1A* locus using multi-guide system with two gRNA molecules (Supplementary Table S3). For both experiments, cells were collected on 5th, 8th, 11th, 20th and 30th day after the transfection, while non-targeting gRNA plasmids (NT-gRNA; gRNA with no target sequence in the human genome) and mock-transfected cells were used as controls.

Analyses of activity profiles for C- and N-terminal fusions of DNMT3A and TET1 catalytic domains to dCas9 from *Streptococcus pyogenes* and *Staphylococcus aureus*

All experiments were done with plasmids co-expressing certain dCas9 fusion and chimeric gRNA in biological duplicates, where control groups were catalytically active constructs co-expressing non-targeting gRNA and mock-transfected cells. For obtaining the activity profile of dSpCas9-TET1 C-terminal fusion, the *MGAT3* promoter was targeted using eight gRNAs: MGAT3-sg01 to MGAT3-sg08 and *LAMB1* promoter was targeted using two gRNAs (Supplementary Table S3), while the activity profile of dSpCas9-DNMT3A C-terminal fusion was described previously (27). In experiments with N-terminal fusions of dCas9 (from either *Streptococcus pyogenes*, dSpCas9 or *Staphylococcus aureus*, dSaCas9) with DNMT3A or TET1 catalytic domains, specific gRNAs targeting the *BACH2* and *MGAT3* promoters were used. The DNMT3A-dSpCas9 was targeted with seven gRNAs and DNMT3A-dSaCas9 with six gRNAs to unmethylated region of the *BACH2* promoter. The TET1-dSpCas9 and TET1-dSaCas9 fusions were targeted with five gRNAs to highly methylated region of the *MGAT3* promoter (Supplementary Table S3).

Simultaneous manipulations of two different gene loci using dCas9-based tools with antagonistic functionalities in HEK293 and BG1 cells

In experiments of simultaneous methylation and demethylation in HEK293 cells, we targeted gene pairs *BACH2–HNF1A* and *IL6ST–MGAT3*. The *MGAT3* and *HNF1A* genes were targeted with TET1-dSaCas9 N-terminal fusions containing mClover3 as fluorescence marker, while *IL6ST* and *BACH2* were targeted with DNMT3A-dSpCas9 N-terminal fusions with mRuby3. All gene loci were targeted using the multi-guide module: the *MGAT3* was targeted with six gRNAs, *HNF1A* with two gRNAs, while *IL6ST* and *BACH2* were targeted each with four gRNAs (Figures 4 and 5; Supplementary Table S3). Cells were collected on 8th day following transfection for subsequent DNA and RNA isolation. We also performed time-course experiment for monitoring CpG methylation and gene expression with the same setup as described above and collected total DNA and RNA on 5th, 8th, 9th and 12th day after cell transfection (Supplementary Figure S6). In BG1 cells, we targeted *HNF1A–MGAT3* gene pair for simultaneous methylation and demethylation. The TET1-dSaCas9 along with mClover3 was targeted to highly methylated region of the *HNF1A* promoter using the multi-guide module containing two gRNAs, while the DNMT3A-dSpCas9 fusion along with mRuby3 was targeted to unmethylated TIS region of the *MGAT3* promoter with one gRNA (Figure 6A and Supplementary Table S3). BG1 cells were collected on 7th day following transfection for subsequent DNA, RNA and protein isolation. All dCas9 fusions had puromycin resistance, which was added using the dual-marker system (Figure 3C). Three types of negative controls were used in experiments: (i) plasmids with inactive DNMT3A and TET1 catalytic domains; (ii) plasmids with non-targeting gRNAs (NT-gRNA); and (iii) mock-transfected cells. Transfections were done in nine and three biological replicates for experiments performed on HEK293 and BG1 cells, respectively.

Analysis of total *N*-glycome in BG1 cells

The dried protein pellets were resuspended in 30 μl of 1.33% (wt./vol.) SDS (Invitrogen) and incubated at 65°C for 10 min. Subsequently, 10 μl of 4% Igepal-CA630 (Sigma Aldrich) and 1.2 U of PNGase F (Promega, Madison, WI, USA) in 10 μl 5 × PBS were added to each sample and incubated overnight at 37°C to allow release of N-glycans. The released N-glycans were labeled with procainamide (Sigma Aldrich), followed by hydrophilic interaction liquid chromatography solid-phase extraction (HILIC-SPE) clean-up. Fluorescently labeled and purified N-glycans were separated by HILIC on Waters Acuity ultra-performance liquid chromatography (UPLC) instrument (Waters, Milford, MA, USA) consisting of a quaternary solvent manager, sample manager and a fluorescence detector set with excitation and emission wavelengths of 310 and 370 nm, respectively. The instrument was under the control of Empower 2 software, build 2145 (Waters). Plasma N-glycans were separated on a Waters bridged ethylene hybrid (BEH) Glycan column, 150 × 2.1 mm, 1.7 μm BEH particles, with 100 mmol/l ammonium formate, pH 4.4, as solvent A and acetonitrile as solvent B. The separation method used a linear gradient of 70–53% acetonitrile (vol./vol.) at flow rate of 0.561 ml/min in a 25 min analytical run. Data were processed using an automated integration method and the chromatograms were all separated in the same manner into 31 peaks (GP1–GP31), where the content of glycans in each peak was expressed as a percentage of the total integrated area. All glycan structures were annotated with MS/MS analysis by HILIC-UPLC coupled with a Synapt G2-Si ESI-QTOF-MS system (Waters). The instrument was under the control of MassLynx v.4.1 software (Waters). MS conditions were set as follows: positive ion mode, capillary voltage 3 kV, sampling cone voltage 30 V, source temperature 120°C, desolvation temperature 350°C, desolvation gas flow 800 l/h. Mass spectra were recorded from 500 to 3000 m/z at a frequency of 1 Hz. MS/MS experiments were performed in a data-dependent acquisition (DAD) mode. Spectra were first acquired from 500 to 3000 m/z and then two precursors with the highest intensities were selected for CID fragmentation (m/z 100 to 3000 was recorded). A collision energy ramp was used for the fragmentation (LM CE Ramp Start 7 V, LM CE Ramp End 13 V, HM CE Ramp Start 97 V, HM CE Ramp End 108 V).

Time-course experiment using VPR-dSpCas9 and TET1-dSaCas9 in HEK293 cells

Highly methylated regulatory region of *HNF1A* was simultaneously targeted with TET1-dSaCas9 and the VPR-dSpCas9 N-terminal fusions in HEK293 cells. Two gRNAs using multi-guide system were used to target TET1-dSaCas9 (as in Figure 4B), while VPR-dSpCas9 was targeted with one gRNA (Supplementary Figure S8 and Supplementary Table S3). The VPR-dSpCas9 construct had mRuby3, while TET1-dSaCas9 had mClover3 as fluorescent marker and both constructs had puromycin resistance, which was added using the dual-marker system. One

transfection was done with the mixture of VPR-dSpCas9 and TET1-dSaCas9 fusions, while the other two transfections were done with individual VPR-dSpCas9 and TET1-dSaCas9 fusion to assess the contribution of each domain separately. Cells were collected on 5th, 8th, 11th, 20th and 30th day after the transfection for subsequent DNA, RNA and protein isolation. Transfections were done in seven biological replicates with NT-gRNA and mock as negative controls.

Time-course evaluation of the TET1-dSaCas9 presence in HEK293 cells using quantitative real-time PCR

To determine presence of TET1-dSaCas9 fusion in cells at level of DNA (presence of plasmid DNA), RNA (presence of mRNA transcript) and protein (dCas9), we performed qPCR analysis and western blotting. HEK293 cells were transfected with the TET1-dSaCas9 fusion and targeted to the *HNF1A* regulatory region using the multi-guide module containing two gRNAs. Cells were collected on 5th, 8th, 11th, 20th and 30th day after transfection. The amount of the TET1-dSaCas9 plasmid was determined relative to the genomic *RAG1* sequence (31), while the amount of mRNA transcript was normalized to the *GAPDH* gene expression using the comparative C_t method (30). Reactions were done using Power SYBR Green PCR Master Mix (Applied Biosystems), 50 nM of primers and 4 ng of total isolated DNA or 10 ng of cDNA. Reaction conditions were as following: 10 min at 95°C followed by 40 cycles at 95°C for 15 s and 60°C for 1 min (for detection of TET1-dSaCas9 plasmid) or 54°C for 1 min (for detection of TET1-dSaCas9 cDNA), followed by melt curve analysis of amplified targets. All primers used for real-time qPCR are listed in Supplementary Table S5.

Detection of dCas9 protein by western blotting

We also checked the presence of TET1-dSaCas9 protein at each time point: 5th, 8th, 11th, 20th and 30th day after HEK293 cell transfection. To prove that the weak EFS promoter, in fusions containing a 'secondary cassette' (for separate expression of antibiotic resistance marker under the stronger SV40 promoter, see Supplementary Figure S5), directly limits the amount of fusion proteins (DNMT3A-dSpCas9 and TET1-dSpCas9) compared to dCas9 fusions where the strong CBh promoter drives expression of both the fusion protein and the puromycin resistance, we compared the amount of the fusion proteins at 4th and 8th day after cell transfection. In both experiments, cells were lysed in RIPA buffer supplemented with Protease Inhibitor Cocktail (cOmplete™ ULTRA Tablets, Roche) and sonicated in Sonorex Super Ultrasonic bath (Bandelin, Berlin, Germany). Total protein concentration was determined with BCA Protein Assay Kit (Santa Cruz Biotechnology, Dallas, TX, USA) according to the manufacturer's protocol and 40 μg of protein extract was used for standard immunoblotting with β-tubulin as a loading control. Rabbit anti-SaCas9 1:5000 (ab203943, Abcam, Cambridge, UK), mouse anti-SpCas9 1:1000 (ab191468, Abcam), rabbit anti- β-Tubulin 1:2000 (ab6046, Abcam), goat anti-Rabbit IgG H&L (HRP) 1:200 000 (ab6721, Abcam) and goat Anti-Mouse IgG H&L (HRP) 1:200 000 (ab205719, Abcam) antibodies were used. The antibody-tagged protein bands were visualized by Immobilon Western Chemiluminescent HRP Substrate (Milipore, Burlington, MA, USA) according to the manufacturer's protocol. ECL signals captured on Alliance Q9 Advanced Chemiluminescence Imager (Clever Scientific, Warwickshire, UK) were quantified using the ImageJ software (32).

Whole-genome methylation analysis by Infinium MethylationEPIC (850K) array

To compare the effect of two types of dCas9 fusions (one with weak EFS promoter and another with strong CBh promoter, driving the expression of both dCas9 fusion protein and puromycin resistance gene) on global off-target activity, the whole-genome methylation analysis was performed. The DNMT3A-dSpCas9 was targeted to the *IL6ST* promoter using the multi-guide module containing four gRNAs and the TET1-dSpCas9 was targeted to *MGAT3* using the multi-guide module consisting of five gRNAs (Figure 8 and Supplementary Table S3). For each locus, two types of dCas9 fusion constructs were used: (i) the 'primary cassette' construct, whereby both the fusion protein and the selection marker (PuroR) were under the strong constitutive CBh promoter; (ii) the 'secondary cassette', whereby the dCas9 fusion construct was expressed under the weak EFS promoter, while maintaining efficient selection of transfected cells with PuroR under the strong SV40 promoter. Transfection of HEK293 cells was done in biological and technical duplicates and cells were harvested 8th day after transfection for DNA isolation. DNA (600 ng) was subjected to bisulfite conversion and then methylome profiling was done using the Infinium MethylationEPIC (850K) array. Of the array containing total of 866 091 probes, 793 038 probes remained for further analysis after removing cross-reactive probes, probes overlapping genetic variants at targeted CpG sites, probes with genetic variants overlapping the body of the probe and XY chromosome probes, as well as probes reporting missing values in more than 5% of all samples.

Statistical analysis and data representation

Changes in CpG methylation and expression levels, as well as changes in glycan structure ratios, were analyzed using the Mann–Whitney test. All data were visualized using the R Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, Austria). Images from western blotting experiments were processed and quantitated using the ImageJ program (32). Pre-processing and normalization of whole methylome data were done using the 'minfi' Bioconductor package (33). PCA analysis followed by Wilcoxon Rank Sum Tests was conducted to identify possible confounding variables. SVA was performed to correct for potential batch effects. We converted β-values converted to *M*-values, and a linear regression model with adjustments for Sentrix ID was used to identify differentially methylated probes (DMPs) using the Bioconductor 'limma' package (34). Statistically significant DMPs were defined as those with false discovery rate (FDR)-adjusted *P*-values of < 0.05, and delta-beta values of > 0.05 or < -0.05.

The 'IlluminaHumanMethylationEPICanno.ilm10b2.hg19' annotation package was used to map DMPs to CpG annotations, i.e. CpG islands, CpG shores, CpG shelves or open sea regions. The 'annotatr' Bioconductor package was used to map the DMP regions to chromatin states based on annotations defined by chromHMM in Hmec cells. Fold enrichment of each category based on relationship to CpG islands and regulatory regions were calculated by dividing the fraction of DMPs in each annotation category with the fraction of all filtered probes mapped to the respective category, correcting for differences in representation of each annotation category on the Infinium MethylationEPIC array.

RESULTS

Active fusions of DNMT3A and TET1 catalytic domains with the dCas9 orthologs

Based on our successfully constructed C-terminal dSpCas9-DNMT3A fusion (27), we aimed to create a modular toolbox with exchangeable dCas9 orthologs and fused various effector domains (DNMT3A, TET1, VPR and KRAB). To this end, we needed to test each configuration for activity in cell lines, i.e. verify that the fusions show the intended activity. First, we fused the TET1 catalytic domain to the C-terminus of dSpCas9 (analogous fusion to dSpCas9-DNMT3A; (27)) and targeted the *MGAT3* promoter with eight gRNAs and the *LAMB1* promoter with two gRNAs. The dSpCas9-TET1 fusion showed very similar activity profile to the analogous dSpCas9-DNMT3A fusion (Figure 1 and (27)). Then, we tested C-terminal fusions of the DNMT3A and TET1 catalytic domains to the dCas9 from *S. aureus*. Both C-terminal fusions revealed a poor activity in HEK293 transfected cells (data not shown). We speculated that it had been caused by steric hindrances or incorrect orientation of the catalytic domain, so we went on to test the effect of the tryptophan zipper motif (35) or the Inntag (36) as linkers, none of which restored catalytic activity (data not shown). Next, we tested whether a C-terminal nuclear localization signal (NLS) on dSaCas9 in the fusion protein was rendered inaccessible after appending an effector domain (DNMT3A) to the C-terminus, thus interfering with import into the nucleus and consequently lowering the measured activity of the effector domain. Indeed, the dSaCas9-DNMT3A fusion acquired activity when an additional nucleoplasmin NLS was appended to the C-terminus of the catalytic domain (Figure 2). Although the dSpCas9-DNMT3A fusion has already shown appreciable methylation activity (27), its activity was further enhanced after addition of another nucleoplasmin NLS on the C-terminus of the catalytic domain (Figure 2).

The fusion of DNMT3A with dCas9 from *S. aureus*, with an additional nucleoplasmin NLS added to the catalytic domain, still achieved only 2% to maximally 24% of methylation level increase, depending on CpG site (Figure 2). In comparison with fusion of DNMT3A with dCas9 from *S. pyogenes*, which achieved up to 60% increase of methylation level (27), the activity of the dSaCas9-DNMT3A fusion seemed unsatisfactory. Therefore, the next step was to fuse the catalytic domain to N-terminus of dCas9 in search for a universal approach compatible with all effector domains. Indeed, more robust activity of both DNMT3A and TET1 catalytic domains in N-terminal fusions with dCas9 from *S. aureus* (as well as with *S. pyogenes*) was achieved. Therefore, this configuration was selected for the modular approach. Comparing the two orthologs, the dSpCas9 N-terminal fusions consistently showed stronger activity than dSaCas9 N-terminal fusions (Figure 1). For clarity, in further text we designated C-terminal fusions as dSpCas9-DNMT3A and dSaCas9-TET1, and N-terminal fusions as DNMT3A-dSpCas9 and TET1-dSaCas9.

Activity profiles and time course of activity for the C- and N-terminal fusions of DNMT3A and TET1 catalytic domains to dCas9 from *S. pyogenes* and *S. aureus*

The individual profiles showing activity of N-terminal fusions of DNMT3A and TET1 catalytic domains with dCas9 from *S. pyogenes* and *S. aureus*, as well as C-terminal fusion of TET1 and DNMT3A catalytic domains to dCas9 from *S. pyogenes* are presented in Figure 1 along with the summary activity profile. The profiles show average activity of DNMT3A or TET1 (seen as the level of CpG methylation or demethylation, and named % of change) with respect to the gRNA-binding site. The activity profiles were created by collecting data from many experiments, where each data point represents absolute increase or decrease of CpG methylation percentage, positioned relative to the gRNA-binding site; finally, a smoothing curve was drawn to represent the summary activity profile. The activity profiles of all N- and C-terminal fusions were similar (Figure 1), except for C-terminal fusion of DNMT3A and TET1 to dCas9 from *S. aureus*, which showed very poor activity (data not shown). In addition, C-terminal fusions with dCas9 from *S. pyogenes* had their activity concentrated mainly downstream from the targeted sequence, while the N-terminal fusions had activity distributed equally up- and downstream from the gRNA-binding site. The activity profiles of all dCas9 fusions, along a linear DNA molecule, revealed a peak of methylation change about 30 bp downstream from the gRNA-binding site, with some additional activity peaks detected around 180–200 bp upstream or downstream from the gRNA-binding site, indicating possible involvement of adjacent nucleosomes. Similarity among all six profiles demonstrates that effector domains fused to either end of dCas9 contact DNA in a similar manner.

A time-course experiment for dSpCas9-TET1 revealed maximum of the demethylation change within 7–8 days after transfection, with the effect slowly declining afterwards, but remaining substantial at the main activity site (30 bp downstream from gRNA-binding site) even after 30 days (up to 19% for *MGAT3* and up to 30% for *LAMB1*); meanwhile, the distal site at 120 bp downstream from gRNA reverted to its initial methylation level after 20–24 days (Supplementary Figure S1). A time course for the N-terminal fusions of DNMT3A-dSpCas9 and TET1-dSaCas9 targeting the *IL6ST* and *HNF1A* genes, respectively, is given in Supplementary Figure S2. The maximal effect of DNMT3A-dSpCas9 and TET1-dSaCas9 fusions was detected 8th and 11th day following transfection, while the effect of methylation/demethylation remained altered

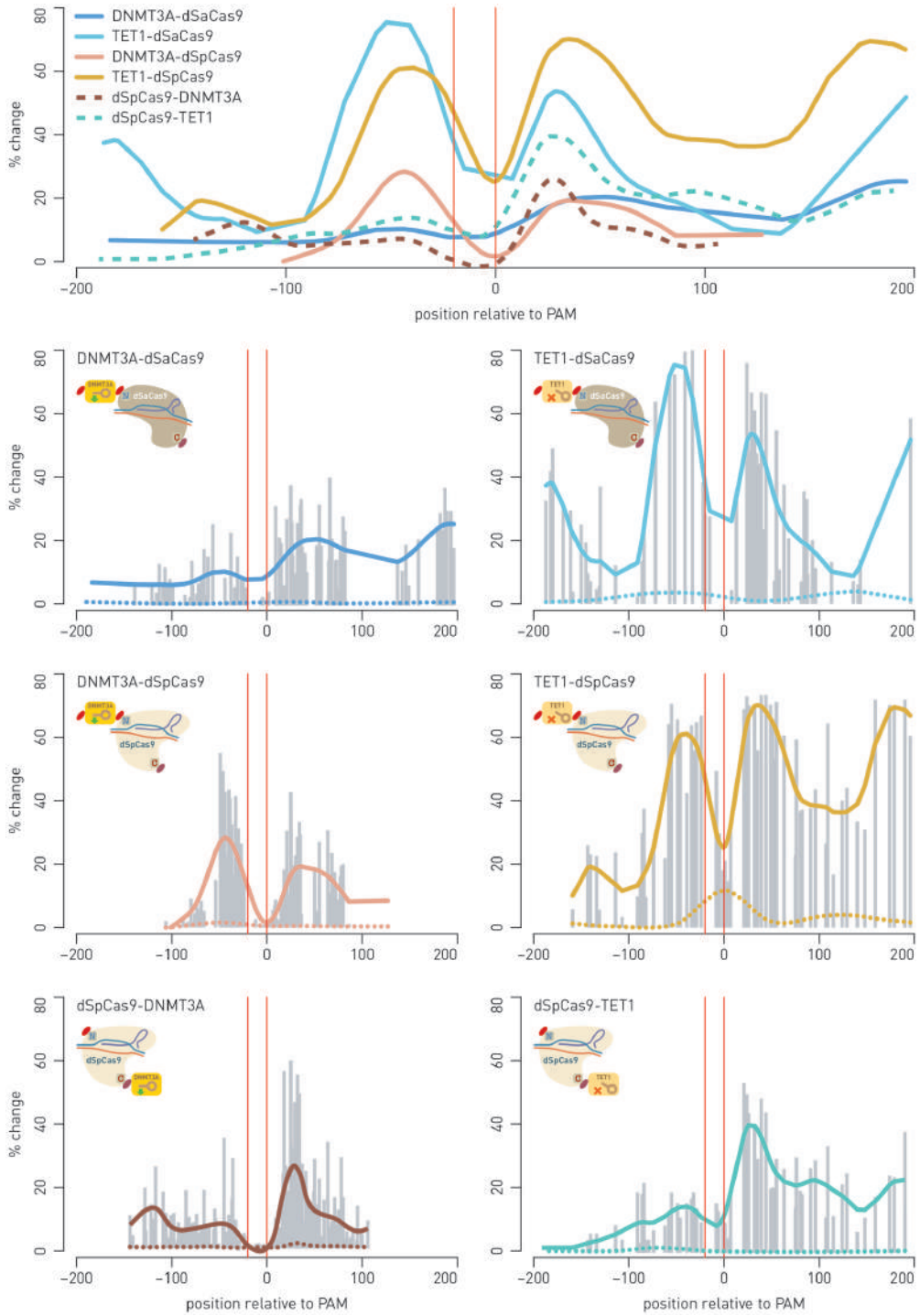


Figure 1. N- and C-terminal fusions of DNMT3A and TET1 catalytic domains to dCas9 orthologs from *Streptococcus pyogenes* and *Staphylococcus aureus* show similar activity profiles. Top panel shows summary of similar activity profiles for all combinations of effector domains (catalytic domains of DNMT3A or TET1) to C- or N-terminus of either dSpCas9 or dSaCas9 orthologs. The dSpCas9 C-terminal fusions show activity concentrated mainly downstream from the targeted sequence, while all N-terminal fusions had activity distributed equally up- and downstream from the gRNA binding site. The peak of change in CpG methylation appears at around 30 bp on either side of the dCas9-binding site (indicated by vertical red lines). The percentage of change is defined as absolute increase (for fusions with DNMT3A) or decrease (for fusions with TET1) in CpG methylation level (averaged from multiple independent experiments). Individual figures represent profiles of individual dCas9-fusions and summarize multiple experiments as a result of targeting certain loci using specific gRNA molecules: fusions DNMT3A-dSpCas9 and DNMT3A-dSaCas9 were targeted to *BACH2* using six gRNAs specific for dSpCas9, or five gRNAs specific for dSaCas9, and one dual gRNA for both orthologs; fusions TET1-dSpCas9 and TET1-dSaCas9 were targeted to *MGAT3* using four gRNAs specific for dSpCas9 and four gRNAs specific for dSaCas9, and one dual gRNA for both orthologs; fusion dSpCas9-TET1 was targeted to *MGAT3* using eight gRNAs, and fusion dSpCas9-DNMT3A was targeted to *BACH2* and *IL6ST* using eight gRNAs and four gRNAs, respectively (27). Gray bars show data points from individual experiments and represent methylation change at a certain distance from the gRNA binding site. Dotted lines in individual figures of the activity profiles show that catalytically inactive fusions elicit no changes of methylation level.

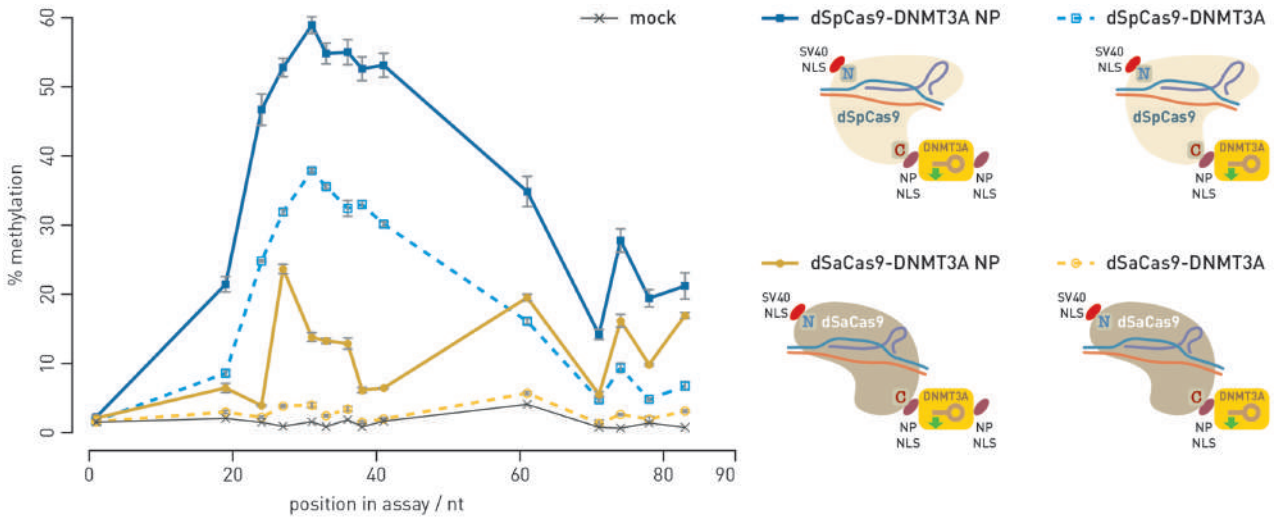


Figure 2. Nuclear localization signal (NLS) is critical for activity of an effector domain fused to C-terminus of dCas9. An additional nucleoplasmin NP-NLS appended to the C-terminus of the catalytic domain of DNMT3A clearly showed an increase in activity of the dCas9 fusion (solid lines) compared to activity obtained using the dCas9 fusions without additional NP-NLS (dotted lines). Each data point represents one of the 14 CpG sites analyzed for methylation level in the BACH2-A2 assay.

from 30% to 40% even 30 days following transfection of HEK293 cells.

Modular CRISPR/dCas9-based toolbox for epigenetic modulation and gene regulation

Based on the activity profiles of C- and N-terminal fusions of effector domains to orthologous dCas9 proteins, we have developed a modular CRISPR/dCas9-based system for rapid assembly of desired constructs based on Golden Gate cloning (Figure 3A). A minimal bacterial backbone carrying pUC19 origin of replication and ampicillin resistance received six or seven modules with functional units released using the BsaI enzyme in a single Golden Gate reaction (Supplementary Figure S3). Non-palindromic four nucleotide (nt) overhangs linking functional units ensured accurate assembly, while blue–white selection enabled rapid identification of correctly assembled clones. The overnight reaction had nearly 100% efficiency, while a shortened reaction time (2 h) assembly was sufficient when used in conjunction with blue–white selection. The final constructs comprised a gRNA unit (either a single gRNA under the U6 promoter or a multi-guide system module), the eukaryotic promoter (CBh or EFS), an effector domain fused N-terminally to dCas9 ortholog of choice, a marker (or the dual marker system) linked using a self-cleaving 2A peptide and a eukaryotic transcriptional terminator.

For targeting, a 20 nt long gRNA fragment complementary to the genomic target was cloned into an appropriate gRNA module containing a U6 promoter and gRNA scaffold for either *S. pyogenes* or *S. aureus* using the BpiI Golden Gate enzyme and red–white selection for rapid identification of positive clones (Supplementary Figure S4). A gRNA could be cloned into the module either before or after assembly. For inclusion of up to six gRNA expression cassettes (comprising U6 promoter, cloning site, gRNA scaffold and U6 terminator) in a single construct, the multi-guide system was developed (Figure 3B). A special multi-gRNA module allowing for red–white screening and insertion of a fixed number of gRNA cassettes (1–6) can be included into the assembly. Target specificity was given to individual gRNA modules for positions 1–6 by inserting the 20 nt complementary sequence using the same BpiI strategy as for single gRNAs. Individual gRNA-preloaded modules were linked together in a secondary Golden Gate assembly with the Esp3I enzyme. The Esp3I assembly step was successfully used both for creation of a multi-guide module and for direct cloning into the fusion construct following a BsaI assembly of functional units. Both *S. pyogenes* and *S. aureus* scaffolds have independent modules for all six positions; they can be mixed and matched in a single multi-guide assembly. Each gRNA is specific for its cognate ortholog and can guide dCas9 when expressed *in trans*. For concurrent antibiotic selection and fluorescent protein (FP) tracking, we developed a dual marker system (Figure 3C). In addition to individual marker modules for antibiotic resistance (PuroR) and fluorescent proteins mClover3 (a GFP variant) and mRuby3 (an RFP variant), we added another optional non-palindromic 4 nt overhang for joining a T2A-linked antibiotic resistance gene to a P2A-linked fluorescent protein. The two 2A self-cleaving peptides enabled co-expression of both markers with the selected dCas9 fusion (Figure 3C). Although the system was initially designed for construction of a special dual marker module, we found the Golden Gate cloning approach to be sufficiently robust for direct assembly of all seven modules (instead of six 'standard' modules) in a single BsaI reaction.

Simultaneous epigenetic manipulation of two genes using dCas9 fusions with antagonistic activities in HEK293 cells

To demonstrate that two different dCas9 fusions with antagonistic activities can operate simultaneously within the same cell, we chose candidate genes in the context

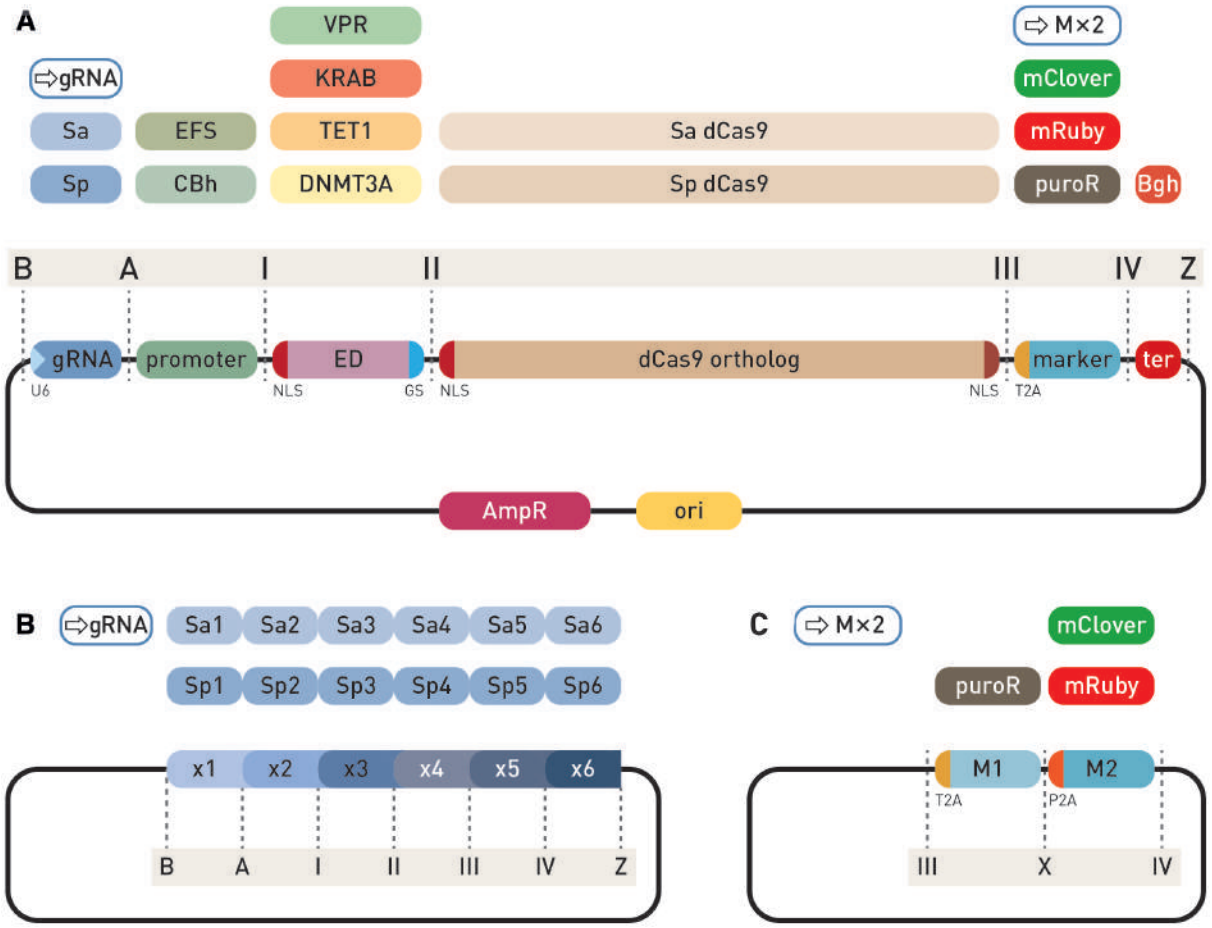


Figure 3. Modular system for CRISPR/dCas9-based epigenetic editing and direct gene regulation. (**A**) Individual modules of an expression cassette for N-terminal dCas9 fusions are assembled into the backbone vector using Golden Gate cloning with the BsaI type IIS restriction enzyme. The backbone plasmid has ends compatible with 'B' and 'Z' type module ends. The first position ('B' to 'A') receives a gRNA expression module with SaCas9 or SpCas9 scaffold, either containing a pre-inserted variable gRNA region, an empty module for gRNA cloning with red-white selection or a multi-guide module for a second step insertion of up to six gRNAs. Next position ('A' to 'I') is for insertion of a eukaryotic promoter, followed by the effector domain ('I' to 'II') containing an N-terminal NLS and a short G_4S linker to a dCas9 ortholog ('II' to 'III'), followed by a selection marker (fluorescence or antibiotic resistance, 'III' to 'IV') linked via the self-cleaving T2A peptide, which can be substituted with the dual-marker system. Finally, a module for eukaryotic transcription terminator is inserted between ends 'IV' and 'Z'. (**B**) The multi-guide system accepts up to six gRNA modules for either dSpCas9 or dSaCas9 protein. Individual modules require the variable part of gRNA pre-cloned for the second step of assembly, facilitated by the type IIS restriction enzyme Esp3I and red–white selection. (**C**) The dual marker system enables addition of both antibiotic resistance and a fluorescent protein. The dual modules have T2A and P2A self-cleaving peptides for equimolar expression with dCas9.

of studies of IgG glycosylation and chronic inflammatory diseases (29): *BACH2*, the key transcription factor important for B-lymphocyte maturation and differentiation (37); *IL6ST*, signal transducer for many cytokines; *MGAT3*, a mannosyl(β-1,4-)-glycoprotein β-1,4-N-acetylglucosaminyltransferase, responsible for addition of bisecting GlcNAc to the three-mannose core of *N*-glycan structures and involved in inflammation (38); and *HNF1A*, a transcription factor, a master regulator of protein fucosylation (39). We used our new CRISPR/dCas9 modular toolbox to manipulate gene pairs *BACH2–HNF1A* and *IL6ST–MGAT3*, where we targeted *BACH2* and *IL6ST* with DNMT3A-dSpCas9 for hypermethylation, and *HNF1A* and *MGAT3* with TET1-dSaCas9 for demethylation. The candidate loci were either almost completely unmethylated (*BACH2*, *IL6ST*) or highly methylated (*HNF1A*, *MGAT3*) in untransfected HEK293 cells. By using two fluorescent markers, we confirmed that the bulk of the cells were co-transfected with both dCas9 fusions. Of the 257 cells analyzed, 59.1% were positive for both fluorescent signals, while 30% were positive only for red mRuby3 and 10.9% only for green mClover3 fluorescence signal. Since fluorescent signals originated from the same transcripts that directed synthesis of dCas9 orthologues with effector domains (fluorescent proteins were linked by 2A-type self-cleaving peptides in frame with dCas9), the yellow fluorescent signal indicated that both dCas9 fusions entered the same cell. In addition, this showed that the bulk of antagonistic activities (methylation versus demethylation) co-occurred in doubly transfected cells and that there is no interference of the two different effector domains that were directed by dSpCas9 and dSaCas9 orthologs and their cognate gRNAs (Figures 4 and 5). Changes in methylation level ranged from 20% to 80% compared to mock transfected cells, in the opposite

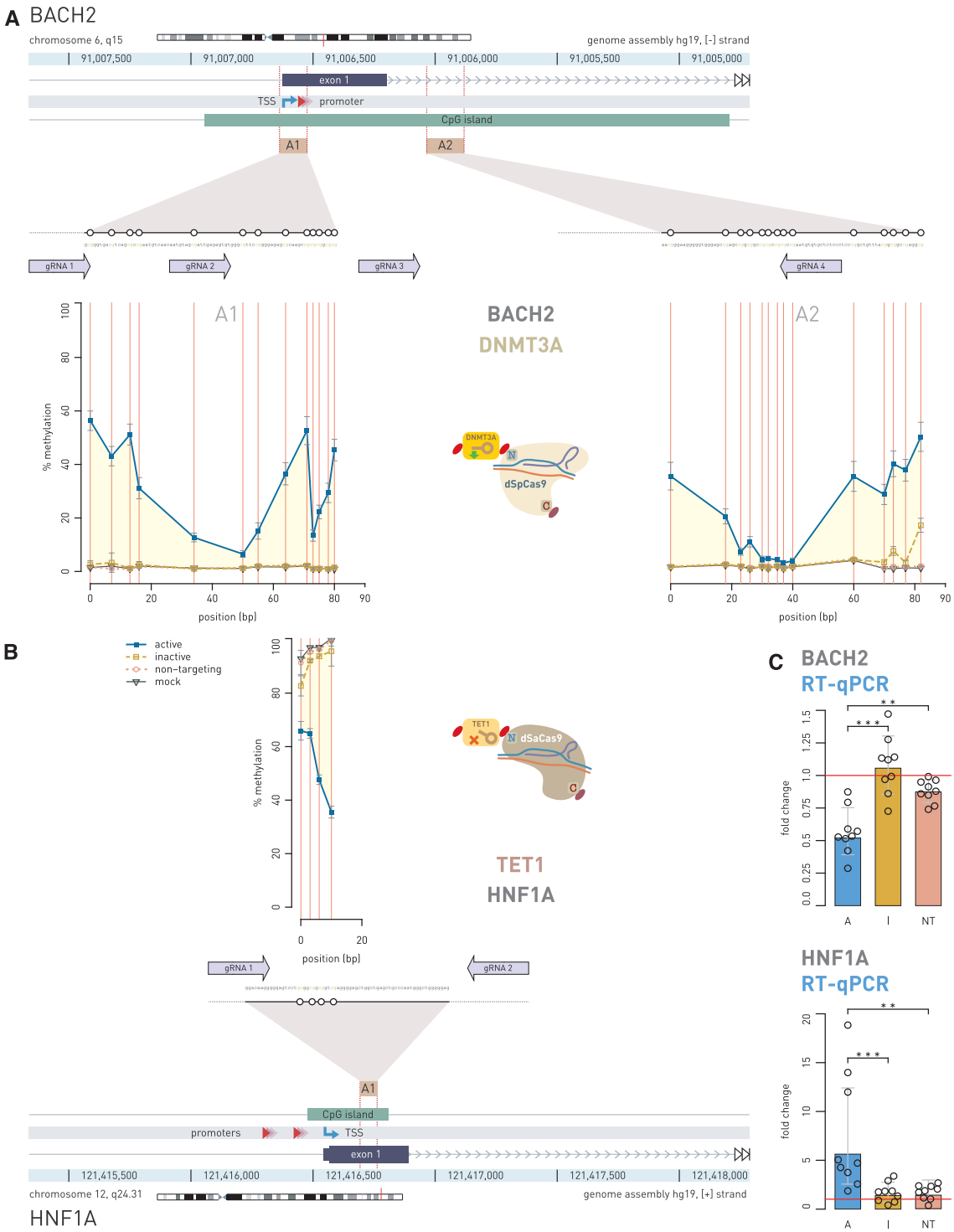


Figure 4. Simultaneous epigenetic manipulation of the *BACH2* and *HNF1A* loci using DNMT3A-dSpCas9 and TET1-dSaCas9 in HEK293 cell line. (**A**) Twenty seven CpG sites in CpG island of the *BACH2* promoter which was targeted with DNMT3A-dSpCas9 guided by co-expression of four different gRNA molecules using the multi-guide system, led to increase of methylation level ranging from 2% to 55% depending on the CpG site. (**B**) At the same time in the same cells, two gRNAs targeted TET1-dSaCas9 to four CpG sites located in the first exon of *HNF1A*, which led to decrease in methylation level of 27%, 32%, 49% and 65% for CpG1, CpG2, CpG3 and CpG4, respectively. (**C**) Induced changes in CpG methylation level were followed by changes in gene transcript level of both genes: *HNF1A* showed 5.614-fold change ($P = 0.0005$ to inactive TET1) and *BACH2* showed 0.568-fold change ($P = 0.0003$ to inactive DNMT3A). The human *BACH2* locus (**A**) and *HNF1A* locus (**B**) are shown with positions of the pyrosequencing assays (A1, A2) indicated by yellow rectangles. Magnified insets show individual CpG sites (white circles) targeted by dCas9 fusions and subsequently analyzed for methylation level. Arrows aligned to the magnified regions indicate 20 bp binding sites of gRNAs used to guide the dCas9 fusions. Arrows point toward the PAM sequence. Annotated promoter is represented by shaded red triangles; TSS, transcription start site (kinked arrow); A, dCas9 fusion with active catalytic domain; I, dCas9 fusions with inactive catalytic domain; NT, active dCas9 fusion with non-targeting gRNA. ** $P < 0.01$, *** $P < 0.001$.

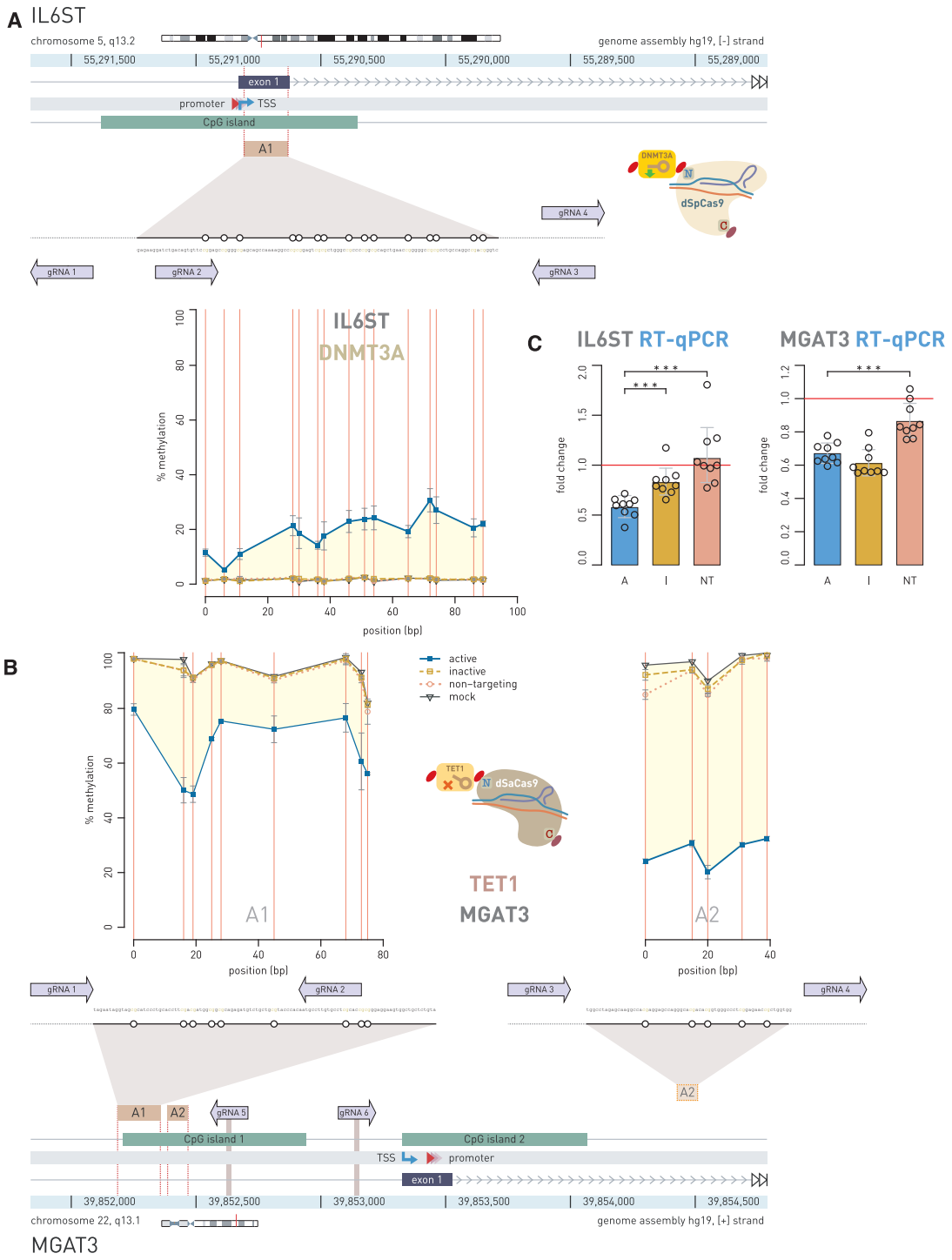


Figure 5. Simultaneous epigenetic manipulation of the *IL6ST* and *MGAT3* loci using DNMT3A-dSpCas9 and TET1-dSaCas9 in HEK293 cell line. (**A**) Fifteen CpG sites within CpG island of the *IL6ST* promoter were targeted with DNMT3A-dSpCas9, guided by the multi-guide system including four gRNAs, while 14 CpG sites located in CpG island 1 of the *MGAT3* promoter were targeted with TET1-dSaCas9, guided by the multi-guide system consisting of six gRNAs. (**B**) Doubly transfected cells with the two dCas9 fusions showed concurrent increase and decrease of methylation level at the targeted CpG sites: 3–29% methylation change for *IL6ST*, and 19–72% methylation change for *MGAT3*, depending on the CpG site. (**C**) While change of methylation level in *IL6ST* promoter was followed with 0.583-fold change in transcript level ($P = 0.0002$ to inactive DNMT3A), the hypomethylation of 14 CpG sites in CpG island 1 of *MGAT3* induced no increase of gene transcription. The human *IL6ST* locus (A) and *MGAT3* locus (B) are shown with positions of the pyrosequencing assays (A1, A2) indicated by yellow rectangles. Magnified insets show individual CpG sites (white circles) targeted by dCas9 fusions and subsequently analyzed for methylation level. Arrows, aligned to the magnified regions, indicate 20 bp binding sites of gRNAs used to guide the dCas9-fusions. Arrows point toward the PAM sequence. Annotated promoter is represented by shaded red triangles; TSS, transcription start site (kinked arrow); A, dCas9 fusions with active catalytic domain; I, dCas9 fusions with inactive catalytic domain; NT, active dCas9 fusions with non-targeting gRNA. *** $P < 0.001$.

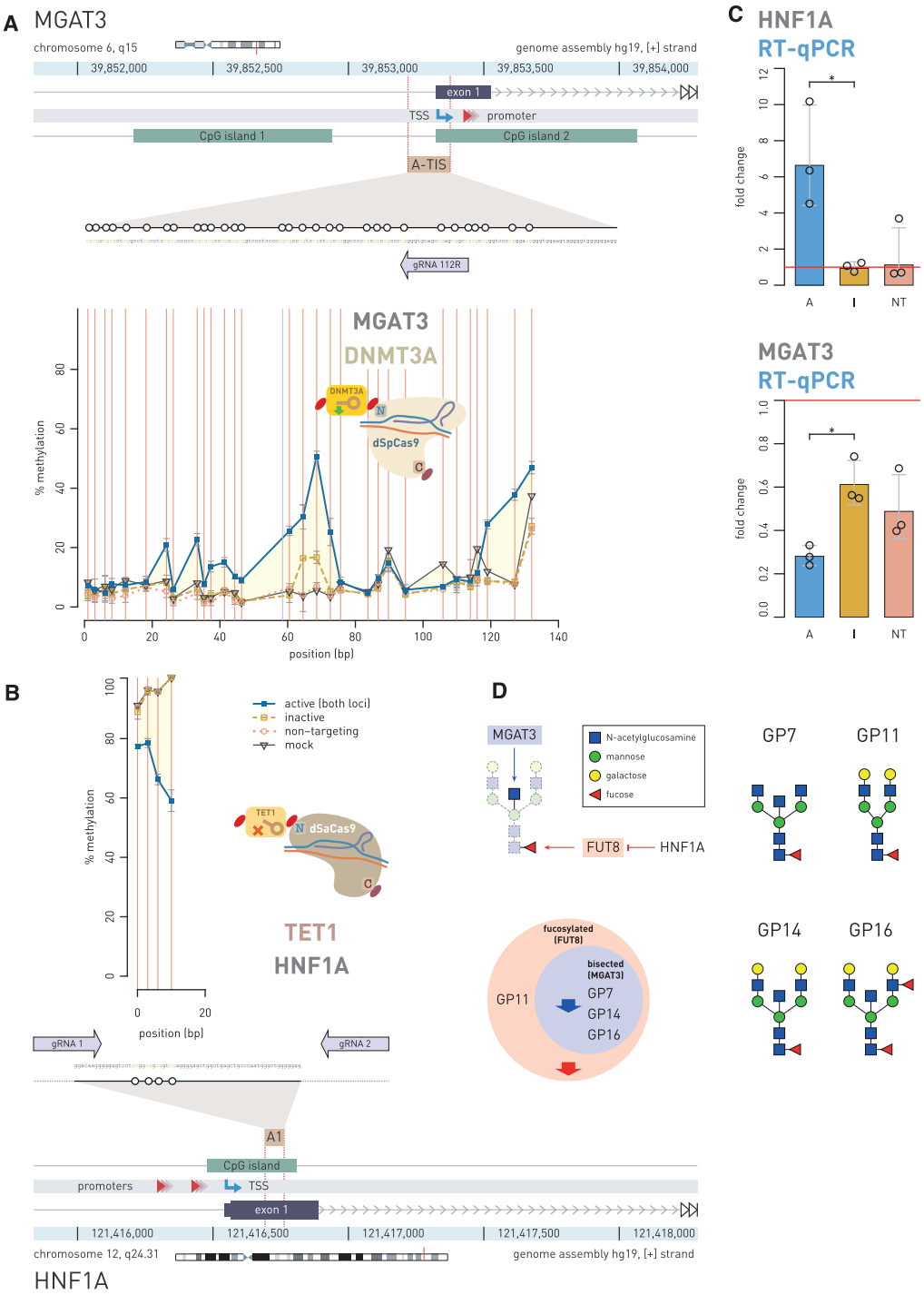


Figure 6. Simultaneous epigenetic manipulation of the *MGAT3* and *HNF1A* loci using DNMT3A-dSpCas9 and TET1-dSaCas9 showed an effect on total N-glycome in BG1 cells. (**A**) Thirty CpG sites within 200 bp wide putative regulatory region, residing in CpG island 2 of the *MGAT3* promoter and centered about 100 bp upstream from the TSS, were targeted with DNMT3A-dSpCas9 using a single gRNA. (**B**) At the same time, four regulatory CpG sites in the first exon of the *HNF1A* promoter was targeted with TET1-dSaCas9 using two gRNAs. (**C**) Induced hyper- and hypo-methylation resulted in significant change in gene transcript level: *MGAT3* showed 0.28-fold change ($P = 0.004$) and *HNF1A* showed 6.64-fold change ($P = 0.004$). A, dCas9 fusions with active catalytic domain; I, dCas9 fusions with inactive catalytic domain; NT, active dCas9 fusions with non-targeting gRNA. (**D**) The *MGAT3* gene downregulation led to a statistically significant decrease ($P = 0.005$ compared to inactive DNMT3A-dSpCas9; $P = 0.032$ compared to mock-transfected cells) of the glycan structures with bisecting GlcNAc (GP7, GP14 and GP16). The *HNF1A* gene upregulation resulted in decrease ($P = 0.005$ compared to inactive TET1-dSaCas9, $P = 0.028$ to mock-transfected cells) of core fucose in the glycan structures with and without bisecting *N*-GlcNAc, probably through a known effect of *HNF1A* on the fucosyltransferase FUT8. Effects of *HNF1A* on fucosylation were represented by the sum of abundance of all complex core-fucosylated structures that are products of FUT8: GP7, GP11, GP14, GP16. Analogously, the effect of GNT-III enzyme, encoded by *MGAT3*, was represented by the sum of abundance of all structures containing bisecting *N*-GlcNAc: GP7, GP14, GP16. Thick blue and red arrows on the Venn diagram show the direction of the glycan change (a decrease of certain glycan structures in the total *N*-glycome of BG1 cells). * $P < 0.05$.

direction for each locus within a pair (methylation versus demethylation). The level of methylation change was specific for each CpG site, but the methylation profiles for all analyzed CpG sites were consistent across experiments. We were interested if imposed CpG methylation and demethylation have an effect on transcript level of the targeted genes. Therefore, gene activity was always measured 8th day following transfection, based on the results of the time-course experiment, which had showed that gene expression profiles (fold change) closely followed CpG methylation profiles (% methylation change) for both dCas9-fusion with DNMT3A (27) and TET1 (Supplementary Figure S6).

The most pronounced effect of altered CpG methylation on RNA transcripts level was observed for the *HNF1A* gene where we targeted four CpG dinucleotides (Figure 4B) for which we have had previous indication to had putative regulatory role from a correlation study (40). The TET1-dSaCas9 fusion decreased methylation level 27%, 32%, 49% and 65% at CpG1, CpG2, CpG3 and CpG4, respectively, which was associated with 5.614-fold change ($P = 0.0005$ to inactive control, I; $P = 0.0012$ to non-targeting control, NT) of RNA transcript level (Figure 4C). Its gene partner in pair, *BACH2*, was at the same time targeted with DNMT3A-dSpCas9 using a multi-guide module comprising four different gRNAs (Figure 4A); randomly chosen 27 CpG sites located within CpG island (some of them positioned near transcription start site, TSS) were targeted. The methylation level change varied from 2–5% to maximally 55%, depending on the CpG site. However, despite the wider region of the *BACH2* promoter was methylated using DNMT3A-dSpCas9 the associated fold change of the gene transcription was modest (0.568-fold change, $P = 0.0003$ to I; $P = 0.0012$ to NT). In case of the *MGAT3–IL6ST* gene pair, randomly chosen CpG sites positioned within CpG islands of these gene promoters were manipulated using the multi-guide system containing four gRNAs for *IL6ST* targeting and six gRNAs for *MGAT3* targeting (Figure 5). The *IL6ST* region of choice, comprising 15 randomly chosen CpG sites, was targeted using DNMT3A-dSpCas9 and methylation change varied from 3% to 29% depending on an individual CpG site. The TET1-dSaCas9 fusion was used in the same experiment for targeting 14 randomly chosen CpG sites dispersed within CpG island 1 of the *MGAT3* gene. Corresponding decrease in methylation level varied between 19% and 72% depending on an individual CpG site (Figure 5). The externally introduced hypermethylation in the *IL6ST* promoter was followed by modest decrease (0.583-fold change; $P = 0.0002$ to I, $P < 0.0001$ to NT) of transcriptional activity, while induced hypomethylation of *MGAT3* did not associate with increased gene transcript level (0.671-fold change; $P = 0.06$ to I). Furthermore, construct with an inactive TET1 catalytic domain (negative control) also decreased *MGAT3* gene expression level, possibly due to CRISPR interference.

Simultaneous epigenetic manipulation of the *MGAT3* and *HNF1A* genes using DNMT3A-dSpCas9 and TET1-dSaCas9 induced changes in the *N*-glycome of BG1 cells

The regulatory impact of DNA methylation on *MGAT3* gene expression was previously shown for several ovarian cancer cell lines, including the BG1 cell line. Furthermore, *MGAT3* gene expression in these cell lines correlates with the presence of N-glycan structures with bisecting GlcNAc (41). Because we had previous knowledge about putative regulatory CpG sites within the *MGAT3* promoter and the effect of DNA methylation and expression of this gene on the glycan phenotype of BG1 cells, we went for epigenetic manipulation of the *MGAT3* promoter in this cell line. Since the TSS region of the *MGAT3* promoter is unmethylated and the gene is expressed in BG1 cells, we targeted 30 CpG sites, located within 200 bp wide putative regulatory region (41) in CpG island 2, and centered about 100 bp upstream from the TSS, with DNMT3A-dSpCas9 using a single gRNA (Figure 6A). At the same time, we targeted four regulatory CpG sites in the *HNF1A* promoter with TET1-dSaCas9 using two gRNAs, because we had confirmed in HEK293 cells that epigenetic manipulation of these CpGs affected gene transcription. We checked the efficacy of double transfection by detection of fluorescent markers—out of 173 cells analyzed, 84.4% were positive for yellow color (showing of co-localization of two fluorescent proteins), while 8.1% were positive only for red mRuby3 and 7.5% only for green mClover3 fluorescence signal. Altered methylation level led to changes in expression, with *HNF1A* upregulated 6.64-fold and *MGAT3* downregulated to 0.28-fold expression level of mock-transfected cells (Figure 6C). To investigate the effect of simultaneous *MGAT3* downregulation and *HNF1A* upregulation on the glycan phenotype, we analyzed total *N*-glycome of BG1 cells (Supplementary Figure S7 and Supplementary Table S6). We detected changes in glycan structures with core fucose and bisecting GlcNAc (Figure 6D). Direct effect of downregulated GNT-III enzyme (encoded by *MGAT3*), responsible for the addition of *N*-acetylglucosamine to the three-mannose core of glycan structures, led to a statistically significant decrease ($P = 0.0050$ compared to inactive DNMT3A-dSpCas9; $P = 0.0324$ compared to mock-transfected cells) in structures with bisecting *N*-GlcNAc (glycan peaks GP7, GP14 and GP16; Figure 6D, and Supplementary Figure S7). Upregulation of *HNF1A*, probably via its known regulatory role on the FUT8 enzyme (39), led to a statistically significant decrease in complex glycan structures with core-fucose (GP7, GP11, GP14 and GP16; $P = 0.0053$ compared to inactive TET1-dSaCas9, $P = 0.0281$ compared to mock-transfected cells) where GP11 is a fucosylated glycan structure without a bisecting GlcNAc (Figure 6D and Supplementary Figure S7). The decrease in GP11 alone was statistically significant ($P = 0.0122$ compared to inactive TET1-dSaCas9, $P = 0.0161$ compared to mock-transfected cells).

Time course of simultaneous manipulation of the *HNF1A* regulatory region using the VPR-dSpCas9 and TET1-dSaCas9 fusions

Results of the previous experiments of targeted demethylation of four CpG sites, located in the first exon near TSS in *HNF1A* gene, using the TET1-dSaCas9 fusion in HEK293 and BG1 cells, confirmed that these CpGs were regulatory, since the gene transcript level significantly changed following the epigenetic manipulation. Next, we compared the effects of using TET1-dSaCas9, VPR-dSpCas9 (for direct

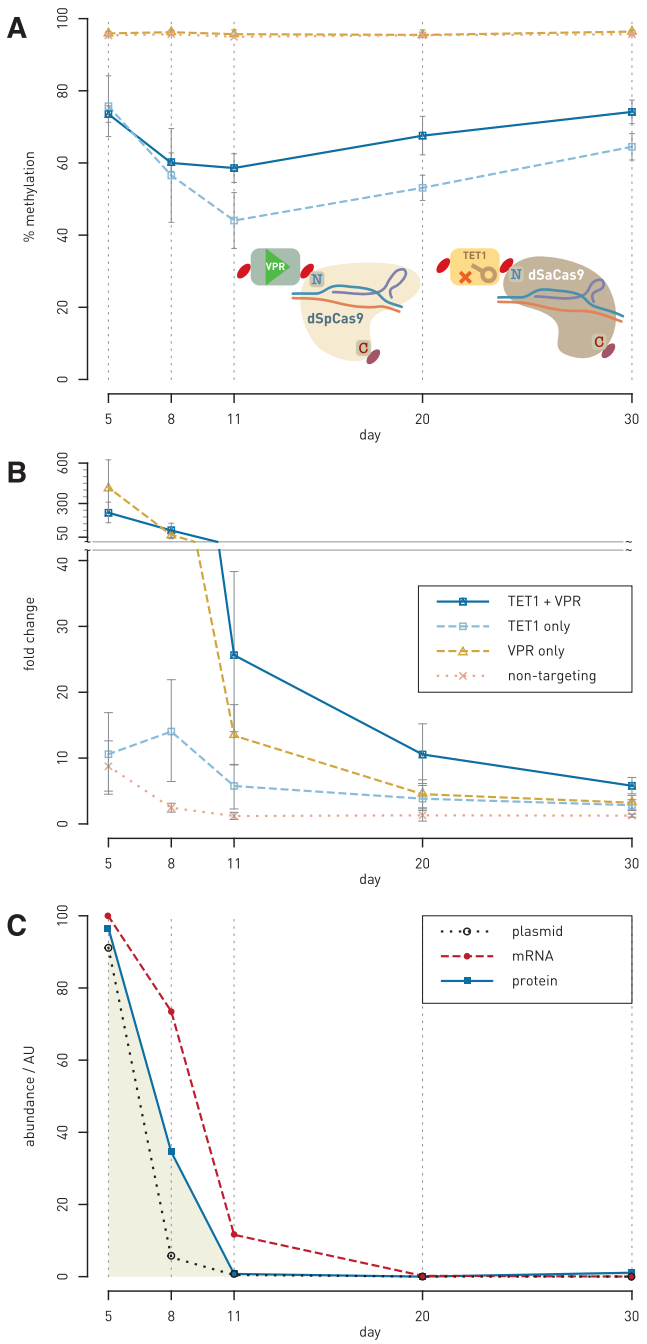


Figure 7. Time-course experiment of simultaneous manipulation of the *HNF1A* regulatory region using VPR-dSpCas9 and TET1-dSaCas9. (**A** and **B**) The *HNF1A* regulatory region was targeted with VPR-dSpCas9 for gene activation and TET1-dSaCas9 for targeted CpG demethylation, either using individual dCas9-fusions or both dCas9 fusions at the same time in HEK293 cells. Methylation level at four regulatory CpG sites (**A**) as well as gene transcript level (**B**) were monitored for 30 days following transfection. Synergistic effect of the VPR-dSpCas9 and TET1-dSaCas9 fusions on gene transcription level was achieved already on 11th day following transfection: the increase in gene transcript level was 26 ± 12-fold change when compared to 13 ± 4-fold change using only VPR-dSpCas9 ($P = 0.017$). The synergistic effect persisted until 20th and 30th day of gene expression monitoring ($P = 0.004$; $P = 0.007$). Percentage of methylation was shown as average methylation level on four regulatory CpG sites. (**C**) Time course of the level of plasmid encoding TET1-dSaCas9, transcribed mRNA and the expressed dCas9 fusion protein in HEK293 cells after transfection. Plasmid and RNA were quantitated by RT-qPCR, with their abundance calculated from delta-C_t relative to internal controls (*RAG1* for plasmid DNA and *GAPDH* for transcript level). Level of the fusion protein TET1-dSaCas9 was determined by quantitation of signals obtained by western blotting (pictures of membranes are given as Supplementary Figure S9A). Abundance of the TET1-dSaCas9 was normalized to β-tubulin for each lane. No protein could be detected by western blotting at day 11 and later.

gene activation) and simultaneous transfection of both fusions in modulating *HNF1A* transcriptional expression. We conducted single and double transfections in HEK293 cells using these two dCas9 fusions and monitored CpG methylation and gene transcript levels over a period of 30 days. If we observe the effect of the dCas9-fusions used either together or individually on CpG methylation only, the TET1-dSaCas9 fusion used alone induced a greater decrease of methylation level than if the fusions were used together (Figure 7A), and this could be explained by possible interference between the two effector domains (i.e. VPR and TET1). In terms of *HNF1A* transcriptional expression, direct activation by VPR-dSpCas9 alone induced high level of expression on the 5th day following transfection (411 ± 209-fold change). This effect however was reduced by the 8th day (62 ± 35-fold change) post-transfection. However, simultaneous targeting of the VPR-dSpCas9 and TET1-dSaCas9 fusions to the four CpG sites induced persistent upregulation of *HNF1A*, exceeding the effect of VPR only from 11th day post-transfection, and persisting until the 30th day following transfection (Figure 7B). On the 11th, 20th and 30th days, respectively, the VPR-dSpCas9 and TET1-dSaCas9 fusions together induced 26 ± 12, 11 ± 5 and 6 ± 1-fold change, while the VPR-dSpCas9 alone induced 13 ± 4, 5 ± 2 and 3 ± 1-fold change ($P = 0.017$, $P = 0.004$ and $P = 0.007$, respectively). Furthermore, to determine whether the stability of demethylation through time is due to the inheritance of induced epigenetic change or due to persistent expression and activity of the fusion proteins, we checked for the presence of TET1-dSaCas9 at DNA (plasmid DNA), RNA (transcript) and protein levels. The amount of TET1-dSaCas9 at DNA and RNA level decreased sharply on 11th day after the transfection while the presence of TET1-dSaCas9 protein was undetectable from 11th day after transfection (Figure 7C and Supplementary Figure S9A).

Whole-genome methylation analysis reveals reduced off-target effect when DNMT3A- and TET1-dCas9 fusions are expressed under a weak promoter

Recent reports indicate that exogenous expression of dCas9 fusions with different DNA methyltransferases often induces extensive off-target DNA methylation (for a review, see Tadić *et al.* (42) and references therein), underscoring the need for further improvements. Therefore, we aimed to assess the specificity of our modular dCas9-based system and to investigate the effect of modulating fusion protein expression relative to selection marker expression on off-target dCas9 activity. The DNMT3A-dSpCas9 fusions were targeted to *IL6ST*, and TET1-dSpCas9 fusions were targeted to *MGAT3*. Two versions of each DNMT3A-

dSpCas9 and TET1-dSpCas9 fusions were designed: (i) a version in which the effector proteins (dCas9 in fusion with either DNMT3A or TET1) and the puromycin-resistance selection marker were driven under the same strong, constitutive CBh promoter (denoted as the 'primary cassette') and (ii) a version in which the effector proteins were expressed under the weak EFS promoter, with independent SV40-driven expression of puromycin resistance selection marker (denoted as the 'secondary cassette'). Stronger expression of dCas9 fusions in the 'primary cassette' compared to much weaker expression in the 'secondary cassette' was confirmed by western blotting (Supplementary Figure S9B).

To assess DNA methylome-wide on- and off-target effects, each transfection condition was analyzed using Illumina Infinium MethylationEPIC BeadChip arrays (see Supplementary Figure S10, Supplementary Tables S7 and S8 for quality control). We observed that expressing dSpCas9-fused effector proteins under the 'secondary cassette' reduced off-target DNA methylation without a significant reduction in on-target effects. The introduction of the DNMT3A-dSpCas9 expressed in the 'primary cassette' resulted in 71 910 ± 2367 differentially methylated probes (DMPs) (with the cut-off for significance set as FDR-adjusted P-values < 0.05 and effect size > 5%), accounting for 9.1% ± 0.3 of all filtered probes (Figure 8A and G). The DNMT3A-dSpCas9 fusion expressed in the 'secondary cassette' reduced the average number of DMPs to 41 894 ± 3939 probes, accounting for 5.3% ± 0.5 of all filtered probes, though the difference between the 'primary' and 'secondary cassettes' was not significant when tested by Student–Newman–Keuls (SNK) post-ANOVA multiple comparison tests. The TET1-dSpCas9 fusion expressed under the 'primary cassette' induced differential methylation in 240 991 ± 55 452 probes accounting for 30.4% ± 6.99 of filtered probes, while the TET1-dSpCas9 expressed under the secondary cassette significantly reduced differential methylation in 135 697 ± 13 079 probes, accounting for 17.1% ± 1.65 of the filtered probes ($P = 0.022$ by SNK post-ANOVA multiple comparison test). As expected, 98.7% of the DMPs induced using the DNMT3A-dSpCas9 fusions with either 'primary' or 'secondary cassettes' were hypermethylated, whereas 97.2% and 93.2% of the reported DMPs were hypomethylated using TET1-dSpCas9 with either 'primary' or 'secondary cassettes', respectively (Figure 8B). Interestingly, the DMPs induced by the 'primary' and 'secondary' cassettes of the DNMT3A and TET1 fusion proteins were largely overlapping (Figure 8G), suggesting that off-target fusion protein activity occurs in a non-stochastic manner.

To confirm the on-target activity of the dCas9-based modular system, methylation in the targeted regions of the *IL6ST* and *MGAT3* loci was analyzed in DNMT3A-dSpCas9 and TET1-dSpCas9 samples (Figure 8C and E). The Infinium MethylationEPIC array contains two probes in the targeted *IL6ST* region (Figure 8D) and one probe in the targeted *MGAT3* region (Figure 8F). Expression of DNMT3A-dSpCas9 from both 'primary' and 'secondary cassettes' induced significant hypermethylation in the probes cg21950518 and cg15219433, whereas TET1-dSpCas9 expressed from either 'primary' or 'secondary cassettes' induced significant hypomethylation in the probe cg21461856. The targeted probes of DNMT3A-dSpCas9 were not affected by TET1-dSpCas9 and vice versa (Figure 8D and F). Interestingly, we observed that hypermethylation is induced to a greater extent at CpG sites flanking the targeted region (±2500 bp) in DNMT3A-dSpCas9 samples (Figure 8C). In contrast, hypomethylation effect by TET1-dSpCas9 was strongest at the target sequences and *cis*-sites, with less pronounced effects observed at the relatively distal flanking regions (±2500 bp) (Figure 8E). It should be noted that to allow clearer visualization, Figure 8C and E was plotted using methylation (β) values that were adjusted by Surrogate Variable Analysis (SVA) and log-transformed into *M*-values, which do not directly reflect the real biological % DNA methylation. Similar graphs depicting actual methylation values (unadjusted betas normalized by the R package 'funnorm') are included in Supplementary Figure S11.

DISCUSSION

Various platforms and protein scaffold systems such as SunTag, SAM and MoonTag have been developed for direct transcriptional modulation (14,22,43–46). Also, individual fusions of dCas9 to various epigenetic writers and erasers are available for targeted changes of epigenetic marks. Here, we have created and validated a CRISPR/dCas9-based modular system that can be efficiently assembled and easily reconfigured for a specific experimental setup. This system represents a complete cloning platform, to which various functional modules can be added easily. It has the possibility to exchange Cas9 orthologs, effector domains and selection markers as well as to be combined with a multi-guide system, which ensures simultaneous delivery of up to six gRNAs into the same cell. We validated the modular CRISPR/dCas9-based toolbox in multiple configurations regarding fusions to either C- or N-terminus of a dCas9 protein as well as use of different orthologous dCas9 (from either *S. pyogenes* or *S. aureus*). The system based on direct fusions between a dCas9 protein and an effector domain has the minimal number of components that need to be delivered on separate vectors, making experimental design simpler and more robust compared to other systems where dCas9 and effector domains are expressed separately, and the proteins need to form an active complex within the cell (for an overview, see Tadić *et al.* (42) and references therein). Further, a multi-guide system can be fully integrated, facilitating multiple targeting with up to six gRNAs. Finally, different dCas9 orthologs show no interference when multiple functionalities are independently targeted to different loci.

Recently, Gao *et al.* designed a flexible dCas9-based platform for inducible orthogonal gene regulation (12). Their platform uses the dCas9 protein fused with KRAB and VPR effector domains and shows robust function in these simple fusions in direct gene repression/activation within mammalian cells (12). On the other hand, activity of the DNMT3A and TET1 catalytic domains critically depends on the configuration of the fusion (N-terminal versus C-terminal), the type of the linker and/or on NLS availability, which we have demonstrated in this work. While fusions of DNMT3A and TET1 with dCas9 have already been described (23–28), there was no universal and robust approach

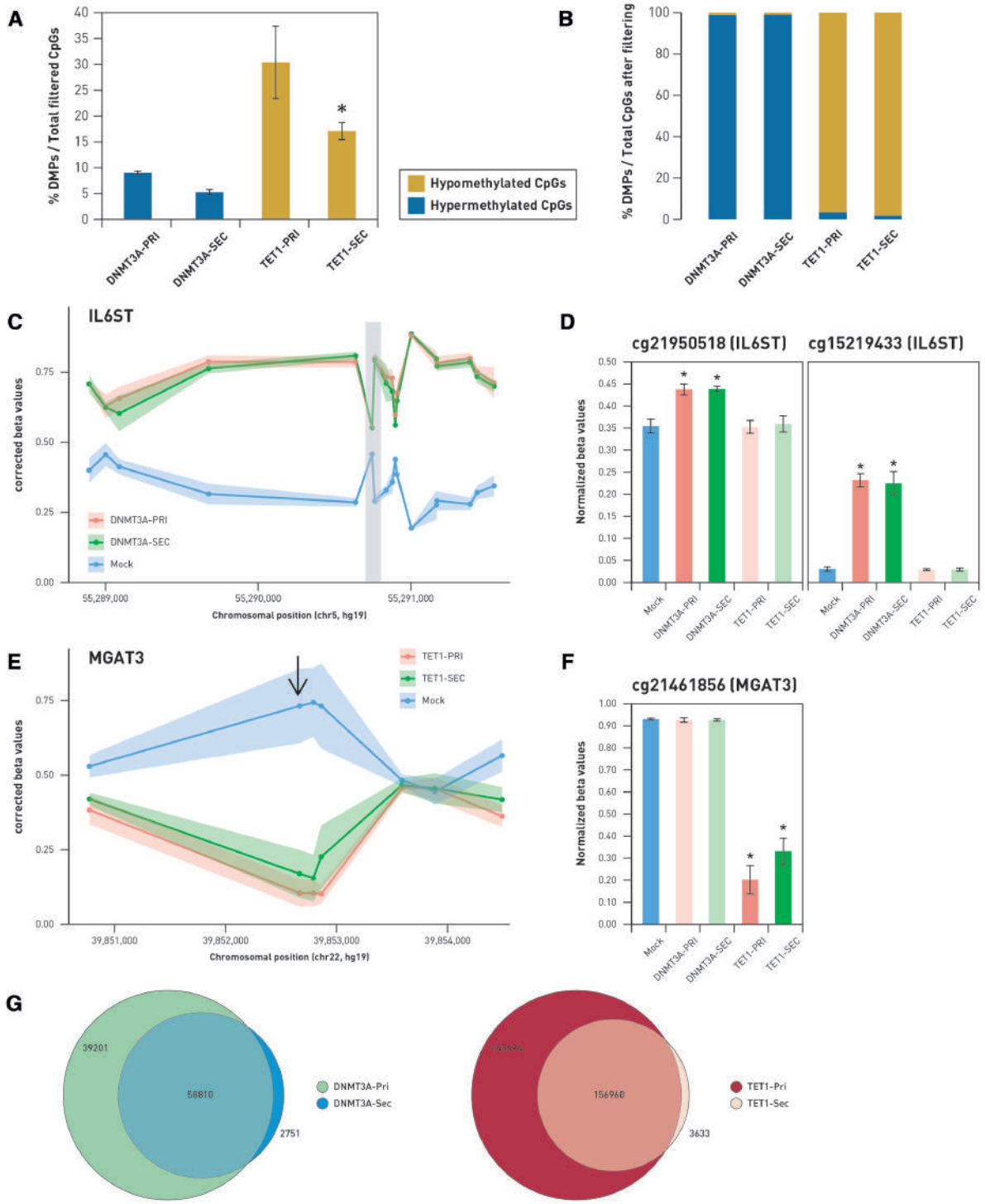


Figure 8. Analysis of whole-genome methylation following epigenetic modulation of the *IL6ST*/*MGAT3* promoters using DNMT3A-dSpCas9 and TET1-dSpCas9 fusion constructs. (**A**) Fraction of differentially methylated probes compared to mock controls against total filtered probes analyzed on the Infinium MethylationEPIC array. Labels DNMT3A-PRI/TET1-PRI and DNMT3A-SEC/TET1-SEC indicate 'primary' and 'secondary cassette', respectively. (**B**) Fraction of differentially methylated probes, which were hypermethylated or hypomethylated, compared to mock controls in cells transfected with DNMT3A-dSpCas9 and TET1-dSpCas9 fusions. (**C** and **E**) Dots represent CpG sites targeted by DNMT3A-dSpCas9 and TET1-dSpCas9 fusion constructs. Lines show pattern of methylation (C) and demethylation (E) change at CpG island of *IL6ST* and *MGAT3* genes. Comparison of CpG methylation patterns induced by DNMT3A-dSpCas9 expressed under strong ('primary cassette') or weak ('secondary cassette') promoter on probes located 2500 bp up- and downstream of the targeted region (shaded in gray) of *IL6ST* is shown on (C). Changes to β-values in the two probes lying within this region are illustrated on (**D**). (**E**) Comparison of DNA methylation patterns induced by TET1-dSpCas9 primary and secondary cassettes 2500 bp up- and downstream of the targeted region of the *MGAT3* promoter region. The probe cg21461856 is the only one lying within this region (indicated with the arrow). Changes to β-values in cg21461856 are illustrated in the graph (**F**). (**G**) Venn diagrams illustrating overlap between DMPs induced by the primary and secondary cassettes of DNMT3A-dSpCas9 and TET1-dSpCas9 fusions. * $P < 0.05$.

for creating active fusions of DNMT3A and TET1 with different dCas9 orthologs, which would work within the same cell simultaneously and perform antagonistic activities (i.e. DNA methylation and demethylation). In order to create a modular system capable of antagonistic and/or synergistic activities, we tested both C- and N-terminal configuration between dCas9 orthologs (*S. pyogenes* and *S. aureus*) and DNMT3A or TET1 catalytic domains for activity in two human cell lines. We consistently observed more robust activity of both DNMT3A and TET1 catalytic domains when fused N-terminally to either dSpCas9 or dSaCas9. Although C-terminal fusions to dCas9 from *S. pyogenes* performed satisfactorily when proper targeting to the nucleus was assured by an additional NLS nucleoplasmin (NLS-NP) at the C-terminus of the catalytic domain, we chose N-terminal fusions as a streamlined and universal approach. On the other hand, C-terminal fusions of catalytic domains with dCas9 from *S. aureus* suffered poor performance even when properly targeted to the nucleus by addition of NLS-NP to the both sides of the construct (Figure 2) or when different linkers were used in fusions (Supplementary Table S4). Though the Cas9 from *S. aureus* is smaller and has improved level of specificity in comparison with Cas9 from *S. pyogenes*, and both SpCas9 and SaCas9 show similar kinetics and activity as nucleases (47), in our hands dSaCas9 C-terminal fusions did not work either for DNMT3A or TET1 catalytic domains. This lack of activity may be due to steric hindrances between dSaCas9 and the effector domains or to misfolding and degradation of the fusion proteins.

The activity profiles of C- and N-terminal fusions of dSpCas9 with both DNMT3A and TET1 appeared quite similar, with similar position of peak activity relative to the gRNA-binding site. Also, N-terminal fusions of DNMT3A and TET1 to dSaCas9 were similar in activity profile to all dSpCas9 C- and N-terminal fusions with DNMT3A and TET1 effector domains. This shows that, at the scale of the effector domain interaction with DNA, configuration of the fusion construct (C- or N-terminal) does not significantly contribute to the activity profile, although the effect of fusions with C-terminal effector domains was mainly concentrated downstream from the binding site, while N-terminal domains demonstrated a strong upstream activity as well (Figure 1). The activity of all dCas9 N-terminal fusions could be detected on both sides of the gRNA-binding site. The additional activity peaks, reproducibly detected further upstream or downstream (180–200 bp) from the gRNA-binding site, were consistent with possible contact of effector domains with DNA on adjacent nucleosomes. When comparing the activity of N-terminally fused effector domains to either dSaCas9 or dSpCas9, both orthologs performed satisfactorily, though dSpCas9 fusions consistently showed stronger activity (Figure 1; top panel).

We also demonstrated the application of this molecular toolbox to simultaneous targeting of several pairs of gene loci using fusions of dCas9 with effector domains conferring antagonistic or synergistic activities, at the same time within the same cells. We epigenetically manipulated *BACH2–HNF1A* and *IL6ST–MGAT3* gene pairs in HEK293 cells, where we targeted *BACH2* and *IL6ST* with DNMT3A-dSpCas9 fusion and *HNF1A* and *MGAT3* with TET1-dSaCas9 fusion. As expected, changes in cytosine methylation went in the opposite direction for each locus within a pair (methylation versus demethylation) as a result of manipulation of the two dCas9 fusions with antagonistic activities. Imposed change of cytosine methylation was shown to change transcription level of the manipulated genes at different scale, depending on the region within the promoter that we targeted with a dCas9 fusion. The effect on the transcript level of the *BACH2* and *IL6ST* genes was moderate (0.568-fold change, $P = 0.0003$ and 0.583; $P = 0.002$, respectively), although wide regions spanning about 90 bp encompassing 27 CpGs (for *BACH2*) to 15 CpGs (for *IL6ST*) were targeted using DNMT3A-dSpCas9 (Figures 4 and 5). On the other hand, TET1-dSaCas9 targeted demethylation of only four CpG sites, located in the first exon of the *HNF1A* gene strongly influenced the expression level of this gene (5.614-fold change; $P = 0.0012$). In contrast, TET1-dSaCas9 induced demethylation at 14 CpG sites located in the first CpG island 1 of *MGAT3* promoter did not increase transcription level of this gene in HEK293 cells. However, when we targeted the DNMT3A-dSpCas9 fusion to 30 CpG sites positioned near TIS region in the CpG island 2 of the *MGAT3* promoter, which was previously suggested as putative regulatory region in several ovarian cell lines including BG1 (41), we achieved considerable effect on *MGAT3* transcription (0.28-fold change, $P = 0.004$) (Figure 6C). Although the same CpG sites are unmethylated in the HEK293 cell line, expression level of the *MGAT3* gene is low. Therefore, we reasoned that we could not observe any effect of externally imposed hypomethylation on *MGAT3* expression level in HEK293 cells, which suggests that other factors besides CpG methylation are involved in regulation of *MGAT3* expression and that relevant regulatory CpG sites are probably cell line specific.

Further, we manipulated the same CpG sites in another cell line, BG1, using TET1-dSaCas9, and again achieved an increase of the transcript level (6.64-fold change, $P = 0.004$). At the same time *MGAT3* gene was manipulated in same cells, since we had information about putative regulatory region within the *MGAT3* promoter, too (41). Both genes are involved in protein *N*-glycosylation: *MGAT3* is a classical glyco-gene coding for glycosyltransferase, which produces *N*-glycan structures with bisecting *N*-GlcNAc, while *HNF1A* is a master regulator of both core and antennary fucosylation of *N*-glycan structures (39). In addition, the relationship between cytosine methylation in CpG island 2 of the *MGAT3* gene promoter, its expression and the presence of bisecting GlcNAc on N-glycans in ovarian cancer cell lines were already demonstrated, although the methylation change was obtained using 5-aza-deoxycitidine treatment (41). Here, we targeted 30 putative regulatory CpG sites, centered about 100 bp upstream from the TSS, with DNMT3A-dSpCas9 and downregulated *MGAT3* to 0.28-fold change following epigenetic manipulation. In addition, simultaneous epigenetic manipulation of the *MGAT3* and *HNF1A* genes resulted in a decrease of *N*-glycan structures with bisecting GlcNAc (the result of *MGAT*3 downregulation) and a decrease of complex glycan structures with core-fucose (an indirect effect of *HNF1A* upregulation). The effect of *HNF1A* on several fucosyltransferases, such as FUT3, FUT5, FUT6, FUT7, FUT8, FUT9, FUT10 and FUT11, has been previously demonstrated by RNA inter-

ference, where *HNF1A* appeared to downregulate FUT8 (39), an enzyme that adds a fucose on the three-mannose core of N-glycan structures.

Our results demonstrate that proper targeting of functional CpG sites or regulatory region is necessary to obtain an effect on gene transcription. Therefore, the difference in the effect of externally induced hyper- and hypomethylation we achieved on transcript level of the candidate genes was due to the relevance of the targeted CpG sites for transcriptional regulation. In our earlier correlation study between CpG methylation and gene expression in different cell lines, we identified four putative regulatory CpG sites in the first exon of the *HNF1A* gene (40). In the present study, using dCas9 fusions we confirmed unambiguously that cytosine methylation at these sites plays a regulatory role, because it significantly altered *HNF1A* transcript level. Therefore, as few as four adjacent CpG sites spanning a 13 bp short region can regulate transcription. Similar effect was reported for the *Ascl1* locus when TALE fusions with TET1 or DNMT3A were used (48). However, although Lo *et al.* (48) searched for putative transcription factor binding sites that included the manipulated CpG dinucleotides, the achieved influence on *Ascl1* gene expression was relatively weak (1.58-fold increase in CpG methylation that was associated with a 1.75-fold decrease in gene expression) and the level of regulation was in line with the results observed when targeting random CpG sites scattered throughout a wider region within the *BACH2* and *IL6ST* promoters in our study. It is also worth noting that we deliberately selected almost fully methylated or unmethylated regions for targeted manipulation, which was done in order to make any induced changes in expression obvious. However, it is still not clear in terms of dose–response how DNA methylation affects gene expression. In our earlier work, we observed that multiple targeting by co-expression of several gRNAs can have a synergistic effect on site-specific CpG methylation (27). This knowledge proved valuable in case of *BACH2* and *IL6ST*, where DNMT3A-dSpCas9 in combination with multi-guide system was used for a greater increase of methylation levels at the targeted regions. This is particularly useful when we do not have previous knowledge of functional CpG sites, because methylation of a wide region tends to spread along a CpG island and eventually reaches the several CpG sites directly responsible for transcriptional control. In this case, CpG methylation change probably induces changes in chromatin conformation and DNA accessibility, triggering multiple processes such as histone modifications and nucleosome remodeling (49).

We were also interested in the effect on gene expression from simultaneous targeting of the VPR-dSpCas9 (for direct gene activation) and TET1-dSaCas9 fusions (for hypomethylation to demethylation) to the *HNF1A* regulatory region. Therefore, we double transfected HEK293 cells using these two dCas9 fusions and monitored CpG methylation and gene transcript levels during 30 days. Although direct gene activation showed a stronger effect on transcription than the combined VPR/TET1 activation on the 5th day after transfection, the effects of the VPR-dSpCas9 fusion was relatively transient, and was surpassed by the VPR/TET1 combination after 11th day post-transfection, and *HNF1A* upregulation persisted until the 30th day (Figure 7B). If we observe the effect of the dCas9 fusions used either together or individually on CpG methylation only, the TET1-dSaCas9 fusion used alone induced a greater reduction of CpG methylation level than if the fusions were used together, and this was expected because the binding of VPR-dSpCas9 possibly interfered with binding of the TET1-dSaCas9 construct. In our previous and present work, we demonstrated that externally induced hyper- and hypo-methylated state at targeted CpG sites, using simple C-terminal fusions dSpCas9-DNMT3A (27) and dSpCas9-TET1 (Supplementary Figure S1), persisted through mitotic divisions, reaching the maximum level 7–9 days after transfection and persisting for at least 15 days. While the initial change in methylation level was rapid, the return to the baseline was gradual, with a detectable change (around 20% in average) persisting for up to 30 days after transfection. In this study, we performed a time-course experiment targeting the *IL6ST* and *HNF1A* genes independently, using the N-terminal fusions DNMT3A-dSpCas9 and TET1-dSaCas9. These two fusions were constructed using the new modular approach. During 30 days of monitoring CpG methylation, we found that the effect of methylation and demethylation remained stable for a longer time than in case when we used C-terminal fusions were used. Methylation levels stay altered up to 30% to 40% even 30 days after transfection of HEK293 cells compared to initial level of methylation (mock) at the targeted CpG sites within the promoters of *IL6ST* and *HNF1A* (Supplementary Figure S8). Furthermore, we found that almost no plasmid was present on 8th day after the transfection, while the dCas9 protein was undetectable after 11th day following transfection and any lingering mRNA expression extinguished by day 20 after cell transfection (Figure 7C and Supplementary Figure S9A). Therefore, we can be fairly confident that around day 11 there was no significant new demethylation induced by TET1-dSaCas9 fusion construct. Yet, the effect on CpG methylation persisted until the day 30 after transfection (Figure 7A and Supplementary Figure S8). A possible explanation for the long-lasting effect on cytosine methylation (especially compared to the earlier experiments involving C-terminal fusion dSpCas9-DNMT3A) is the larger initial effect achieved by the new tools on DNA methylation levels, which might have activated other epigenetic mechanisms (such as histone modifications), which subsequently consolidated and reinforced the induced epigenetic changes. In addition, somewhat higher activity of N-terminal fusions might be explained by the additional NLS compared to C-terminal fusions: while the N-terminal fusion has exposed NLS at both termini, the C-terminal fusion has an exposed NLS only at the N-terminus of dCas9 portion of fusion construct (see schematics in Figure 1). Targeting a wider region using the multi-guide system may have contributed to an increased overall effect on methylation. Similarly, Liu *et al.* demonstrated that the effect of demethylation induced by dCas9-TET1 lasted for 14 days, and restored *FMR1* gene expression during that time in the absence of constitutive presence of dCas9-TET1, suggesting sustained gene reactivation through cell divisions (22). It would be interesting to analyze how forced demethylation and imposed methylation affect other epigenetic mechanisms, notably histone modifications, at the short region

of *HNF1A* targeted by TET1-dSaCas9, in comparison with larger region of *IL6ST* promoter targeted by DNMT3A-dSpCas9, and whether those other modifications contribute to long-lasting change of CpG methylation that we observed in the time-course experiments.

In contrast to applications of Cas9 for generating knockouts, where every productive binding to genomic DNA has a potential to introduce a mutation, and where strategies for avoiding off-target effects focus on achieving high fidelity of Cas9/DNA interaction (50), the problem with epigenome editing tools is fundamentally different (for a review, see Tadić *et al.* (42) and references therein). Here, passing interactions will leave only a minimal mark; more importantly, the bulk of unspecific activity does not come from unspecific Cas9/DNA interactions, but from unguided activity of the linked catalytic domains, such as DNMT3A and TET1. Indeed, Galonska *et al.* (51) demonstrated that the dCas9-DNMT3A fusion induced global increase in CpG methylation without gRNA and that the off-target activity was mostly random and remained unchanged even when using single or multiple gRNAs. In an attempt to mitigate the off-target activity, we opted for down-regulating expression of the dCas9 fusion construct by placing it under a weaker promoter, with the aim to achieve a more favorable on-target to off-target ratio. The selection marker expression (puromycin resistance) was driven by a strong promoter in a separate expression cassette on the backbone plasmid, thus ensuring efficient selection. This 'secondary cassette' approach proved successful, as evidenced by whole-genome methylation analysis, which showed decreased off-target activity with practically unaffected on-target activity when dSpCas9 N-terminal fusions with DNMT3A and TET1 were put under a weaker promoter. Whole-genome methylation analysis using the Illumina 850K platform also revealed details about functioning of our new and improved dCas9-based tools for targeted CpG methylation and demethylation. We have demonstrated that reducing dCas9 fusion protein expression relative to the selection marker can substantially reduce off-target activity. This complements the findings of a separate study, which demonstrates that modulating DNMT3A expression relative to dCas9 results in a reduction in off-target DNA methylation while maintaining high on-target DNA methylation (52). However, it should be noted that while our approach has reduced off-target hypomethylation induced by TET1-dSpCas9, further improvements need to be made before the TET1-dSpCas9 fusion can be considered reliable and used as a specific tool for inducing site-specific hypomethylation. Interestingly, the DMPs induced by the 'primary' and 'secondary' cassettes of the DNMT3A-dSpCas9 and TET1-dSpCas9 fusion proteins were largely overlapping (Figure 8G), suggesting that off-target activity occur in a non-stochastic manner. Annotation of DMPs by their relationship to CpG islands indicated that hypermethylation induced by DNMT3A-dSpCas9 was primarily enriched in the CpG-rich CpG island-to-shore regions (Supplementary Figure S12A) while the hypomethylation induced by TET1-dSpCas9 occurred mostly at the less dense CpG shore-to-shelf regions (Supplementary Figure S12B). The reverse was observed in less CpG-rich shelves and open sea regions, which were depleted in hypermethylated DMPs of DNMT3A-dSpCas9 and TET1-dSpCas9 samples. This could be explained by the fact that CpG-density negatively correlates with DNA methylation levels, as reported in multiple studies (53–55). Thus, CpG islands will have low methylation levels, and hence, will be more susceptible to methylation by DNMT3A than to further demethylation by TET1. The reverse was observed in less CpG-rich shelves and open sea regions, which were depleted in hypermethylated DMPs of the dCas9-DNMT3A and dCas9-TET1 samples. Hypermethylation of these regions would be less dynamic, due to their tendency toward relative hypermethylation at baseline (56). Similarly, in terms of regulatory regions, DNMT3A-dSpCas9 induced hypermethylation preferentially targeted promoters (strong or weak), which are often CpG-rich (Supplementary Figure S13A), while the TET1-dSpCas9 induced hypomethylation was enriched at enhancers (strong or weak), repressed regions and transcription regulatory sites (Supplementary Figure S13B). Some hypermethylation effect by DNMT3A-dSpCas9 was also seen, though to a lesser extent, at strong enhancers and insulator regions.

This study provides a proof of concept that externally imposed CpG methylation in promoters of the candidate genes, relevant for protein glycosylation, can change their transcription level and consequently have an effect on the glycan phenotype. The newly developed CRISPR/dCas9-based modular toolbox is therefore very useful for studying gene regulatory networks that regulate IgG glycosylation involved in inflammation, which is essential for our effort within the frame of the H2020 flagship interdisciplinary consortium SYSCID ('A systems medicine approach to chronic inflammatory disease') with the goal of understanding molecular mechanism underlying chronic inflammatory disease. In general, such molecular toolbox could have a great potential in therapeutic strategies for disorders that involve epigenetic silencing. Finally, a molecular toolbox with exchangeable effector domains, Cas9 orthologs, selection markers and weak/strong promoters provides an extensible platform suited for resolving open questions in the field of targeted epigenome editing as well as fundamental questions related to the role of epigenetic marks in gene regulation. For instance, the use of xCas9 variants with broad PAM recognition and higher DNA specificity (28,42) could further reduce off-target activity. The use of some other epigenetic 'writers' and 'erasers' for simultaneous epigenetic editing could reveal causal relationship between directly manipulated individual epigenetic marks other than CpG methylation and gene transcription, thus helping in better understanding of the link between the complex chromatin layer, transcriptional regulation and cell function (21).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Francis Jacob for providing the BG1 cell line. We acknowledge anonymous reviewers for constructive criticism, which helped us to significantly improve the revised version of the manuscript.

FUNDING

Croatian National Science Foundation Project EpiGlycoIgG [Epigenetic regulation of IgG glycosylation; contract #3361]; IRI grant from the European Structural and Investment Funds [contract #KK.01.2.1.01.0003]; Croatian National Centre of Research Excellence in Personalized Healthcare [contract #KK.01.1.1.01.0010]; ICGEB CRP grant [Comprehensive Toolbox for Epigenetic Modulation of Gene Expression; Contract no. CRP/17/006; ICGEB Ref. No. CRP/ HRV17-03]; European Horizon2020 project SYSCID, SC1-2016-2017 [A systems medicine approach to chronic inflammatory disease; contract #733100]. Funding for open access charge: All funders/projects declared above can be used to cover page charges; there is enough funding from each of them. The work carried out at IARC was partially supported by: Institut National du Cancer (INCa, France), Foundation ARC pour la Recherche sur le Cancer (France), INCa-DGOS-Inserm (grant #12563), EC FP7 Marie Curie Actions – People – Co-funding of regional, national and international programmes (COFUND).
Conflict of interest statement. None declared.

REFERENCES

1. Chen,B., Gilbert,L.A., Cimini,B.A., Schnitzbauer,J., Zhang,W., Li,G.W., Park,J., Blackburn,E.H., Weissman,J.S., Qi,L.S. *et al.* (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, **155**, 1479–1491.
2. Ma,H., Naseri,A., Reyes-Gutierrez,P., Wolfe,S.A., Zhang,S. and Pederson,T. (2015) Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 3002–3007.
3. Ma,H., Tu,L.C., Naseri,A., Huisman,M., Zhang,S., Grunwald,D. and Pederson,T. (2016) Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat. Biotechnol.*, **34**, 528–530.
4. Nelles,D.A., Fang,M.Y., O'Connell,M.R., Xu,J.L., Markmiller,S.J., Doudna,J.A. and Yeo,G.W. (2016) Programmable RNA tracking in live cells with CRISPR/Cas9. *Cell*, **165**, 488–496.
5. Wang,H., La Russa,M. and Qi,L.S. (2016) CRISPR/Cas9 in genome editing and beyond. *Annu. Rev. Biochem.*, **85**, 227–264.
6. Cheng,A.W., Wang,H., Yang,H., Shi,L., Katz,Y., Theunissen,T.W., Rangarajan,S., Shivalila,C.S., Dadon,D.B. and Jaenisch,R. (2013) Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.*, **23**, 1163–1171.
7. Gilbert,L.A., Larson,M.H., Morsut,L., Liu,Z., Brar,G.A., Torres,S.E., Stern-Ginossar,N., Brandman,O., Whitehead,E.H., Doudna,J.A. *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, **154**, 442–451.
8. Maeder,M.L., Linder,S.J., Cascio,V.M., Fu,Y., Ho,Q.H. and Joung,J.K. (2013) CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods*, **10**, 977–979.
9. Perez-Pinera,P., Kocak,D.D., Vockley,C.M., Adler,A.F., Kabadi,A.M., Polstein,L.R., Thakore,P.I., Glass,K.A., Ousterout,D.G., Leong,K.W. *et al.* (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods*, **10**, 973–976.
10. Polstein,L.R. and Gersbach,C.A. (2015) A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.*, **11**, 198–200.
11. Zetsche,B., Volz,S.E. and Zhang,F. (2015) A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat. Biotechnol.*, **33**, 139–142.
12. Gao,Y., Xiong,X., Wong,S., Charles,E.J., Lim,W.A. and Qi,L.S. (2016) Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nat. Methods*, **13**, 1043–1049.
13. Wangensteen,K.J., Wang,Y.J., Dou,Z., Wang,A.W., Mosleh-Shirazi,E., Horlbeck,M.A., Gilbert,L.A., Weissman,J.S., Berger,S.L. and Kaestner,K.H. (2018) Combinatorial genetics in liver repopulation and carcinogenesis with a in vivo CRISPR activation platform. *Hepatology*, **68**, 663–676.
14. Zhou,H., Liu,J., Zhou,C., Gao,N., Rao,Z., Li,H., Hu,X., Li,C., Yao,X., Shen,X. *et al.* (2018) In vivo simultaneous transcriptional activation of multiple genes in the brain using CRISPR-dCas9-activator transgenic mice. *Nat. Neurosci.*, **21**, 440–446.
15. Chatterjee,A., Rodger,E.J. and Eccles,M.R. (2018) Epigenetic drivers of tumourigenesis and cancer metastasis. *Semin. Cancer Biol.*, **51**, 149–159.
16. Absher,D.M., Li,X., Waite,L.L., Gibson,A., Roberts,K., Edberg,J., Chatham,W.W. and Kimberly,R.P. (2013) Genome-wide DNA methylation analysis of systemic lupus erythematosus reveals persistent hypomethylation of interferon genes and compositional changes to CD4+ T-cell populations. *PLoS Genet.*, **9**, e1003678.
17. Glossop,J.R., Nixon,N.B., Emes,R.D., Haworth,K.E., Packham,J.C., Dawes,P.T., Fryer,A.A., Mattey,D.L. and Farrell,W.E. (2013) Epigenome-wide profiling identifies significant differences in DNA methylation between matched-pairs of T- and B-lymphocytes from healthy individuals. *Epigenetics*, **8**, 1188–1197.
18. Rakyan,V.K., Beyan,H., Down,T.A., Hawa,M.I., Maslau,S., Aden,D., Daunay,A., Busato,F., Mein,C.A., Manfras,B. *et al.* (2011) Identification of type 1 diabetes-associated DNA methylation variable positions that precede disease diagnosis. *PLoS Genet.*, **7**, e1002300.
19. Ventham,N.T., Kennedy,N.A., Adams,A.T., Kalla,R., Heath,S., O'Leary,K.R., Drummond,H., consortium,I.B., consortium,I.C., Wilson,D.C. *et al.* (2016) Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. *Nat. Commun.*, **7**, 13507.
20. Tough,D.F., Tak,P.P., Tarakhovsky,A. and Prinjha,R.K. (2016) Epigenetic drug discovery: breaking through the immune barrier. *Nat. Rev. Drug Discov.*, **15**, 835–853.
21. Pulecio,J., Verma,N., Mejia-Ramirez,E., Huangfu,D. and Raya,A. (2017) CRISPR/Cas9-Based engineering of the epigenome. *Cell Stem Cell*, **21**, 431–447.
22. Liu,X.S., Wu,H., Krzisch,M., Wu,X., Graef,J., Muffat,J., Hnisz,D., Li,C.H., Yuan,B., Xu,C. *et al.* (2018) Rescue of fragile X syndrome neurons by DNA methylation editing of the FMR1 gene. *Cell*, **172**, 979–992.
23. Chaudhary,K., Chattopadhyay,A. and Pratap,D. (2018) The evolution of CRISPR/Cas9 and their cousins: hope or hype? *Biotechnol. Lett.*, **40**, 465–477.
24. Liu,P., Chen,M., Liu,Y., Qi,L.S. and Ding,S. (2018) CRISPR-Based chromatin remodeling of the endogenous Oct4 or Sox2 locus enables reprogramming to pluripotency. *Cell Stem Cell*, **22**, 252–261.
25. Liu,X.S., Wu,H., Ji,X., Stelzer,Y., Wu,X., Czauderna,S., Shu,J., Dadon,D., Young,R.A. and Jaenisch,R. (2016) Editing DNA methylation in the mammalian genome. *Cell*, **167**, 233–247.
26. Stepper,P., Kungulovski,G., Jurkowska,R.Z., Chandra,T., Krueger,F., Reinhardt,R., Reik,W., Jeltsch,A. and Jurkowski,T.P. (2017) Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. *Nucleic Acids Res.*, **45**, 1703–1713.
27. Vojta,A., Dobrinic,P., Tadic,V., Bockor,L., Korac,P., Julg,B., Klasic,M. and Zoldos,V. (2016) Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res.*, **44**, 5615–5628.
28. Xu,X., Tao,Y., Gao,X., Zhang,L., Li,X., Zou,W., Ruan,K., Wang,F., Xu,G.L. and Hu,R. (2016) A CRISPR-based approach for targeted DNA demethylation. *Cell Discov.*, **2**, 16009.
29. Klasic,M., Markulin,D., Vojta,A., Samarzija,I., Birus,I., Dobrinic,P., Ventham,N.T., Trbojevic-Akmacic,I., Simurina,M., Stambuk,J. *et al.* (2018) Promoter methylation of the MGAT3 and BACH2 genes correlates with the composition of the immunoglobulin G glycome in inflammatory bowel disease. *Clin. Epigenet.*, **10**, 75.
30. Schmittgen,T.D. and Livak,K.J. (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.*, **3**, 1101–1108.
31. van Dongen,J.J., Langerak,A.W., Bruggemann,M., Evans,P.A., Hummel,M., Lavender,F.L., Delabesse,E., Davi,F., Schuuring,E., Garcia-Sanz,R. *et al.* (2003) Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations:

report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*, **17**, 2257–2317.
32. Schneider,C.A., Rasband,W.S. and Eliceiri,K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, **9**, 671–675.
33. Aryee,M.J., Jaffe,A.E., Corrada-Bravo,H., Ladd-Acosta,C., Feinberg,A.P., Hansen,K.D. and Irizarry,R.A. (2014) Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, **30**, 1363–1369.
34. Ritchie,M.E., Phipson,B., Wu,D., Hu,Y., Law,C.W., Shi,W. and Smyth,G.K. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.*, **43**, e47.
35. Cochran,A.G., Skelton,N.J. and Starovasnik,M.A. (2001) Tryptophan zippers: stable, monomeric beta -hairpins. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 5578–5583.
36. Georgieva,M.V., Yahya,G., Codo,L., Ortiz,R., Teixido,L., Claros,J., Jara,R., Jara,M., Iborra,A., Gelpi,J.L. *et al.* (2015) Inntags: small self-structured epitopes for innocuous protein tagging. *Nat. Methods*, **12**, 955–958.
37. Igarashi,K., Ochiai,K., Itoh-Nakadai,A. and Muto,A. (2014) Orchestration of plasma cell differentiation by Bach2 and its gene regulatory network. *Immunol. Rev.*, **261**, 116–125.
38. Gudelj,I., Lauc,G. and Pezer,M. (2018) Immunoglobulin G glycosylation in aging and diseases. *Cell Immunol.*, **333**, 65–79.
39. Lauc,G., Essafi,A., Huffman,J.E., Hayward,C., Knezevic,A., Kattla,J.J., Polasek,O., Gornik,O., Vitart,V., Abrahams,J.L. *et al.* (2010) Genomics meets glycomics-the first GWAS study of human N-Glycome identifies HNF1alpha as a master regulator of plasma protein fucosylation. *PLoS Genet.*, **6**, e1001256.
40. Zoldos,V., Horvat,T., Novokmet,M., Cuenin,C., Muzinic,A., Pucic,M., Huffman,J.E., Gornik,O., Polasek,O., Campbell,H. *et al.* (2012) Epigenetic silencing of HNF1A associates with changes in the composition of the human plasma N-glycome. *Epigenetics*, **7**, 164–172.
41. Kohler,R.S., Anugraham,M., Lopez,M.N., Xiao,C., Schoetzau,A., Hettich,T., Schlotterbeck,G., Fedier,A., Jacob,F. and Heinzelmann-Schwarz,V. (2016) Epigenetic activation of MGAT3 and corresponding bisecting GlcNAc shortens the survival of cancer patients. *Oncotarget*, **7**, 51674–51686.
42. Tadic,V., Josipovic,G., Zoldos,V. and Vojta,A. (2019) CRISPR/Cas9-based epigenome editing: An overview of dCas9-based tools with special emphasis on off-target activity. *Methods*, **164-165**, 109–119.
43. Chavez,A., Tuttle,M., Pruitt,B.W., Ewen-Campen,B., Chari,R., Ter-Ovanesyan,D., Haque,S.J., Cecchi,R.J., Kowal,E.J.K., Buchthal,J. *et al.* (2016) Comparison of Cas9 activators in multiple species. *Nat. Methods*, **13**, 563–567.
44. Liao,H.K., Hatanaka,F., Araoka,T., Reddy,P., Wu,M.Z., Sui,Y., Yamauchi,T., Sakurai,M., O'Keefe,D.D., Nunez-Delicado,E. *et al.* (2017) In vivo target gene activation via CRISPR/Cas9-Mediated Trans-epigenetic modulation. *Cell*, **171**, 1495–1507.
45. Yeo,N.C., Chavez,A., Lance-Byrne,A., Chan,Y., Menn,D., Milanova,D., Kuo,C.C., Guo,X., Sharma,S., Tung,A. *et al.* (2018) An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat. Methods*, **15**, 611–616.
46. Zalatan,J.G., Lee,M.E., Almeida,R., Gilbert,L.A., Whitehead,E.H., La Russa,M., Tsai,J.C., Weissman,J.S., Dueber,J.E., Qi,L.S. *et al.* (2015) Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell*, **160**, 339–350.
47. Cebrian-Serrano,A. and Davies,B. (2017) CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm. Genome*, **28**, 247–261.
48. Lo,C.L., Choudhury,S.R., Irudayaraj,J. and Zhou,F.C. (2017) Epigenetic editing of Ascl1 gene in neural stem cells by optogenetics. *Sci. Rep.*, **7**, 42047.
49. Allis,C.D. and Jenuwein,T. (2016) The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.*, **17**, 487–500.
50. Hu,J.H., Miller,S.M., Geurts,M.H., Tang,W., Chen,L., Sun,N., Zeina,C.M., Gao,X., Rees,H.A., Lin,Z. *et al.* (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*, **556**, 57–63.
51. Galonska,C., Charlton,J., Mattei,A.L., Donaghey,J., Clement,K., Gu,H., Mohammad,A.W., Stamenova,E.K., Cacchiarelli,D., Klages,S. *et al.* (2018) Genome-wide tracking of dCas9-methyltransferase footprints. *Nat. Commun.*, **9**, 597.
52. Pflueger,C., Tan,D., Swain,T., Nguyen,T., Pflueger,J., Nefzger,C., Polo,J.M., Ford,E. and Lister,R. (2018) A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. *Genome Res.*, **28**, 1193–1206.
53. Varriale,A. and Bernardi,G. (2010) Distribution of DNA methylation, CpGs, and CpG islands in human isochores. *Genomics*, **95**, 25–28.
54. Mugal,C.F. and Ellegren,H. (2011) Substitution rate variation at human CpG sites correlates with non-CpG divergence, methylation level and GC content. *Genome Biol.*, **12**, R58.
55. Chuang,T.-J., Chen,F.-C. and Chen,Y.-Z. (2012) Position-dependent correlations between DNA methylation and the evolutionary rates of mammalian coding exons. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 15841.
56. Ziller,M.J., Gu,H., Müller,F., Donaghey,J., Tsai,L.T.Y., Kohlbacher,O., De Jager,P.L., Rosen,E.D., Bennett,D.A., Bernstein,B.E. *et al.* (2013) Charting a dynamic DNA methylation landscape of the human genome. *Nature*, **500**, 477–481.

SUPPLEMENTARY MATERIAL

Supplementary methods

Construction of plasmids for modular assembly

The backbone plasmid (pBackBone-BZ) was derived from pUC19 (Addgene plasmid #50005). First, the undesired BsaI restriction site in the beta-lactamase ORF was removed by site-directed mutagenesis. Next, an 1800 bp fragment containing the beta lactamase expression cassette and the origin of replication was amplified by PCR from pUC19 using primers pUC19-FW and pUC19-RE. The PCR product was joined, using Acc65I and AgeI restriction endonucleases, with a custom-synthesized gene fragment "Gblock-BB" amplified using primers C9seq1 and C9seq2 (Supplementary Table 4) and cut with type IIS restriction endonuclease Esp3I to generate compatible ends. SV40 origin of replication was removed by restriction with NgoMIV and re-ligation of the larger fragment. The final construct contained a DTS region (TF binding region from SV40) for efficient import of the plasmid into the nucleus (Dean, 1997) and a section with outward-facing BsaI restriction sites creating "B" and "Z" type ends for golden-gate assembly (Supplementary Figure S3). This section also contained an expression cassette for LacZα with lac promoter and the L3S1P13 synthetic terminator (Chen et al., 2013) for blue-white selection of undigested or re-ligated backbone vector during golden-gate assembly.

Modules were cloned into the pUK21 vector (Addgene plasmid #49788) or our derived vector pUK21gg (optimized for Golden Gate cloning using type IIS restriction endonucleases). Each module was amplified by PCR using primers with appropriate restriction sites compatible with pUK21 MCS (Supplementary Table 4). The primers also contained inward-facing BsaI recognition sequences that upon digestion leave appropriate four bp ends used for subsequent Golden Gate assembly (type dependent on module position; Supplementary Figure S3). All undesired BpiI, BsaI and Esp3I restriction sites were removed by site-directed mutagenesis (Supplementary Table 4).

Module for VPR domain was amplified from the plasmid SP-dCas9-VPR (Addgene plasmid #63798); catalytically inactive dSaCas9 was amplified from pX603-AAV-CMV::NLS-dSaCas9(D10A,N580A)-NLS-3xHA-bGHpA (Addgene plasmid #61594); EFS promoter was amplified from AAV:ITR-U6-sgRNA(backbone)-pEFS-Rluc-2A-Cre-WPRE-hGHpA-ITR (Addgene

plasmid #60226); fluorescence marker mRuby3 was amplified from the plasmid pKanCMV-mClover3-mRuby3 (Addgene plasmid #74252); puromycin resistance gene was amplified from lentiCRISPR v2 (Addgene plasmid #52961); TET1 domain was amplified from pJFA344C7 (Addgene plasmid #49236). Catalytically inactive TET1 (as negative control) was created by site-directed mutagenesis by introducing H1671Y and D1673A mutations in TET1 active site (see annotations in pFD-TET1_NCTRL plasmid sequence). All other modules were amplified from plasmids published in our earlier work (Vojta et al., 2016). Coding sequences for modules containing fluorescent and antibiotic resistance markers (PuroR, mRuby3, mClover3) were cloned using a streamlined strategy, with the NcoI restriction site around the start codon and the KasI site encoding the last two amino acids (Gly-Ala) of the coding sequence. All NcoI and KasI restriction sites were removed from the coding sequence of markers by introducing silent mutations by site-directed mutagenesis (Supplementary Table 4).

Destination plasmid for selection modules (fluorescent and antibiotic resistance markers) was pUK21_FP_entry that contains custom synthesized gene fragment pFP-Entry cloned between PstI and XhoI restriction sites in pUK21 vector lacking KasI restriction site. KasI restriction site was removed by inserting short oligonucleotide (pUK21_KasI-BamHI) between KasI and BamHI restriction sites in pUK21 vector. Fragment pFP-Entry contains eukaryotic promoter EFS, as well as prokaryotic Lac promoter for validation of fluorescence signal in animal and bacteria cells respectively. In addition, two BpiI restriction sites in fragment can later release fluorescence module under the Lac promoter for its insertion into gRNA modules and use in red-white selection. PCR amplified fluorescence modules mRuby3 and mClover3 were then cloned using NcoI and KasI strategy into pUK21_FP_entry, resulting with plasmids pUK21_FP_mClover3 and pUK21_FP_mRuby3, that can be used for the construction of single or dual marker module (described later).

For red-white selection, coding sequence for mRuby3 under the Lac promoter was cut from the finished mRuby3 module with BpiI restriction enzyme (described above) and inserted into the pSg-Sp module (gRNA module for SpCas9) within two BpiI restriction sites used for insertion of gRNA variable part. Module pSg-Sa (gRNA module for SaCas9) was amplified from the pX601-AAV-CMV::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA (Addgene plasmid #61591); original Esp3I sites for gRNA cloning were replaced with BpiI sites and mRuby3 cassette for red-white selection was added analogously to pSg-Sp.

Modules dSpCas9 and dSaCas9, initially made for C-terminal effector domain fusion, were repurposed for the better-performing N-terminal fusion. Specific Golden Gate ends “I” and “II” were changed into “II” and “III”. To do that, specific annealed oligonucleotide (N-C9), defining the new ends, was cloned between HindIII and XhoI restriction sites in the pUK21gg vector. Modules dSpCas9 and dSaCas9 were then cut out using BsaI and cloned into two BpiI restriction sites in pUK21gg_N-C9.

Functional modules for N-terminal fusion with dCas9 module were made analogously. Initially, ends of types “II” and “III” were changed into types “I” and “II” by cloning DNMT3A, TET1 and VPR modules into two BpiI restriction sites in pUK21gg_N-FD.

In C-terminal fusions with dCas9, the nucleoplasmin NLS might be covered by effector domain (DNMT3A). Thus, to test the effect of additional nucleoplasmin NLS on C-terminus of DNMT3A effector domain, specific annealed oligonucleotide FD_c-NLS-NP was cloned between KasI and XhoI restriction sites in the pUK21gg vector. The oligonucleotide contains two Esp3I restriction sites for DNMT3A cloning along with nucleoplasmin NLS downstream from cloning site. DNMT3A effector domain was then cut out from finished module vector, initially made for C-terminal fusion with dCas9, with BsaI restriction enzyme and cloned into pFD_c-NLS-NP vector.

Construction of the multi-guide system

First, the "individual" modules for golden-gate assembly of gRNAs at positions 1-6 in the multi-guide system were created as follows. Undesired NgoMIV and XhoI restriction sites were removed from the plasmid pFUS_A (Addgene plasmid #31028) and a new NgoMIV restriction site was introduced downstream from the multiple cloning site by site-directed mutagenesis (Supplementary Table 4). Multiple cloning site from the plasmid pUK21 (Addgene plasmid #49788) was cut out with SapI and NgoMIV and inserted into altered pFUS_A, yielding the plasmid pUS21gg. Six different pairs of annealed oligonucleotides (sgM1-6) were then cloned into the KasI and XhoI restriction sites, each carrying two Esp3I restriction sites that create different four nt 5' protruding ends labelled with roman numerals I to VII, corresponding to the ends B-Z in the system for assembly of fusion constructs. An adapter (XbaI_B_C9seq1_A_NcoI) was oligo-annealed and inserted between XbaI and NcoI sites, facilitating BsaI cloning of the finished gRNA cassette (including mRuby3 expression for red-white selection) from the modules pSg-Sa and pSg-Sp, previously created

for single gRNA cloning in the system for dCas9 fusions (described in previous section). Two sets of six different plasmids for cloning of individual gRNA molecules for SaCas9 and SpCas9 were made (pSgMxA represent plasmids for SaCas9 gRNA molecules while pSgMxG represent plasmids for SpCas9 gRNA molecules, x = (1, 2, 3, 4, 5, 6)), that enable assembly (using Esp3I type IIS enzyme) of up to six different gRNA modules, each carrying its gRNA expression cassette.

Next, "multi-guide" modules were created, which replace the gRNA module in the assembly of core fusion constructs and allow cloning of one to six "individual" gRNA modules of the multi-guide system. Kanamycin resistance gene and origin of replication were amplified from pUK21 (Addgene plasmid #49788) with primers that enabled ligation with the multiple cloning site cut out from the same plasmid with NgoMIV and SapI, which resulted in the removal of two undesired Esp3I restriction sites. The third Esp3I restriction site (located within kanamycin resistance gene) was removed by site-directed mutagenesis (Supplementary Table 4), yielding the plasmid pUK21gg. A pair of annealed oligonucleotides (SgMult) containing one restriction site for Esp3I that defines the cohesive end "I" needed for the assembly of the first gRNA module and BsaI restriction site that defines the sticky end named “B” needed for the assembly of functional construct for epigenetic modulation was cloned within KasI and NcoI restriction sites. Six different pairs of annealed oligonucleotides (SgMx1-6), each containing Esp3I restriction site defining the other sticky end named “2-7” that determines the capacity of accepting up to six different gRNA modules assembled and BsaI restriction site that defines the sticky end named “A” are then cloned within NcoI and XhoI sites. In the final step, mRuby3 fluorescence marker under the Lac promoter was cut out from the plasmid containing mRuby3 module (described in previous section) with BpiI and cloned between two BpiI restriction sites that were introduced with the cloning of first pair of annealed oligonucleotides (SgMult). The mRuby3 marker was then used for red-white selection of correct gRNA modules following assembly with Esp3I restriction enzyme. The final products of this step were plasmids (pSg-x1 to pSg-x6) that have capacity to receive up to six gRNA modules with Esp3I-mediated assembly.

Dual marker modules

Plasmids for the dual marker system (Figure 3C in the main text) at the first position with ends "III" to "X" were generated from the corresponding modules with ends of type "III" and

"IV" by amplifying the T2A region of the plasmid pUK21_FP_T2A (not deposited) using primers T2A_X-FW and T2A_X-RE (Supplementary Table 4), which added the non-complementary extensions needed to convert the end "IV" into "X". The PCR product was then cloned between HindIII and XhoI restriction sites into the plasmid pUK21_noKasI (lacking KasI restriction site; described in section: Construction of plasmids for modular assembly) and the intermediary plasmid was named pM2-FP_T2A-X.

For the second position, an "empty" plasmid pM2-FP_X-P2A was generated by amplifying custom-synthesized DNA fragment Gblock-BB (also used for backbone plasmid construction; described in section: Construction of plasmids for modular assembly) using primers C9seq3 and C9seq4 and cloning the PCR product into the pUK21_FP_T2A vector between the NcoI and HindIII restriction sites.

Marker coding sequences for antibiotic resistance or fluorescent proteins (from previously constructed modules) were subsequently inserted into pM2-FP_T2A-X and pM2-FP_X-P2A using the streamlined cloning strategy: coding sequences were flanked by NcoI and KasI restriction sites, which were used to seamlessly shift the existing coding sequences into new empty module vectors.

Oligo cloning of variable parts into gRNA modules

Variable parts of gRNA were cloned into appropriate modules essentially as described in (Cong et al., 2013). Briefly, oligonucleotides designed to form dsDNA with overhangs compatible with gRNA modules (Supplementary Figure S4) were custom-synthesized. They were phosphorylated at the 5' end and annealed in a reaction containing 100 pmol of each oligonucleotide, 1× T4 ligation buffer (TaKaRa) and 5 U of T4 polynucleotide kinase (NEB). Phosphorylation was done at 37°C for 30 min, followed by denaturation at 95°C for 5 min. Oligonucleotides were annealed in a thermocycler by gradual decrease of temperature by 5°C per minute, from 95°C down to 25°C. Phosphorylated and annealed oligonucleotides were then cloned into module plasmids in a single reaction containing 1× Buffer G (Thermo Fisher Scientific), 0.5 mM DTT (Thermo Fisher Scientific) 0.5 mM ATP, 350 U of T4 DNA Ligase and 10 U of type IIS restriction enzyme BpiI (Thermo Fisher Scientific). Reaction conditions consisted of six cycles at 37°C for 5 min and 23°C for 5 min. Exonuclease V treatment was performed by directly adding 10 U of enzyme into the reaction mix along with additional 0.5 mM ATP and incubating for 30 min at 37°C to remove any remaining linear DNA. After

bacterial transformation, white colonies were selected for further verification (red colonies represented uncut or empty backbone vector).

Testing of different linkers between dSaCas9 and effector domains

The length between effector domains linked to the C-terminus of *Staphylococcus aureus* Cas9 was varied to find the optimal spacing that would enable the fusion construct to access the target DNA. First, dSaCas9 along with C-terminal nucleoplasmin NLS was amplified from pX603-AAV-CMV:NLS-dSaCas9(D10A,N580A)-NLS-3xHA-bGHpA (Addgene plasmid #61594). Alternatively, dSaCas9 with three tandem HA epitope tags located downstream from the nucleoplasmin NLS, which extend the spacing between dSaCas9 and effector domain, was amplified from the same plasmid (Supplementary Table 4). Both versions of dSaCas9 were then cloned into pUK21 (Addgene plasmid #49788) between XbaI and XhoI restriction sites in a version of modules for fusions of effector domains at the C-terminus of dSaCas9.

To test trip-zip linker (TZ2) (Cochran, Skelton, & Starovasnik, 2001), it was added to the dSaCas9 module by oligo annealing of primers LNK_tripzip2_G4Sx2-S and LNK_tripzip2_G4Sx2-A and cloning into a dSaCas9 module (for assembly of effector domains fused to the C-terminus of dCas9) using BamHI and XhoI restriction enzymes, thus generating SaCas9_TZ2_2xG4S module (not deposited).

Confirming dual transfection using fluorescence markers

We monitored transfection of HEK293 and BG1 cells with plasmids for expression of fusion proteins by observing red mRuby3 and yellow-green mClover3 (Bajar et al., 2016) fluorescence of the fused fluorescent proteins, translated in the same reading frame as the dCas9 fusion construct and linked via self-cleaving 2A peptides. One day after transfection, images were acquired under the same conditions using an Olympus IX73 microscope. Cells positive for mClover3 and mRuby3 fluorescence were counted using Object Count tool in Olympus cellSens Standard software. Fluorescence was imaged using Olympus filter sets U-FRFP (mRuby3) and U-FYFP (mClover3), which ensured good separation of fluorescent signals.

Plasmids for whole-genome methylation analysis

The SV40 promoter and terminator were amplified from the plasmid pLVET-tTR-KRAB (Addgene Plasmid #11644) and cloned into the pUK21 (Addgene plasmid #49788) multiple cloning site. Both pairs of primers contained BsaI restriction sites that can release modules from pUK21 with specific four nt overhangs for joining with the puromycin resistance module. For that purpose, we changed the puromycin resistance module type "III" overhangs into type "I" to enable its BsaI assembly with SV40 promoter, while other overhang of type "IV" is necessary for assembly with terminator. A pair of annealed oligonucleotides (M14) that contained two BsaI restriction sites defining the overhangs type "I" and "IV" was cloned between KasI and XhoI restriction sites in pUK21. Puromycin resistance module was then cut out from the finished plasmid (as described in Construction of plasmids for modular assembly) with NcoI and KasI and cloned into the pUK21_M14.

To create a backbone with the secondary puromycin expression cassette, we inserted a pair of annealed oligonucleotides (BB_2nd_Cassette) into the plasmid pBackBone-BZ upstream of the DTS nuclear import sequence between AatII and SacI restriction sites. The insert contained two Esp3I sites needed for separate assembly of SV40 promoter, puromycin resistance gene and SV40 terminator. For that purpose, modules were cut out with BsaI, while the backbone was cut with Esp3I, which left compatible overhangs for assembly of the secondary expression cassette. The prepared backbone and module fragments were gel-purified and ligated.

Four different gRNA molecules for *IL6ST* and five for *MGAT3* locus (Supplementary Table 3) were cloned using the multi-guide system and the final BsaI assembly was done with both the standard backbone (pBackBone-BZ) and the backbone with she secondary puromycin expression cassette (described above).

Supplementary material describing Illumina 850k analysis

Pre-processing and quality control

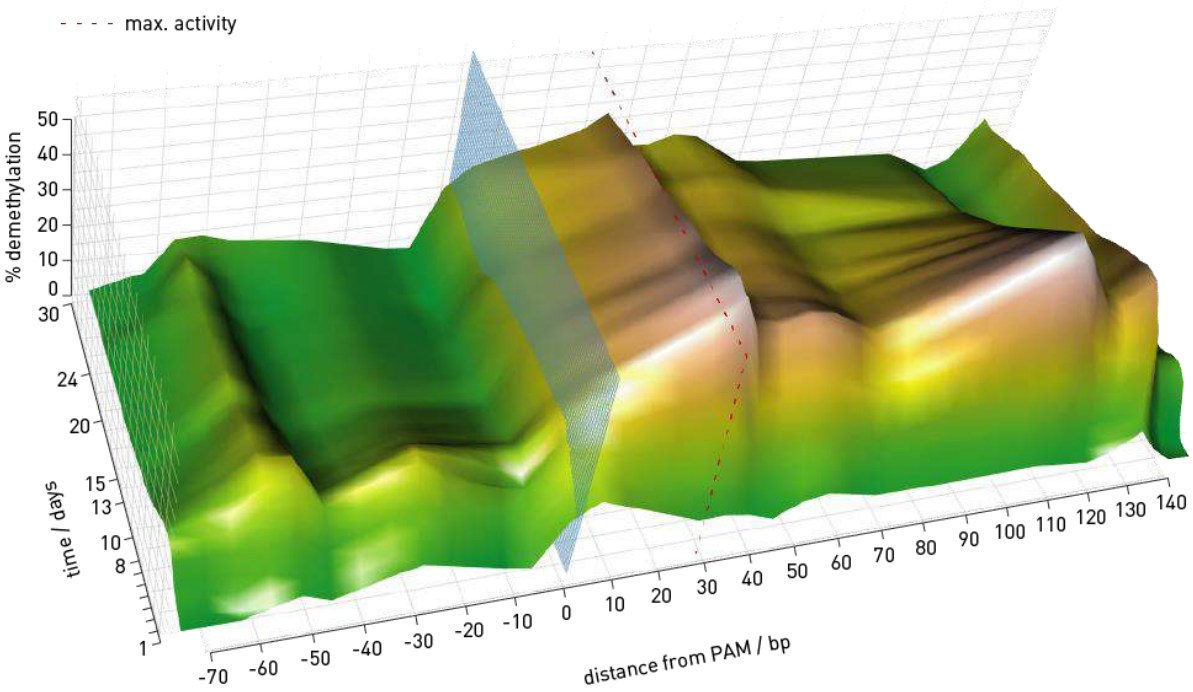
Preliminary quality control indicated that all samples were of good quality (Supplementary Figure S10A), and that there were no great differences in red and blue signal intensities between the samples (Supplementary Figure S10B). After normalization, multidimensional scaling was used to visualize the similarity of individual cases within the dataset. Here, it was observed that there was a horizontal effect segregating biological replicates 1 and 2, as well

as a vertical effect segregating the TET1-dSpCas9 and TET1-dSpCas9 secondary cassette samples from the samples in the Mock, DNMT3A-dSpCas9 and DNMT3A-dSpCas9 secondary cassette groups. The variation between the biological replicates could be attributed to the fact that each set of technical (and hence biological) replicates were run on separate chips (Sentrix_IDs). Indeed, PCA analyses indicated that Sentrix_ID, Biological Replicate and Technical Replicate contributed significantly to principal component 2 (Supplementary Table 7). These effects were abrogated after SVA correction, as evidenced by the PCA analyses (Supplementary Table 8) and by the multidimensional clustering of samples by experimental condition rather than by biological/technical replication (Supplementary Figure S10 C and D, that is, the horizontal effect segregating biological replicates disappeared after SVA

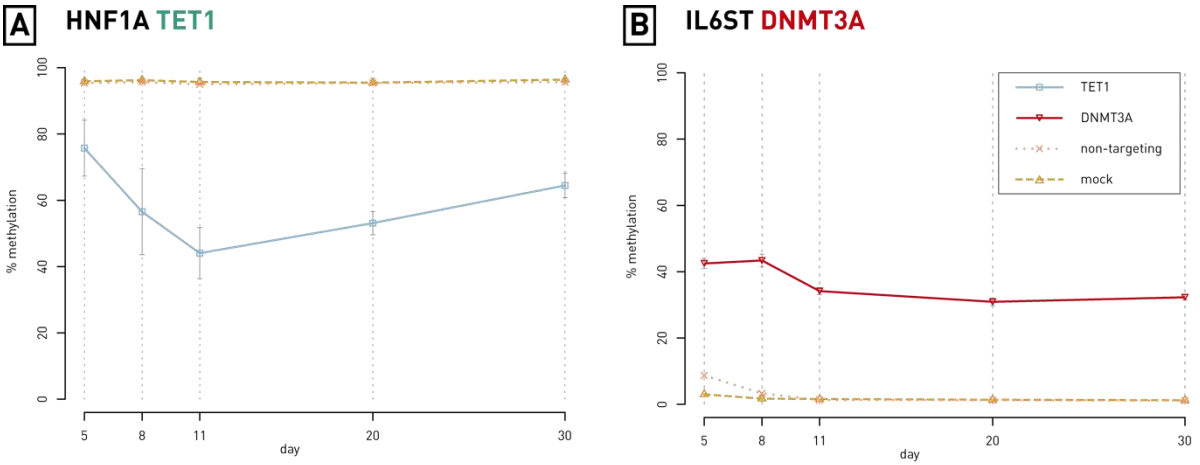
Methylome-wide variations by DNMT3A-dSpCas9 and TET1-dSpCas9

Unsupervised multidimensional scaling of methylome-wide variations identified three distinct clusters, clearly distinguishing each of DNMT3A-dSpCas9, TET1-dSpCas9 and mock-transfected cells, respectively (Supplementary Figure S10D). Moreover, the constructs with DNMT3A-dSpCas9 expressed under a strong (primary cassette, marker expressed fused via 2A self-cleaving peptide) or a weak promoter (secondary cassette, marker expressed separately under a strong promoter) were clearly distinguishable within their relevant cluster, and this was also seen, though to a lesser extent, with the primary and secondary constructs of TET1-dSpCas9 (Supplementary Figure S10D). Biological and technical replicates clustered together, showing low experimental variation (Supplementary Figure S10D). Overall, the largest observed methylome variations were induced by each of DNMT3A-dSpCas9 and TET1-dSpCas9 relative to mock, with lesser but evident effects seen between the constructs expressed from strong vs. weak promoter. Those effects were stronger than the basal experimental variations between technical or biological replicates.

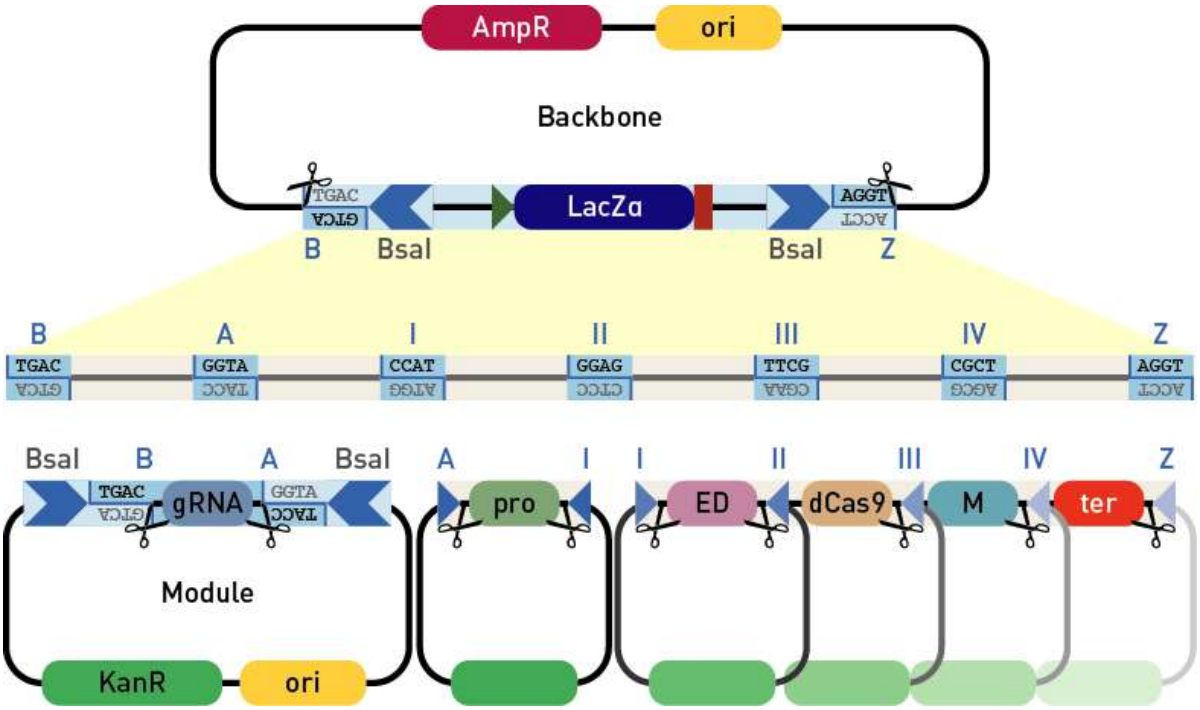
Supplementary figures



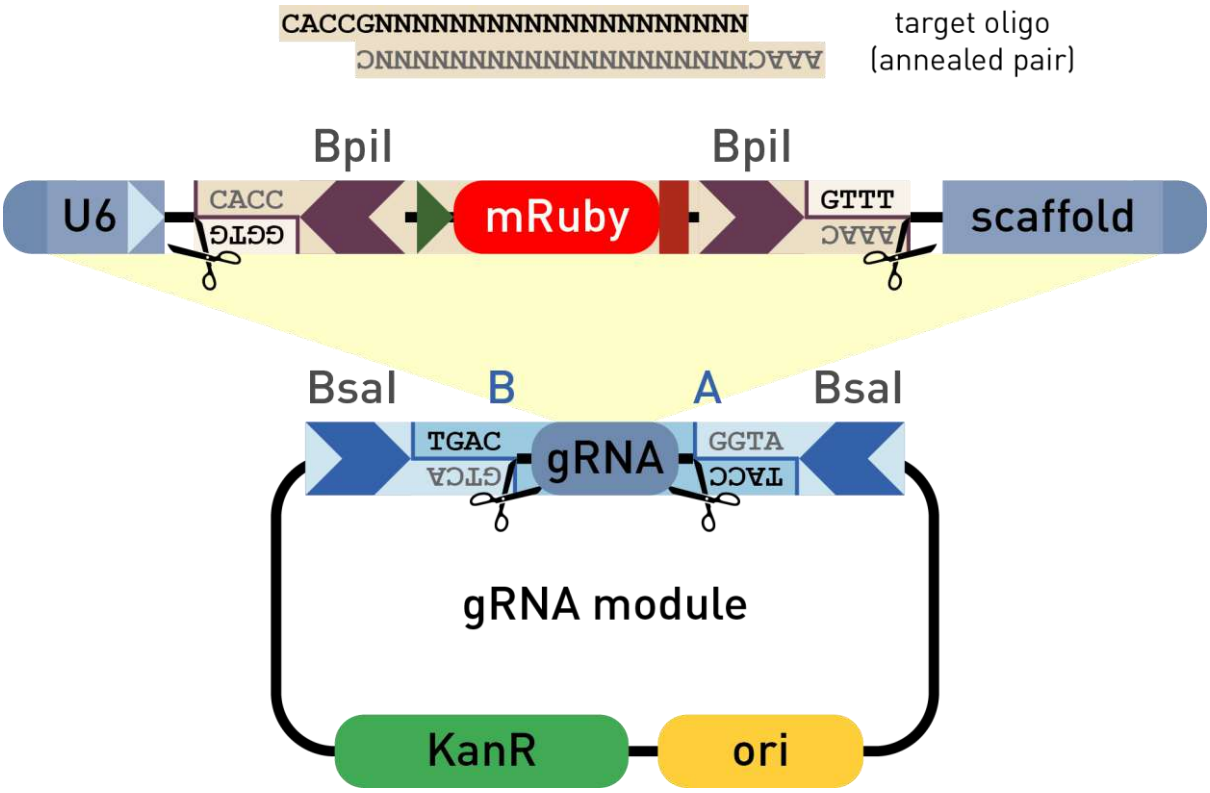
Supplementary Figure S1. Time course of dSpCas9-TET1 demethylation activity. Data from several experiments – loci *MGAT3* (2 assays: A1 and A2, 9 + 5 CpG sites; 8 gRNAs) and *LAMB1* (one assay, 6 CpG sites; 2 gRNAs) – provide a comprehensive picture of TET1 activity when fused to C-terminus of dSpCas9. The *x*-axis shows distance from gRNA binding site (oriented relative to PAM sequence, represented by the blue plane at position zero), while the *y*-axis shows time in days; the *z*-axis ("height") corresponds to CpG demethylation activity at a given position at selected time points. The activity peak is clearly distinguishable at about 30 bp from the PAM sequence, with minor other satellite peaks some distance apart, which might represent contact with adjacent nucleosomes. Thirty days following transfection, the main peak remained fairly stable, while the satellite peaks reverted almost to their original methylation level.



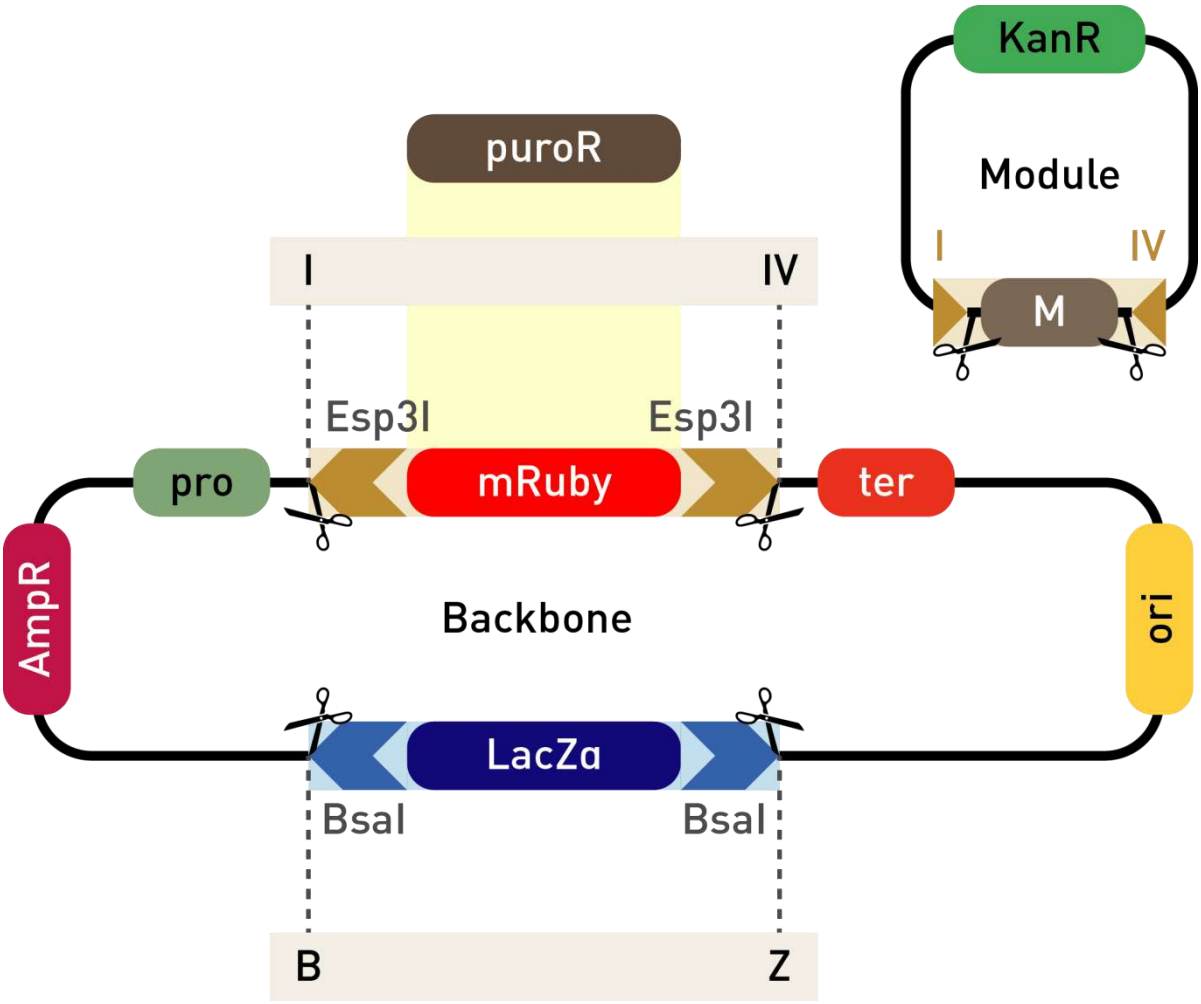
Supplementary Figure S2. Time course of targeted methylation and demethylation. Effect of demethylation of *HNF1A* by TET1-dSaCas9 (A) and methylation of *IL6ST* by DNMT3A-dSpCas9 (B) was followed during 30 days; most of the change in methylation persisted throughout the whole period.



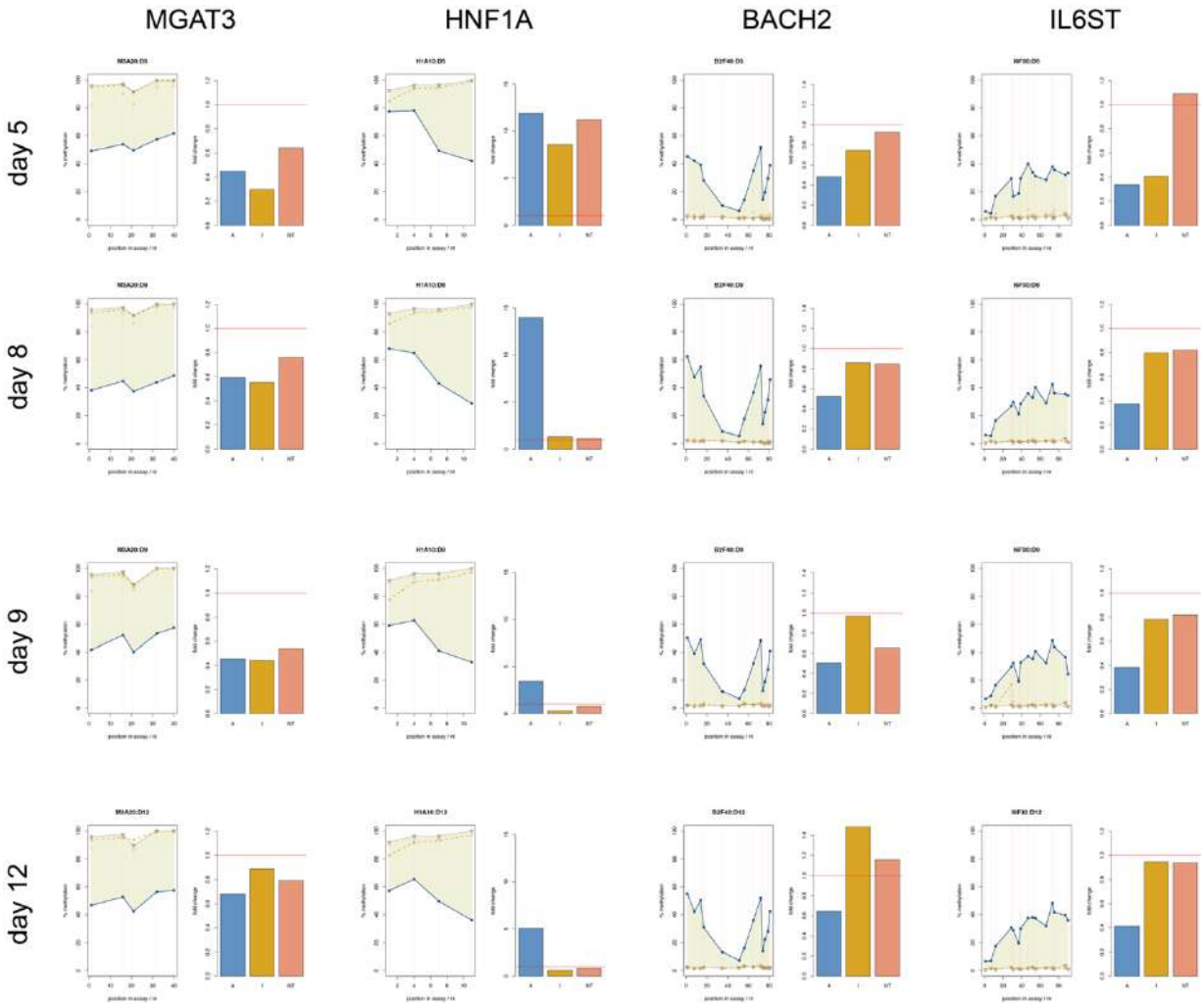
Supplementary Figure S3 Golden Gate assembly of individual modules into a cassette for eukaryotic expression. The components are joined via non-palindromic four nt cohesive ends generated by type IIS restriction enzymes (Golden Gate cloning), which enables efficient assembly of up to eight DNA fragments. The backbone plasmid contains a lacZα cassette for blue-white selection of bacterial clones with uncut/re-ligated backbone, while bacteria taking up only module plasmids are eliminated by counter-selection with ampicillin. White colonies contain the correctly assembled eukaryotic expression cassette for the selected dCas9 fusion construct. If the gRNA module has no pre-cloned variable part or accepts multiple modules for a second round of multi-guide assembly, the positive colonies are red and the gRNA (or multi-guide) cloning is facilitated by red-white selection.



Supplementary Figure S4. The 20 bp gRNA variable region is cloned into the module vector by oligo annealing. Every individual gRNA module contains a U6 promoter, gRNA scaffold for either SaCas9 or SpCas9 and a fragment to be excised by the type IIS endonuclease BpiI at the 5' end of the gRNA scaffold, which prepares the vector for accepting the 20 bp variable gRNA fragment (inserted by oligo cloning) conferring specificity for a particular genomic region. The excised part encodes a bacterial cassette for expression of the mRuby3 fluorescent protein, which gives colonies a red color thus facilitating red-white selection. Alternatively, the gRNA module can be substituted for a multi-guide module accepting one to six gRNA modules for a second round of assembly.

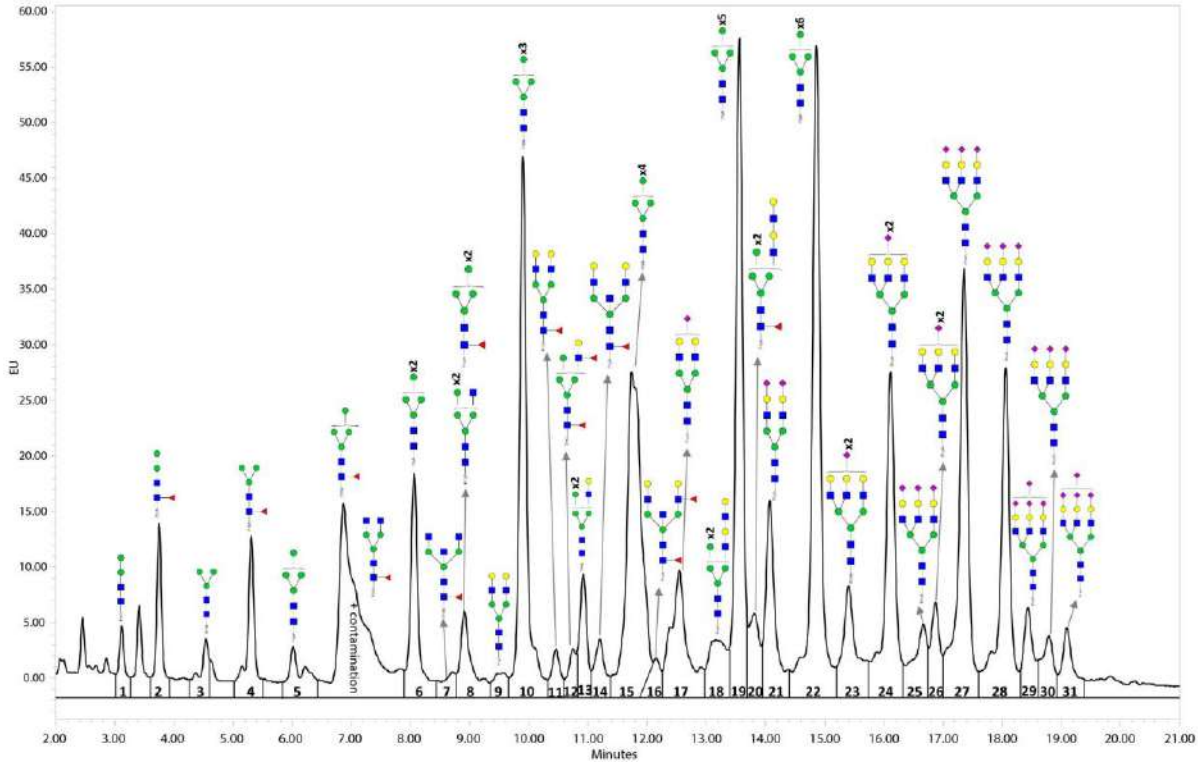


Supplementary Figure S5. Secondary cassette for expression of selection markers separately from the dCas9/ED fusion. Alternative backbone accepts an additional module for the marker protein via Esp3I Golden Gate cloning. Afterwards, the marker is expressed under a strong (SV40) promoter independently of the main fusion construct assembled using the standard protocol. This enables strong selection (especially by puromycin) while giving the flexibility to choose a weaker promoter for the main dCas9 fusion cassette, thereby enabling fine tuning needed for controlling the off-target effect.

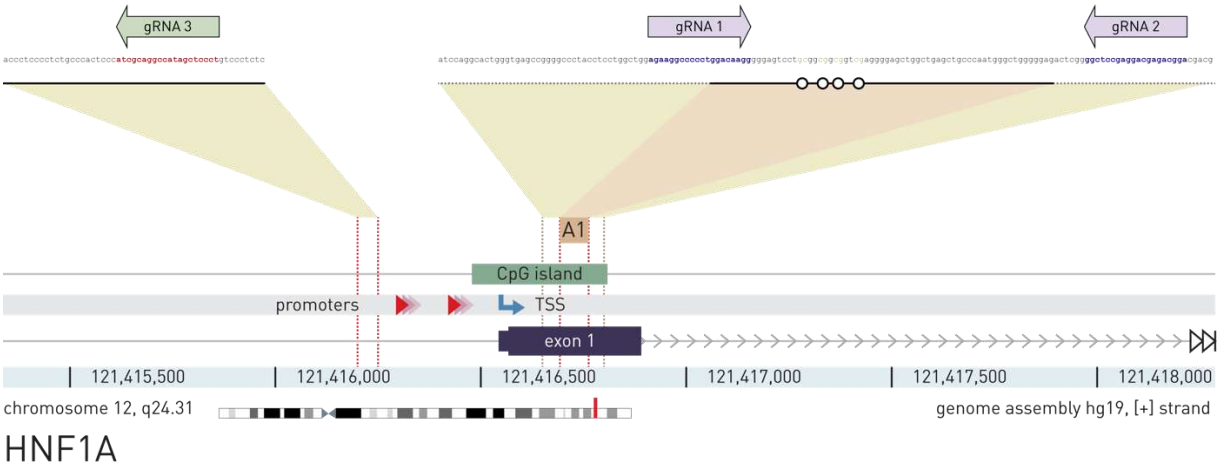


Supplementary Figure S6. Time course of simultaneous methylation and demethylation of two loci using dCas9-based tools with antagonistic activities in HEK293 cells.

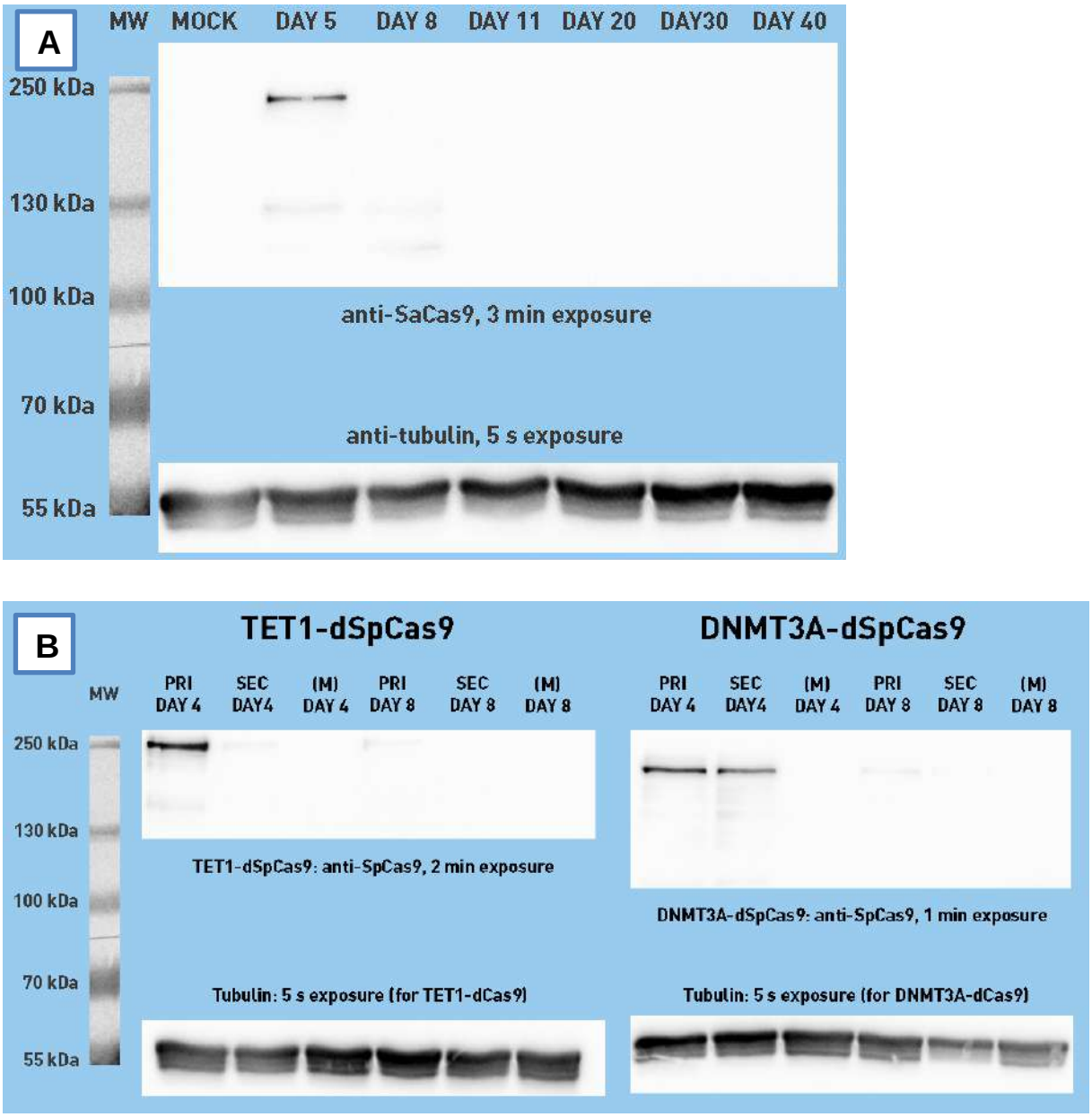
Time course experiment shows how altered gene expression reflects changes in CpG methylation levels. We confirmed temporal activity profiles of both DNMT3A and TET1 fusions to dCas9 (both dSpCas9 and dSaCas9) to be similar to the profiles obtained for dSpCas9-DNMT3A alone (Vojta et al., 2016): a rapid rise of methylation activity after transfection up to day 8, with slowly diminishing effect afterwards. The expression profile closely followed CpG methylation profile. Thus, with this experiment we verified that the peak of methylation and consequent change in expression falls at the day 8 after transfection. A – dCas9 fusions with active catalytic domain; I – dCas9 fusions with inactive catalytic domain; NT – active dCas9 fusions with non-targeting gRNA.



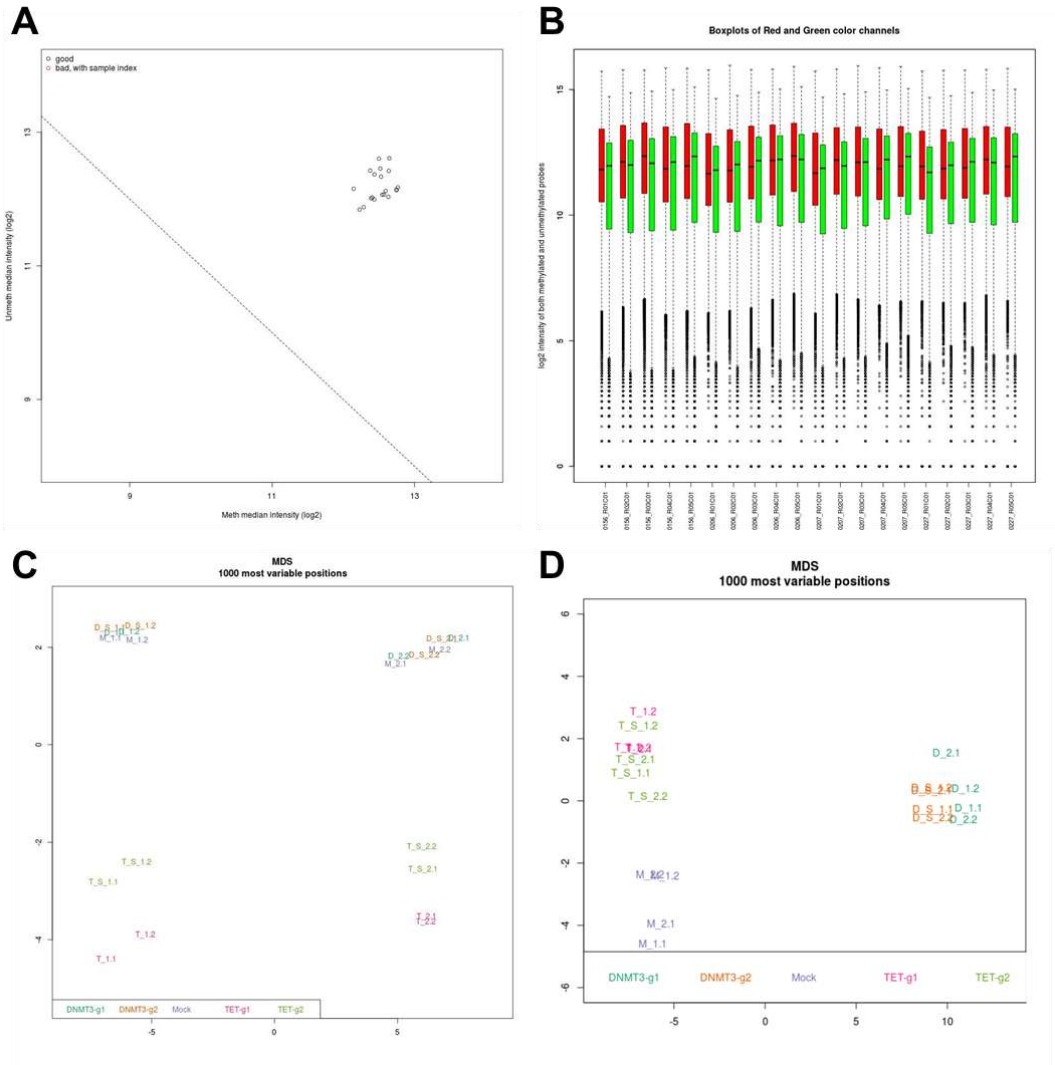
Supplementary Figure S7. Annotated glycan peaks. Each glycan peak in the chromatogram has been assigned corresponding structures based on mass spectrometry analysis. See Supplementary Table 6 for glycan peak annotations.



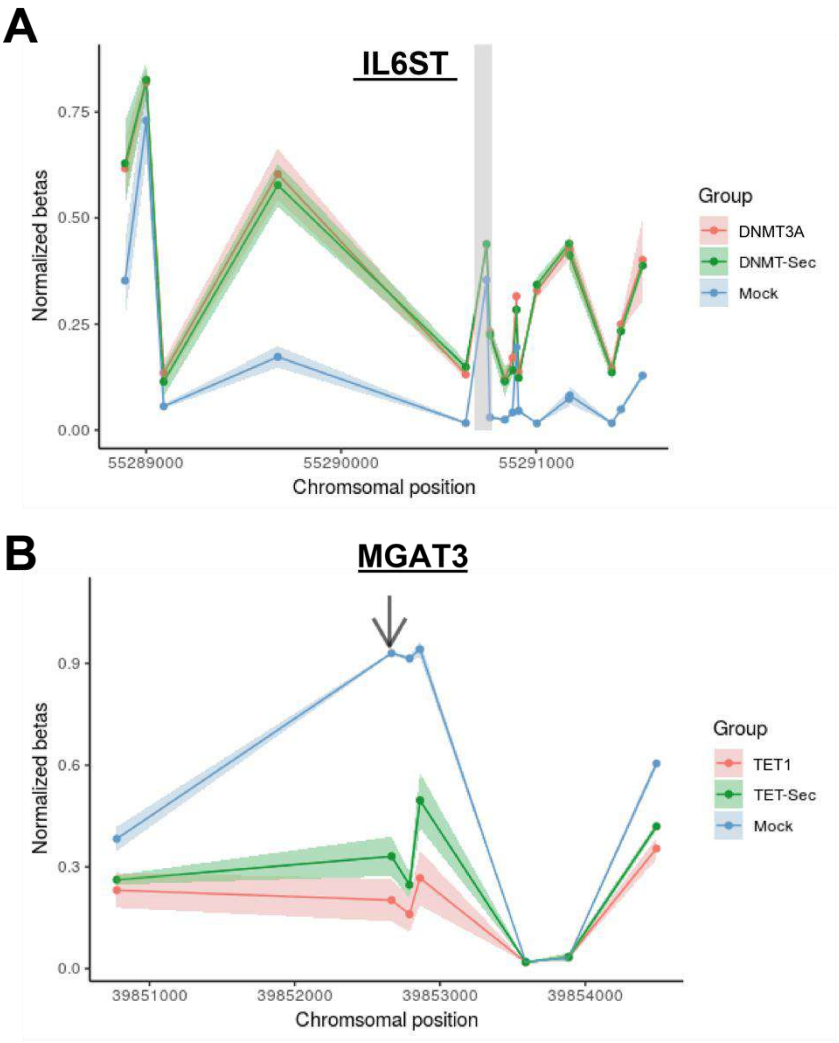
Supplementary Figure S8. Position of gRNAs used for synergistic activation of *HNF1A* by TET1-dSaCas9 and VPR-dSpCas9. The lavender colored arrows show positions of gRNA1 and gRNA2 for guidance of TET1-dSpCas9, while the green arrow shows the gRNA3 for guidance of VPR-dSpCas9. Arrows point in the direction of the PAM sequence. A1 is the pyrosequencing assay for *HNF1A*.



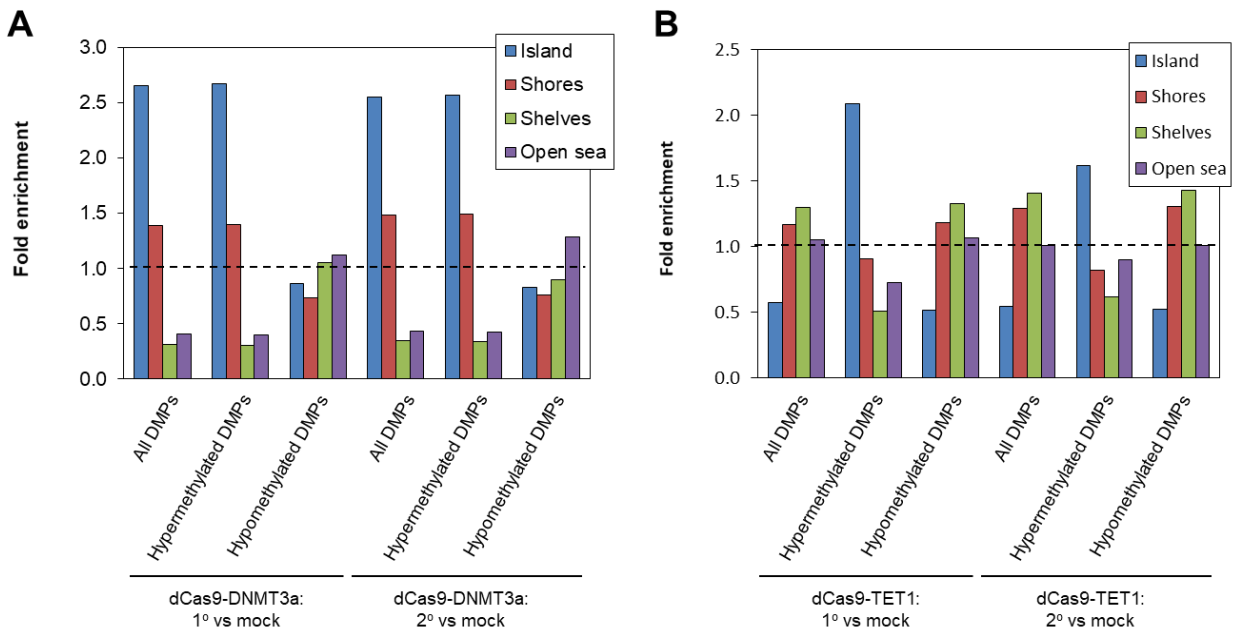
Supplementary Figure S9. Original western blots for protein presence (dCas9) in the time course (A) and "strong vs. weak promoter" (CBh vs. EFS) (B) experiments. The presented pictures were used for quantitation of protein presence; each lane was normalized to tubulin in order to control for variation in gel loading. (A) The time course shows that no protein was detectable by day 11 after transfection. (B) Stronger expression of the "primary cassette" construct (dCas9 fusions under CBh promoter, marker in frame joined by 2A peptides) compared to the "secondary cassette" construct (dCas9 fusions under the weak EFS promoter, marker under the stronger SV40 promoter) is apparent at both timepoints (days 4 and 8 after transfection). After quantitation, the ratio of expression from primary to secondary cassette is: for TET1-dSpCas9, 32.6 on day 4 and 6.8 on day 8; for DNMT3A-dSpCas9, 1.4 on day 4 and 1.2 on day 8.



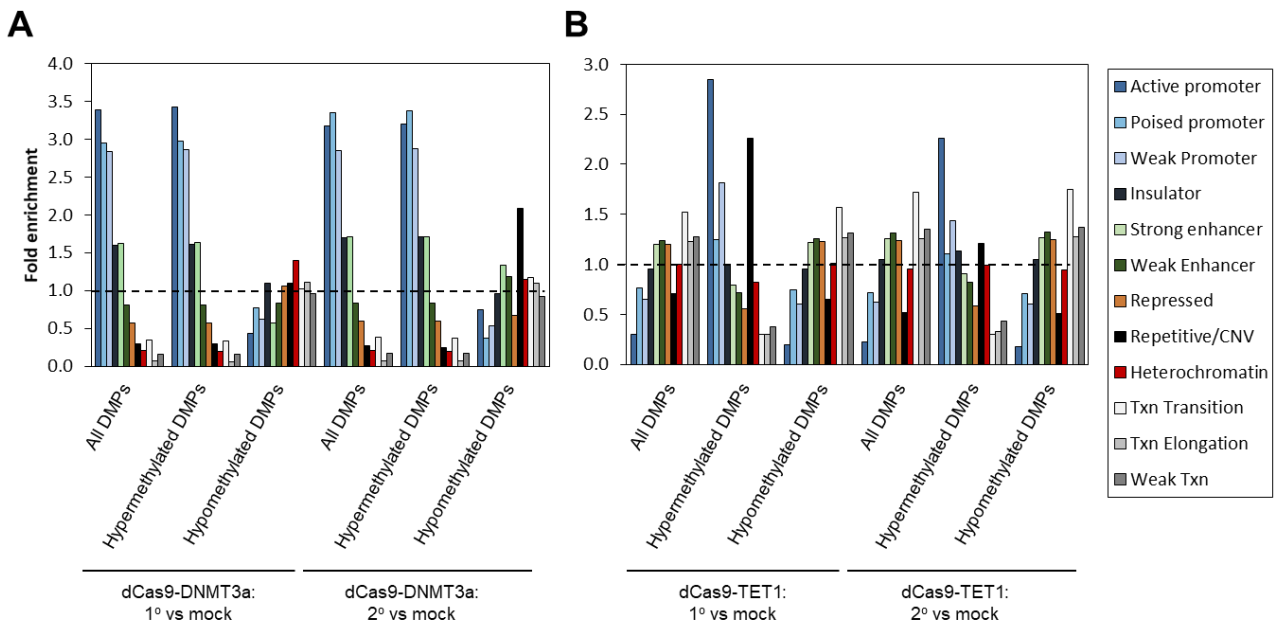
Supplementary Figure S10. Pre-processing and quality control plots. (A) QC plot indicating the log median intensity in both methylated and unmethylated channels. High median intensities and clustering of samples is an indication of good quality. (B) Boxplot showing the spread of log2 intensity in both methylated and unmethylated channels. (C and D) Multidimensional scaling (MDS) plots (C) before and (D) after SVA correction.



Supplementary Figure S11. Targeted modification to DNA methylation patterns can be induced by DNMT3A-dSpCas9 and TET1-dSpCas9 constructs. (A) Comparison of DNA methylation patterns induced by DNMT3A-dSpCas9 primary and secondary cassettes on probes 2500 bp up- and downstream of the targeted region (shaded in grey) of *IL6ST*. (B) Comparison of DNA methylation patterns induced by TET1-dSpCas9 primary and secondary cassettes 2500 bp up- and downstream of the targeted region of *MGAT3*. The probe cg21461856 is the only probe lying within this region (indicated with an arrow). Normalized betas were used to plot the above graphs, as opposed to SVA-corrected betas in Supplementary Figure S10.



Supplementary Figure S12. Distribution of DMPs by relationship to CpG islands for (A) DNMT3A-dSpCas9 primary and secondary cassettes and (B) TET1-dSpCas9 primary and secondary cassettes. Fold enrichment was calculated by dividing the fraction of DMPs in each annotation category against the fraction of all filtered probes mapped to the same category, correcting for differences in representation of each annotation category on the Infinium MethylationEPIC array.



Supplementary Figure S13. Distribution of DMPs across various regulatory regions for (A) DNMT3A-dSpCas9 primary and secondary cassettes and (B) TET1-dSpCas9 primary and secondary cassettes. Fold enrichment was calculated by dividing the fraction of DMPs in each annotation category against the fraction of all filtered probes mapped to the same category, correcting for differences in representation of each annotation category on the Infinium MethylationEPIC array.

Supplementary Table 1. Sequences of pyrosequencing assays used for methylation analysis

Assay name	Assay sequence 5'→3' (analyzed CpG sites are underlined)
MGAT3-A1	**CG**CATCCCTGCACCTT**CG**A**CG**ATGG**CG**G**CG**CAGAGATGTCTGCTG**CG**TACCCACAAT GCCTTGTGCCT**CG**CAC**CGCG**GGA
MGAT3-A2	CCTAGAGCAAGGCCA**CG**AGGAGCCAGGGCA**CG**ACA**CG**GTGGGCCCT**CG**GAGAAC**C G**CTG
MGAT3 A-TIS	**CGCG**C**CGCG**CT**CG**GCCT**CG**CCTC**CGCG**CCCCC**CGCGCG**CC**CG**C**CGCG**GTCCCTCCCC**C GCG**CCCGTC**CG**CT**CG**C**CG**GGCCCC**CG**C**CG**C**CG**CCC**CG**GGGTGCAGC**CG**AG**CG**GC**CG CG**C**CG**GGTCCC**CG**GGA**CG**GGGTG
BACH2-A1	G**CG**GGTGA**CG**TCAG**CG**C**CG**AATGTCAACAATGTAG**CG**ATTGAGAGTGTGGG**CG**TTC **CG**GGGAGAG**CG**CAAGC**CGCGCG**G**CGCG**
BACH2-A2	AA**CG**GGAAGGGGGTGGGAGC**CG**AGC**CG**G**CG**GC**CGCG**C**CGCG**C**CG**AATGTGTGCTC TCCCTCC**CG**GCTGTTTA**CG**G**CG**GC**CG**AGG**CG**
HNF1A-A1	TG**CG**G**CG**G**CG**GT**CG**AGG
IL6ST-A1	**CG**GAGC**CG**GGG**CG**AGCAGCCAAAAGGCC**CGCG**GAGT**CGCG**CTGGGC**CG**CCC**CG**G**C G**CAGCTGAAC**CG**GGGGC**CGCG**CCTGCCAGGC**CG**A**CG**GG

Supplementary Table 2. Sequences of PCR primers and pyrosequencing primers

Primer name	Sequence (5' → 3')	Use
MGAT3-A1-Fw	GTTGGGATATAGAATAGGTAG	PCR amplification of MGAT3-A1 fragment, forward primer was used for pyrosequencing
MGAT3-A1-Rev	[Btn]ACCATTCCTCTCAAAACTCA	
MGAT3-A2-Fw	GTTTTTGAGTTTTGAGAGGAATGG	PCR amplification of MGAT3-A2 fragment, forward primer was used for pyrosequencing
MGAT3-A2-Rev	[Btn]ACCCTCTTAAACCTACTCTCCTAC	
MGAT3-F1_BS-alt-v2	[Btn]GGAGGGGGTTGYGGAGGGGG	PCR amplification of MGAT3 A-TIS fragment, with only one biotinylated primer (forward or reverse)
MGAT3-R_BS	[Btn]CCTCCCCCCACCCCCACTTC	
MGAT3-F1_BS-v4	GGAGYGGGTATTTTTG	Pyrosequencing of MGAT3 A-TIS
MGAT3-R_BS	CCTCCCCCCACCCCCACTTC	
LAMB1-A1-Fw	GAAGTGGAGGGTTTAT	PCR amplification of LAMB1-A1 fragment
LAMB1-A1-Rev	[Btn]ATCAAATCTATCCAACAA	
LAMB1-A1-seq	TTGATTAGGGTGGG	Pyrosequencing of LAMB1-A1
HNF1A-A1-Fw	GGATAAGGGGGAGTTTTG	PCR amplification of HNF1A-A1 fragment
HNF1A-A1-Rev	[Btn]CTCCCCCAACCCATTAAA	
HNF1A-A1-seq	AAGGGGGAGTTT	Pyrosequencing of HNF1A-A1
BACH2-A1-Fw	TTATTGTGAATGGGGA	PCR amplification of BACH2-A1, forward primer was used for pyrosequencing
BACH2-A1-Rev	[Btn]ACTACTACTACTAAAAC	
BACH2-A2-Fw	GTTTTTATGGTATTTTTTAGG	PCR amplification of BACH2-A2, forward primer was used for pyrosequencing
BACH2-A2-Rev	[Btn]TCCCTCTACTATTCCAAAA	
IL6ST-A1_Fw	GAGAAGGATTTGATAGTGT	PCR amplification of IL6ST-A1 fragment
IL6ST-A1-Rev	[Btn]CCTCCTCACCTCAAAC	
IL6ST-A1-seq	AAGGATTTGATAGTGTTT	Pyrosequencing of IL6ST-A1

Supplementary Table 3. Sequences of gRNA molecules for each experiment.

gRNA molecule	Target site sequence (5'→3') including PAM (underlined)	Experiment
MGAT3-sg01	CATTCGCTGGGATATAGAATAGG	Activity profile of C-terminal fusion dSpCas9-TET1
MGAT3-sg02	CCCTGCACCTTCGACGATGGCGG	
MGAT3-sg03	ATGCCTTGTGCCTCGCACCGCGG	
MGAT3-sg04	CTGCTCTGTAGGCCCCAGAACGG	
MGAT3-sg05	GGACGCCTCTGAGCCCTGAGAGG	
MGAT3-sg06	TGGCCTAGAGCAAGGCCACGAGG	
MGAT3-sg07	CGGCACCGTGCACACATCACAGG	
MGAT3-sg08	AATCCCGGCCCAGGTTCACGCGG	
LAMB1-sg01	ACACATCCACCCTTTGTTGGGGG	
LAMB1-sg02	AGCAGCGAGAGCCTCCCTCCCGG	
NT-gRNA	GTAGGCGCGCCGCTCTCTAC	
BACH2-F3f	GCATTTTCTAGGAACGGGAAGGG	Localization by NLS
MGAT3-sg03	ATGCCTTGTGCCTCGCACCGCGG	Time-course evaluation of C-terminal fusion dSpCas9-TET1
LAMB1-sg01	ACACATCCACCCTTTGTTGGGGG	
IL6ST-gRNA1	GCCACCCCAGTCCCGCGGCGGGG	Time-course evaluation of N-terminal fusion DNMT3A-dSpCas9
IL6ST-gRNA2	ATCTGACAGTGTTCCGGAGCCGG	
IL6ST-gRNA3	CGCACGAACCCCTTGGCGCCAGG	
IL6ST-gRNA4	GCCAAGGGGTTCGTGCGCTGTGG	
NT-gRNA	GTAGGCGCGCCGCTCTCTAC	
HNF1A-gRNA1	AGAAGGCCCCCTGGACAAGGGGGAGT	Time-course evaluation of N-terminal fusion TET1-dSaCas9
HNF1A-gRNA2	TCCGTCTCGTCCTCGGAGCCCCGAGT	
NT-gRNA	GTAGGCGCGCCGCTCTCTAC	
BACH2-sg01 (Sp)	TGTATTTTGCTGGCGTCGAAGGG	Activity profile of N-terminal fusions of DNMT3A and TET1 catalytic domains with dSpCas9 and dSaCas9
BACH2-sg02 (Sp)	CTCTCCCTCCCGGCTGTTTACGG	
BACH2-sg03 (Sp)	GGGAGAGCACACATTCGGCGCGG	
BACH2-sg06 (Sp)	CCCATTCACAATAACTTTACGGG	
BACH2-sg07 (Sp)	AGTTATTGTGAATGGGGAGCGGG	
BACH2-sg08 (Sp)	AATGTAGCGATTGAGAGTGTGGG	
BACH2-sgF3f (Sp)	GCATTTTCTAGGAACGGGAAGGG	
BACH2-sg08 (Sa)	AATGTCAACAATGTAGCGATTGAGAGT	
BACH2-sgF3f (Sa)	GCATTTTCTAGGAACGGGAAGGGGGT	
BACH2-sg01 (Sa)	AAAGTTATTGTGAATGGGGAGCGGGT	
BACH2-sg06 (Sa)	GCCGGGCCGGGGGCAGGGCCGCGGGT	
BACH2-sg08 (Sa)	ATTGTATTTTGCTGGCGTCGAAGGGT	
BACH2-sg09 (Sa)	CAGCAGAGGGAGGAGGAGCAGAGAGT	
MGAT3-sg01 (Sp)	CATTCGCTGGGATATAGAATAGG	
MGAT3-sg02 (Sp)	CCCTGCACCTTCGACGATGGCGG	
MGAT3-sg03 (Sp)	ATGCCTTGTGCCTCGCACCGCGG	
MGAT3-sg04 (Sp)	CTGCTCTGTAGGCCCCAGAACGG	
MGAT3-sg05 (Sp)	GGACGCCTCTGAGCCCTGAGAGG	
MGAT3-sg01 (Sa)	TGGAGCACATTCGCTGGGATATAGAAT	
MGAT3-sg05 (Sa)	GGACGCCTCTGAGCCCTGAGAGGAAT	
MGAT3-sg01 (Sa)	GTAGACCAGCCCTAGGCAGCCCGGAT	
MGAT3-sg02 (Sa)	GGTGCGAGGCACAAGGCATTGTGGGT	
MGAT3-sg05 (Sa)	TGGCAGGAGAGTAGGCTCAAGAGGGT	
NT-gRNA	GTAGGCGCGCCGCTCTCTAC	
HNF1A-gRNA1	AGAAGGCCCCCTGGACAAGGGGGAGT	Simultaneous methylation and

HNF1A-gRNA2	TCCGTCTCGTCCTCGGAGCCCCGAGT	demethylation in HEK293 cells
MGAT3-gRNA1	GGAGCACATTCGCTGGGATATAGAAT	
MGAT3-gRNA2	GGTGCGAGGCACAAGGCATTGTGGGT	
MGAT3-gRNA3	GGACGCCTCTGAGCCCTGAGAGGAAT	
MGAT3-gRNA4	TGGCAGGAGAGTAGGCTCAAGAGGGT	
MGAT3-gRNA5	TCTGTGTGTCTGCTTGGGGCGTGGGT	
MGAT3-gRNA6	CCGGCTGGCGGGGGAGGGGAGGGGGT	
BACH2-gRNA1	AGTTATTGTGAATGGGGAGCGGG	
BACH2-gRNA2	AATGTAGCGATTGAGAGTGTGGG	
BACH2-gRNA3	CCGCGCCCTGCCCGCTTTTATGG	
BACH2-gRNA4	GGGAGAGCACACATTCGGCGCGG	
IL6ST-gRNA1	GCCACCCCAGTCCCGCGGCGGGG	
IL6ST-gRNA2	ATCTGACAGTGTTCCGGAGCCGG	
IL6ST-gRNA3	CGCACGAACCCCTTGGCGCCAGG	
IL6ST-gRNA4	GCCAAGGGGTTCGTGCGCTGTGG	
NT-gRNA	GTAGGCGCGCCGCTCTCTAC	
HNF1A-gRNA1	AGAAGGCCCCCTGGACAAGGGGGAGT	Simultaneous methylation and demethylation in BG1 cells
HNF1A-gRNA2	TCCGTCTCGTCCTCGGAGCCCCGAGT	
MGAT3_112R	GGCCGCTCGGCTGCACCCCGGGG	
NT-gRNA	GTAGGCGCGCCGCTCTCTAC	
HNF1A-gRNA1 (TET1)	AGAAGGCCCCCTGGACAAGGGGGAGT	Synergistic epigenetic regulation of *HNF1A*
HNF1A-gRNA2 (TET1)	TCCGTCTCGTCCTCGGAGCCCCGAGT	
HNF1A-gRNA3 (VPR)	AGGGAGCTATGGCCTGCGATGGG	
NT-gRNA	GTAGGCGCGCCGCTCTCTAC	
IL6ST-gRNA1	GCCACCCCAGTCCCGCGGCGGGG	Whole-genome methylation analysis
IL6ST-gRNA2	ATCTGACAGTGTTCCGGAGCCGG	
IL6ST-gRNA3	CGCACGAACCCCTTGGCGCCAGG	
IL6ST-gRNA4	GCCAAGGGGTTCGTGCGCTGTGG	
MGAT3-gRNA1	CATTCGCTGGGATATAGAATAGG	
MGAT3-gRNA2	CCCTGCACCTTCGACGATGGCGG	
MGAT3-gRNA3	ATGCCTTGTGCCTCGCACCGCGG	
MGAT3-gRNA4	CTGCTCTGTAGGCCCCAGAACGG	
MGAT3-gRNA5	GGACGCCTCTGAGCCCTGAGAGG	

Supplementary Table 4. Sequences of primers and custom oligonucleotides used for construction of modular toolbox

Primer/custom oligonucleotide	Sequence (5'→3')	Use
	sgRNA modules	
tNS-Sg_G_XbaI-FW	TAATCTCTAGAGGTCTCATGACGAGGGCCTATTTCCCATGATTCCTTC	Amplification of gRNA module for SpCas9
tNS-Sg_G_NcoI-RE	TTACTCCATGGTCTCATACCTCTCGAATTCAAAAAAGCACCGACTCGG	
tNS-Sg_G_XbaI-FW	TAATCTCTAGAGGTCTCATGACGAGGGCCTATTTCCCATGATTCCTTC	Amplification of gRNA module for SaCas9
tNS-Sg_A_NcoI-RE	TTACTCCATGGTCTCATACCCAAAAATCTCGCCAACAAGTTG	
	Eukaryotic promoters	

tNS-Pro-C_PstI-Fw	TTAATCTGCAGGGTCTCAGGTACAGACAAATGGCTCTAGAGGTACCCGTTACATAACTT	Amplification of CAG promoter
tNS-Pro-C_XhoI-RE	TTACTCTCGAGGGTCTCTATGGTGGCAGCGCTCTAGAACCGGTCTGAAAAAAAGTGATTTCAGGCAGGTGCTCCAGG	
tNS-Pro_SV40_PstI-FW	TAATCTGCAGGGTCTCAGGTAGCTGTGGAATGTGTGTCAGTTAGGGTG	Amplification of SV40 promoter
tNS-Pro_SV40_XhoI-RE	TAATCTCGAGGGTCTCTATGGTGGCAGCGCTCTAGAACCGGTGCTTTTTGCAAAAGCCTAGGCCTCC	
tNS-Pro_E_PstI-FW	TAATCTGCAGGGTCTCGGGTATAGGTCTTGAAAGGAGTGGGAATTGGC	Amplification of EFS promoter
tNS-Pro_E_XhoI-RE	TTACTCTCGAGGGTCTCAATGGTGGCAGCGCTCTAGAACCGGTCCTGTGTTCTGG	
	Effector domains	
tNS-ED_D_Hind3-FW	TTAAAAGCTTGGTCTCCGGAGGCGGGAGCGGATCCCCCTC	Amplification of DNMT3A effector domain
tNS-ED_D_SacI-RE	TTACGAGCTCGGTCTCTCGAATGGCCGGCCGGACACACACG	
FD_c-NLS-NP-S	GCGCGGTCTCAGGAGAGAGACGTTGCGTCTCATTCGGTTCCGGAAAGAGGCCAGCAGCTACAAAGAAAGCTGGACAGGCAAAAAAGAAAAAGTCAAGCTTCGAGAGACC	Construction of C-terminal NLS on effector domains
FD_c-NLS-NP-A	TCGAGGTCTCTCGAAGCTTGACTTTTTCTTTTTTGCCTGTCCAGCTTTCTTTGTAGCTGCTGGCCTCTTTCCGGAACCGAATGAGACGCAACGTCTCTCTCCTGAGACC	
tNS-ED_T_XbaI-FW	TTAATCTAGAGGTCTCCGGAGGCGGGAGCGGATCCCTGC	Amplification of TET1 effector domain
tNS-ED_T_Acc65I-RE	TTACGGTACCGGTCTCTCGAATGGCCGGCCGACCCAATGG	
TET1_noBbsI-1_S	CTTCTCCTGGTCCCCAAAGACTGCTTCAGCC	BbsI restriction sites mutagenesis in TET1
TET1_noBbsI-1_A	GGCTGAAGCAGTCTTTGGGGACCAGGAGAAG	
TET1_noBbsI-2_S	GATGCCTTCGGGAAGGCTCAGTGGTGCCAAT	
TET1_noBbsI-2_A	ATTGGCACCACTGAGCCTTCCCGAAGGCATC	
TET1_noBsaI-S	CCAACCTTAGGGAGTAACACTGAAACCGTGCAACCT	BsaI restriction site mutagenesis in TET1
TET1_noBsaI-A	AGGTTGCACGGTTTCAGTGTTACTCCCTAAGGTTGG	
TET1_H1671Y_D1673A_S	GACTTCTGTGCTCATCCCTACAGGGCCATTCACAACATGAATAA	Mutagenesis of active site in TET1 (H1671Y, D1673A)
TET1_H1671Y_D1673A_A	TTATTCATGTTGTGAATGGCCCTGTAGGGATGAGCACAGAAGTC	
tNS-ED_V_Hind3-FW	ATTAAAAGCTTGGTCTCAGGAGGCGGCAGCGAGGCCAGCGGTTCCGGACG	Amplification of VPR activation domain
tNS-ED_V_XhoI-RE	TTACTCTCGAGGGTCTCACGAAAAACAGAGATGTGTCGAAGATGGACAGTCCTGTGC	
VPR-noEsp3I- S	ATCCCGATGAAGAGACAAGCCAGGCTGTCAAAG	Esp3I restriction site mutagenesis in VPR
VPR_noEsp3I-A	CTTTGACAGCCTGGCTTGTCTCTTCATCGGGAT	
N-FD-S	AGCTTGGTCTCACCATGGTGCCAAAAAAGAAGAGAAAGGTAGGCGGAGTAGTCTTCTATAGAAGACATTTCGGTGGAGCGAGACCC	Construction of effector domains for N-terminal fusion with dCas9
N-FD-A	TCGAGGGTCTCGCTCCACCGAAATGTCTTCTATAGAAGACTACTCCGCCTACCTTTCTCTTCTTTTTTGGCACCATGGTGAGACCA	
	dCas9 modules	
tNS-Cas_N_Hind3-FW	ATTACAAGCTTGGTCTCACCATGGACTATAAGGACCACGACGG	Amplification of dSpCas9 module
tNS-Cas_N_XhoI-RE	TTACTCTCGAGGGTCTCACTCCGCCCTTTTTCTTTTTGGCCTGTCCGG	
tNS-Cas_A_XbaI-FW	ATTACTCTAGAGGTCTCACCATGGCCCCAAAGAAG	Amplification of dSaCas9 module
tNS-Cas_A_XhoI-RE	TTACTCTCGAGGGTCTCACTCCGCCCTTTTTCTTTTTGCCTGGCCGG	

tNS-Cas_A_XbaI-FW	ATTACTCTAGAGGTCTCACCATGGCCCCAAAGAAG	Amplification of dSaCas9 module with 3×HA-FLAG
tNS-Cas_A_XhoI+FLAG-RE	TTACTCTCGAGGGTCTCACTCCGCCGAGCTCTAGGAATTCTTCAGC GTA	
LNK_trpzip2_G4Sx2-S	GATCCAGCTGGACGTGGGAGAACGGGAAATGGACCTGGAAGGCT AGCGGCGGAGGTGGCAGCGGTGGAGTGAGACCC	Construction of Trip-zip linker (TZ2)
LNK_trpzip2_G4Sx2-A	TCGAGGGTCTCACTCCACCGCTGCCACCTCCGCCGCTAGCCTTCCA GGTCCATTTCCCGTTCTCCCACGTCCAGCTG	
N-C9-S	AGCTTGGTCTCTGGAGGTGGCTCCATGTGTCTTCGATGCAGAAGA CGTGGAGATTTCGTGAGACCC	Construction of dCas9 module for N-terminal fusion
N-C9-A	TCGAGGGTCTCACGAAATCTCCACGTCTTCTGCATCGAAGACACAT GGAGCCACCTCCAGAGACCA	
	Selection markers and dual marker system	
AR_Puro_NcoI_Fw	GAGAGAGCCATGGTGACCGAGTACAAGCCCACGG	Amplification of PuroR module
AR_Puro_KasI_Re	GAGAGAGGCGCCGGCACCGGGCTTGCG	
PuroM_C210T_S	GCGGACGACGGCGCTGCGGTGGCGGTCTGG	KasI restriction sites mutagenesis in PuroR module
PuroM_C210T_A	CCAGACCGCCACCGCAGCGCCGTCGTCCGC	
PuroM_G339T_S	GGAAGGCCTCCTGGCTCCGCACCGGCCCAAG	
PuroM_G339T_A	CTTGGGCCGGTGCGGAGCCAGGAGGCCTTCC	
Puro_noBsaI-S	CTTCCTGGAGACATCCGCGCCCCGC	BsaI restriction site mutagenesis in PuroR module
Puro_noBsaI-A	GCGGGGCGCGGATGTCTCCAGGAAG	
FP_Ruby_NcoI_Fw	GAGAGAGCCATGGTGTCTAAGGGCGAAGAGCTGA	Amplification of mRuby3 module
FP_KasI_Rev	GAGAGAGGGCGCCCTTGTACAGCTCGTCCATGCC	
G414A_A	CCAATGGTCCCGTGATGCAGAAAAAGACCAAGGGT	BpiI restriction site mutagenesis in mRuby3 module
G414A_S	ACCCTTGGTCTTTTTCTGCATCACGGGACCATTGG	
FP_CC_NcoI_Fw	GAGAGAGCCATGGTGAGCAAGGGCG	Amplification of mClover3 module
FP_KasI_Rev	GAGAGAGGGCGCCCTTGTACAGCTCGTCCATGCC	
pFP-Entry_S	ATCATACTGCAGTAGGTCTTGAAAGGAGTGGGAATTGGCTCCGGT GCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAG TTGGGGGGAGGGGTCGGCAATTGATCCGGTGCCTAGAGAAGGTG GCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCT TTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGC CGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGCA CCAGGTCTTCAGGAGGTAAAACGACGGCCAGTAATTAATGTGAGT TAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGC TCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGG AGGCTGCCACCATGGGAGACCGATATCCGCTCTAGAACTAGTGGA TCGGTCTCGGCGCCGGCTAGCTTGAGTAACTAGGACGAACAATAA GGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAAGTCATAG CTGTTTCCTGCCGTGGAAGACCTGTTTGCAGCCTCGACTGTGCCTT CTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTG ACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGG AAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGG TGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAGAAT AGCAGGCATGCTCGAGACTAAC	Construction of pUK21_FP_entry plasmid for cloning of fluorescence and selection modules
pFP-Entry_A	GTTAGTCTCGAGCATGCCTGCTATTCTCTTCCCAATCCTCCCCCTTG CTGTCCTGCCCCACCCCACCCCCCAGAATAGAATGACACCTACTCA GACAATGCGATGCAATTTCCTCATTTTATTAGGAAAGGACAGTGG	

	GAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCA AACAACAGATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGCAA ACAGGTCTTCCACGGCAGGAAACAGCTATGACTTTTGTTATCAATA AAAAAGGCCCCCCGTTAGGGAGGCCTTATTGTTCGTCCTAGTTACT CAAGCTAGCCGGCGCCGAGACCGATCCACTAGTTCTAGAGCGGAT ATCGGTCTCCCATGGTGGCAGCCTCCTGTGTGAAATTGTTATCCGC TCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAG CCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTACTGGCCG TCGTTTTACCTCCTGAAGACCTGGTGCCTGTGTTCTGGCGGCAAAC CCGTTGCGAAAAAGAACGTTCACGGCGACTACTGCACTTATATACG GTTCTCCCCCACCCTCGGGAAAAAGGCGGAGCCAGTACACGACAT CACTTTCCCAGTTTACCCCGCGCCACCTTCTCTAGGCACCGGATCAA TTGCCGACCCCTCCCCCCAACTTCTCGGGGACTGTGGGCGATGTGC GCTCTGCCCACTGACGGGCACCGGAGCCAATTCCCACTCCTTTCAA GACCTACTGCAGTATGAT	
pFP-T2A_S	ATCTCATACGTCTCCAGCTCTGCAGAGGAGGACAAGCTTATGGTCT CATTCGGCAGTGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGT GACGTCGAGGAGAATCCTGGCCCAGCCATGGTGTCTTCGATATCC GCTCTAGAACTAGTGGATCTGAAGACCGGCGCCGGCTAGCATCTC GAGTAACGCTAGAGACCGAGCTCTCGATGAGACGACTACTAC	Construction of pUK21_FP_T2A
pFP-T2A_A	GTAGTAGTCGTCTCATCGAGAGCTCGGTCTCTAGCGTTACTCGAGA TGCTAGCCGGCGCCGGTCTTCAGATCCACTAGTTCTAGAGCGGATA TCGAAGACACCATGGCTGGGCCAGGATTCTCCTCGACGTCACCGC ATGTTAGCAGACTTCCTCTGCCCTCTCCACTGCCGAATGAGACCAT AAGCTTGTCCTCCTCTGCAGAGCTGGAGACGTATGAGAT	
T2A_FW	ATCTCATACGTCTCCAGCTCTGC	Amplification of pFP-T2A
T2A_RE	GTAGTAGTCGTCTCATCGAGAGCTC	
T2A_X-FW	TTCAGTCTCGTCTCCAGCTCTGCAGAGGAGGACAAGC	Amplification of pFP-T2A for dual marker system
T2A_X-RE	ACTTGATACGTCTCATCGAGGTCTCTAGTACGGCGCCGGTCTTCAG ATCCAC	
C9seq3	GGACATCCAGAAAGCCCAGG	Amplification of X-P2A from Gblock-BB
C9seq4	GAAGTCCAAGCTGGTGTCCG	
pUK21_KasI-BamHI-S	GCGCATGTAAAACGACGGCCAGTTAATACGACTCACTATAGGTAG	KasI restriction site removal from pUK21gg
pUK21_KasI-BamHI-A	GATCCTACCTATAGTGAGTCGTATTAACTGGCCGTCGTTTTACAT	
	Eukaryotic terminators	
tNS-Ter-H_Hind3-FW	TTAATAAGCTTGGTCTCTCGCTAGCCTCGACTGTGCCTTCTAGTTGC C	Amplification of Bgh terminator
tNS-Ter-H_NcoI-RE	TTACTCCATGGTCTCAACCTCATGCCTGCTATTCTCTTCCCAATCCTC	
tNS-Ter-SV40_PstI-FW	TAATCTGCAGAAGCTTGGTCTCTCGCTGTTGTTAACTTGTTTATTGC AGCTTATAATGG	Amplification of SV40 terminator
tNS-Ter-SV40_XhoI-RE	TAATCTCGAGGTCTCAACCTCCAGTTGATCCAGACATGATAAGATA CATTG	
	Backbone vector	
pUC19-FW	ATCGGTACCGGAACCCCTATTTGTTTATTTTTCT	Construction of backbone plasmid (pBackBone-BZ)
pUC19-RE	TATACCGGTCATGTGAGCAAAAGGCCAGC	
Gblock-BB_S	TACATGGATCCTGAACCCCGACAACAGCGACCGTCTCAGTACACTG CAGGCCGGCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTC CGCCTCATGGCTGACTAATTTTTTGCGTCTTCGGTACCGAAGACGC TTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTAT TCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCGGCGACGTCTAC ATGAGCTCATTGAGATGCATGCTTTGCATACTTCTGCCTGCTGGGG AGCCTGGGGACTTTCCACACCTGGTTGCTGACTAATTGAGATCAGG AGGTAAAACGACGGCCAGTGGTACCTATGAACAGTTACTAAGAGT	

	CGACTGACTGAGACCAATTAATGTGAGTTAGCTCACTCATTAGGCA CCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAAT TGTGAGCGGATAACAATTTCACACAGGAGGCTGCCACCATGGTGA TGATTACGGATTCACTGGCAGTGGTCCTGCAACGTCGTGACTGGG AAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCGCC TTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCC TTCCCAACAGCTGCGCAGCCTGAATGGCGCCTAAAGCTAGCTTGA GTAACTAGGACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTT TATTGATAACAAAAGGTCTCTAGGTTCTAGATATACTTCATGTAAG ACACTCGAGGTCATAGCTGTTTCCTGCCGTGGTCCGGTGAGACGG GGATGCGGAAGGTCAGGATGGACATCCAGAAAGCCCAGGGAAGA CTAAGCTTATGGTCTCATACTTGGCAGTGGAGCCACGAACTTCTCT CTGTTAAAGCAAGCTGGCGACGTGGAAGAAAACCCCGGTCCTGCC ATGGTGTCTTCCGGACACCAGCTTGGACTTCAGATCTTACAT	
Gblock-BB_A	ATGTAAGATCTGAAGTCCAAGCTGGTGTCCGGAAGACACCATGGC AGGACCGGGGTTTTCTTCCACGTCGCCAGCTTGCTTTAACAGAGAG AAGTTCGTGGCTCCACTGCCAAGTATGAGACCATAAGCTTAGTCTT CCCTGGGCTTTCTGGATGTCCATCCTGACCTTCCGCATCCCCGTCTC ACCGGACCACGGCAGGAAACAGCTATGACCTCGAGTGTCTTACAT GAAGTATATCTAGAACCTAGAGACCTTTTGTTATCAATAAAAAAGG CCCCCCGTTAGGGAGGCCTTATTGTTCGTCCTAGTTACTCAAGCTA GCTTTAGGCGCCATTCAGGCTGCGCAGCTGTTGGGAAGGGCGATC GGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGCGGATG TGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTC ACGACGTTGCAGGACCACTGCCAGTGAATCCGTAATCATCACCATG GTGGCAGCCTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACAC AACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAA TGAGTGAGCTAACTCACATTAATTGGTCTCAGTCAGTCGACTCTTA GTAACTGTTCATAGGTACCACTGGCCGTCGTTTTACCTCCTGATCTC AATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCA GGCAGAAGTATGCAAAGCATGCATCTCAATGAGCTCATGTAGACG TCGCCGGCCTCCAAAAAAGCCTCCTCACTACTTCTGGAATAGCTCA GAGGCCGAGGCGGCCTCGGCCTCTGCATAAATAAAAAAGCGTCTT CGGTACCGAAGACGCAAAAAATTAGTCAGCCATGAGGCGGAGAA TGGGCGGAACTGGGCGGAGTTAGGGGCGGGATGCCGGCCTGCAG TGTACTGAGACGGTCGCTGTTGTCGGGGTTCAGGATCCATGTA	
C9seq1	TGAACCCCGACAACAGCGAC	Amplification of Gblock-BB
C9seq2	ATCCTGACCTTCCGCATCCC	
pUC19bla_noBsaI-S	GGAGCCGGTGAGCGTGGAAGCCGCGGTATCATTGCAGC	BsaI restriction site mutagenesis in backbone plasmid (pBackBone-BZ)
pUC19bla_noBsaI-A	GCTGCAATGATACCGCGGCTTCCACGCTCACCGGCTCC	
	Multi-guide system	
pFUS-newNgoMIV-A	GTCGAGGCATTTCTGTGCCGGCTGGTCTAGACGTC	Construction of pUS21gg vector
pFUS-newNgoMIV-S	GACGTCTAGACCAGCCGGCACAGAAATGCCTCGAC	
pFUS-noNgoMIV-A	CTGATACTGGGCTGGCAGGCGCTCC	
pFUS-noNgoMIV-S	GGAGCGCCTGCCAGCCCAGTATCAG	
pFUS-noXhoI-A	GGTCATGGGTGGCTCTAGGGTTATTTGCCGA	
pFUS-noXhoI-S	TCGGCAAATAACCCTAGAGCCACCCATGACC	
KAN_NgoMIV-RE	ATATTGCCGGCTTCGAAAGGGCCTCGTGATACGC	Construction of pUK21gg vector
KAN_SapI-FW	CGTATTGGGCGCTCTTCCG	
KanR_noEsp3I-A	GATCGCGTATTTCGGCTCGCTCAGGCGCA	
KanR_noEsp3I-S	TGCGCCTGAGCGAGCCGAAATACGCGATC	
sgM1-S	GCGCCCGTCTCATGACTTCTAGATGACACATCCATGGAGGTATGAG ACGC	Construction of pSgMxA or

sgM1-A	TCGAGCGTCTCATACCTCCATGGATGTGTCATCTAGAAGTCATGAG ACGG	pSgMxG plasmids (x = 1, 2, 3, 4, 5, 6)
sgM2-S	GCGCCCGTCTCAGGTATTCTAGATGACACATCCATGGACCATTGAG ACGC	
sgM2-A	TCGAGCGTCTCAATGGTCCATGGATGTGTCATCTAGAATACCTGAG ACGG	
sgM3-S	GCGCCCGTCTCACCATTTCTAGATGACACATCCATGGAGGAGTGA GACGC	
sgM3-A	TCGAGCGTCTCACTCCTCCATGGATGTGTCATCTAGAAATGGTGAG ACGG	
sgM4-S	GCGCCCGTCTCAGGAGTTCTAGATGACACATCCATGGATTCGTGA GACGC	
sgM4-A	TCGAGCGTCTCACGAATCCATGGATGTGTCATCTAGAACTCCTGAG ACGG	
sgM5-S	GCGCCCGTCTCATTCGTTCTAGATGACACATCCATGGACGCTTGAG ACGC	
sgM5-A	TCGAGCGTCTCAAGCGTCCATGGATGTGTCATCTAGAACGAATGA GACGG	
sgM6-S	GCGCCCGTCTCACGCTTTCTAGATGACACATCCATGGAAGGTTGAG ACGC	
sgM6-A	TCGAGCGTCTCAACCTTCCATGGATGTGTCATCTAGAAAGCGTGAG ACGG	
XbaI_B_C9seq1_A_NcoI-S	CTAGATGACTGAGACCTGAACCCCGACAACAGCGACGGTCTCTGG TACGC	
XbaI_B_C9seq1_A_NcoI-A	CATGGCGTACCAGAGACCGTCGCTGTTGTCGGGGTTCAGGTCTCA GTCAT	
SgMult-S	GCGCCGGTCTCATGACTGACTGAGACGTCTAGACACCAGGTCTTCT GACACATGAAGACCTGTTTC	Construction of pSgx1-pSgx6 plasmids
SgMult-A	CATGGAAACAGGTCTTCATGTGTCAGAAGACCTGGTGTCTAGACG TCTCAGTCAGTCATGAGACCG	
SgMx1-S	CATGGCGTCTCAGGTAGGTATGAGACCC	
SgMx1-A	TCGAGGGTCTCATACCTACCTGAGACGC	
SgMx2-S	CATGGCGTCTCACCATTGGTATGAGACCC	
SgMx2-A	TCGAGGGTCTCATACCAATGGTGAGACGC	
SgMx3-S	CATGGCGTCTCAGGAGGGTATGAGACCC	
SgMx3-A	TCGAGGGTCTCATACCCTCCTGAGACGC	
SgMx4-S	CATGGCGTCTCATTCGGGTATGAGACCC	
SgMx4-A	TCGAGGGTCTCATACCCGAATGAGACGC	
SgMx5-S	CATGGCGTCTCACGCTGGTATGAGACCC	
SgMx5-A	TCGAGGGTCTCATACCAGCGTGAGACGC	
SgMx6-S	CATGGCGTCTCAAGGTGGTATGAGACCC	
SgMx6-A	TCGAGGGTCTCATACCACCTTGAGACGC	
	Secondary cassette	
M14-S	GCGCAGGTCTCACCATGGCACCAGGTCTTCTGAACCCCGACAACA GCGACGAAGACCTGTTTGGCGCCGGCTAGCCGCTAGAGACC	Construction of secondary cassette
M14-A	TCGAGGTCTCTAGCGGCTAGCCGGCGCCAAACAGGTCTTCGTCGC TGTTGTCGGGGTTCAGAAGACCTGGTGCCATGGTGAGACCT	
BB_2nd_Cassette-S	CGGTACGAGACGTGAACCCCGACAACAGCGACCGTCTCCAGGTGA GCT	
BB_2nd_Cassette-A	CACCTGGAGACGGTCGCTGTTGTCGGGGTTCACGTCTCGTACCGA CGT	

Supplementary Table 5. Quantitative real-time PCR primers used for evaluation of TET1-dSaCas9 presence through time

Primer name	Sequence (5' → 3')	Use
SaCas9_Fw	CCGCCCGGAAAGAGATTATT	Detection of SaCas9
SaCas9_Rev	CGGAGTTCAGATTGGTCAGTT	
SpCas9_Fw	TGCCCCAAGTGAATATCGTG	Detection of SpCas9
SpCas9_Rev	GACTTGCCCTTTTCCACTTTG	
RAG1_Fw	TGTTGACTCGATCCACCCCA	Endogenous control for Cas9 plasmid DNA normalization
RAG1_Rev	TGAGCTGCAAGTTTGGCTGAA	
GAPDH_Fw	AGGGCTGCTTTTAACTCTGGT	Endogenous control for Cas9 expression normalization
GAPDH_Rev	CCCCACTTGATTTTGGAGGGA	

Supplementary Table 6. Annotation of glycan peaks with corresponding glycan structures according to the results of mass spectrometry. A typical chromatogram is given in Supplementary Figure S7. In the first column labelled GP (glycan peak), "CONT" denotes contamination between peaks GP5 and GP6, resulting in the inability to quantitate the affected glycan structures (see also: Supplementary Figure S7).

GP	Structure	RT (min)	Theoretical mass $[MH]^+$	Measured mass $[MH]^+$	MS/ MS	major structure in peak
1	$(Hex)_2$ $(HexNAc)_2$	3.11	968.456	968.461	yes	yes
2	$(Hex)_2$ $(HexNAc)_2$ $(Deoxyhexose)_1$	3.75	1114.514	1114.527	yes	yes
3	(Hex)3 $(HexNAc)_2$	4.54	1130.509	1130.523	yes	yes
4	(Hex)3 $(HexNAc)_2$ $(Deoxyhexose)_1$	5.31	1276.567	1276.581	yes	yes
5	$(Hex)_4$ $(HexNAc)_2$	6.02	1292.562	1292.576	yes	yes
CONT	$(Hex)_4$ $(HexNAc)_2$ $(Deoxyhexose)_1$	6.85	1438.620	1438.627	yes	yes
	$(Hex)_1$ $(HexNAc)_1$ + $(Man)_3(GlcNAc)_2$	7.3	1495.641	1495.646	no	no
	$(HexNAc)_2$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$		1682.726	1682.732	yes	yes
6	$(Hex)_2$ + $(Man)_3(GlcNAc)_2$	8.06	1454.615	1454.620	yes	yes
7	$(HexNAc)_3$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$	8.71	1885.805	1885.819	yes	yes
8	$(Hex)_2$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$	8.91	1600.673	1600.697	yes	yes
	$(Hex)_2$ $(HexNAc)_1$ + $(Man)_3(GlcNAc)_2$		1657.694	1657.699	yes	yes
9	$(HexNAc)_4$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$	9.52	2088.885	2088.915	no	no
	$(Hex)_1$ $(HexNAc)_3$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$		2047.858	2047.867	no	no
	$(Hex)_2$ $(HexNAc)_2$ + $(Man)_3(GlcNAc)_2$		1860.774	1860.783	yes	yes
10	$(Hex)_3$ + $(Man)_3(GlcNAc)_2$	9.9	1616.668	1616.678	yes	yes

11	$(Hex)_2$ $(HexNAc)_2$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$	10.45	2006.832	2006.841	yes	yes
	$(Hex)_2$ $(HexNAc)_3$ + $(Man)_3(GlcNAc)_2$		2063.853	2063.879	yes	no
12	$(Hex)_2$ $(HexNAc)_1$ $(Deoxyhexose)_2$ + $(Man)_3(GlcNAc)_2$	10.74	1949.810	1949.821	yes	yes
13	$(Hex)_3$ $(HexNAc)_1$ + $(Man)_3(GlcNAc)_2$	10.92	1819.747	1819.759	yes	yes
	$(Hex)_2$ $(HexNAc)_2$ $(Deoxyhexose)1$ + $(Man)3(GlcNAc)2$		2006.832	2006.841	no	no
14	$(Hex)2$ $(HexNAc)3$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$	11.2	2209.911	2209.935	yes	yes
	$(Hex)_3$ $(HexNAc)_2$ + $(Man)_3(GlcNAc)_2$		2022.826	2022.851	yes	no
15	$(Hex)_4$ + $(Man)_3(GlcNAc)_2$	11.8	1778.720	1778.737	yes	yes
	$(Hex)_3$ $(HexNAc)_1$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$		1965.805	1965.831	no	no
	$(Hex)_3$ $(HexNAc)_2$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$		2168.884	2168.911	yes	no
16	$(Hex)_2$ $(HexNAc)_3$ $(Deoxyhexose)_2$ + $(Man)_3(GlcNAc)_2$	12.16	2355.969	2356.000	yes	yes
17	$(Hex)_3$ $(HexNAc)_2$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$	12.55	2110.842	2110.867	yes	no
	$(Hex)_2$ $(HexNAc)_2$ $(NeuAc)_1$ + $(Man)_3(GlcNAc)_2$		2151.869	2151.905	yes	yes
18	$(Hex)_4$ $(HexNAc)_2$ + $(Man)_3(GlcNAc)_2$	13.18	2184.879	2184.915	yes	yes
	$(Hex)_3$ $(HexNAc)_2$ $(Deoxyhexose)_2$ + $(Man)_3(GlcNAc)_2$		2314.942	2314.977	yes	yes
	$(Hex)_5$ + $(Man)_3(GlcNAc)_2$		1940.773	1940.803	yes	no
	$(Hex)_2$ $(HexNAc)_2$ $(NeuAc)_2$ + $(Man)_3(GlcNAc)_2$		2442.964	2442.999	no	no
19	$(Hex)_5$ + $(Man)_3(GlcNAc)_2$	13.56	1940.773	1940.803	yes	yes
20	$(Hex)_4$ $(HexNAc)_2$ + $(Man)_3(GlcNAc)_2$	13.82	2184.879	2184.905	no	no
	$(Hex)_4$ $(HexNAc)_2$ $(Deoxyhexose)_1$ + $(Man)_3(GlcNAc)_2$		2330.937	2330.975	yes	yes
21	$(Hex)_2$ $(HexNAc)_2$ $(NeuAc)_2$ + $(Man)_3(GlcNAc)_2$	14.07	2442.964	2442.997	yes	yes
	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_1$ + $(Man)3(GlcNAc)_2$		2517.001	2517.035	yes	no
22	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_1$ + $(Man)_3(GlcNAc)_2$	14.86	2517.001	2517.036	yes	no
	$(Hex)_6$ + $(Man)_3(GlcNAc)_2$		2102.826	2102.867	yes	yes
	$(Hex)_2$ $(HexNAc)_2$ $(NeuAc)_2$ + $(Man)_3(GlcNAc)_2$		2442.964	2443.007	no	no
23	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_2$ + $(Man)_3(GlcNAc)_2$	15.41	2808.097	2808.142	yes	yes
24	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_2$ + $(Man)_3(GlcNAc)_2$	16.11	2808.097	2808.133	yes	yes
	$(Hex)_7$ + $(Man)_3(GlcNAc)_2$		2264.879	2264.919	no	no
25	$(Hex)_4$ $(HexNAc)_4$ $(Deoxyhexose)_2$ + $(Man)_3(GlcNAc)_2$	16.66	2883.154	2883.203	no	no
	$(Hex)3$ $(HexNAc)_3$ $(Deoxyhexose)_1$ $(NeuAc)_2$ + $(Man)_3(GlcNAc)_2$		2954.155	2954.199	no	no
	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_3$ + $(Man)_3(GlcNAc)_2$		3099.192	3099.237	yes	yes
26	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_2$ + $(Man)_3(GlcNAc)_2$	16.87	2808.097	2808.143	yes	yes
27	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_3$ + $(Man)_3(GlcNAc)_2$	17.36	3099.192	3099.247	yes	yes

28	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_3$ + $(Man)_3(GlcNAc)_2$	18.06	3099.192	3099.251	yes	yes
29	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_4$ + $(Man)_3(GlcNAc)_2$	18.44	3390.287	3390.340	yes	yes
30	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_3$ + $(Man)_3(GlcNAc)_2$	18.79	3099.192	3099.244	yes	yes
31	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_4$ + $(Man)_3(GlcNAc)_2$	19.1	3390.287	3390.352	yes	yes

Supplementary Table 7. Wilcoxon rank sum p-values to assess the effect of potential covariates on principal components after probe filtering and normalization. Covariates contributing significant effects [Wilcoxon p-values less than the calculated Bonferroni limit (0.05/40 = 0.0013)] are highlighted in red.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Sentrix ID	0.340	0.002	0.329	0.086	0.559	0.323	0.432	0.472	0.804	0.374
Sentrix Position	0.806	0.998	0.482	0.757	0.362	0.063	0.523	0.822	0.553	0.355
Sample Well	0.457	0.457	0.457	0.457	0.457	0.457	0.457	0.457	0.457	0.457
Biological Replicate	0.123	0.000	0.796	0.123	0.684	0.579	0.684	0.123	0.796	0.393
Technical Replicate	0.340	0.002	0.329	0.086	0.559	0.323	0.432	0.472	0.804	0.374
Group	0.006	0.474	0.047	0.070	0.834	0.546	0.614	0.190	0.632	0.765
Group and Biological replicates	0.047	0.037	0.183	0.218	0.948	0.475	0.350	0.061	0.931	0.655

Supplementary Table 8. Wilcoxon rank sum p-values to assess the effect of potential covariates on principal components after probe filtering, normalization, and SVA correction.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Sentrix ID	0.807	0.424	0.886	0.081	0.925	0.934	0.972	0.617	0.300	0.374
Sentrix Position	0.903	0.955	0.511	0.095	0.604	0.362	0.642	0.402	0.336	0.168
Sample Well	0.457	0.457	0.457	0.457	0.457	0.457	0.457	0.457	0.457	0.457
Biological Replicate	0.971	0.912	0.796	0.739	0.739	0.739	0.684	0.739	0.631	1.000
Technical Replicate	0.807	0.424	0.886	0.081	0.925	0.934	0.972	0.617	0.300	0.374
Group	0.002	0.004	0.712	0.668	0.775	0.998	0.009	0.060	0.958	0.998
Group and Biological replicates	0.048	0.083	0.904	0.655	0.109	1.000	0.065	0.266	0.953	1.000

Supplementary references

Bajar, B. T., Wang, E. S., Lam, A. J., Kim, B. B., Jacobs, C. L., Howe, E. S., . . . Chu, J. (2016). Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting. *Sci Rep, 6*, 20889. doi:10.1038/srep20889

Chen, Y. J., Liu, P., Nielsen, A. A., Brophy, J. A., Clancy, K., Peterson, T., & Voigt, C. A. (2013). Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat Methods, 10*(7), 659-664. doi:10.1038/nmeth.2515

Cochran, A. G., Skelton, N. J., & Starovasnik, M. A. (2001). Tryptophan zippers: stable, monomeric beta -hairpins. *Proc Natl Acad Sci U S A, 98*(10), 5578-5583. doi:10.1073/pnas.091100898

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., . . . Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science, 339*(6121), 819-823. doi:10.1126/science.1231143

Dean, D. A. (1997). Import of plasmid DNA into the nucleus is sequence specific. *Exp Cell Res, 230*(2), 293-302. doi:10.1006/excr.1996.3427

Vojta, A., Dobrinic, P., Tadic, V., Bockor, L., Korac, P., Julg, B., . . . Zoldos, V. (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res, 44*(12), 5615-5628. doi:10.1093/nar/gkw159

2.3 CRISPR/Cas9-based epigenome editing: An overview of dCas9-based tools with special emphasis on off-target activity

Methods 164–165 (2019) 109–119

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

CRISPR/Cas9-based epigenome editing: An overview of dCas9-based tools with special emphasis on off-target activity

Vanja Tadić[1], Goran Josipović[1], Vlatka Zoldoš, Aleksandar Vojta*

University of Zagreb, Faculty of Science, Department of Biology, Division of Molecular Biology, Horvatovac 102a, HR-10000 Zagreb, Croatia

ARTICLE INFO

Keywords:
CRISPR/Cas
dCas9
Genome editing
Epigenome editing
Cas9 orthologs
Cas9 off-target

ABSTRACT

Molecular tools for gene regulation and epigenome editing consist of two main parts: the targeting moiety binding a specific genomic locus and the effector domain performing the editing or regulatory function. The advent of CRISPR-Cas9 technology enabled easy and flexible targeting of almost any locus by co-expression of a small sgRNA molecule, which is complementary to the target sequence and forms a complex with Cas9, directing it to that particular target. Here, we review strategies for recruitment of effector domains, used in gene regulation and epigenome editing, to the dCas9 DNA-targeting protein. To date, the most important CRISPR-Cas9 applications in gene regulation are CRISPR activation or interference, while epigenome editing focuses on targeted changes in DNA methylation and histone modifications. Several strategies for signal amplification by recruitment of multiple effector domains deserve special focus. While some approaches rely on altering the sgRNA molecule and extending it with aptamers for effector domain recruitment, others use modifications to the Cas9 protein by direct fusions with effector domains or by addition of an epitope tag, which also has the ability to bind multiple effector domains. A major barrier to the widespread use of CRISPR-Cas9 technology for therapeutic purposes is its off-target effect. We review efforts to enhance CRISPR-Cas9 specificity by selection of Cas9 orthologs from various bacterial species and their further refinement by introduction of beneficial mutations. The molecular tools available today enable a researcher to choose the best balance of targeting flexibility, activity amplification, delivery method and specificity.

1. Introduction

Different types of molecular tools are used for modification of epigenetic marks (CpG methylation and histone modifications) or direct manipulation of gene promoters, ultimately altering gene expression. Artificial transcription factors give us the ability to precisely modulate expression of specific genes. A typical artificial transcriptional activator is a fusion of a DNA-binding domain and a transcriptional activation domain. Several studies over the past 20 years characterized artificial transcription factors based on zinc finger proteins (ZFP) and TALE (Transcription activator-like effector) fusions with transactivation domains which were proven successful for targeted activation of genes in mammalian cells [1–8]. The concept of sequence-specific guidance of effector domains has been already developed with ZFP and TALE systems, and the accumulated experience with those systems gave a large initial boost to the emerging CRISPR-Cas technology by laying foundations for its most exciting and valuable applications. With the emergence of CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats; CRISPR-associated protein) systems for genome editing, CRISPR activation (CRISPRa) tools have completely dominated the field of targeted gene regulation in the last five years. A typical molecular tool (*i.e.* a plasmid encoding functional components in an appropriate expression cassette) for gene regulation encodes components consisting of a targeting domain that binds a specific sequence in the genome, fused to an effector domain that mediates gene regulation. In this configuration, the Cas9 is a nuclease-null mutant (deactivated/dead Cas9, dCas9), where the two catalytic sites are inactivated by point mutations. Therefore, the dCas9 is unable to cut the DNA molecule, but it retains the strong binding activity programmed by the co-expressed sgRNA molecule. Since binding specificity is essentially unaltered between catalytically active Cas9 nuclease and dCas9, discussion of specificity in the context of epigenome editing takes the advantage of many studies focused primarily on the Cas9 nuclease. The CRISPR-Cas system has become the targeting domain of choice due to its versatility and ease of targeting. Many detailed reviews describing the growing toolbox for dCas9-mediated transcriptional regulators have

* Corresponding author.
E-mail address: vojta@biol.pmf.hr (A. Vojta).
[1] These authors contributed equally to the manuscript.

https://doi.org/10.1016/j.ymeth.2019.05.003
Received 1 February 2019; Received in revised form 29 April 2019; Accepted 2 May 2019
Available online 06 May 2019

been published since [9–13]. Here, we give an overview of CRISPRa efforts, with highlights on several important advances in the field.

2. Discovery and development of CRISPR-Cas9 system and development of novel, powerful molecular tools

A naturally occurring defense system in many bacteria and archaea, the CRISPR-Cas system, protects the integrity of the organisms against foreign invading elements like viruses or plasmids [14–17]. The Cas9 component of the type II CRISPR system from *Streptococcus pyogenes* is widely used as a powerful tool for genome and epigenome editing in diverse organisms. While the CRISPR-Cas9 system in its basic form (with intact nuclease activity) can be used for targeted gene knockout, in this review we focus on the repurposing of dCas9 as a powerful targeting moiety which enables efficient and precise delivery of various effector domains, such as those for gene activation, repression or for editing the epigenome. Three main components are needed for the use of the CRISPR-Cas9 system in eukaryotic cells: (1) the endonuclease Cas9 (SpCas9, 1,368 amino acids), which is guided to the target site in genome by a (2) crRNA molecule (CRISPR RNA), along with its (3) tracrRNA (*trans*-activating CRISPR RNA) molecule, crucial for crRNA maturation in bacteria and for complex assembly with the Cas9 protein [18–20]. This system was simplified by combining the crRNA and tracrRNA molecules into one chimeric sgRNA molecule (single guide RNA) [20], which facilitates the application of the CRISPR-Cas9 in eukaryotic cells. Specificity of Cas9 targeting is determined by the first 20 nucleotides at the 5′ end of an sgRNA molecule, which can be easily changed to target any desired locus. The guide RNA loaded onto Cas9 binds its complementary target sequence only if it is immediately followed by the downstream protospacer-adjacent motif (PAM), which is 5′-NGG-3′ for the Cas9 from *Streptococcus pyogenes* (SpCas9) [20,21]. The PAM-Interacting domain of Cas9 is crucial for PAM specificity and is responsible for initiating contact with DNA, Cas9 activation and target site binding [20,22].

The CRISPR-Cas system is widely used to target single or multiple genes in order to understand their function [23–25]. Thousands of genes can be screened in order to identify candidates contributing to a specific phenotype [26,27]. Unraveling the potential role of epigenetic modification on gene expression regulation became more straightforward. Epigenetic modifying enzymes can be fused to the inactivated dCas9 (D10A mutation in RuvC and H840A in HNH nuclease domain) and used to specifically write or erase certain epigenetic marks at the target location. To date, various “writes” and “erasers” of histone modifications were successfully fused to dCas9, including histone acetyltransferase p300 [28], histone deacetylase (HDAC) [29], Lysine-specific histone demethylase 1 (LSD1) [30], as well as various DNA modifying enzymes like the eukaryotic DNMT3A [31,32] or the prokaryotic MQ1 [33] DNA methyltransferases and DNA demethylases from the TET family [31,34]. One of the newest applications of the CRISPR-Cas system also found its place in DNA base editing [35–37] to specifically correct point mutations that are the most common trigger of various diseases. Furthermore, the use of CRISPR-Cas system was raised to the next level by harnessing its ability to specifically identify attomolar quantities of DNA or RNA viruses [38,39], as well as to target and knockdown RNA molecules, repair RNA bases or to monitor RNA trafficking and localization in a cell [40,41].

3. CRISPR-Cas for gene activation (CRISPRa)

3.1. Commonly used transactivation domains

Transactivation domains in a (epi)genome editing toolbox (a set of components for expression of sgRNA, dCas9 and effector domains) need to be modular, in the sense that they retain their function in fusions with different DNA-binding moieties of such gene activation systems. Several potent activators have met the criteria and are commonly used for artificial transcriptional regulation. The first and the simplest activation domain is the virion protein 16 (VP16) from the herpes simplex virus type 1 transactivation domain, which has been successfully used for transcriptional activation in mammalian cells [42–44]. The VP64 domain is a tetramer of the minimal activator VP16 (four VP16 linked with Gly-Ser linkers), which was shown to be a stronger activator than a single VP16 and has been routinely used for artificial transcriptional activation ever since it was described [45]. Another commonly used transcriptional activation domain is the p65, the larger subunit of the NF-kappa B transcription factor, known to induce strong gene transactivation in mammalian cells [46]. The Rta domain is another transcriptional activator in this portfolio, encoded by Kaposi’s sarcoma associated herpesvirus/human herpesvirus 8 (KSHV/HHV-8). It activates the expression of viral genes in the lytic cycle [47,48]. The activation domain from human heat-shock factor 1 (HSF1) is another transactivation domain in the growing molecular toolbox. The HSF1 initiates transcription of chaperons in response to cellular stress induced by protein misfolding in human cells [49]. Finally, the VPR transactivator is a chimeric unit composed of the activation domains of VP64, p65 and Rta. It was recently developed for more robust gene activation by using several activation units, exhibiting a strong synergistic effect [50].

3.2. CRISPRa tools

CRISPRa tools consist of dCas9 fused in various configurations with the described transcriptional activation domains or their combinations. In 2013, Gilbert and collaborators created dCas9-VP64 and dCas9-p65 fusion proteins and achieved targeted reporter gene activation in HEK293 reporter cell line [51]. At the same time, another group reported successful activation of endogenous genes in HEK93 cells with their dCas9-VP64 fusion protein [52]. They compared dCas9-VP64 induced gene activation effects with previously reported effects achieved using activators where TALE (used for tatgeting) was fused to the VP64 domain (used for transcriptional activation) at the same endogenous loci [3]. Stronger activation was observed in the experiments using TALE activators, indicating that the designed CRISPRa tool needed improvement. In the same year, another group developed and tested the dCas9-VP64 activator, with which they obtained a substantial increase in gene activation when simultaneously targeting the same promoter region with multiple sgRNAs [53]. The VPR tripartite activation domain was designed and well characterized by George Church’s group and it was proven successful for robust multi-locus activation, drastically outperforming VP64 alone [50]. They tested the ability of Cas9 to simultaneously induce genomic modifications and modulate transcription in an approach that utilizes the Cas9-VPR fusion protein guided with sgRNAs of variable length, in order to determine whether the molecular tool will act as a targeted “scissors” or an activator of transcription [54]. Therefore, the irreversible double-strand cleavage would lead to functional gene silencing, while activation would proceed without cleavage when using truncated sgRNA. Loss of Cas9 nuclease function was observed when ≤ 16-nt sgRNA was used. Therefore, truncated sgRNAs were used for CRISPRa and full-length sgRNAs (20-nt) were used when targeted genome editing was needed. The Cas9-VPR targeted gene activation achieved 40% of the dCas9-VPR 20-nt sgRNA-guided gene activation (transcript level). For some of the loci analyzed for off-target cleavage activity, indels were observed with 16-nt sgRNAs, so off-target activity was clearly a concern in this approach [54]. Therefore, the dCas9-VPR approach seems more efficient, and the truncated sgRNA option remains an alternative when switching between DNA cleavage and transcriptional activation is needed.

Gao and collaborators suggested a combined approach using CRISPRa tools and TALE-based activators for more potent transcriptional activation [55]. Several other interesting CRISPRa studies reported successful gene activation using dCas9-based tools and mainly confirmed more robust transcriptional activation with multiple sgRNA targeting [56–62].

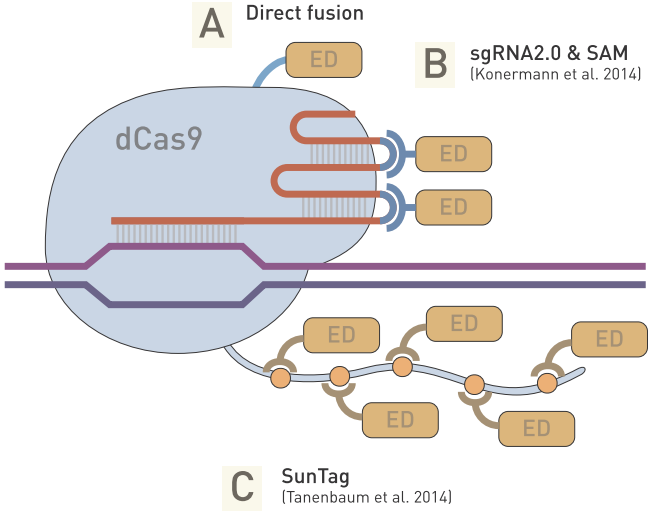


Fig. 1. Targeting of effector domains by dCas9 and amplification of their activity. The binding of dCas9 to its DNA target can be used to recruit effector domains (ED) for gene activation, repression or epigenetic modifications. Direct fusions of effector domains with dCas9 (A) can be used in conjunction with RNA aptamers on a modified sgRNA (B), which recruit effector domains fused to an aptamer-binding moiety of a phage protein, thus constituting the sgRNA2.0 system, or the SAM system when direct fusions and aptamers are used together (Konermann et al. [63]). In the SunTag system (C), the dCas9 is extended with a peptide containing multiple antibody binding sites, and a single chain antibody is used to recruit multiple effector domains (Tanenbaum et al. [64]).

3.3. Strategies for increasing the gene activation potential

Instead of applying multiple sgRNA targeting the same locus as a strategy to amplify CRISPRa effects, structural modifications of the core dCas9 systems were made in efforts to simultaneously amplify the effect as well as to simplify the delivery and experimental design for CRISPRa utilizing a single sgRNA molecule [63,64]. Systems for targeting effector domains and amplification of their activity are summarized in Fig. 1.

3.3.1. Synergistic activation mediator

Konermann and collaborators were first to introduce the sgRNA2.0 concept for improved CRISPRa in an attempt to achieve robust up-regulation of targeted genes with a single sgRNA [63]. They made structural changes to dCas9 system with sgRNA scaffold modification. Based on the previously determined crystal structure of the dSpCas9 in complex with sgRNA and targeted DNA [65], the authors concluded that some sgRNA loop structures do not affect dCas9 functions and could potentially tolerate structural changes while not interfering with dCas9 binding. The goal was to create a versatile sgRNA, capable of recruiting additional activation domains to dCas9 target site, so robust gene up-regulation could be achieved with a single targeting sgRNA through synergistic effect of multiple activation domains. Hairpins which form loops in the sgRNA that protrude from the Cas9 structure were extended with aptamers – stretches of RNA capable of binding other molecules, in this particular case two copies of aptamers binding bacteriophage MS2 coat proteins (MCPs). Effector domains fused with the MS2 protein were then recruited to such Cas9/sgRNA complex. This approach enabled recruitment of multiple effector domains (thus amplifying activity) and it was named “sgRNA2.0”. Examples of effector domains include KRAB and VPR [66]; the principle could be extended to include others. MCPs can dimerize and bind specific RNA hairpin aptamers, so transactivation domains fused to MCPs can bind these “dCas9-sgRNA2.0” complexes for more robust gene activation. The strongest transcriptional activation was achieved using the dCas9-VP64 targeting with a single sgRNA2.0 and the MS2-p65-HSF1 activation helper fusion. The system has been named synergistic activation mediator (SAM). The SAM demonstrated more potent gene activation than targeting of a conventional dCas9-VP64 tool combined with a pool of eight classic sgRNAs [67]. Conceptually, the selection of functionality was shifted from binding or fusion to the Cas9 protein to binding with the gRNA molecule. Alternatively, both approaches were used for their synergistic effect.

3.3.2. The SunTag activator system

A similar concept relying on multiple binding sites added to the Cas9 protein was introduced by Tannenbaum and collaborators [64]. They developed the SunTag system employing a recombinant peptide array (named SunTag) which possesses the ability to recruit multiple antibody-fusion proteins. The system was first used for improved fluorescent visualization via recruitment of multiple fluorescent protein molecules upon target binding [64]. Since the system turned out to be efficient for amplification of fluorescence signal, it has also found use as a transcription activating system by amplification of signals necessary for gene transcription. The SunTag activator system relies on the recruitment of VP64 transactivation domains bound to specifically modified NLS-tagged (for import into the nucleus) GCN4 antibodies to a single dCas9-SunTag molecule. Robust transcriptional activation was successfully achieved with SunTag-mediated recruitment of VP64 domains to dCas9 at the endogenous promoter of the *CXCR4* gene in K562 cell line using a single sgRNA molecule for targeting. Gene expression changes induced with dCas9-SunTag-VP64 were up to 25 times higher in comparison with conventional dCas9-VP64 mediated activation of the same locus. Thus, the SunTag activator system, as well as SAM, demonstrates an enormous potential for CRISPRa through activation domain synergy. Very recently, a modular dCas9-SunTag that can recruit multiple DNMT3A catalytic domains for targeted DNA methylation was developed. The authors demonstrated much higher induction of DNA methylation at target sites than achieved with dCas9-DNMT3A fusion, and importantly they showed minimal off-target protein binding and activity of DNMT3A [68].

3.3.3. Optogenetic CRISPR/dCas9 systems for CRISPRa

Optogenetic systems for gene expression modulation based on ZFP or TALE DNA targeting combined with light-inducible heterodimerizing proteins have been previously developed for light-controlled transcriptional manipulation [69,70]. Basic principle is that of two interaction partners, which heterodimerize upon blue light irradiation, one is fused to a DNA-binding protein and the other to a transcriptional activation domain. Recently, optogenetic systems which utilize dCas9 and photoinducible dimerization domains have been developed for precisely controlled modulation of gene expression in mammalian cells [71–73]. These systems exploit the properties of light-inducible heterodimerizing cryptochrome 2 (CRY2) protein and calcium and integrin-binding protein 1 (CIB1) from *Arabidopsis thaliana*, which heterodimerize upon light excitation [74]. The CRY2 and CIB1, or their truncated forms, are fused to different counterparts of a conventional CRISPRa tool – dCas9 and an activation domain. With a simple sgRNA for guidance and some blue light, targeted gene activation is immediately triggered by such systems.

Nihongaki and collaborators used a full length CIB1 protein fused to dCas9 for targeting endogenous genes via sgRNA guidance and a freely-floating fusion of CRY2 photolyase homology region (CRY2PHR) with an activation domain (VP64 or p65) and successfully induced gene activation upon blue light irradiation in HEK293T cells [71]. Three hours of blue light irradiation were sufficient for this system to induce a 10-fold increase in ASCL1 mRNA transcription, and transcription level was readily reduced to the baseline after incubation in the dark for 18 h. Upon repeated blue light irradiation, gene transcription rate was elevated again. This system achieved robust multiplexed gene photoactivation with multiple sgRNA targeting. Polstein and Gersbach characterized the light-activated CRISPR/Cas9 effector (LACE) system,

fusion of CIB1 N-terminal fragment (CIBN) and dCas9 paired with either full length or truncated CRY2 fused to VP64 activation domain [73]. In this approach, the CIBN was fused both C- and N-terminally to dCas9 in a single construct (CIBN-dCas9-CIBN) in order to increase the number of activation domains for synergistic enhancement of transcriptional activation.

3.4. CRISPR-Cas for gene silencing

CRISPR interference system (CRISPRi) was first described by Qi and collaborators [62]. CRISPRi utilizes dCas9 and sgRNAs complementary to transcription start sites of targeted genes and blocks transcription by interfering with transcription factor or RNA polymerase binding. Such CRISPRi was successful in repression of targeted genes [62]. The approach was further improved by recruitment of repression domains (such as the KRAB domain) to dCas9, which resulted in stronger repression [75]. Although CRISPRi has been used successfully to silence gene expression, its effects usually lead to a moderate reduction of transcript level, typically about 50%. For that reason, CRISPRi is discussed in less detail compared to CRISPR activation, which typically achieves increase in transcription tens to hundreds of times the original level. While the efficiency of the CRISPRi approach is comparable to that of RNAi, its fidelity seems superior [76].

The most commonly used repression domain is the Krüppel-associated box (KRAB), a highly conserved amino-terminal region of many Krüppel-class Cys2His2 zinc finger proteins with repressive function [77–79]. The KRAB domain acts with a corepressor KAP1, which recruits the heterochromatin protein 1 (HP1). The HP1 usually recognizes and binds methylated H3K9, achieved by histone methyltransferase activity (Suv39h1/2 and SETDB1) and in turn recruits more HMTs, which results in chromatin condensation of neighboring nucleosomes [80,81]. Chromo Shadow Domain (CSD) of HP1, a potent repression domain [82], is also a domain that strongly represses gene expression when targeted to promoter regions. The mSin3 interaction domain (SID), conserved in many transcriptional repressor proteins, with function to recruit histone deacetylases [83] is also frequently used for CRISPRi.

Gilbert and collaborators tested the ability of several fusions of dCas9 and different repression domains (including KRAB and CSD) to repress endogenous genes [51]. They used RNA-seq analysis to confirm that CRISPRi-mediated repression is highly specific. The dCas9-KRAB fusion has been tested extensively, and it seems that the dCas9-KRAB fusion protein can successfully recruit chromatin-modifying complexes for amplified CRISPRi effects. Lentivirus-mediated dCas9-KRAB expression tested in the same study resulted in stable and efficient knockdown of the *CD71* and *CXCR4* genes in HeLa cells. Other efficient fusions of dCas9 and KRAB have been developed for gene silencing by many groups in the past few years [84–86]. One group created a modular dCas9 platform for chemically inducible gene activation and repression and demonstrated simultaneous regulation of two genes using two different dCas9 orthologs for orthogonal gene regulation [66]. They used fusions of dCas9 from *Streptococcus pyogenes* (dSpCas9) with KRAB and dCas9 from *Staphylococcus aureus* (dSaCas9) with VPR in a dual reporter cell line for simultaneous activation and repression of reporter genes. The dSpCas9-VPR fusion targeted the *CD95* gene and the dSaCas9-VPR targeted the *CXCR4* gene simultaneously to activate expression of both genes in HEK293T cells upon co-expression with their respective sgRNAs.

George Church's group developed an improved dCas9-based repressor consisting of two repression domains fused C-terminally to dCas9, KRAB and MeCP2 (methyl CpG binding protein 2), yielding the dCas9-KRAB-MeCP2 fusion protein [87]. The system demonstrated superior gene silencing in comparison with both KRAB and MeCP2 single or double fusions to dCas9. The authors explained this robust silencing effect as the combined influence of the diverse repression domains. Systems like SAM and SunTag could also be repurposed for CRISPRi and could potentially enhance gene repression owing to their ability to recruit multiple effector domains (such as KRAB-MeCP2) to the targeted site in the genome.

4. Off-target effect of CRISPR-Cas based molecular tools

4.1. Why an off-target effect occurs in the CRISPR-Cas9 system?

Despite all advantages of the CRISPR-Cas9 technology, its main drawback is the still unresolved off-target effect. Strong interest in the off-target effect of Cas9 started when it was used in the genome editing context, *i.e.* when Cas9 was primarily used as a nuclease. Numerous genome engineering studies demonstrated off-target binding of the Cas9, which then cleaves DNA molecules at non-perfect matching sites in genome. Further, some sgRNA molecules were shown to be more specific than others; however, reasons for this are not entirely clear [88,89]. One study showed that 98.4% of 124,793 guide RNA molecules for SpCas9 targeting promoter or exon regions of genes have one or more off-target sites and that only 1.6% of them have perfect match [90]. This demonstrates that some off-target binding is almost inevitable in large and complex genomes. Indeed, some of the most serious concerns about the use of CRISPR-Cas9 system for therapeutic purposes are due to Cas9 promiscuity and toleration of imperfect matches.

An sgRNA molecule consists of a seed region located immediately upstream from the PAM sequence and a non-seed region located further to the 5′ end. Several studies showed that base pairing in the seed region is essential for proper targeting; the length of this critical region varies from 8 to 13 nucleotides [18,20,22,91]. Another study [92] reported that only 5 nucleotides immediately upstream from the PAM sequence can define the seed region, although it can be extended if PAM proximal region is U-rich, which largely influences thermodynamic stability of sgRNA binding. The same study showed that each sgRNA molecule has thousands of potential “seed-PAM” sites in the genome, and that less than 1% are bound by SpCas9, which is largely affected by chromatin accessibility. The correlation with chromatin accessibility showed that most of the off-target binding sites are located in actively transcribed genome regions, within promoter, enhancer or gene body regions [92].

Initially, it was thought that perfect complementary match between the first 20 nt of an sgRNA molecule along with the PAM sequence immediately downstream from the target site tightly controls Cas9 binding and target site cleavage. Several studies reported that the 5′ portion of sgRNA molecules can tolerate mismatches to a greater extent than the seed region and therefore the 5′ region largely contributes to the off-target effect. *In vitro* screening (outside of a cellular context) for Cas9 cleavage showed that up to 7 mismatches in sgRNA can be tolerated [93]. Off-target binding and cleavage with Cas9 was recorded at sites containing up to 5 mismatched nucleotides relative to the target site in human cells [88,89,93]. However, a more recent study showed that Cas9 can cleave off-target sites even with a 6 nt mismatch *in vivo* (in cultured cells) [94]. Testing Cas9 sensitivity to mismatches in the seed region revealed a “core” region, located from 4 to 7 nt upstream relative to the PAM sequence, to be extremely sensitive to mismatches, leading to abolished target cleavage [95].

Online tools for *in silico* design of guide RNA molecules can rank them according to sequence similarity to non-target sequences in the genome and can predict potential off-target sites. However, *in vivo* and *in vitro* experiments revealed more off-target sites. Some of these off-target sites can be explained by Cas9 binding to non-canonical PAM sequences. Even though a typical PAM sequence for SpCas9 is 5′-NGG-3′, several studies showed that 5′-NAG-3′ can be considered as an alternative PAM sequence [89,93,96,97]. In addition, 5′-NNG-3′ and 5′-NGN-3′ PAM variants were also tolerated for *in vitro* cleavage when SpCas9 was present in excess [93]. The limitations of the relatively simple sgRNA binding models have been recognized as a severe

impediment to successful application of *in silico* analysis for accurate Cas9 targeting and assessing efficacy of designed sgRNA molecules [98]. Moreover, the sgRNA efficacy (as opposed to expression level) has been shown to play a key role in determining successful application of Cas9-based molecular tools [99]. New highly successful algorithms for design of specific and efficient sgRNA molecules show a clear trend towards machine learning and data-driven approaches, which learn from experimentally validated examples and encapsulate the internal complexity of the model [100–102].

Major concern appeared after the discovery of a new class of off-target sites with extra bases (resulting in DNA bulges) or missing bases (resulting in RNA bulges). It was demonstrated experimentally that single base guide RNA bulges are well tolerated at many positions, due to the flexibility of RNA molecules, except for the PAM proximal position where they abolish Cas9 activity. Also, tolerance for up to 4 bp long RNA bulges were proved in HEK293T cells. In contrast to RNA bulges, only single base DNA bulges were tolerated at few positions, PAM proximal, middle and PAM distal position. It was also demonstrated that Cas9 tolerates single base DNA bulge coupled with up to three mismatches between target DNA and guide RNA [103].

Many present studies put effort into solving off-target effect of Cas9. No approach has shown complete absence of off-target effect, so its minimization and the balance between on-target *vs.* off-target effect is the main goal in experiments using CRISRP-Cas9 technology. It is worth noting that the off-target effect affects the outcome differently for cleavage by the Cas9 nuclease, where a specific locus is either cut or not, and scenarios where dCas9 is used for targeting effector domains, in which case the off-target effect shows a gradual, quantitative response with strength proportional to the degree of off-target binding. Therefore, gene activation, repression and epigenome editing, which are inherently gradual in their effect, seem to be more "forgiving" compared to the Cas9 nuclease, if the off-target effect is slight. Newer published literature supports this view [104]. Since CRISPR-Cas9 technology is very promising in therapeutic applications, solving of the off-target effect is of utmost importance in future research.

4.2. Strategies to minimize off-target effect of the CRISPR-Cas9 system

4.2.1. Changing the effective concentration and delivery of sgRNA-Cas9 complexes

In vitro and *in vivo* experiments showed that the concentration of sgRNA-Cas9 complexes affects non-specific binding. At larger effective concentrations of complexes relative to DNA, nuclease Cas9 can tolerate more mismatches. Transfecting a lower amount of plasmids coding for the nuclease Cas9 and its sgRNA is one potential strategy to minimize non-specific binding [89,93], so titration of Cas9-sgRNA concentration is needed. If direct transfection of cells is performed, *e.g.* if cells are transfected with purified ribonucleoprotein Cas9-sgRNA complexes, the enzymes are immediately active, and much more unstable than when integrated within a plasmid – they have shorter lifetimes and are present in lower intracellular concentration than complexes derived from plasmids. Several studies reported that such an approach can significantly decrease off-target effects of Cas9 complexes [105,106].

Another way to affect the concentration of Cas9-sgRNA ribonucleoprotein complexes is to reduce the amount of sgRNA molecules in a cell, because it can directly limit the concentration of functional complexes and reduce off-target effect. To do that, an sgRNA molecule module can be moved to another plasmid and its concentration can be reduced in transfection [92], so that the optimal concentration for the best on-target *vs* off-target ratio can be determined by titration. Also, the use of alternative weaker promoters of RNA polymerase III for expressing sgRNA would directly affect the amount of sgRNA transcription, number of functional sgRNA-Cas9 complexes and consequently the off-target effect. The weaker H1 promoter has been reported to have greater specificity than the commonly used U6 promoter, while only slightly reducing on-target cleavage [107]. In line with this, another study showed that the H1 promoter produced a lower amount of guide RNA transcripts than did the U6 promoter [108]. The H1 promoter can also function as an efficient RNA polymerase II promoter for robust mRNA transcription [109], and was able to drive transcription of Cas9 in amounts comparable to those driven by the CBh promoter [108]. A very recent study by Gao and collaborators showed that simultaneous expression of both components of CRISPR-Cas9 system, sgRNA molecule and the Cas9, from a single H1 promoter, made the system more specific due to reduced expression of sgRNA molecule (up to 7-fold) and Cas9 (up to 11-fold) when compared with individual components expressed from two different promoters [108].

Different inducible or repressible systems for functional Cas9 expression or assembly can also reduce binding at off-target sites, due to temporal control of functional Cas9 effective concentration in a cell. Cas9 expression induced by small molecules that easily diffuse into a cell has successfully been used in a doxycycline-inducible Cas9 system [110]. Other approaches that control functional Cas9 assembly, like 4-hydroxytamoxifen-inducible intein-Cas9 [111] and rapamycin-inducible split Cas9 systems [112], also utilize small molecules as triggers. To avoid the use of chemical control for Cas9 induction, a novel approach with blue light-inducible paCas9-2 system makes the control of functional Cas9 assembly even more non-invasive (*i.e.* activated by light only) and reversible by optical control [113]. In contrast to inducible systems, the self-restricted CRISPR system [114] utilizes an additional sgRNA that targets Cas9 itself, thereby limiting its expression, while performing successful genome editing with significantly improved specificity. Among these strategies, doxycycline-inducible Cas9 and self-restricted CRISPR system showed the on-target activity comparable to constitutively expressed wild type SpCas9 [110,114].

4.2.2. Rational design of sgRNA molecules

Rational design and *in silico* verification of sgRNA can significantly reduce the off-target effect. Running simulations of sgRNA binding across all possible sites within the targeted genome facilitates design of specific sgRNA molecules that are likely to bind a unique target and perform as intended. Poly T repeats should be avoided in guide RNA due to the possibility of premature termination of U6-driven transcription. GC content should be balanced (neither too low nor too high) to maximize cleavage efficiency [27,115] as well as to minimize occurrence of DNA and RNA bulges [103]. Change in the sgRNA length can also affect the level of off-target activity. Initially, the GN_{19}-NGG rule was used to design guide RNA molecules for SpCas9. By simply adding two extra guanines (G) at the 5′ end (GGN_{20}-NGG or $GGGN_{19}$-NGG), sgRNA molecules become much more specific, while maintaining on-target activity, even though it can be slightly lower for some sgRNA molecules [116]. Another approach using truncated sgRNAs (tru-sgRNAs) showed that GN_{17}-NGG design is the best for on-target *vs* off-target effect. For efficient on-target effect, SpCas9 requires a minimum of 17 nt to bind target sites, while the addition of one extra mismatched guanine (G) at 5′ end of sgRNA makes an sgRNA less tolerant to any mismatch because it largely affects the binding energy and stability of the SpCas9-sgRNA complex. Thus, the extra 5′ guanine that was initially added only to facilitate efficient transcription from the U6 promoter was found to improve specificity by slightly destabilizing the target binding process; the gains in specificity usually outweight the loss of on-target efficiency [117]. With further reduction of guide RNA length, Cas9 activity decreases dramatically or its expression even becomes undetectable [118]. The hypothesis that Cas9 lacks its nuclease activity but retains DNA binding activity if sgRNA molecules 14-15 nt in length are used for guidance has been proven by Kiani and collaborators [54]. The advantage of "dead" sgRNAs was used in the same study to induce transcriptional activation with SpCas9-VPR fusion, where SpCas9 was catalytically active. Using sgRNA molecules 20 and 18 nt in length (even 16 nt for some genomic loci), the SpCas9-VPR fusion showed robust nuclease activity at target sites, comparable to wild-type SpCas9, with minimal gene activation. Gene activation reached its

Table 1
Properties of various Cas9 orthologs and their variants derived by mutagenesis.

Cas9 orthologs and variants	Bacterial species	Size (aa)	PAM sequence (5′-3′)	sgRNA	Strengths	Limitations	Refs.
SpCas9	*Streptococcus pyogenes*	1368	NGG	20 nt	High activity, well characterized, active fusion constructs with different effector domains, works with truncated sgRNA molecules	Large size, high off-target effect	Cong et al. [18], Mali et al. [61]
xCas9 3.7			NGG/NG/GAA/GAT	20 nt (with matched first G)	Expanded targeting range, high activity at canonical and wide range of non-canonical PAM sequences, lower off-target effect than wild-type Cas9	Large size,On-target activity attenuated significantly with truncated sgRNA molecules	Hu et al. [127]
Sniper-Cas9			NGG	19–21 nt	High activity with full length, truncated and extended sgRNA molecules, highest specificity	Large size	Lee et al. [129]
SaCas9	*Staphylococcus aureus*	1053	NNGRRT	20–24 nt	Activity comparable with SpCas9, lower off-target activity, smaller size	Narrow targeting range due to complex PAM requirements	Nishimasu et al. [135], Ran et al. [136], Ye et al. 137)
SaCas9-KKH			NNNRRT	21–22 nt	Expanded targeting range, similar to wild-type SaCas9 regarding the number of off-target sites	Slightly lower on-target activity at some target sites with NNGRRT PAM than wild-type SaCas9, narrow targeting range due to complex PAM requirements	Kleinstiver et al. [151]
St1Cas9	*Streptococcus thermophilus*	1122	NNAGAAW	19–20 nt	More specific than SpCas9	Lower on-target activity than SpCas9, narrow targeting range due to complex PAM requirements	(Cong et al. [18], Esvelt et al. [138], Müller et al. [139]
St3Cas9		1393	NGGNG		Greater targeting efficiency than St1Cas9, more specific than SpCas9		Glemzaite et al. [139], Müller et al. [140], Xu et al. 141
NmCas9	*Neisseria meningitidis*	1081	NNNNGATT/NNNNGHTT	23–24 nt for use with crRNA and tracrRNA, 21 nt for sgRNA	More specific than SpCas9, potential for temporal control of NmCas9 activity with anti-CRISPR proteins	Lower on-target activity than SpCas9 (poor DNA unwinding), narrow targeting range due to complex PAM requirements	Esvelt et al. [138], Hou et al. [142], Lee et al., [143], Pawluk et al. [156]
FnCas9	*Francisella novicida*	1629	NGG	20 nt	DNA and RNA targeting activity	Large size, low on-target activity due to poor DNA unwinding ability	Chen et al. [157], Hirano et al. [158], Price et al. (2015), Zhang et al. [161]
FnCas9-RHA			YG		Expanded targeting range		Hirano et al. [157]
CjCas9	*Campylobacter jejuni*	984	NNNNACAC/NNNNRYAC	22 nt	Smallest Cas9 ortholog, on-target activity comparable with SpCas9, reduced off-target binding	Narrow targeting range due to complex PAM requirements	Kim et al. [144]

maximum if the SpCas9-VPR fusion was guided with a 14 nt long sgRNA, while the SpCas9 nuclease activity was abolished. Mismatch tolerance for 14 nt long sgRNA molecules was decreased compared to full-length 20 nt sgRNA molecules [54]. This approach was successfully used to perform simultaneous gene activation and knock-out at different gene loci using the same construct coupled with sgRNA molecules of different length [54,119].

Except for varying the length of sgRNA molecules, different chemical modifications can also decrease off-target effect [120,121]. A novel approach with chimeric guide RNAs, which contain partial DNA nucleotides at their 5′ end, can also decrease off-target effect. Due to the lower thermodynamic stability of a DNA-DNA duplex, chimeric sgRNAs or crRNAs with up to 10 DNA nucleotides at 5′end, are less tolerant to mismatches [122].

4.2.3. Engineering and rational design of new Cas9 variants

Nowadays, engineering and rational design of new Cas9 variants is the modern approach to increase CRISPR-Cas9 specificity. Many mutants are available which show significantly lower off-target effect. According to structural studies, manipulation of SpCas9-DNA contacts affects the stability of SpCas9-sgRNA complex with DNA and its toleration to mismatches. Alteration of the key amino acid residues (N497, R661, Q695, Q926) in SpCas9, responsible for direct contact with the target strand in the SpCas9-HF1 (“high fidelity”) variant, greatly reduces non-specific DNA binding, while maintaining at least 70% of on-target activity compared to the wild-type SpCas9 [123]. Alternatively, changes of amino acids involved in non-target strand binding in the mutant eSpCas9(1.1) (enhanced specificity SpCas9 version 1.1) encourage DNA re-hybridization. On-target effect remains comparable with wild-type SpCas9, while reducing off-target effect significantly [124]. Novel structural studies revealed that the REC3 domain of SpCas9 is necessary for guiding the RNA-target DNA heteroduplex interaction, and that the REC3 domain keeps SpCas9 in an inactive form if mismatches are present. Mutations in the REC3 region (N692A, M694A, Q695A, H698A) in the mutant HypaCas9 makes heteroduplex binding and verification of target site binding stricter, resulting in minimization of the off-target effect. This variant performs even better than the previously used mutants, while maintaining high on-target activity [125].

Directed evolution of SpCas9 was performed by screening random mutations in the REC3 domain, to find key mutations that confer even better specificity, which resulted in the discovery of the evoCas9 mutant (M495V, Y515N, K526E, R661Q). On-target activity is slightly lower compared to wild-type SpCas9, while specificity is largely improved, up to 4-fold compared to the SpCas9-HF1 and eSpCas9(1.1) [126]. Recently, by the use of phage-assisted continuous evolution (PACE), a novel mutant xCas9 3.7 showed much lower off-target effect compared to wild type SpCas9, with the ability to recognize a wide range of non-NGG PAM sequences, such as 5′-NG-3′, 5′-GAA-3′ and 5′-GAT-3′. The xCas9 3.7 mutant outperformed the wild type Cas9 when used in fusion with the transcription activator VPR at all non-NGG PAM proximal sites, as well as at NGG PAM proximal sites. When considering genomic DNA, nuclease activity of the xCas9 3.7 slightly outperformed the wild-type SpCas9 at NGG PAM proximal sites, while for all non-NGG PAM proximal sites its activity was significantly higher [127]. More recently, using direct evolution in *E. coli* combined with positive and negative selection, the Sniper-Cas9 mutant was developed. The Sniper-Cas9 shows even higher specificity ratio, and can be efficiently combined with truncated and extended sgRNA molecules, in contrast to previously described mutants SpCas9-HF1, eSpCas9(1.1), HypaCas9, evoCas9 and xCas9 3.7, all of which showed the best activity only with GN19 sgRNAs where first guanine (G) is matched to the target site [128–130].

Ribonucleoprotein delivery (RNP) is a commonly used method to reduce off-target effect [131–133]. Previously mentioned engineered SpCas9 variants have lower affinity for their DNA targets due to extensive mutations of amino acids which are responsible for direct contacts with the target or the non-target strand. Therefore, RNP delivery of such mutants often shows reduced on-target effect. In contrast to the RNP delivery, plasmid delivery by transfection overcomes this problem, due to the presence of higher effective concentration of sgRNA-Cas9 complexes after transcription. A new SpCas9 variant, the Hifi Cas9, shows the best activity when RNP is delivered to primary cells, stem cells or progenitor cells. The Hifi Cas9 combines two different approaches to minimize the off-target effect – the advantages of RNP delivery combined with the mutation R691A in REC3 domain, which inhibits Cas9 activation if mismatches are present, and at the same time maintains a high level of on-target activity. In direct comparison with the wild-type SpCas9 along with the mutants eSpCas9(1.1), SpCas9-HF1 and HypaCas9, the Hifi Cas9 shows comparable but slightly lower on-target activity, but is much more active than all other mutants tested [134]. An overview of Cas9 mutants can be found in Table 1.

4.2.4. Use of Cas9 orthologs

At this moment, a broad range of Cas9 orthologs is available for CRISPR/Cas9-based genome and epigenome editing. By requiring different, often more complex PAM sequences, they may possess greater specificity in contrast to the wild-type SpCas9. More complex PAM sequences appear less frequently in genome than the 5′-NGG-3′ PAM sequence for SpCas9, thereby minimizing the number of potential off-target sites in genome, but also inconveniently reducing the number of potential on-target sites. Various orthologs have shown their potential to be used in fusions for manipulation of human cells: *Staphylococcus aureus* Cas9 (SaCas9) [135–137], *Streptococcus thermophilus* Cas9 (St1Cas9 [18,138,139], St3Cas9 [139–141]), *Neisseria meningitidis* Cas9 (NmCas9) [138,142,143], *Campylobacter jejuni* Cas9 (CjCas9) [144], as well as different classes of Cas12a family (also known as Cpf1), *Lachnospiraceae* bacterium ND2006 Cas12a (LbCas12a) and *Acidaminococcus* sp. BV3L6 Cas12a (AsCas12a) [145–147]. Another interesting CRISPR effector, the Cas12b (formerly known as C2c1), has been characterized recently [148]. The ortholog from *Bacillus hisashii* (BhCas12b) is compact (1,108 amino acids) and exhibits greater specificity than the SpCas9, which makes its engineered version a candidate for genome and epigenome editing in human cells [148]. Applications and further development of the new discoveries in the exciting CRISPR field are very promising; however, at this point, only the SaCas9 and the CjCas9 have verified on-target activity comparable with that of the SpCas9 [136,144]. In addition to being one of the smallest Cas9 orthologs, the CjCas9 also showed specificity superior to that of the SaCas9 and the SpCas9, which makes it a promising candidate for therapeutic purposes. An overview of Cas9 orthologs and their properties is given in Table 1.

4.2.5. Class II (Type II) Cas9 orthologs

The SaCas9, derived from the bacterium *Staphylococcus aureus*, is among the smallest Cas9 orthologs, containing only 1053 amino acids. It recognizes a more complex PAM sequence compared to SpCas9 (5′-NNGRRT-3′, where N represents any base and R represents purine bases). Its targeting activity and efficiency is comparable to that of SpCas9 and can be directly tested due to their PAM sequences that are not mutually exclusive. Maximum of activity in eukaryotic cells is achieved with guide RNA molecules from 20 to 24 nt in length, while guide RNAs shorter than 17 nt abolish SaCas9 activity. Off-target activity is lower than that of SpCas9, especially when targeting with guide RNA molecules 20 nt in length, which led to the conclusion that the SaCas9 is more specific than the SpCas9. Another advantage of using the SaCas9 is its smaller size, which enables packaging into single adeno-associated virus (AAV) particles, along with U6-driven guide RNA, which creates an opportunity for *in vivo* delivery [136,149,150]. The main disadvantage of using the wild-type SaCas9 is its limited targeting range due to the complex and extended PAM sequence. Even though it was shown that thymine is preferred at 6th position in the

PAM sequence, the SaCas9 can cleave (at least to some extent) target sites with NNGRR PAM sequence [136]. To expand SaCas9 targeting range and flexibility for two-to four-fold, the SaCas9-KKH variant (E782K, N968K, R1015H) has been evolved, which recognizes an altered 5′-NNNRRT-3′ PAM sequence. The SaCas9-KKH variant has preserved the robust on-target activity in eukaryotic cells, while the off-target activity remained at the same level as for the wild-type SaCas9 [151]. Recently, nuclease-null dSaCas9 (D10A and N580A) was fused with the KRAB domain to perform direct gene repression. Gene expression of the *Pcsk9* was successfully reduced using the dSaCas9-KRAB fusion construct, which resulted in reduced levels of LDL cholesterol in mice [152].

To further expand the targeting range, other orthologs with different PAM requirements can be utilized. The St1Cas9 and St3Cas9 proteins, derived from *Streptococcus thermophilus*, are encoded by different loci, *CRISPR1* and *CRISPR3*. The St1Cas9 specifically recognizes 5′-NNAGAAW-3′, while the St3Cas9 recognizes 5′- NGGNG-3′ PAM sequence (N represents any base, W represents adenine (A) or thymine (T)) [153,154]. Although it was shown that the St3Cas9 possesses greater targeting efficiency when compared to the St1Cas9 at some human loci, both orthologs have lower targeting activity when compared with the SpCas9. In contrast to targeting efficiency, when considering off-target activity they are more specific than the SpCas9, probably due to the requirement for a more complex PAM sequence [136,139,141]. Both orthologs require 20 nt long guide RNA for proper targeting. Reduction of its length can be tolerated only for 1 nucleotide, while further reduction almost completely abolishes the St1Cas9 and St3Cas9 activity [139].

The NmCas9 derived from *Neisseria meningitidis* is another small Cas9 ortholog (1081 amino acids) which specifically recognizes the 5′-NNNNGATT-3′ PAM sequence [97]. It has been experimentally established that an adenine (A) is not favored as the first base in its PAM sequence, and a thymine (T) is not favored as the first PAM proximal base in the protospacer. Several studies showed that an alternative 5′-NNNNGHTT-3′ PAM sequence (H represents adenine, thymine and cytosine) is compatible with high on-target activity [138,143]. When the NmCas9 system is used in eukaryotic cells along with crRNA and tracrRNA, separately expressed from different U6 promoters, longer guide RNA length (23–24 nt) showed better activity, whereas use of the NmCas9 in combination with sgRNA prefers shorter guide RNA length (21 nt). When compared with SpCas9, off-target activity was generally lower, but also was the on-target activity at all tested loci [143]. Several studies demonstrated the successful use of the NmCas9 system for genome editing in human HEK293T cells as well as in human induced pluripotent stem cells (iPSC) [138,142,143]. Recently, the NmCas9 along with its U6-sgRNA was packaged into AAV particles, and it successfully targeted the *Pcsk9* gene in mice [155]. The authors showed the presence of naturally occurring anti-CRISPR proteins that can specifically bind to the NmCas9, thus preventing its target site binding and activity. Their ability to inhibit the NmCas9 activity was demonstrated in human HEK293 cells with anti-CRISPR protein AcrIIC3$_{Nme}$, showing the ability to completely inhibit the NmCas9 [156], which opens the new possibilities to temper NmCas9 activity.

One of the largest Cas9 orthologs is the FnCas9 (1,629 amino acids) derived from *Francisella novicida*. Its experimentally determined PAM sequence is 5′-NGG-3′ (N represents any base), but it also shows slight activity with adenine at the second and third position. Based on structural studies, the FnCas9-RHA variant with substitutions E1369R, E1449H and R1556A was evolved, expanding the targeting range by recognizing the alternative 5′-YG-3′ PAM sequence (Y represents cytosine or thymine) while maintaining activity comparable to that of the wild-type FnCas9. Functional validation failed when expressed in human 293FT cells. In contrast, pre-assembled RNP complexes induced high activity at the Tet1EX4 locus when injected into mouse zygotes [157]. Later, it was clarified that low activity of the FnCas9 in human cells is mainly due to its inability to access target sites in complex chromatin structures, which was proven by another study showing variable activity of FnCas9 in human cells and poor activity at most of the target sites. The limitation imposed by the chromatin environment was successfully removed by the “proxy-CRISPR” approach, where catalytically inactive SpCas9 was targeted near the FnCas9 target site, ostensibly opening the local chromatin structure and thus facilitating the FnCas9 nuclease activity at that site [158]. In addition to DNA targeting, the FnCas9 possesses RNA targeting activity with a distinct PAM independent mechanism containing small CRISPR/Cas-associated RNA (scaRNA):tracrRNA complex, independent of *RuvC-I and the HNH nuclease domains* [159]. RNA targeting activity of the FnCas9 was successfully utilized to target hepatitis C virus (HCV) in infected human hepatocellular carcinoma cells (Huh-7.5) [160], as well as to establish RNA virus resistance to tobacco mosaic virus (TMV) or cucumber mosaic virus (CMV) in plants [161].

One of the smallest Cas9 orthologs is the CjCas9 (984 amino acids) derived from *Campylobacter jejuni*. Using the PAM discovery assay, 5′-NNNNACAC-3′ and 5′-NNNNRYAC-3′ (N represents any base, R represents purines, Y represents pyrimidines) were determined as PAM sequences. It was found that the optimal PAM sequence in human cells is 5′-NNNNACAC-3′. The GX$_{22}$ rule for sgRNA design exhibited the best activity, while shorter sgRNAs (20 or 19 nt in length) showed lower efficiency or completely abolished CjCas9 activity at some tested loci. In comparison with the SpCas9 and SaCas9 orthologs, the CjCas9 is more specific, while having comparably high on-target activity. The CjCas9 was efficiently packaged into AAV particles along with its U6-sgRNA, and successfully targeted to several loci in muscle and retinal cells in mice [144]. Its small size leaves more space for other elements to be packaged into AAV particles. Combined with its high on-target activity and significantly reduced off-target effect, this gives the CjCas9 a great potential for use in therapeutic settings.

When considering dCas9-based tools for epigenome editing, catalytic domains of epigenetic editors in such fusions can also contribute to off-target activity, which is independent of dCas9 binding promiscuity, but instead originates from the presence of an overexpressed catalytic domain. Dose-response experiments are extremely important in such cases, because they are essentially the only means to reduce non-specific activity due to catalytic domain overexpression [32].

5. Conclusions and future directions

CRISPR/(d)Cas9 molecular tools have already started to fundamentally change the way of examining eukaryotic genome and epigenome. The field of epigenetics is turning from being correlative (*i.e.* using an approach to correlate epigenetic modifications with status of gene transcription in different experiments) to a field which can directly infer the causal relationship between the epigenetic and transcriptional status of a gene through direct manipulation of specific genome regions. This can be performed within a living cell or a whole organism by the ability to introduce (d)Cas9-based tools directly to the cell nucleus. More importantly, genome and epigenome engineering have an immense potential to revolutionize the approach to cure human diseases. Aberrant epigenetic signatures are almost always associated with complex diseases [162,163]. Unlike mutations, epigenetic modifications which define specific gene states (active *vs.* silenced) are reversible, which offers exciting possibilities to reprogram the diseased to a healthy cell state [12]. Use of CRISPR/dCas9 for epigenetic editing has a great potential in development of precise medicine [31,164]. However, this task still requires more work on improving the Cas9-based tools because of the two most important challenges: (1) safe and precise delivery of sophisticated Cas9-tools to a target tissue/organ; (2) a highly specific activity of the imported tools at the targeted locus/genome region. Finally, there are ethical concerns with using CRISPR-Cas and related technologies in medicine, best illustrated by the recent controversies around “CRISPR babies”, which can serve as a cautionary tale urging careful translation of new methods into clinical settings

[165,166]. Nevertheless, CRISPR/(d)Cas9-based tools open exciting new possibilities to better understand the molecular basis not only of cancer, but of many complex diseases, as well as to facilitate the design of new causative, precise therapies.

References

[1] R.R. Beerli, C.F. Barbas, Engineering polydactyl zinc-finger transcription factors, Nat. Biotechnol. 20 (2) (2002) 135–141.
[2] A.J. Keung, C.J. Bashor, S. Kiriakov, J.J. Collins, A.S. Khalil, Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation, Cell 158 (1) (2014) 110–120.
[3] M.L. Maeder, S.J. Linder, D. Reyon, J.F. Angstman, Y. Fu, J.D. Sander, J.K. Joung, Robust, synergistic regulation of human gene expression using TALE activators, Nat. Methods 10 (3) (2013) 243–245.
[4] P. Perez-Pinera, D.G. Ousterout, J.M. Brunger, A.M. Farin, K.A. Glass, F. Guilak, G.E. Crawford, A.J. Hartemink, C.A. Gersbach, Synergistic and tunable human gene activation by combinations of synthetic transcription factors, Nat. Methods 10 (3) (2013) 239–242.
[5] X. Wang, P. Wang, Z. Fu, H. Ji, X. Qu, H. Zeng, X. Zhu, J. Deng, P. Lu, S. Zha, Z. Song, H. Zhu, Designed transcription activator-like effector proteins efficiently induced the expression of latent HIV-1 in latently infected cells, AIDS Res. Human Retrovir. 31 (1) (2015) 98–106.
[6] A. Wilber, U. Tschulena, P.W. Hargrove, Y.-S. Kim, D.A. Persons, C.F. Barbas, A.W. Nienhuis, A zinc-finger transcriptional activator designed to interact with the γ-globin gene promoters enhances fetal hemoglobin production in primary human adult erythroblasts, Blood blood-2009 (2010).
[7] B. Zhang, S. Xiang, Q. Zhong, Y. Yin, L. Gu, D. Deng, The p16-specific reactivation and inhibition of cell migration through demethylation of CpG islands by engineered transcription factors, Human Gene Ther. 23 (10) (2012) 1071–1081.
[8] F. Zhang, L. Cong, S. Lodato, S. Kosuri, G.M. Church, P. Arlotta, Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription, Nat. Biotechnol. 29 (2) (2011) 149–153.
[9] S. Baliou, M. Adamaki, A.M. Kyriakopoulos, D.A. Spandidos, M. Panayiotidis, I. Christodoulou, V. Zoumpourlis, Role of the CRISPR system in controlling gene transcription and monitoring cell fate, Mol. Med. Rep. 17 (1) (2018) 1421–1427.
[10] R.M. Genga, N.A. Kearns, R. Maehr, Controlling transcription in human pluripotent stem cells using CRISPR-effectors, Methods 101 (2016) 36–42.
[11] Marie F. La Russa, Lei S. Qi, The new state of the Art: Cas9 for Gene activation and repression, Mol. Cell. Biol. 35 (22) (2015) 3800 LP – 3809.
[12] J. Pulecio, N. Verma, E. Mejía-Ramírez, D. Huangfu, A. Raya, CRISPR/Cas9-based engineering of the epigenome, Cell Stem Cell 21 (4) (2017) 431–447.
[13] P.I. Thakore, J.B. Black, I.B. Hilton, C.A. Gersbach, Editing the epigenome: technologies for programmable transcription and epigenetic modulation, Nat. Methods 13 (2) (2016) 127–137.
[14] R. Barrangou, C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D.A. Romero, P. Horvath, CRISPR provides acquired resistance against viruses in prokaryotes, Science (2007).
[15] G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria, Proceed. Natl. Acad. Sci. (2012).
[16] P. Horvath, R. Barrangou, CRISPR/Cas, the immune system of bacteria and archaea, Science 327 (5962) (2010) 167–170.
[17] M.P. Terns, R.M. Terns, CRISPR-based adaptive immune systems, Curr. Opin. Microbiol. (2011).
[18] L. Cong, F.A. Ran, D. Cox, S. Lin, R. Barretto, P.D. Hsu, X. Wu, W. Jiang, L.A. Marraffini, Multiplex genome engineering using CRISPR/Cas systems, Science 339 (6121) (2013) 819–823.
[19] E. Deltcheva, K. Chylinski, C.M. Sharma, K. Gonzales, Y. Chao, Z.A. Pirzada, M.R. Eckert, J. Vogel, E. Charpentier, CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III, Nature (2011).
[20] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A programmable dual-RNA–guided dna endonuclease in adaptive bacterial immunity, Science 337 (August) (2012) 816–822.
[21] F.J.M. Mojica, C. Díez-Villaseñor, J. García-Martínez, C. Almendros, Short motif sequences determine the targets of the prokaryotic CRISPR defence system, Microbiology (2009).
[22] S.H. Sternberg, S. Redding, M. Jinek, E.C. Greene, J.A. Doudna, DNA interrogation by the CRISPR RNA-guided endonuclease Cas9, Nature (2014).
[23] Y. Feng, S. Sassi, J.K. Shen, X. Yang, Y. Gao, E. Osaka, J. Zhang, S. Yang, C. Yang, H.J. Mankin, F.J. Hornicek, Z. Duan, Targeting Cdk11 in osteosarcoma cells using the CRISPR-cas9 system, J. Orthop. Res. (2015).
[24] M. Lopez-Obando, B. Hoffmann, C. Gery, A. Guyon-Debast, E. Teoule, C. Rameau, S. Bonhomme, F. Nogue, Simple and efficient targeting of multiple genes through CRISPR-Cas9 in physcomitrella patens, G3, Genes|Genomes|Genetics (2016).
[25] D. Seruggia, A. Fernandez, M. Cantero, P. Pelczar, L. Montoliu, Functional validation of mouse tyrosinase non-coding regulatory DNA elements by CRISPR-Cas9-mediated mutagenesis, Nucl. Acids Res. (2015).
[26] O. Shalem, N.E. Sanjana, E. Hartenian, X. Shi, D.A. Scott, T.S. Mikkelsen, D. Heckl, B.L. Ebert, D.E. Root, J.G. Doench, F. Zhang, Genome-scale CRISPR-Cas9 knockout screening in human cells, Science (2014).
[27] T. Wang, J.J. Wei, D.M. Sabatini, E.S. Lander, Genetic screens in human cells using the CRISPR-Cas9 system, BMJ Support. Palliat. Care (2012).
[28] I.B. Hilton, A.M. D'Ippolito, C.M. Vockley, P.I. Thakore, G.E. Crawford, T.E. Reddy, C.A. Gersbach, Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers, Nat. Biotechnol. (2015).
[29] D.Y. Kwon, Y.T. Zhao, J.M. Lamonica, Z. Zhou, Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC, Nat. Commun. (2017).
[30] N.A. Kearns, H. Pham, B. Tabak, R.M. Genga, N.J. Silverstein, M. Garber, R. Maehr, Functional annotation of native enhancers with a Cas9-histone demethylase fusion, Nat. Methods (2015).
[31] X.S. Liu, H. Wu, X. Ji, Y. Stelzer, X. Wu, S. Czauderna, J. Shu, D. Dadon, R.A. Young, R. Jaenisch, Editing DNA methylation in the mammalian genome, Cell (2016).
[32] A. Vojta, P. Dobrinic, V. Tadic, L. Bockor, P. Korac, B. Julg, M. Klasic, V. Zoldos, Repurposing the CRISPR-Cas9 system for targeted DNA methylation, Nucl. Acids Res. 44 (12) (2016) 5615–5628.
[33] Y. Lei, X. Zhang, J. Su, M. Jeong, M.C. Gundry, Y.H. Huang, Y. Zhou, W. Li, M.A. Goodell, Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein, Nat. Commun. (2017).
[34] X. Xu, Y. Tao, X. Gao, L. Zhang, X. Li, W. Zou, K. Ruan, F. Wang, G.L. Xu, R. Hu, A CRISPR-based approach for targeted DNA demethylation, Cell Discovery (2016).
[35] N.M. Gaudelli, A.C. Komor, H.A. Rees, M.S. Packer, A.H. Badran, D.I. Bryson, D.R. Liu, Programmable base editing of A • T to G • C in genomic DNA without DNA cleavage, Nature 551 (7681) (2017) 464–471.
[36] A.C. Komor, Y.B. Kim, M.S. Packer, J.A. Zuris, D.R. Liu, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage, Nature 533 (7603) (2016) 420–424.
[37] A.C. Komor, K.T. Zhao, M.S. Packer, N.M. Gaudelli, A.L. Waterbury, L.W. Koblan, Y.B. Kim, A.H. Badran, D.R. Liu, Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T: A base editors with higher efficiency and product purity, Sci. Adv. 3 (8) (2017) 1–9.
[38] J.S. Chen, E. Ma, L.B. Harrington, M. Da Costa, X. Tian, J.M. Palefsky, J.A. Doudna, CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity, Science (2018).
[39] J.S. Gootenberg, O.O. Abudayyeh, M.J. Kellner, J. Joung, J.J. Collins, F. Zhang, Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a and Csm6, Science (2018).
[40] O.O. Abudayyeh, J.S. Gootenberg, P. Essletzbichler, S. Han, J. Joung, J.J. Belanto, V. Verdine, D.B.T. Cox, M.J. Kellner, A. Regev, E.S. Lander, D.F. Voytas, A.Y. Ting, F. Zhang, RNA targeting with CRISPR–Cas13, Nature (2017).
[41] D.B.T. Cox, J.S. Gootenberg, O.O. Abudayyeh, B. Franklin, M.J. Kellner, J. Joung, F. Zhang, RNA editing with CRISPR-Cas13, Science (2017).
[42] W. Cress, S. Triezenberg, Critical structural elements of the VP16 transcriptional activation domain, Science 251 (4989) (1991) 87–90.
[43] I. Sadowski, J. Ma, S. Triezenberg, M. Ptashne, GAL4-VP16 is an unusually potent transcriptional activator, Nature 335 (6190) (1988) 563–564.
[44] K. Seipel, O. Georgiev, W. Schaffner, Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions, EMBO J. 11 (13) (1992) 4961–4968.
[45] R.R. Beerli, D.J. Segal, B. Dreier, C.F. Barbas 3rd, Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks, Proceed. Natl. Acad. Sci. U.S.A. 95 (25) (1998) 14628–14633.
[46] M.L. Schmitz, P.A. Baeuerle, The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B, EMBO J. 10 (12) (1991) 3805–3817.
[47] L. Gradoville, J. Gerlach, E. Grogan, D. Shedd, S. Nikiforow, C. Metroka, G. Miller, Kaposi's sarcoma-associated herpesvirus open reading frame 50/Rta protein activates the entire viral lytic cycle in the HH-B2 primary effusion lymphoma cell line, J. Virol. 74 (13) (2000) 6207–6212.
[48] R. Sun, S.-F. Lin, L. Gradoville, Y. Yuan, F. Zhu, G. Miller, A viral gene that activates lytic cycle expression of Kaposi's sarcoma-associated herpesvirus, Proceed. Natl. Acad. Sci. 95 (18) (1998) 10866–10871.
[49] D.W. Neef, A. Jaeger, R. Gomez-Pastor, F. Willmund, J. Frydman, D.J. Thiele, A direct regulatory interaction between chaperonin TRiC and stress responsive transcription factor HSF1, Cell Reports 9 (3) (2014) 955–966.
[50] A. Chavez, J. Scheiman, S. Vora, B.W. Pruitt, M. Tuttle, E.P.R. Iyer, S. Lin, S. Kiani, C.D. Guzman, D.J. Wiegand, D. Ter-Ovanesyan, J.L. Braff, N. Davidsohn, B.E. Housden, N. Perrimon, R. Weiss, J. Aach, J.J. Collins, G.M. Church, Highly efficient Cas9-mediated transcriptional programming, Nat. Methods 12 (4) (2015) 326–328.
[51] L.A. Gilbert, M.H. Larson, L. Morsut, Z. Liu, G.A. Brar, S.E. Torres, N. Stern-Ginossar, O. Brandman, E.H. Whitehead, J.A. Doudna, W.A. Lim, J.S. Weissman, L.S. Qi, CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes, Cell 154 (2) (2013) 442–451.
[52] M.L. Maeder, S.J. Linder, V.M. Cascio, Y. Fu, Q.H. Ho, J.K. Joung, CRISPR RNA-guided activation of endogenous human genes, Nat. Methods 10 (10) (2013) 977–979.
[53] P. Perez-Pinera, D.D. Kocak, C.M. Vockley, A.F. Adler, A.M. Kabadi, L.R. Polstein, P.I. Thakore, K.A. Glass, D.G. Ousterout, K.W. Leong, F. Guilak, G.E. Crawford, T.E. Reddy, C.A. Gersbach, RNA-guided gene activation by CRISPR-Cas9-based transcription factors, Nat. methods 10 (10) (2013) 973–976.
[54] S. Kiani, A. Chavez, M. Tuttle, R.N. Hall, R. Chari, D. Ter-Ovanesyan, J. Qian, B.W. Pruitt, J. Beal, S. Vora, J. Buchthal, E.J.K. Kowal, M.R. Ebrahimkhani, J.J. Collins, R. Weiss, G. Church, Cas9 gRNA engineering for genome editing, activation and repression, Nat. Methods 12 (11) (2015) 1051–1054.
[55] X. Gao, J.C.H. Tsang, F. Gaba, D. Wu, L. Lu, P. Liu, Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers, Nucl. Acids Res. 42 (20) (2014) e155.
[56] A.W. Cheng, H. Wang, H. Yang, L. Shi, Y. Katz, T.W. Theunissen, S. Rangarajan, C.S. Shivalila, D.B. Dadon, R. Jaenisch, Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system, Cell Res. 23 (10) (2013) 1163–1171.
[57] F. Farzadfard, S.D. Perli, T.K. Lu, Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas, ASC Synth. Biol 2 (2013) 604–613.
[58] S.-M. Ho, B.J. Hartley, E. Flaherty, P. Rajarajan, R. Abdelaal, I. Oboriah, N. Barretto, H. Muhammad, H.P. Phatnani, S. Akbarian, K.J. Brennand, Evaluating synthetic activation and repression of neuropsychiatric-related genes in hiPSC-

derived NPCs, Neurons, and Astrocytes, Stem Cell Reports 9 (2) (2017) 615–628.
[59] J. Hu, Y. Lei, W.-K. Wong, S. Liu, K.-C. Lee, X. He, W. You, R. Zhou, J.-T. Guo, X. Chen, X. Peng, H. Sun, H. Huang, H. Zhao, B. Feng, Direct activation of human and mouse Oct4 genes using engineered TALE and Cas9 transcription factors, Nucl. Acids Res. 42 (7) (2014) 4375–4390.
[60] N.A. Kearns, R.M.J. Genga, M.S. Enuameh, M. Garber, S.A. Wolfe, R. Maehr, Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells, Development (Cambridge, England) 141 (1) (2014) 219–223.
[61] P. Mali, J. Aach, B.P. Stranges, K.M. Esvelt, M. Moosburner, S. Kosuri, L. Yang, G.M. Church, CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering, Nat. Biotechnol. 31 (9) (2013) 833–838.
[62] L.S. Qi, M.H. Larson, L.A. Gilbert, J.A. Doudna, J.S. Weissman, A.P. Arkin, W.A. Lim, Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression, Cell 152 (5) (2013) 1173–1183.
[63] S. Konermann, M.D. Brigham, A.E. Trevino, J. Joung, O.O. Abudayyeh, C. Barcena, P.D. Hsu, N. Habib, J.S. Gootenberg, H. Nishimasu, O. Nureki, F. Zhang, Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex, Nature 517 (7536) (2014) 583–588.
[64] M.E. Tanenbaum, L.A. Gilbert, L.S. Qi, J.S. Weissman, R.D. Vale, A protein-tagging system for signal amplification in gene expression and fluorescence imaging, Cell 159 (3) (2014) 635–646.
[65] H. Nishimasu, F.A. Ran, P.D. Hsu, S. Konermann, S.I. Shehata, N. Dohmae, R. Ishitani, F. Zhang, O. Nureki, Crystal structure of Cas9 in complex with guide RNA and target DNA, Cell 156 (5) (2014) 935–949.
[66] Y. Gao, X. Xiong, S. Wong, E.J. Charles, W.A. Lim, L.S. Qi, Complex transcriptional modulation with orthogonal and inducible dCas9 regulators, Nat. Methods 13 (2016) 1043–1049.
[67] Y. Zhang, C. Yin, T. Zhang, F. Li, W. Yang, R. Kaminski, P.R. Fagan, R. Putatunda, W.B. Young, K. Khalili, W. Hu, CRISPR/gRNA-directed synergistic activation mediator (SAM) induces specific, persistent and robust reactivation of the HIV-1 latent reservoirs, Sci. Rep. 5 (2015) 16277.
[68] C. Pflueger, D. Tan, T. Swain, T. Nguyen, J. Pflueger, C. Nefzger, J.M. Polo, E. Ford, R. Lister, A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs, Genome Res. 28 (8) (2018) 1193–1206.
[69] S. Konermann, M.D. Brigham, A. Trevino, P.D. Hsu, M. Heidenreich, L. Cong, R.J. Platt, D.A. Scott, G.M. Church, F. Zhang, Optical control of mammalian endogenous transcription and epigenetic states, Nature 500 (7463) (2013) 472–476.
[70] L.R. Polstein, C.A. Gersbach, Light-inducible spatiotemporal control of gene activation by customizable zinc finger transcription factors, J. Am. Chem. Soc. 134 (40) (2012) 16480–16483.
[71] Y. Nihongaki, S. Yamamoto, F. Kawano, H. Suzuki, M. Sato, CRISPR-Cas9-based photoactivatable transcription system, Chem. Biol. 22 (2) (2015) 169–174.
[72] G.P. Pathak, J.I. Spiltoir, C. Höglund, L.R. Polstein, S. Heine-Koskinen, C.A. Gersbach, J. Rossi, C.L. Tucker, Bidirectional approaches for optogenetic regulation of gene expression in mammalian cells using Arabidopsis cryptochrome 2, Nucl. Acids Res. 45 (20) (2017) e167.
[73] L.R. Polstein, C. Gersbach, A light-inducible CRISPR/Cas9 system for control of endogenous gene activation, Nat. Chem. Biol. 11 (3) (2015) 198–200.
[74] M.J. Kennedy, R.M. Hughes, L.A. Peteya, J.W. Schwartz, M.D. Ehlers, C.L. Tucker, Rapid blue-light–mediated induction of protein interactions in living cells, Nat. Methods 7 (2010) 973–975.
[75] J.G. Zalatan, M.E. Lee, R. Almeida, L.A. Gilbert, E.H. Whitehead, M. La Russa, J.C. Tsai, J.S. Weissman, J.E. Dueber, L.S. Qi, W.A. Lim, Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds, Cell 160 (1–2) (2015) 339–350.
[76] L. Stojic, A.T.L. Lun, J. Mangei, P. Mascalchi, V. Quarantotti, A.R. Barr, C. Bakal, J.C. Marioni, F. Gergely, D.T. Odom, Specificity of RNAi, LNA and CRISPRi as loss-of-function methods in transcriptional analysis, Nucl. Acids Res. 46 (12) (2018) 5950–5966.
[77] E.J. Bellefroid, D.A. Poncelet, P.J. Lecocq, O. Revelant, J.A. Martial, The evolutionarily conserved Krüppel-associated box domain defines a subfamily of eukaryotic multifingered proteins, Proceed. Natl. Acad. Sci. 88 (9) (1991) 3608–3612.
[78] J.F. Margolin, J.R. Friedman, W.K. Meyer, H. Vissing, H.J. Thiesen, F.J. Rauscher 3rd, Krüppel-associated boxes are potent transcriptional repression domains, Proceed. Natl. Acad. Sci. U.S.A. 91 (10) (1994) 4509–4513.
[79] R. Witzgall, E. O'Leary, A. Leaf, D. Onaldi, V.J. Bonventre, The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression, Proceed. Natl. Acad. Sci. 91 (10) (1994) 4514–4518.
[80] L. Fritsch, P. Robin, J.R. Mathieu, M. Souidi, H. Hinaux, C. Rougeulle, A. Harel-Bellan, M. Ameyar-Zazoua, S. Ait-Si-Ali, A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex, Mol. Cell 37 (1) (2010) 46–56.
[81] M. Lachner, D. O'Carroll, S. Rea, K. Mechtler, T. Jenuwein, Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins, Nature 410 (6824) (2001) 116–120.
[82] M.S. Lechner, G.E. Begg, D.W. Speicher, F.J. Rauscher, Molecular Determinants for targeting heterochromatin protein 1-mediated gene silencing: direct chromoshadow domain–KAP-1 corepressor interaction is essential, Mol. Cell. Biol. 20 (17) (2000) 6449–6465.
[83] J.-S. Zhang, M.C. Moncrieffe, J. Kaczynski, V. Ellenrieder, F.G. Prendergast, R. Urrutia, A conserved α-helical motif mediates the interaction of Sp1-like transcriptional repressors with the corepressor mSin3A, Mol. Cell. Biol. 21 (15) (2001) 5041–5049.
[84] A. Amabile, A. Migliara, P. Capasso, M. Biffi, D. Cittaro, L. Naldini, A. Lombardo, Inheritable silencing of endogenous genes by hit-and-run targeted epigenetic editing, Cell 167 (1) (2016) 219–232.
[85] M.A. Mandegar, N. Huebsch, E.B. Frolov, E. Shin, A. Truong, M.P. Olvera, A.H. Chan, Y. Miyaoka, K. Holmes, C.I. Spencer, L.M. Judge, D.E. Gordon, V.T. Eskildsen, J.E. Villalta, M.A. Horlbeck, L.A. Gilbert, N.J. Krogan, S.P. Sheikh, J.S. Weissman, L.S. Qi, P.-L. So, B.R. Conklin, CRISPR interference efficiently induces specific and reversible gene silencing in human iPSCs, Cell Stem Cell 18 (4) (2016) 541–553.
[86] P.I. Thakore, A.M. D'Ippolito, L. Song, A. Safi, N.K. Shivakumar, A.M. Kabadi, T.E. Reddy, G.E. Crawford, C.A. Gersbach, Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements, Nat. Methods 12 (2015) 1143.
[87] N.C. Yeo, A. Chavez, A. Lance-Byrne, Y. Chan, D. Menn, D. Milanova, C.-C. Kuo, X. Guo, S. Sharma, A. Tung, R.J. Cecchi, M. Tuttle, S. Pradhan, E.T. Lim, N. Davidsohn, M.R. Ebrahimkhani, J.J. Collins, N.E. Lewis, S. Kiani, G.M. Church, An enhanced CRISPR repressor for targeted mammalian gene regulation, Nat. Methods 15 (8) (2018) 611–616.
[88] Y. Fu, J.A. Foden, C. Khayter, M.L. Maeder, D. Reyon, J.K. Joung, J.D. Sander, High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells, Nat. Biotechnol. (2013).
[89] P.D. Hsu, D.A. Scott, J.A. Weinstein, F.A. Ran, S. Konermann, V. Agarwala, Y. Li, E.J. Fine, X. Wu, O. Shalem, T.J. Cradick, L.A. Marraffini, G. Bao, F. Zhang, DNA targeting specificity of RNA-guided Cas9 nucleases, Nat. Biotechnol. (2013).
[90] M.F. Bolukbasi, A. Gupta, S.A. Wolfe, Creating and evaluating accurate CRISPR-Cas9 scalpels for genomic surgery, Nat. Methods (2015).
[91] W. Jiang, D. Bikard, D. Cox, F. Zhang, L.A. Marraffini, RNA-guided editing of bacterial genomes using CRISPR-Cas systems, Nat. Biotechnol. (2013).
[92] X. Wu, D.A. Scott, A.J. Kriz, A.C. Chiu, P.D. Hsu, D.B. Dadon, A.W. Cheng, A.E. Trevino, S. Konermann, S. Chen, R. Jaenisch, F. Zhang, P.A. Sharp, Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells, Nat. Biotechnol. (2014).
[93] V. Pattanayak, S. Lin, J.P. Guilinger, E. Ma, J.A. Doudna, D.R. Liu, High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity, Nat. Biotechnol. (2013).
[94] S.Q. Tsai, Z. Zheng, N.T. Nguyen, M. Liebers, V.V. Topkar, V. Thapar, N. Wyvekens, C. Khayter, A.J. Iafrate, L.P. Le, M.J. Aryee, J.K. Joung, GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases, Nat. Biotechnol. (2015).
[95] T. Zheng, Y. Hou, P. Zhang, Z. Zhang, Y. Xu, L. Zhang, L. Niu, Y. Yang, D. Liang, F. Yi, W. Peng, W. Feng, Y. Yang, J. Chen, Y.Y. Zhu, L.H. Zhang, Q. Du, Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence, Scient. Rep. 7 (2017) 1–8.
[96] F.A. Ran, P.D. Hsu, C.Y. Lin, J.S. Gootenberg, S. Konermann, A.E. Trevino, D.A. Scott, A. Inoue, S. Matoba, Y. Zhang, F. Zhang, Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity, Cell (2013).
[97] Y. Zhang, N. Heidrich, B.J. Ampattu, C.W. Gunderson, H.S. Seifert, C. Schoen, J. Vogel, E.J. Sontheimer, Processing-independent CRISPR RNAs limit natural transformation in Neisseria meningitidis, Mol. Cell (2013).
[98] M. Labuhn, F.F. Adams, M. Ng, S. Knoess, A. Schambach, E.M. Charpentier, A. Schwarzer, J.L. Mateo, J.H. Klusmann, D. Heckl, Refined sgRNA efficacy prediction improves large- and small-scale CRISPR-Cas9 applications, Nucl. Acids Res. 46 (3) (2018) 1375–1385.
[99] G. Yuen, F.J. Khan, S. Gao, J.M. Stommel, E. Batchelor, X. Wu, J. Luo, CRISPR/Cas9-mediated gene knockout is insensitive to target copy number but is dependent on guide RNA potency and Cas9/sgRNA threshold expression level, Nucl. Acids Res. 45 (20) (2017) 12039–12053.
[100] G. Chuai, H. Ma, J. Yan, M. Chen, N. Hong, D. Xue, C. Zhou, C. Zhu, K. Chen, B. Duan, F. Gu, S. Qu, D. Huang, J. Wei, Q. Liu, DeepCRISPR: optimized CRISPR guide RNA design by deep learning, Gen. Biol. 19 (1) (2018) 80.
[101] H. Peng, Y. Zheng, Z. Zhao, T. Liu, J. Li, Recognition of CRISPR/Cas9 off-target sites through ensemble learning of uneven mismatch distributions, Bioinformatics 34 (17) (2018) i757–i765.
[102] L. Xue, B. Tang, W. Chen, J. Luo, Prediction of CRISPR sgRNA activity using a deep convolutional neural network, J. Chem. Inf. Model 59 (1) (2019) 615–624.
[103] Y. Lin, T.J. Cradick, M.T. Brown, H. Deshmukh, P. Ranjan, N. Sarode, B.M. Wile, P.M. Vertino, F.J. Stewart, G. Bao, CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences, Nucl. Acids Res. 42 (11) (2014) 7473–7485.
[104] C. Moses, F. Nugent, C.B. Waryah, B. Garcia-Bloj, A.R. Harvey, P. Blancafort, Activating PTEN tumor suppressor expression with the CRISPR/dCas9 system, Mol. Ther. Nucl. Acids 14 (2019) 287–300.
[105] S. Kim, D. Kim, S.W. Cho, J. Kim, J.S. Kim, Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins, Gen. Res. (2014).
[106] S. Ramakrishna, A.B. Kwaku Dad, J. Beloor, R. Gopalappa, S.K. Lee, H. Kim, Gene, disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA, Gen. Res. (2014).
[107] V. Ranganathan, K. Wahlin, J. Maruotti, D.J. Zack, Expansion of the CRISPR-Cas9 genome targeting space through the use of H1 promoter-expressed guide RNAs, Nat. Commun. (2014).
[108] Z. Gao, E. Herrera-Carrillo, B. Berkhout, A single H1 promoter can drive both guide RNA and endonuclease expression in the CRISPR-Cas9 system, Mol. Ther. – Nucl. Acids 14 (2019) 32–40.
[109] Z. Gao, E. Herrera-Carrillo, B. Berkhout, RNA polymerase II activity of type 3 Pol III promoters, Mol. Ther. – Nucl. Acids (2018).
[110] J. Cao, L. Wu, S.M. Zhang, M. Lu, W.K.C. Cheung, W. Cai, M. Gale, Q. Xu, Q. Yan, An easy and efficient inducible CRISPR/Cas9 platform with improved specificity for multiple gene targeting, Nucl. Acids Res. 44 (19) (2016) 1–10.
[111] K.M. Davis, V. Pattanayak, D.B. Thompson, J.A. Zuris, D.R. Liu, Small molecule-triggered Cas9 protein with improved genome-editing specificity, Nat. Chem. Biol. (2015).
[112] B. Zetsche, S.E. Volz, F. Zhang, A split-Cas9 architecture for inducible genome editing and transcription modulation, Nat. Biotechnol. (2015).

[113] Y. Nihongaki, F. Kawano, T. Nakajima, M. Sato, Photoactivatable CRISPR-Cas9 for optogenetic genome editing, Nat. Biotechnol. (2015).
[114] Y. Chen, X. Liu, Y. Zhang, H. Wang, H. Ying, M. Liu, D. Li, K.O. Lui, Q. Ding, A Self-restricted CRISPR system to reduce off-target effects, Mol. Ther. (2016).
[115] L.A. Gilbert, M.A. Horlbeck, B. Adamson, J.E. Villalta, Y. Chen, E.H. Whitehead, C. Guimaraes, B. Panning, H.L. Ploegh, M.C. Bassik, L.S. Qi, M. Kampmann, J.S. Weissman, Genome-scale CRISPR-mediated control of gene repression and activation, Cell (2014).
[116] S.W. Cho, S. Kim, Y. Kim, J. Kweon, H.S. Kim, S. Bae, J.-s. Kim, Analysis of off-target effect of CRISPR/Cas-derived RNA-guided endonucleases and nickases, 1992 0-1.
[117] J. Tycko, V.E. Myer, P.D. Hsu, Methods for optimizing CRISPR-Cas9 genome editing specificity, Mol. Cell 63 (3) (2016) 355–370.
[118] Y. Fu, J.D. Sander, D. Reyon, V.M. Cascio, J.K. Joung, Improving CRISPR-Cas nuclease specificity using truncated guide RNAs, Nat. Biotechnol. 32 (3) (2014) 279–284.
[119] J.E. Dahlman, O.O. Abudayyeh, J. Joung, J.S. Gootenberg, F. Zhang, S. Konermann, Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease, Nat. Biotechnol. (2015).
[120] A. Hendel, R.O. Bak, J.T. Clark, A.B. Kennedy, D.E. Ryan, S. Roy, I. Steinfeld, B.D. Lunstad, R.J. Kaiser, A.B. Wilkens, R. Bacchetta, A. Tsalenko, D. Dellinger, L. Bruhn, M.H. Porteus, Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells, Nat. Biotechnol. (2015).
[121] M. Rahdar, M.A. McMahon, T.P. Prakash, E.E. Swayze, C.F. Bennett, D.W. Cleveland, Synthetic CRISPR RNA-Cas9–guided genome editing in human cells, Proceed Natl. Acad. Sci. (2015).
[122] H. Yin, C.Q. Song, S. Suresh, S.Y. Kwan, Q. Wu, S. Walsh, J. Ding, R.L. Bogorad, L.J. Zhu, S.A. Wolfe, V. Koteliansky, W. Xue, R. Langer, D.G. Anderson, Partial DNA-guided Cas9 enables genome editing with reduced off-target activity, Nat. Chem. Biol. 14 (3) (2018) 311–316.
[123] B.P. Kleinstiver, V. Pattanayak, M.S. Prew, S.Q. Tsai, N.T. Nguyen, Z. Zheng, J.K. Joung, High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects, Nature 529 (7587) (2016) 490–495.
[124] I.M. Slaymaker, L. Gao, B. Zetsche, D.A. Scott, W.X. Yan, F. Zhang, Rationally engineered Cas9 nucleases with improved specificity, Science 351 (6268) (2016) 84–88.
[125] J.S. Chen, Y.S. Dagdas, B.P. Kleinstiver, M.M. Welch, A.A. Sousa, L.B. Harrington, S.H. Sternberg, J.K. Joung, A. Yildiz, J.A. Doudna, Enhanced proofreading governs CRISPR-Cas9 targeting accuracy, Nature 550 (7676) (2017) 407–410.
[126] A. Casini, M. Olivieri, G. Petris, C. Montagna, G. Reginato, G. Maule, F. Lorenzin, D. Prandi, A. Romanel, F. Demichelis, A. Inga, A. Cereseto, A highly specific SpCas9 variant is identified by in vivo screening in yeast, Nat. Biotechnol. 36 (3) (2018) 265–271.
[127] J.H. Hu, S.M. Miller, M.H. Geurts, W. Tang, L. Chen, N. Sun, C.M. Zeina, X. Gao, H.A. Rees, Z. Lin, D.R. Liu, Evolved Cas9 variants with broad PAM compatibility and high DNA specificity, Nature 556 (7699) (2018) 57–63.
[128] S. Kim, T. Bae, J. Hwang, J.S. Kim, Rescue of high-specificity Cas9 variants using sgRNAs with matched 5' nucleotides, Gen. Biol. (2017).
[129] J.K. Lee, E. Jeong, J. Lee, M. Jung, E. Shin, Y.H. Kim, K. Lee, I. Jung, D. Kim, S. Kim, J.S. Kim, Directed evolution of CRISPR-Cas9 to increase its specificity, Nat. Commun. 9 (1) (2018).
[130] D. Zhang, H. Zhang, T. Li, K. Chen, J.L. Qiu, C. Gao, Perfectly matched 20-nucleotide guide RNA sequences enable robust genome editing using high-fidelity SpCas9 nucleases, Gen. Biol. (2017).
[131] D.P. Dever, R.O. Bak, A. Reinisch, J. Camarena, G. Washington, C.E. Nicolas, M. Pavel-Dinu, N. Saxena, A.B. Wilkens, S. Mantri, N. Uchida, A. Hendel, A. Narla, R. Majeti, K.I. Weinberg, M.H. Porteus, CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells, Nature 539 (7629) (2016) 384–389.
[132] M.A. DeWitt, W. Magis, N.L. Bray, T. Wang, J.R. Berman, F. Urbinati, S.J. Heo, T. Mitros, D.P. Munoz, D. Boffelli, D.B. Kohn, M.C. Walters, D. Carroll, D.I.K. Martin, J.E. Corn, Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells, Sci. Transl. Med. 8 (360) (2016) 360ra134-360ra134.
[133] X. Liang, J. Potter, S. Kumar, Y. Zou, R. Quintanilla, M. Sridharan, J. Carte, W. Chen, N. Roark, S. Ranganathan, N. Ravinder, J.D. Chesnut, Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection, J. Biotechnol. (2015).
[134] C.A. Vakulskas, D.P. Dever, G.R. Rettig, R. Turk, A.M. Jacobi, M.A. Collingwood, N.M. Bode, M.S. McNeill, S. Yan, J. Camarena, C.M. Lee, S.H. Park, V. Wiebking, R.O. Bak, N. Gomez-Ospina, M. Pavel-Dinu, W. Sun, G. Bao, M.H. Porteus, M.A. Behlke, A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells, Nat. Med. 24 (8) (2018) 1216–1224.
[135] H. Nishimasu, L. Cong, W.X. Yan, F.A. Ran, B. Zetsche, Y. Li, A. Kurabayashi, R. Ishitani, F. Zhang, O. Nureki, Crystal structure of Staphylococcus aureus Cas9, Cell (2015).
[136] F.A. Ran, L. Cong, W.X. Yan, D.A. Scott, J.S. Gootenberg, A.J. Kriz, B. Zetsche, O. Shalem, X. Wu, K.S. Makarova, V.E. Koonin, P.A. Sharp, F. Zhang, In vivo genome editing using Staphylococcus aureus Cas9, Nature (2015).
[137] L. Ye, J. Wang, Y. Tan, A.I. Beyer, F. Xie, M.O. Muench, Y.W. Kan, Genome editing using CRISPR-Cas9 to create the HPFH genotype in HSPCs: an approach for treating sickle cell disease and β-thalassemia, Proceed. Natl. Acad. Sci. (2016).
[138] K.M. Esvelt, P. Mali, J.L. Braff, M. Moosburner, S.J. Yaung, G.M. Church, Orthogonal Cas9 proteins for RNA-guided gene regulation and editing, Nat. Methods (2013).
[139] M. Müller, C.M. Lee, G. Gasiunas, T.H. Davis, T.J. Cradick, V. Siksnys, G. Bao, T. Cathomen, C. Mussolino, Streptococcus thermophilus CRISPR-Cas9 systems enable specific editing of the human genome, Mol. Ther. 24 (3) (2016) 636–644.
[140] M. Glemzaite, E. Balciunaite, T. Karvelis, G. Gasiunas, M.M. Grusyte, G. Alzbutas, A. Jurcyte, E.M. Anderson, E. Maksimova, A.J. Smith, A. Lubys, L. Zaliauskienelolita, V. Siksnys, Targeted gene editing by transfection of in vitro reconstituted streptococcus thermophilus cas9 nuclease complex, RNA Biol. (2015).
[141] K. Xu, C. Ren, Z. Liu, T. Zhang, T. Zhang, D. Li, L. Wang, Q. Yan, L. Guo, J. Shen, Z. Zhang, Efficient genome engineering in eukaryotes using Cas9 from Streptococcus thermophilus, Cell. Mol. Life Sci. (2015).
[142] Z. Hou, Y. Zhang, N.E. Propson, S.E. Howden, L.F. Chu, E.J. Sontheimer, J.A. Thomson, Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis, Proceed. Natl. Acad. Sci. (2013).
[143] C.M. Lee, T.J. Cradick, G. Bao, The neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells, Mol. Ther. (2016).
[144] E. Kim, T. Koo, S.W. Park, D. Kim, K. Kim, H.Y. Cho, D.W. Song, K.J. Lee, M.H. Jung, S. Kim, J.H. Kim, J.H. Kim, J.S. Kim, In, vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni, Nat. Commun. (2017).
[145] D. Kim, J. Kim, J.K. Hur, K.W. Been, S.H. Yoon, J.S. Kim, Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells, Nat. Biotechnol. (2016).
[146] H.K. Kim, M. Song, J. Lee, A.V. Menon, S. Jung, Y.M. Kang, J.W. Choi, E. Woo, H.C. Koh, J.W. Nam, H. Kim, In vivo high-throughput profiling of CRISPR-Cpf1 activity, Nat. Methods (2017).
[147] B. Zetsche, J.S. Gootenberg, O.O. Abudayyeh, I.M. Slaymaker, K.S. Makarova, P. Essletzbichler, S.E. Volz, J. Joung, J. Van Der Oost, A. Regev, V.E. Koonin, F. Zhang, Cpf1 Is a Single RNA-guided endonuclease of a class 2 CRISPR-Cas system, Cell (2015).
[148] J. Strecker, S. Jones, B. Koopal, J. Schmid-Burgk, B. Zetsche, L. Gao, K.S. Makarova, E.V. Koonin, F. Zhang, Engineering of CRISPR-Cas12b for human genome editing, Nat. Commun. 10 (1) (2019) 212.
[149] A.E. Friedland, R. Baral, P. Singhal, K. Loveluck, S. Shen, M. Sanchez, E. Marco, G.M. Gotta, M.L. Maeder, E.M. Kennedy, A.V.R. Kornepati, A. Sousa, M.A. Collins, H. Jayaram, B.R. Cullen, D. Bumcrot, Characterization of Staphylococcus aureus Cas9: A smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications, Gen. Biol. (2015).
[150] Z. Wu, H. Yang, P. Colosi, Effect of genome size on AAV vector packaging, Mol. Ther. (2010).
[151] B.P. Kleinstiver, M.S. Prew, S.Q. Tsai, N.T. Nguyen, V.V. Topkar, Z. Zheng, J.K. Joung, Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition, Nat. Biotechnol. (2015).
[152] P.I. Thakore, J.B. Kwon, C.E. Nelson, D.C. Rouse, M.P. Gemberling, M.L. Oliver, C.A. Gersbach, RNA-guided transcriptional silencing in vivo with S. aureus CRISPR-Cas9 repressors, Nat. Commun. (2018).
[153] J.E. Garneau, M.È. Dupuis, M. Villion, D.A. Romero, R. Barrangou, P. Boyaval, C. Fremaux, P. Horvath, A.H. Magadán, S. Moineau, The CRISPR/cas bacterial immune system cleaves bacteriophage and plasmid DNA, Nature 468 (7320) (2010) 67–71.
[154] A.H. Magadán, M.È. Dupuis, M. Villion, S. Moineau, Cleavage of phage DNA by the Streptococcus thermophilus CRISPR3-Cas system, PLoS ONE (2012).
[155] R. Ibraheim, C.Q. Song, A. Mir, N. Amrani, W. Xue, E.J. Sontheimer, All-in-one adeno-associated virus delivery and genome editing by Neisseria meningitidis Cas9 in vivo, Gen. Biol. (2018).
[156] A. Pawluk, N. Amrani, Y. Zhang, B. Garcia, Y. Hidalgo-Reyes, J. Lee, A. Edraki, M. Shah, E.J. Sontheimer, K.L. Maxwell, A.R. Davidson, Naturally occurring off-switches for CRISPR-Cas9, Cell (2016).
[157] H. Hirano, J.S. Gootenberg, T. Horii, O.O. Abudayyeh, M. Kimura, P.D. Hsu, T. Nakane, R. Ishitani, I. Hatada, F. Zhang, H. Nishimasu, O. Nureki, Structure and Engineering of Francisella novicida Cas9, Cell (2016).
[158] F. Chen, X. Ding, Y. Feng, T. Seebeck, Y. Jiang, G.D. Davis, Targeted activation of diverse CRISPR-Cas systems for mammalian genome editing via proximal CRISPR targeting, Nat. Commun. (2017).
[159] T.R. Sampson, S.D. Saroj, A.C. Llewellyn, Y.L. Tzeng, D.S. Weiss, A CRISPR/Cas system mediates bacterial innate immune evasion and virulence, Nature (2013).
[160] A.A. Price, T.R. Sampson, H.K. Ratner, A. Grakoui, D.S. Weiss, Cas9-mediated targeting of viral RNA in eukaryotic cells, Proceed. Natl. Acad. Sci. (2015).
[161] T. Zhang, Q. Zheng, X. Yi, H. An, Y. Zhao, S. Ma, G. Zhou, Establishing RNA virus resistance in plants by harnessing CRISPR immune system, Plant Biotechnol. J. (2018).
[162] E. Baxter, K. Windloch, F. Gannon, J.S. Lee, Epigenetic regulation in cancer progression, Cell Biosci. 4 (2014) 45.
[163] M. Berdasco, M. Esteller, Aberrant epigenetic landscape in cancer: how cellular identity goes awry, Dev. Cell 19 (5) (2010) 698–711.
[164] C.A. Scacheri, P.C. Scacheri, Mutations in the noncoding genome, Curr. Opin. Pediatr. 27 (6) (2015) 659–664.
[165] D. Cyranoski, CRISPR-baby scientist fails to satisfy critics, Nature 564 (7734) (2018) 13–14.
[166] D. Dickenson, M. Darnovsky, Did a permissive scientific culture encourage the 'CRISPR babies' experiment? Nat. Biotechnol. 37 (4) (2019) 355–357.

3 RASPRAVA

Razumijevanje epigenetičkog statusa promotora gena te funkcionalne organizacije različitih dijelova genoma temeljeno je na podacima velikih konzorcija kao što su *NIH Roadmap Epigenomics Mapping Consortium*[203], *Blueprint Epigenome*[204], ENCODE[205] te *International Human Epigenome Consortium* (IHEC)[206]. Funkcionalni značaj epigenetičkih oznaka te njihovo međudjelovanje s proteinima uključenima u kompleksnu mrežu regulacije ekspresije gena i dalje nije potpuno razjašnjen. Stoga su molekularni alati pomoću kojih se ciljano može promijeniti određena epigenetička oznaka od velikog značaja u razjašnjavanju mehanizama i uloge epigenetičkih oznaka u regulaciji ekspresije gena. Naime, dosadašnje studije temeljene su na korelacijama između prisustva pojedinih epigenetičkih oznaka i razine ekspresije gena što ne daje jasne odgovore o funkcionalnom značaju istih. Najveći pomak u razumijevanju funkcionalne važnosti epigenetičkih oznaka krenuo je s razvojem tehnologije CRISPR/dCas9 te vezanjem tzv. epigenetičkih „brisača“ ili „pisača“ na inaktiviranu nukleazu dCas9. Primjenom takvih fuzijskih konstrukata moguće je ciljano uvoditi ili uklanjati određenu epigenetičku oznaku u molekuli DNA[134–137] ili histonskim proteinima[138–140] na točno određenom mjestu u genomu, te odrediti njezinu važnost u regulaciji ekspresije gena.

Tijekom izrade doktorske disertacije razvio sam modularni sustav temeljen na tehnologiji CRISPR/dCas9 u svrhu dodatnog pojednostavljenja sklapanja različitih fuzijskih konstrukata ovisno o potrebama specifičnog eksperimenta[196]. Modularni sustav može se koristiti u svrhu ciljane metilacije i demetilacije molekule DNA, kao i u svrhu direktne regulacije transkripcijske aktivnosti ciljanih gena te je lagano nadogradiv s novim modulima u svrhu različitih primjena. Također, u ovakvo smišljenom modularnom sustavu omogućena je primjena dva različita ortologna proteina Cas9 iz vrsta *Streptococcus pyogenes* (SpCas9) i *Staphylococcus aureus* (SaCas9), čime je proširena mogućnost ciljanja regija u genomu s različitim epigenetičkim efektorskim domenama, direktnim aktivatorima ili represorima genske aktivnosti sinergističkog ili antagonističkog djelovanja. Osnovni način sklapanja modularnog sustava temeljio sam na reakciji „Golden-Gate“ kloniranja više različitih modula pomoću restrikcijskog enzima BsaI, koji stvara nepalindromske ljepljive krajeve dužine 4 nt, čime je osigurano pravilno sklapanje konstrukata. Prvi modul u sustavu je sgRNA ekspresijski modul za ortologni protein SpCas9 ili SaCas9. Osnovne komponente modula su promotor RNA polimeraze III (U6 promotor), okosnica sgRNA i U6 transkripcijski terminator. Okosnica sgRNA sadrži selekcijski biljeg mRuby3, potreban za crveno-bijelu selekciju pozitivnih

bakterijskih kolonija, okružen s dva restrikcijska mjesta BpiI koja se koriste za kloniranje varijabilnog djela duljine 20 nt na 5' kraj molekule sgRNA. Drugi modul je promotor RNA polimeraze II, slabi promotor EFS ili jaki promotor CBh. Potom slijedi željena efektorska domena (DNMT3A, TET1, VPR ili KRAB) spojena preko fleksibilnog peptidnog linkera sastava Gly_4Ser na inaktivirani ortologni protein dCas9 (dSpCas9 ili dSaCas9). Funkcija Gly_4Ser linkera je optimalno razdvajanje i pravilno sklapanje funkcionalne efektorske domene i inaktiviranog ortologa dCas9[207]. Nakon ortologa dCas9 slijedi modul za selekciju stanica vezan preko samo-cijepajućeg peptida T2A (engl. *Thoseaasigna virus 2A*), koji osigurava ekvimolarnu translaciju dva različita polipeptida na jednom policistronskom transkiptu[208]. Selekcijski biljeg uključuje rezistenciju na antibiotik Puromicin (PuroR) i/ili fluorescencijski biljeg mClover3 ili mRuby3. Na kraju slijedi eukariotski transkripcijski terminator Bgh (engl. *Bovine growth hormone terminator*)[196]. Pokazao sam kako je sklapanje šest ili sedam različitih modula uspješno u jednoj reakciji „Golden-Gate" kloniranja pomoću enzima BsaI te dodatno olakšao probir pozitivnih bakterijskih kolonija pomoću plavo-bijele selekcije. U minimalni destinacijski plazmid, koji se koristi u tu svrhu, ubacio sam α podjedinicu β-galaktozidaze okruženu s dva BsaI restrikcijska mjesta te SV40 jezgrin ciljajući slijed (engl. *SV40 nuclear DNA-targeting sequence*, SV40-DTS). Prijašnjim studijama je pokazano kako dodavanje slijeda SV40-DTS, kojeg prepoznaju i vežu različiti transkripcijski faktori, pomaže u aktivnom transportu plazmida u jezgru, što olakšava transfekciju stanica koje se ne dijele[209]. Dodavanje slijeda SV40-DTS unutar modularnog sustava, koji sam dizajnirao u ovom doktorskom radu, omogućilo je brži transport plazmida u jezgru stanica neovisan o staničnoj diobi, što je u konačnici povećalo ekspresiju s plazmidne DNA. Isti efekt pokazan je u studijama drugih istraživačkih grupa[210–212]. Modul za selekciju stanica sam dodatno nadogradio u svrhu istovremene antibiotske i fluorescencijske selekcije transfeciranih stanica. Na modul za antibiotsku rezistenciju (PuroR) dodao sam novi ljepljivi nepalindromski kraj koji povezuje modul za fluorescencijski protein preko samo-cijepajućeg peptida P2A (engl. *Porcine Teschovirus-1 2A*)[196].

Pokazano je kako ciljanje promotorske regije određenog genskog lokusa fuzijom dSpCas9-DNMT3A s više molekula sgRNA istovremeno može povećati stupanj metilacije DNA na široj regiji te posljedično ima jači učinak na transkripcijsku aktivnost tog gena[136]. Također, pokazano je kako korištenje više molekula sgRNA istovremeno za navođenje fuzijskih konstrukata u svrhu ciljane aktivacije ili represije transkripcije gena djeluje sinergistički[133,142,144,168]. U tu svrhu ekspresijski sgRNA modul dodatno sam nadogradio tako

da može primiti do šest različitih molekula sgRNA[196]. Varijabilni dijelovi molekula sgRNA prvo se kloniraju pomoću restrikcijskog enzima BpiI u zasebne ekspresijske module, koji sadrže U6 promotor, okosnicu sgRNA za određeni ortologni protein Cas9 (SpCas9 ili SaCas9), U6 terminator, te koji su okruženi s dva restrikcijska mjesta Esp3I. Nepalindromski ljepljivi krajevi dobiveni restrikcijskim enzimom Esp3I potom osiguravaju pravilno sklapanje više ekspresijskih modula sgRNA u jedan modul, tzv. modul „multi-guide“ koji može primiti do šest različitih molekula sgRNA. Prilikom konstrukcije fuzijskih CRISPR/dCas9 konstrukata potrebno je koristiti modul „multi-guide“ kako bi se omogućilo kloniranje do šest različitih molekula sgRNA. Selekcijski biljeg mRuby3 koristio sam i u ovoj reakciji za crveno-bijelu selekciju bakterijskih kolonija.

Kako bih razvio modularni sustav temeljen na dva ortologna proteina Cas9 konstruirao sam aktivne fuzije proteina Cas9 iz dviju različitih vrsta bakterija, *Staphylococcus aureus* i *Streptococcus pyogenes*. Ortolog SaCas9 jedan je od najmanjih ortologa proteina Cas9 koji pokazuje aktivnost usporedivu s ortologom SpCas9, koji se najčešće koristi, uz veću specifičnost[198–200]. U studijama drugih istraživačkih grupa uspješno su konstruirane aktivne fuzije ortologa SaCas9 s direknim aktivatorom VPR[197] i represorom KRAB[201]. Aktivne fuzije s epigenetičkim efektorskim proteinima u svrhu metilacije i demetilacije molekule DNA nisu do sada uspješno konstruirane, stoga su prvi puta konstruirane i testirane u ovom doktorskom radu. Protein dSaCas9 sam stoga vezao s epigenetičkim efektorskim proteinima DNMT3A u svrhu ciljane metilacije DNA odnosno TET1 u svrhu demetilacije molekule DNA[196,213].

U svrhu konstrukcije aktivnih fuzija s ortologom SaCas9 prvo sam testirao vezanje katalitičkih domena DNMT3A i TET1 na C-terminalni kraj katalitički inaktiviranog ortologa dSaCas9 preko Gly_4Ser linkera[196,213]. Primjena takvog konstrukta u svrhu promjene stupnja metilacije DNA nije bila uspješna. Razlog tome može biti neefikasan ulazak fuzijskog konstrukta u jezgru stanica zbog nedostupnosti signala NLS (engl. *nuclear localization signal*). U takvom fuzijskom konstruktu, na N-terminalnom kraju nalazio se samo jedan signal za unos u jezgru (SV40 NLS), dok se nukleoplazmin NLS nalazio na C-terminalnom kraju proteina dSaCas9, između samog terminusa i vezane efektorske domene. Iako neke studije pokazuju da je jedan NLS dovoljan za ulazak proteina SpCas9 u jezgru[214], druge pak pokazuju kako dodavanje još jednog signala NLS pojačava ulazak konstrukta u jezgru stanica[30]. Stoga sam dodao još jedan nukleoplazmin NLS-na C-terminalni kraj efektorske domene DNMT3A unutar fuzijskog konstrukta s ortologom dSpCas9, te sam pokazao kako je ovakav konstrukt efikasniji u odnosu na fuzijski konstrukt bez dodatnog signala NLS – stupanj ciljane metilacije DNA

povećao se za oko 20%. Također, fuzijski konstrukt s ortologom dSaCas9 pokazao se isto efikasnijim ako se za efektorsku domenu DNMT3A na C-terminusu vezao dodatni signal NLS - na određenim CpG dinukleotidima došlo je do povećanja stupnja metilacije do 24%. Prema kristalnoj strukturi proteina Cas9, C- i N-terminali krajevi nisu u neposrednoj blizini vezane DNA[215]. U ovoj doktorskoj disertaciji stoga je bilo potrebno testirati različite konfiguracije fuzijskih konstrukata vezanjem efektorske domene na C- ili N-terminalni kraj dCas9. Različite studije pokazuju kako fuzijski konstrukti imaju bolju aktivnost u određenoj konfiguraciji[164,216]. U ovoj doktorskoj disertaciji pokazao sam da vezanjem efektorskih domena DNMT3A i TET1 na N-terminalni kraj ortologa dSaCas9 i dSpCas9 mogu postići robusniji i bolji efekt na ciljanu metilaciju i demetilaciju DNA[196,213]. Prema tome, ovu sam konfiguraciju koristio kao univerzalni pristup prilikom konstrukcije modularnog sustava.

Za fuzijske konstrukte, u kojima su efektorske domene DNMT3A i TET1 vezane na N-terminalni kraj proteina dSaCas9, izradio sam profile aktivnosti kako bih okarakterizirao njihovo djelovanje u odnosu na mjesto vezanja molekule sgRNA. Fuzijski konstrukti navođeni različitim molekulama sgRNA pokazuju drugačije profile metilacije. Poznato je da različite molekule sgRNA mogu imati različitu efikasnost[217,218] pa čak i kada se cilja isti genski lokus[136,219]. Brojne studije odredile su koje karakteristike molekule sgRNA utječu na njezinu efikasnost. Pokazano je da sam nukleotidni sastav molekule sgRNA[218,220–222], ali i kromatinski status veznog mjesta[223,224] utječu na efikasnost navođenja pomoću molekula sgRNA. Sumarni profil aktivnosti N-terminalnih fuzija efektorskih domena DNMT3A i TET1 na dSaCas9 pokazuje promjenu u metilaciji DNA na obje strane od mjesta vezanja molekule sgRNA uz identificirani maksimum aktivnosti oko 30 pb nizvodno i uzvodno od mjesta vezanja. Također, oko 180 do 200 pb uzvodno i nizvodno od mjesta vezanja može se vidjeti efekt na promjenu u metilaciji DNA što sugerira na interakciju fuzijskog konstrukta sa susjednim nukleosomom. Također, pokazao sam kako fuzijski konstrukti, u kojima su iste efektorske domene (DNMT3A i TET1) vezane na N-terminalni kraj ortologa dSpCas9, pokazuju gotovo iste obrasce aktivnosti kao i fuzije s ortologom dSaCas9. U usporedbi s N-terminalnim fuzijama, fuzije efektorskih domena DNMT3A i TET1 na C-terminalni kraj proteina dSpCas9 pokazuju manju efikasnost u modulaciji metilacije DNA, a maksimum promjene metilacije primijećen je uglavnom nizvodno od mjesta vezanja molekule sgRNA[134,136]. Kada usporedimo C- i N-terminalne fuzije efektorskih domena s različitim ortolozima Cas9, oba tipa fuzija pokazuju jaču efikasnost s ortolognim proteinom SpCas9[196,213]. Liu i suradnici su u svojem istraživanju pokazali da prosječni učinak koji C-terminalna fuzija katalitičke domene TET1 s dSpCas9 postiže iznosi

oko 30%[137]. U svojem istraživanju pokazao sam da se može postići puno veći efekt promjene metilacije DNA ako je katalitička domena TET1 vezana na N-terminalni kraj bilo dSpCas9 ili dSaCas9, što potvrđuje da je ova konfiguracija optimalna. Također, pokazao sam kako je efekt na promjenu metilacije DNA koji postiže katalitička domena DNMT3A slabiji u odnosu na onaj koji postiže katalitička domena TET1, neovisno o konfiguraciji[196]. Poznato je kako je enzim DNMT3L neophodan za vezanje DNMT3A za kromatin, te da pojačava procesivnost same metilitransferaze DNMT3A kroz stvaranje tetramernog kompleksa[63,152]. Studija Stepper i suradnika pokazuje kako vezanje kimernog kompleksa DNMT3A-DNMT3L na dCas9 postiže 4 do 5 puta veći efekt od same katalitičke domene DNMT3A[153].

U mojim istraživanjima zanimalo me je i koliko je stabilna promjena metilacije koju postižem upotrebom fuzijskih konstrukata DNMT3A-dSpCas9 i TET1-dSaCas9, gdje su efektorske domene vezane za N-terminalni kraj dCas9[196]. Profil aktivnosti kroz vrijeme pokazao je maksimalni efekt fuzije DNMT3A-dSpCas9 8. dan nakon prolazne transfekcije stanica, dok je maksimalni efekt fuzije TET1-dSaCas9 zabilježen 11. dan nakon transfekcije. Efekt inducirane promjene metilacije DNA pratio sam do 30. dana nakon prolazne transfekcije stanica te sam identificirao da stupanj metilacije ostaje promijenjen za oko 30 do 40%[196]. U prethodnim istraživanjima grupe za epigenetiku pokazan je maksimalni efekt ciljanja fuzijskog konstrukta dSpCas9-DNMT3A na promotorske regije gena *BACH2* i *IL6ST* 7. dan za gen *BACH2*, odnosno 6. dan za gen *IL6ST*. Efekt inducirane metilacije DNA naglo pada, ali ostaje umjereno promijenjen do 30. dana nakon prolazne transfekcije stanica[136]. U mojem istraživanju pokazao sam da ako N-terminalnu fuziju DNMT3A-dSpCas9 ciljam na genski lokus pomoću više molekula sgRNA istovremeno postižem veći efekt na promjenu metilacije ciljane regije. Genski lokus *IL6ST* ciljao sam istovremeno s četiri molekule sgRNA te pokazao veću inicijalnu promjenu te vremenski duži efekt na promjenu metilacije molekule DNA koji se stabilno nasljeđivao kroz stanične diobe. Različite studije pokazuju kako se maksimalni efekt ciljane demetilacije DNA s enzimom TET1 razlikuje ovisno o genskom lokusu i staničnoj liniji koja se koristi u eksperimentima[134,225]. Ciljanjem N-terminalne fuzije TET1-dSaCas9 pomoću dvije molekule sgRNA istovremeno na genski lokus *HNF1A* pokazao sam promjenu metilacije od oko 50% te njezino stabilno održavanje kroz stanične diobe[196].

Sustav CRISPR/Cas9 je vrlo fleksibilan na način da omogućava više strategija kako se različite efektorske domene mogu istovremeno koristiti u stanici. Strategije uključuju direktnu fuziju više kopija efektorskih domena ili nekoliko različitih efektorskih domena s dCas9[226,227], multimerizaciju pomoću sustava SunTag[163,228], regrutaciju više efektorskih domena pomoću

RNA vezujućih peptida koji prepoznaju određene RNA aptamere[162,229] te primjenu različitih ortologa proteina Cas9[190,230]. U ovoj doktorskoj disertaciji , za izradu modularnog sustava te za ciljanje različitih efektorskih domena istovremeno na različite/iste genske lokuse u modelnim stanicama odabrao sam korištenje dva ortologna proteina Cas9 - SpCas9 i SaCas9. Različiti ortolozi proteina Cas9 razlikuju se u molekuli sgRNA koju specifično vežu kao i u sekvenci PAM koja im je potrebna za vezanje na ciljano mjesto u genomu[190,200,231,232]. Određeni ortolog proteina Cas9 će stoga vezati samo svoju specifičnu sgRNA molekulu, a ne molekulu sgRNA namijenjenu za drugi ortologni protein Cas9, te će se na taj način moći regulirati ciljanje oba ortologna proteina Cas9 istovremeno. U literaturi je pokazana uspješna istovremena primjena ortologa SpCas9 i SaCas9 u svrhu ortogonalne regulacije ekspresije jednog gena, te utišavanja drugog gena pomoću dvolančanog loma (engl. *knockout*)[233]. Također, pokazana je uspješna ortogonalna aktivacija i represija pomoću direktnog aktivatora VPR i represora KRAB koristeći ortologne proteine SpCas9 i SaCas9[202]. Primjena različitih ortologa proteina Cas9 u svrhu antagonističke metilacije *vs*. demetilacije molekule DNA prvi je puta pokazana u ovom istraživanju[196].

Modularni sustav CRISPR/dCas9, koji sam dizajnirao u ovoj doktorskoj disertaciji, upotrijebio sam u svrhu istovremene ortogonalne manipulacije dva različita genska lokusa s efektorskim domenama antagonističkog djelovanja, DNMT3A i TET1, u modelnim stanicama HEK293. Gene kandidate, *BACH2*, *IL6ST*, *MAGT3* i *HNF1A*, odabrao sam jer su oni GWAS (engl. *genome-wide association study*) hitovi za glikozilaciju imunoglogulina G (IgG), a moguće je i da su uključeni i u patogenezu kroničnih upalnih bolesti crijeva[234]. BACH2 glavni je transkripcijski faktor uključen u sazrijevanje i diferencijaciju B-limfocita[235], IL6ST je transmembranski protein koji sudjeluje u vezanju i prijenosu brojnih citokina, MGAT3 je glikoziltransferaza koja dodaje N-acetilglukozamin (GlcNAc) vezom β1-4 na β manozu unutar pentasaharidne srži glikana te stvara rasijecajući GlcNAc[234], dok je transkripcijski faktor HNF1A bitan u mnogim staničnim procesima, uključujući regulaciju fukozilacije glikoproteina[236]. Fuzijskim CRISPR/dCas9 konstruktima konstruiranih pomoću modularnog sustava ciljao sam parove genskih lokusa *BACH2-HNF1A* i *IL6ST-MGAT3*, pri čemu sam genske lokuse *BACH2* i *IL6ST* ciljao fuzijskim konstruktom DNMT3A-dSpCas9, a genske lokuse *HNF1A* i *MGAT3* fuzijskim konstruktom TET1-dSaCas9. Promjene u stupnju metilacije DNA kretale su se od 20 do 80% u odnosu na netransfecirane kontrolne stanice (mock). Ovim sam eksperimentima pokazao da je moguće u istoj stanici specifično primijeniti stupanj metilacije DNA u suprotnom smjeru ciljanjem različitih efektorskih domena vezanih na dva

različita ortologna proteina Cas9. Svaki ortolog Cas9 prepoznao je specifično svoju molekulu sgRNA te nije došlo do krosreaktivnosti. Kako sam u svrhu prolazne transfekcije stanica koristio ekvimolarnu mješavinu plazmida DNMT3A-dSpCas9 i TET1-dSaCas9, koji sadrže isti selekcijski biljeg na antibiotik Puromicin (PuroR), odredio sam postotak dvostruko transfeciranih stanica, tj. stanica koje su primile oba konstrukta, prateći ekspresiju fluorescentnih biljega koji su se nalazili na konstruktima. Pokazao sam da oko 60% stanica sadrži oba fluorescentna biljega (mRuby3 i mClover3), koja se nalaze u istom okviru čitanja kao i fuzijski konstrukt, vezana preko samo-cijepajućeg peptida 2A na modul za antibiotsku rezistenciju, te da dobiveni efekt potječe većinom iz stanica koje su dvostruko transfecirane[196]. Analize u ovom istraživanju provedene su na „miješanoj" populaciji stanica i ovo je jedan od nedostataka ove studije. Naime, jasniji rezultat učinka fuzijskih konstrukata dobio bih da sam analizirao „čistu" populaciju stanica (samo onih koje su dvostruko transfecirane), međutim nisam imao mogućnost selekcije transfeciranih stanica sorterom.

Korelacijske studije pokazuju da metilacija DNA u promotorskoj regiji gena, posebno u CpG otocima, negativno korelira s ekspresijom gena[237]. Međutim, tek primjena molekularnih alata, koji mogu promijeniti metilaciju na ciljanoj regiji promotora, pojačivača i/ili ostalih kontrolnih elemenata, može razotkriti značaj metilacije u regulaciji specifičnog gena. U svojim sam istraživanjima pokazao da je inducirana promjena metilacije na točno određenim citozinima u promotorima parova genskih lokusa, navođenjem fuzijskih konstrukata s efektorskim domenama antagonističkog djelovanja vezanih za različite ortologne proteine Cas9, rezultira promjenom transkripcijske aktivnosti ciljanih gena. Metilaciju DNA i transkripcijsku aktivnost gena analizirao sam u više vremenskih točaka – 5., 8., 9. i 12. dan nakon transfekcije modelnih animalnih stanica HEK293. Ekspresija gena pratila je promjenu stupnja metilacije promotora. Efekt oba fuzijska konstrukta, TET1-dSaCas9 te DNMT3A-dSpCas9, na promjenu metilacije DNA bio je najizraženiji 8. dan nakon transfekcije, a popratio ga je i najizraženiji efekt u promjeni razine transkripcijske aktivnosti istih gena. Razina promjene u ekspresiji ciljanih gena počinje padati 12. dan nakon prolazne transfekcije stanica[196]. Studije drugih autora također pokazuju kako prolazna promjena metilacije DNA, inducirana molekularnim alatima koji se osnivaju na tehnologiji CRISPR/dCas9, ima kratkotrajan efekt na ekspresiju gena koja se nakon određenog vremena vraća na originalno stanje[134,136,225,238]. Razlog tome je gubitak plazmida i ekspresije fuzijskog konstrukta s vremenom pošto prilikom prolazne transfekcije stanica ne dolazi do ugradnje plazmida u genom stanica. Zbog toga sam stupanj metilacije DNA i razinu transkripcijske aktivnosti svih

genskih lokusa analizirao 8. dan nakon transfekcije stanica. Ciljanjem para gena *HNF1A-BACH2* istovremeno s fuzijama TET1-dSaCas9 i DNMT3A-dSpCas9 povećao sam ekspresiju gena *HNF1A* za 5,614 puta, dok sam ekspresiju gena *BACH2* smanjio za 0,568 puta u odnosu na netransfecirane stanice. Ciljanjem para gena *MGAT3-IL6ST* istovremeno fuzijama TET1-dSaCas9 i DNMT3A-dSpCas9 smanjio sam ekspresiju gena *IL6ST* za 0,583 puta u odnosu na netransfecirane stanice, no ekspresija gena *MGAT3* se nije povećala. Štoviše, nakon demetilacije promotorske regije gena *MGAT3* njegova transkripcijska aktivnost se smanjila za 0,671 puta u odnosu na netransfecirane stanice. Bez obzira što je katalitička domena TET1 u fuzijskom konstruktu postigla smanjenje postotka metilacije na pojedinačnim CpG dinukleotidima od 19 do 72%, ekspresija gena *MGAT3* se nije povećala. Također, u stanicama transfeciranim kontrolnim konstruktom koji je sadržavao katalitički inaktivnu efektorsku domenu TET1, stupanj metilacije DNA ostao je nepromijenjen, dok se transkripcijska aktivnost gena *MGAT3* isto smanjila[196]. Razlog smanjenja ekspresije u oba slučaja može biti CRISPR interferencija zbog samog vezanja konstrukata ciljanih sa šest različitih molekula sgRNA istovremeno na promotorsku regiju. Pokazano je kako CRISPR interferencija, tj. vezanje samog kompleksa na promotorsku regiju gena, može blokirati RNA polimerazu i transkripciju ciljanog gena[239].

Najveći efekt na ekspresiju gena nakon epigenetičke manipulacije pokazao sam za gen *HNF1A*. Prijašnja studija je identificirala četiri CpG mjesta u 1. egzonu ovoga gena, koja bi potencijalno mogla imati regulatornu ulogu u njegovoj ekspresiji[240]. Stoga sam na regiju, koja sadrži ova četiri CpG mjesta, ciljao fuzijski konstrukt TET1-dSaCas9 istovremeno s dvije molekule sgRNA. Na 1. CpG mjestu stupanj metilacije se promijenio za 27%, na 2. CpG mjestu 32%, a na zadnja dva 49% i 65%. Značajno povećanje transkripcijske aktivnosti gena *HNF1A*, koje je uslijedilo nakon inducirane promjene metilacije, sugerira da je epigenetička manipulacija u manjoj regiji dovoljna ako ista sadrži regulatorna CpG mjesta. Slabiji efekt inducirane epigenetičke promjene pokazao se na transkripcijsku aktivnost gena *BACH2* i *IL6ST* iako sam oba lokusa ciljao fuzijskim konstruktom DNMT3A-dSpCas9 s ukupno četiri molekule sgRNA na širu regiju unutar CpG otoka ovih gena. Povećanje metilacije DNA kretalo se od 2 do 55% za genski lokus *BACH2*, odnosno 3 do 29% za genski lokus *IL6ST*. Slabije postignuti efekt povećane metilacije DNA na ekspresiju genskih lokusa *BACH2* i *IL6ST* može se objasniti činjenicom da su ciljane regije slučajno odabrane unutar velikih CpG otoka jer ne postoje podaci o potencijalno regulatornim regijama. Maeder i suradnici su pokazali u svojoj studiji da promjena metilacije određenih CpG dinukleotida unutar promotora gena *HBB* i *RHOXF2* utječe

na promjenu u transkipcijskoj aktivnosti istih[241]. Druga studija pokazala je kako promjena metilacije DNA u manjoj regiji unutar CpG otoka gena *CDKN2A* nema efekta na promjenu u ekspresiji gena, te da se efekt na transkripciju tog gena može postići tek ako se cilja cijeli CpG otok s fuzijom dCas9-DNMT3A[151]. Poznato je da nemaju svi geni u svojim promotorskim regijama CpG otoke, te nije u potpunosti jasan mehanizam regulacije ekspresije tih gena[242,243]. Regulacija ekspresije eukariotskih gena kompleksan je proces u koji nije uključena samo metilacija citozina u promotorskoj regiji, već i drugi epigenetički mehanizmi poput post-translacijskih modifikacija histona, ali i neki drugi mehanizami poput djelovanja nekodirajućih molekula RNA. Međudjelovanje različitih epigenetičkih mehanizama rezultira kompaktnijom ili opuštenijom strukturom kromatina[244]. Pokazano je da primjenom inhibitora histonskih deacetilaza, kao što je Trihostatin A, dolazi do reaktivacije gena s hipermetiliranih promotora bez promjene u metilaciji molekule DNA[245–247]. Upravo zbog toga primjena tehnologije CRISPR/dCas9 našla se i u modifikaciji histonskih oznaka, te je pokazano kako i ovaj pristup utječe na ekspresiju gena[138,139,248,249]. Učinak na promjenu u ekspresiji ciljanog gena stoga ovisi o specifičnom genu koji se cilja, staničnoj liniji koja se koristi kao i o međudjelovanju epigenetičkim oznakama.

Kako bih pokazao da je simultana epigenetička manipulacija zaista moguća upotrebom ortolognih proteina Cas9 i vezanjem za njih različitih efektorskih domena, manipulirao sam dodatni par genskih lokusa i u drugoj staničnoj liniji, BG1. U odabranom paru gena *MGAT3*-*HNF1A*, na gen *MGAT3* ciljao sam fuziju DNMT3A-dSpCas9 jednom molekulom sgRNA, a gen *HNF1A* sam ciljao fuzijom TET1-dSaCas9 s dvije molekule sgRNA[196]. Kohler i suradnici su analizom metilacije DNA u više tumorskih stanica jajnika, koje pokazuju različitu ekspresiju gena *MGAT3*, otkrili regiju od 200 pb u neposrednoj blizini TSS-a kao potencijalno regulatornu. U staničnim linijama koje pokazuju visoku ekspresiju gena *MGAT3*, uključujući BG1 stanice, ista regija pokazuje nisku razinu metilacije DNA[250]. Ista regija je u ovoj doktorskoj disertaciji odabrana za manipulaciju fuzijom DNMT3A-dSpCas9. Kao drugi genski lokus u paru odabrao sam *HNF1A* pošto je poznata njegova važnost u regulaciji fukozilacije proteina[236]. Gen *HNF1A* pokazuje visoku metilaciju DNA i nisku razinu ekspresije u stanicama BG1, stoga sam na ovaj genski lokus ciljao fuziju TET1-dSaCas9 vođenu s dvije molekule sgRNA, koje su se i u stanicama HEK293 pokazale uspješnima. Postotak dvostruko transfeciranih stanica, upotrebom transfekcijskog reagensa PEI MAX 40K, bio je 84.4%. Nakon uspješne manipulacije metilacije DNA, pokazao sam i promjenu u transkripcijskoj aktivnosti ciljanih gena, koja je išla u

različitim smjerovima. Razinu ekspresije gena *MGAT3* smanjio sam za 0,28 puta, dok sam ekspresiju gena *HNF1A* povećao za 6,64 puta u odnosu na netransfecirane stanice.

Upotreba molekularnih alata koji se osnivaju na tehnologiji CRISPR/dCas9 u konačnici je važna u svrhu promjene fenotipa. Prva uspješna manipulacija pokazana je u modelnim animalnim stanicama i u mišu gdje je revertiran bolesni fenotip fragilnog X sindroma. Ciljana demetilacija ponavljanja trinukleotida CGG dovela je do reaktivacije gena *FMR1* te posljedično do popravka fenotipa i funkcije neurona. Ovaj pristup pokazao se uspješnim i *in vivo* čime je potvrđen potencijal ciljane epigenetičke promjene pomoću alata CRISPR/dCas9 u terapijske svrhe[155]. Modularni alat izrađen u ovoj doktorskoj disertaciji uspješno je primijenjen u svrhu promjene ekspresije gena, ali je pokazano da takva promjena ima i efekt na mjerljivi fenotip. MGAT3 je glikozil transferaza koja dodaje N-acetilglukozamin (GlcNAc) vezom β1-4 na β manozu unutar pentasaharidne srži glikana te stvara rasijecajući GlcNAc[234], dok je pokazano da je HNF1A glavni regulator fukozilacije, regulira antenarnu i sržnu fukozilaciju glikoproteina kroz regulaciju mnogobrojnih fukozil-transferaza[236]. U stanicama BG1 gdje sam simultano uspješno manipulirao par gena *MGAT3*-*HNF1A* te im promijenjeno ekspresiju, pokazao sam da je posljedično promijenjen i glikanski fenotip. Analiza masenom spektrometrijom ukupnog N-glikoma stanica pokazala je smanjenu razinu struktura s rasijecajućim GlcNAc, što odgovara smanjenju ekspresije gena *MGAT3*, te smanjenu razinu struktura sa sržnom fukozom što odgovara povećanoj ekspresiji gena *HNF1A*. U studiji koju su proveli Lauc i suradnici pokazano je metodom RNA interferencije u stanicama HepG2 da smanjenje ekspresije gena *HNF1A* dovodi do povećanja ekspresije gena *FUT8* zaslužnog za dodavanje sržne fukoze na N-glikane[236]. Povećanje ekspresije gena *HNF1A* nakon ciljane demetilacije pomoću fuzije TET1-dSaCas9, u ovoj studiji, vjerojatno je kroz direktnu regulaciju gena *FUT8* rezultiralo smanjenjem u zastupljenosti struktura sa sržnom fukozom.

Na temelju rezultata antagonističke metilacije-demetilacije u stanicama HEK293 i BG1 pokazao sam važnost manje regije u promotoru gena *HNF1A* u regulaciji njegove ekspresije. Kako bih provjerio može li se ekspresija gena *HNF1A* još više povećati i održati kroz vrijeme testirao sam potencijalno sinergistički učinak istovremenog ciljanja regulatorne regije u promotoru gena *HNF1A* s katalitičkom domenom TET1 (u svrhu demetilacije DNA) i direktnim aktivatorom VPR u stanicama HEK293. Fuziju VPR-dSpCas9 navodio sam na *HNF1A* s jednom molekulom sgRNA, a fuziju TET1-dSaCas9 s dvije molekule sgRNA. Prilikom dizajna molekule sgRNA za ciljanje fuzije VPR-dSpCas9 vodio sam se rezultatima studije Gilbert i suradnika koja je pokazala kako maksimalni učinak na ekspresiju ciljanih gena

pokazuje ciljanje aktivatora na regiju od -400 do -50 pb u odnosu na TSS[163]. Pratio sam promjenu u metilaciji DNA kao i ekspresiju gena *HNF1A* u više vremenskih točaka - 5., 8., 11., 20. i 30. dan nakon prolazne transfekcije stanica. Do demetilacije molekule DNA došlo je i ciljanjem samog TET1-dSaCas9 konstrukta, što je u skladu s prijašnjim eksperimentima, kao i ciljanjem TET1-dSaCas9 u kombinaciji s VPR-dSpCas9. Razina demetilacije molekule DNA bila je veća u slučaju kada sam koristio samo fuzijski konstrukt TET1-dSaCas9, a ne njegova kombinacija s VPR-dSpCas9[196]. Razlog tome može biti interferencija između različitih fuzijskih konstrukata. Nedavna studija Halmai i suradnika pokazala je sličan rezultat u slučaju korištenja kombinacije dCas9-VP64 i dCas9-TET1. Razina metilacije DNA statistički značajno je bila viša kada su koristili kombinaciju fuzijskih konstrukata na genskom lokusu *CDKL5* nego kada su koristili samo dCas9-TET1[251]. Također, u ovoj doktorskoj disertaciji pokazao sam kako sama fuzija VPR-dSpCas9 ne utječe na metilaciju molekule DNA, što je u skladu i sa nedavnom studijom Halmai i suradnika koja je testirala učinak aktivatora VP64 na promjenu u metilaciji DNA[251]. Analizom ekspresije gena *HNF1A* pokazao sam kako aktivator VPR pokazuje naglo povećanje u ekspresiji gena *HNF1A* 5. dan nakon prolazne transfekcije kada se koristi zasebno ili u kombinaciji s TET1. Iako efekt fuzije VPR-dSpCas9 slabi 8. dan nakon prolazne transfekcije stanica, sinergistički učinak kombinacije obje fuzije zabilježen je 11. dan nakon prolazne transfekcije stanica. Nadalje kroz vrijeme, efekt na ekspresiju gena *HNF1A* pada, međutim 30. dan nakon transfekcije stanica kombinacija pokazuje i dalje sinergistički učinak na razinu ekspresije koja ostaje povećana za šest puta u odnosu na netransfecirane stanice. Sama fuzija VPR-dSpCas9 pokazuje 30. dan nakon prolazne transfekcije statistički značajno duplo manji učinak na promjenu ekspresije gena *HNF1A*. Objašnjenje za sinergistički učinak koji se pokazuje 11. dan nakon prolazne transfekcije stanica može biti taj što se do tog dana smanjuje stupanj metilacije DNA (postignut djelovanjem fuzije TET1-dSaCas9) što dovodi do remodeliranja kromatina. Kako bih utvrdio nasljeđuje li se efekt demetilacije DNA na ekspresiju gena *HNF1A* diobama, ili je pak on rezultat konstantne prisutnosti fuzijskog konstrukta, analizirao sam prisutnost TET1-dSaCas9 na DNA, RNA i proteinskoj razini. Prisutnost fuzijskog konstrukta TET1-dSaCas9 naglo pada nakon prolazne transfekcije stanica te se ne može pokazati niti na jednoj razini nakon 11. dana. Ovime sam potvrdio da se inducirana epigenetička promjena nasljeđuje kroz stanične diobe. Brojne studije pokazuju kako status kromatina, tj. opušteno ili kondenzirano stanje kromatina, utječe na efikasnost vezanja proteina Cas9[223,224,252]. Opuštanje kromatina, kao posljedica demetilacije molekule DNA, vjerojatno omogućava lakše vezanje fuzije VPR-dSpCas9 te bolji učinak na ekspresiju gena *HNF1A*. U prilog tome ide studija Halmai i suradnika koja je pokazala da je za statistički

značajnu aktivaciju gena *CDKL5*, lociranog na inaktivnom kromosomu X, potrebno koristiti kombinaciju fuzije dCas9-VP64 s fuzijom dCas9-TET1 koja uzrokuje demetilaciju molekule DNA te posljedično remodeliranje kromatina i time omogućava vezanje fuzije dCas9-VP64[251]. Također, u prilog tome ide i studija Baumann i suradnika koja pokazuje kako kombinacija katalitički inaktivne domene TET1 s VP64 ne uzrokuje statistički značajno povećanje gena *Sox1* za razliku od kombinacije katalitički aktivne domene TET1 s VP64. Također, autori su pokazali kako primjena epigenetičkog inhibitora DNA metiltransferaza, Zebularina, u kombinaciji s dCas9-VP64 ima usporedivi učinak kao i kombinacija dCas9-TET1 s dCas9-VP64[253]. Nedavna studija Morita i suradnika također pokazuje sinergistički učinak na aktivaciju gena pomoću direktnog aktivatora VP64 i epigenetičke efektorske domene TET1[254].

Jedan od glavnih nedostataka sustava CRISPR/Cas9 je njegova nespecifična aktivnost za što je glavni razlog vezanje molekule sgRNA na djelomično komplementarna mjesta u genomu. Pokazano je kako nespecifično vezanje nukleaze Cas9 tolerira do šest nepravilno sparenih baza između molekule sgRNA i veznog mjesta u genomu[177–179,181,255]. Nespecifičnoj aktivnosti pridonosi i prepoznavanje alternativnih PAM sekvenci[35,167,181] kao i insercije i delecije između molekule sgRNA i veznog mjesta u genomu[185]. Kako bi se smanjila nespecifična aktivnost CRISPR/Cas9 sustava testirani su i pokazani brojni pristupi[37]. Smanjenje količine kompleksa sgRNA-Cas9, pravilan dizajn i modifikacije molekula sgRNA, razvoj brojnih varijanti Cas9 s većom specifičnošću, te primjena različitih ortologa proteina Cas9 s kompleksnijim sekvencama PAM glavne su strategije pomoću kojih se smanjuje nespecifičan učinak sustava CRISPR/Cas9[37,187–190,256]. Kromatinskom imunoprecipitacijom pokazano je nespecifično vezanje nukleaze Cas9 duž cijelog genoma[36,191,192]. Također, fuzijski konstrukt dCas9-KRAB pokazuje sličan obrazac nespecifičnog vezanja[191]. Po pitanju ciljanog manipuliranja genoma pokazano je kako ne dolazi do uvođenja dvolančanog loma u molekulu DNA na većini od nespecifično vezanih mjesta koja su detektirana metodom kromatinske imunoprecipitacije. Problem se pojavljuje kod primjene fuzijskih konstrukata s dCas9 u svrhu ciljane manipulacije epigenoma ili direktne regulacije ekspresije gena. Naime, svaka slaba i prolazna interakcija fuzijskog kompleksa s nekim mjestom u genomu može dovesti do nespecifične aktivnosti efektorske domene. Predložen model vezanja nukleaze Cas9 u genomu objašnjava prolazne interakcije. Nukleaza Cas9 prepoznaje i veže sekvencu PAM te započinje lokalno razmotavanje lanaca molekule DNA neposredno uzvodno od sekvence PAM čime je omogućeno vezanje tzv. regije „seed*"*. Pravilnim vezanjem regije „seed" dolazi do potpunog vezanja molekule sgRNA, aktivacije nukleaze Cas9 i uvođenja dvolančanog loma. Ako nema

pravilnog vezanja regije „seed“ dolazi do disocijacije nukleaze Cas9 s tog mjesta[36,257,258]. Smanjenje ekspresije fuzijskog kompleksa ili same efektorske domene glavna je strategija u smanjuju nespecifičnog učinka efektorske domene vezane na dCas9[151,193,195].

U ovoj doktorskoj disertaciji sam pokazao kako razdvajanje fuzijskog kompleksa DNMT3A-dSpCas9 i TET1-dSpCas9 pod slabiji EFS promotor, uz ostavljanje biljega za selekciju transfeciranih stanica pod snažnim SV40 promotorom, dolazi do smanjenja nespecifičnog učinka[196]. Ovaj je rezultat dokazan analizom metilacije CpG dinukleotida na razini genoma što je u skladu s prethodnim istraživanjima[151,193,195]. Fuziju DNMT3A-dSpCas9 ciljao sam na promotorsku regiju gena *IL6ST* koristeći četiri molekule sgRNA istovremeno, dok sam s fuzijom TET1-dSpCas9 ciljao gen *MGAT3* koristeći pet molekula sgRNA istovremeno. Kao kontrolu koristio sam konstrukte kod kojih je fuzijski DNMT3A/TET1-dSpCas9 protein i selekcijski biljeg pod jednim snažnim CBh promotorom. Na razini proteina, metodom „western blot“, potvrdio sam da je korištenjem slabijeg promotora EFS smanjena količina fuzijskih proteina DNMT3A-dSpCas9 i TET1-dSpCas9. Metodom *Infinium MethylationEPIC* (850K) *array* pokazao sam kako slabijom ekspresijom fuzijskih proteina dolazi do smanjena nespecifičnog učinka bez narušavanja razine željenog učinka na ciljanom mjestu u genomu (engl. *on-target effect*). Ekspresija fuzijskog kompleksa DNMT3A-dSpCas9 vođena snažnim CBh promotorom rezultirala je s 71 910 ± 2367 diferencijalno metiliranim pozicijama u genomu (engl. *Differentially Methylated Positions*, DMPs), što odgovara 9,1% ± 0,3 Infinium MethylationEPIC array pozicija, dok je 41 894 ± 3939 diferencijalno metiliranih pozicija (5,3% ± 0,5 Infinium MethylationEPIC array pozicija) dobiveno kada je ekspresija fuzijskog konstrukta vođena slabijim promotorom. Razlika se nije pokazala statistički značajnom kada se analizira neparametrijskim testom Student–Newman–Keuls (SNK), međutim upotrebom *t*-testa za nezavisne uzorke pokazuje se statistička značajnost (p=0,012). Fuzijski konstrukt TET1-dSpCas9 pokazao je veću promjenu u diferencijalno metiliranim pozicijama kada je njegova ekspresija vođena slabijim EFS promotorom. Dobiveno je statistički značajno smanjenje testom SNK (p=0,022) s 240 991 ± 55 452 (30,4% ± 6,99 Infinium MethylationEPIC array pozicija) na 135 697 ± 13 079 (17,1% ± 1,65 Infinium MethylationEPIC array pozicija). U skladu s očekivanjem, 98,7% diferencijalno metiliranih pozicija bilo je hipermetilirano nakon primjene konstrukta DNMT3A-dSpCas9 koji je bio vođen jakim ili slabim promotorom, dok je 97,2% te 93,2% pozicija DMP bilo hipometilirano nakon primjene konstrukta TET1-dSpCas9 čija je ekspresija vođena jakim ili slabim promotorom. Kada gledamo preklapanje pozicija DMP između konstrukata, čija je ekspresija

vođena jakim te slabim promotorom, u oba slučaja i za DNMT3A-dSpCas9 i TET1-dSpCas9 vidimo kako se veći dio pozicija preklapa, što bi ukazivalo na to da nespecifična aktivnost nije potpuno stohastički proces. Međutim, studija Galonska i suradnika pokazuje razlike u nespecifičnoj aktivnosti između različitih klonova stanica čime potvrđuje stohastičku prirodu kontakata efektorske domene i kromatina[194]. Više studija pokazuje kako je nespecifična aktivnost neovisna o molekuli sgRNA[193,194], kao i o broju molekula sgRNA, te da prisutnost same katalitičke domene DNMT3A uzrokuje nespecifično globalno povećanje metilacije DNA[194].

Željeni efekt na ciljanu regiju promotora gena *IL6ST* vidi se na probama cg21950518 i cg15219433 gdje je došlo do statistički značajnog povećanja metilacije. Također, za ciljanu regiju gena *MGAT3*, proba cg21461856 pokazuje statistički značajno smanjenje metilacije DNA u odnosu na netransfecirane stanice. Također, pokazao sam kako se hipermetilacija pomoću fuzijskog konstrukta DNMT3A-dSpCas9 širi na veću regiju ±2500 pb u odnosu na ciljanu regiju. Efekt demetilacije molekule DNA pomoću fuzijskog konstrukta TET1-dSpCas9 maksimalni je na ciljanom mjestu te ima slabi efekt na regijama udaljenima ±2500 pb. Kada gledamo distribuciju inducirane metilacije DNA pomoću konstrukta DNMT3A-dSpCas9, vidimo kako je obogaćena na CpG otocima te regijama udaljenima oko 2 kb od CpG otoka (engl. *CpG shores*). Za konstrukt TET1-dSpCas9, distribucija inducirane hipometilacije DNA obogaćena je na regijama „CpG shores" te regijama udaljenima oko 2 kb od istih (engl. *CpG shelf*). Dobiveni rezultati su u skladu s prethodnim istraživanjima koja pokazuju kako gustoća dinukleotida CpG negativno korelira s metilacijom DNA[259–261]. Stoga, CpG otoci su najčešće hipometilirani i osjetljivi na hipermetilaciju koristeći alat DNMT3A-dSpCas9, dok su hipermetilirane regije („CpG shores" i „CpG shelfs") koje su siromašne dinukleotidima CpG osjetljivije na demetilaciju pomoću TET1-dSpCas9. Kada gledamo raspodjelu nespecifične aktivnosti ovisno o regulatornim regijama gena, pokazao sam kako je inducirana hipermetilacija pomoću DNMT3A-dSpCas9 obogaćena na slabim i jakim promotorskim regijama bogatima dinukleotidima CpG. Nespecifična demetilacija pomoću TET1-dSpCas9 zastupljena je na regijama slabih i jakih pojačivača, utišanim regijama te na regulatornim mjestima transkripcije[196].

U ovom istraživanju uspješno sam konstruirao modularni sustav temeljen na tehnologiji CRISPR/dCas9 u svrhu ciljane metilacije i demetilacije molekule DNA kao i u svrhu direktne modulacije transkripcijske aktivnosti gena. Pokazao sam kako inducirana promjena u metilaciji DNA u kandidat genima uključenima u glikozilaciju proteina ima učinak na njihovu

transkripcijsku aktivnost kao i na glikanski fenotip. Potencijal ovih alata veliki je u pogledu razjašnjavanja komplekse regulacije ekspresije gena uključenih u glikozilaciju IgG-a. Brojne studije pokazuju promijenjenu glikozilaciju IgG-a u autoimunim bolestima, upalnim bolestima crijeva kao i u različitim tumorima[262–267]. Također, alati razvijenu u ovom istraživanju koriste se unutar projekta Horizon 2020 (SYSCID), grupe za epigenetiku, s ciljem razumijevanja molekularnih mehanizama uključenih u patogenezu kroničnih upalnih bolesti. Potencijal sustava CRISPR/dCas9 veliki je kao terapijska strategija kod bolesti izazvanih epigenetičkim promjena.

4 ZAKLJUČAK

- Dizajnirao sam i konstruirao modularni sustav CRISPR/dCas9 koji se sastoji od modula: destinacijskog plazmida, promotora RNA polimeraze II, ortolognih proteina dSpCas9 ili dSaCas9, efektorske domene, selekcijskog biljega, eukariotskog transkripcijskog terminatora te modula koji može primiti do šest različitih molekula sgRNA u svrhu simultanog ciljanja i navođenja fuzijskih konstrukata u genomu.
- Pokazao sam da vezanje efektorskih domena DNMT3A i TET1 isključivo na N-terminalni kraj ortologa dSaCas9 omogućava aktivnost fuzijskih konstrukata. Također, N- terminalne fuzije efektorskih domena DNMT3A i TET1 na ortolog dSpCas9 pokazale su znatno jaču aktivnost u odnosu na C-terminalne fuzije istih efektorskih domena.
- Dodatak signala NLS iznimno je važan za efikasan ulazak CRISPR/dCas9 fuzijskih konstrukata u jezgru i postizanje željenog efekta.
- Sumarni profili aktivnosti N-terminalnih fuzija efektorskih domena DNMT3A i TET1 s dSaCas9 pokazuju da do promjene u stupnju metilacije DNA dolazi na obje strane od mjesta vezanja molekule sgRNA, uz maksimum aktivnosti oko 30 pb nizvodno i uzvodno od mjesta vezanja. Promjena u stupnju metilacije DNA detektirana je i na udaljenosti od 180 do 200 pb uzvodno i nizvodno od mjesta vezanja.
- Maksimum aktivnosti fuzije DNMT3A-dSpCas9 zabilježen je 8. dan kada je došlo do povećanja stupnja metilacije pojedinačnih CpG dinukleotida do 40%, a maksimum aktivnosti fuzije TET1-dSaCas9 zabilježen je 11. dan nakon prolazne transfekcije stanica kada je došlo do smanjenja stupnja metilacije pojedinačnih CpG dinukleotida za 60%.
- Efekt inducirane promjene stupnja metilacije molekule DNA, dobiven upotrebom N-terminalnih fuzija DNMT3A-dSpCas9 i TET1-dSaCas9 uz primjenu više molekula sgRNA istovremeno, stabilno se nasljeđuje kroz stanične diobe. Promjena stupnja metilacije od oko 30 do 40% održava se do 30. dana od prolazne transfekcije stanica.
- Primjena ortologa SpCas9 i SaCas9 pokazala se uspješnom u istovremenoj manipulaciji metilacije promotora više gena u istoj stanici. Ciljanjem parova genskih lokusa *BACH2-HNF1A* i *IL6ST-MGAT3* s fuzijama antagonističkog djelovanja u stanicama HEK293 postignute su promjene u stupnju metilacije DNA za 20 do 80% u odnosu na netransfecirane stanice.

- Promjena stupnja CpG metilacije, postignuta ciljanjem parova genskih lokusa fuzijama DNMT3A-dSpCas9 (hipermetilacija) i TET1-dSaCas9 (hipometilacija), rezultirala je promjenom razine transkripcijske aktivnosti tih gena.
- Postignuto veliko povećanje razine transkripcije gena *HNF1A* nakon ciljane demetilacije četiri CpG mjesta u 1. egzonu tog gena, upotrebom fuzijskog konstrukta TET1-dSaCas9 u stanicama HEK293 i BG1, ukazuje na to da je epigenetička manipulacija na manjoj genomskoj regiji dovoljna ako ista sadrži potencijalno regulatorna CpG mjesta.
- Pokazao sam da epigenetičkom manipulacijom genskih lokusa *MGAT3-HNF1A*, upotrebom molekularnih alata CRISPR/dCas9, mogu promijeniti glikanski fenotip u ljudskim stanicama raka jajnika, BG1. Hipermetilacija i posljedično smanjenje transkripcijske aktivnosti gena *MGAT3* rezultiralo je smanjenom razinom glikanskih struktura s rasijecajućim GlcNAc, dok je hipometilacija i posljedično povećana transkripcijska aktivnost gena *HNF1A* dovela do smanjene razine glikanskih struktura sa sržnom fukozom.
- Pokazao sam da mogu postići sinergistički učinak na ekspresiju gena *HNF1A* istovremenim ciljanjem fuzijskih konstrukata TET1-dSaCas9 (u svrhu ciljane demetilacije molekule DNA) i VPR-dSpCas9 (u svrhu direktne aktivacije ekspresije gena), te sam pokazao da je ovaj učinak postojan kroz vremenski period od najmanje 30 dana.
- Razdvajanjem fuzijskog kompleksa DNMT3A-dSpCas9 i TET1-dSpCas9 od biljega za selekciju transfeciranih stanica, na način da su kontrolirani promotorima različite jakosti, postigao sam manji nespecifični učinak efektorskih domena što je pokazano analizom CpG metilacije na razini cijelog genoma upotrebom metode Infinium Methylation EPIC (850K) array.

5 LITERATURA

1. Thomas KR, Folger KR, Capecchi MR. High frequency targeting of genes to specific sites in the mammalian genome. *Cell*. 1986; 44(3):419-428.

2. Capecchi M. Altering the genome by homologous recombination. *Science*. 1989; 244(4910):1288-1292.

3. Rudin N, Sugarman E, Haber JE. Genetic and physical analysis of double-strand break repair and recombination in Saccharomyces cerevisiae. *Genetics*. 1989; 122(3):519-534.

4. Sussman D, Chadsey M, Fauce S, et al. Isolation and Characterization of New Homing Endonuclease Specificities at Individual Target Site Positions. *J Mol Biol*. 2004; 342(1):31-41.

5. Rouet P, Smih F, Jasin M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol*. 1994; 14(12):8096-8106.

6. Seligman LM. Mutations altering the cleavage specificity of a homing endonuclease. *Nucleic Acids Res*. 2002; 30(17):3870-3879.

7. Klug A, Rhodes D. Zinc Fingers: A Novel Protein Fold for Nucleic Acid Recognition. *Cold Spring Harb Symp Quant Biol*. 1987; 52:473-482.

8. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci*. 1996; 93(3):1156-1160.

9. Bibikova M, Carroll D, Segal DJ, et al. Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Mol Cell Biol*. 2001; 21(1):289-297.

10. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet*. 2010; 11(9):636-646.

11. Deng D, Yan C, Pan X, et al. Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. *Science*. 2012; 335(6069):720-723.

12. Boch J, Scholze H, Schornack S, et al. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science*. 2009; 326(5959):1509-1512.

13. Miller JC, Tan S, Qiao G, et al. A TALE nuclease architecture for efficient genome

editing. *Nat Biotechnol*. 2011; 29(2):143-148.

14. Christian M, Cermak T, Doyle EL, et al. Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics*. 2010; 186(2):757-761.

15. Mussolino C, Alzubi J, Fine EJ, et al. TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic Acids Res*. 2014; 42(10):6762-6773.

16. Mak AN-S, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. The Crystal Structure of TAL Effector PthXo1 Bound to Its DNA Target. *Science*. 2012; 335(6069):716-719.

17. Lamb BM, Mercer AC, Barbas CF. Directed evolution of the TALE N-terminal domain for recognition of all 5′ bases. *Nucleic Acids Res*. 2013; 41(21):9779-9785.

18. Deng D, Yin P, Yan C, et al. Recognition of methylated DNA by TAL effectors. *Cell Res*. 2012; 22(10):1502-1504.

19. Gaj T, Sirk SJ, Shui S, Liu J. Genome-Editing Technologies: Principles and Applications. *Cold Spring Harb Perspect Biol*. 2016; 8(12):a023754.

20. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol*. 1987; 169(12):5429-5433.

21. Mojica FJM, Díez-Villaseñor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol*. 2000; 36(1):244-246.

22. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*. 2005; 151(8):2551-2561.

23. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol*. 2005; 60(2):174-182.

24. Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for

evolutionary studies. *Microbiology*. 2005; 151(3):653-663.

25. Jansen R, Embden JDA van, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol*. 2002; 43(6):1565-1575.

26. Makarova KS, Wolf YI, Alkhnbashi OS, et al. An updated evolutionary classification of CRISPR–Cas systems. *Nat Rev Microbiol*. 2015; 13(11):722-736.

27. Shmakov S, Smargon A, Scott D, et al. Diversity and evolution of class 2 CRISPR–Cas systems. *Nat Rev Microbiol*. 2017; 15(3):169-182.

28. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. 2012; 337(6096):816-821.

29. Deltcheva E, Chylinski K, Sharma CM, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*. 2011; 471(7340):602-607.

30. Cong L, Ran FA, Cox D, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*. 2013; 339(6121):819-823.

31. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*. 2014; 507(7490):62-67.

32. Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014; 346(6213):1258096.

33. Hsu PD, Lander ES, Zhang F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*. 2014; 157(6):1262-1278.

34. Szczelkun MD, Tikhomirova MS, Sinkunas T, et al. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc Natl Acad Sci*. 2014; 111(27):9798-9803.

35. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol*. 2013; 31(3):233-239.

36. Wu X, Scott DA, Kriz AJ, et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol*. 2014; 32(7):670-676.

37. Tadić V, Josipović G, Zoldoš V, Vojta A. CRISPR/Cas9-based epigenome editing: An overview of dCas9-based tools with special emphasis on off-target activity. *Methods*.

2019; 164-165:109-119.

38. Cai P, Gao J, Zhou Y. CRISPR-mediated genome editing in non-conventional yeasts for biotechnological applications. *Microb Cell Fact*. 2019; 18(1):63.

39. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci*. 2012; 109(39):E2579-E2586.

40. Chen H, Choi J, Bailey S. Cut Site Selection by the Two Nuclease Domains of the Cas9 RNA-guided Endonuclease. *J Biol Chem*. 2014; 289(19):13284-13294.

41. Lieber MR. The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annu Rev Biochem*. 2010; 79(1):181-211.

42. Jiang F, Zhou K, Ma L, Gressel S, Doudna JA. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science*. 2015; 348(6242):1477-1481.

43. Bestor TH, Edwards JR, Boulard M. Notes on the role of dynamic DNA methylation in mammalian development. *Proc Natl Acad Sci*. 2015; 112(22):6796-6799.

44. Schübeler D. Function and information content of DNA methylation. *Nature*. 2015; 517(7534):321-326.

45. Ehrlich M, Gama-Sosa MA, Huang L-H, et al. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. *Nucleic Acids Res*. 1982; 10(8):2709-2721.

46. Duncan BK, Miller JH. Mutagenic deamination of cytosine residues in DNA. *Nature*. 1980; 287(5782):560-561.

47. Jones PA. The Role of DNA Methylation in Mammalian Epigenetics. *Science*. 2001; 293(5532):1068-1070.

48. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci*. 2006; 103(5):1412-1417.

49. Dahl C, Guldberg P. DNA methylation analysis techniques. *Biogerontology*. 2003; 4(4):233-250.

50. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation

patterns in plants and animals. *Nat Rev Genet*. 2010; 11(3):204-220.

51. Varley KE, Gertz J, Bowling KM, et al. Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res*. 2013; 23(3):555-567.

52. Lister R, Pelizzola M, Dowen RH, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009; 462(7271):315-322.

53. Patil V, Ward RL, Hesson LB. The evidence for functional non-CpG methylation in mammalian cells. *Epigenetics*. 2014; 9(6):823-828.

54. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci*. 2006; 31(2):89-97.

55. Liang G, Chan MF, Tomigahara Y, et al. Cooperativity between DNA Methyltransferases in the Maintenance Methylation of Repetitive Elements. *Mol Cell Biol*. 2002; 22(2):480-491.

56. Chen T, Ueda Y, Dodge JE, Wang Z, Li E. Establishment and Maintenance of Genomic Methylation Patterns in Mouse Embryonic Stem Cells by Dnmt3a and Dnmt3b. *Mol Cell Biol*. 2003; 23(16):5594-5605.

57. Goll MG, Kirpekar F, Maggert KA, et al. Methylation of tRNA Asp by the DNA Methyltransferase Homolog Dnmt2. *Science*. 2006; 311(5759):395-398.

58. Kaiser S, Jurkowski TP, Kellner S, Schneider D, Jeltsch A, Helm M. The RNA methyltransferase Dnmt2 methylates DNA in the structural context of a tRNA. *RNA Biol*. 2017; 14(9):1241-1251.

59. Gowher H, Liebert K, Hermann A, Xu G, Jeltsch A. Mechanism of Stimulation of Catalytic Activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *J Biol Chem*. 2005; 280(14):13341-13348.

60. Chedin F, Lieber MR, Hsieh C-L. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc Natl Acad Sci*. 2002; 99(26):16916-16921.

61. Ooi SKT, Qiu C, Bernstein E, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*. 2007; 448(7154):714-717.

62. Kareta MS, Botello ZM, Ennis JJ, Chou C, Chédin F. Reconstitution and Mechanism of

the Stimulation of de Novo Methylation by Human DNMT3L. *J Biol Chem*. 2006; 281(36):25893-25902.

63. Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*. 2007; 449(7159):248-251.

64. Jurkowska RZ, Anspach N, Urbanke C, et al. Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. *Nucleic Acids Res*. 2008; 36(21):6656-6663.

65. Moore LD, Le T, Fan G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology*. 2013; 38(1):23-38.

66. Jeltsch A, Jurkowska RZ. Allosteric control of mammalian DNA methyltransferases – a new regulatory paradigm. *Nucleic Acids Res*. 2016; 44(18):8556-8575.

67. Wu H, Zhang Y. Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions. *Cell*. 2014; 156(1-2):45-68.

68. Bhutani N, Burns DM, Blau HM. DNA Demethylation Dynamics. *Cell*. 2011; 146(6):866-872.

69. Ito S, D'Alessio AC, Taranova O V., Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 2010; 466(7310):1129-1133.

70. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science*. 2009; 324(5929):930-935.

71. Ito S, Shen L, Dai Q, et al. Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science*. 2011; 333(6047):1300-1303.

72. He Y-F, Li B-Z, Li Z, et al. Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. *Science*. 2011; 333(6047):1303-1307.

73. Baylin SB. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum Mol Genet*. 2001; 10(7):687-692.

74. Robertson KD. DNA methylation, methyltransferases, and cancer. *Oncogene*. 2001;

20(24):3139-3155.

75. Jones PA, Laird PW. Cancer-epigenetics comes of age. *Nat Genet*. 1999; 21(2):163-167.

76. Eden S, Hashimshony T, Keshet I, Cedar H, Thorne AW. DNA methylation models histone acetylation. *Nature*. 1998; 394(6696):842-842.

77. Mohn F, Weber M, Rebhan M, et al. Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. *Mol Cell*. 2008; 30(6):755-766.

78. Irizarry RA, Ladd-Acosta C, Wen B, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet*. 2009; 41(2):178-186.

79. Watt F, Molloy PL. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev*. 1988; 2(9):1136-1143.

80. Kovesdi I, Reichel R, Nevins JR. Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. *Proc Natl Acad Sci*. 1987; 84(8):2180-2184.

81. Nan X, Ng H-H, Johnson CA, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*. 1998; 393(6683):386-389.

82. Nan X, Campoy FJ, Bird A. MeCP2 Is a Transcriptional Repressor with Abundant Binding Sites in Genomic Chromatin. *Cell*. 1997; 88(4):471-481.

83. Boyes J, Bird A. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell*. 1991; 64(6):1123-1134.

84. Aran D, Toperoff G, Rosenberg M, Hellman A. Replication timing-related and gene body-specific methylation of active human genes. *Hum Mol Genet*. 2011; 20(4):670-680.

85. Ball MP, Li JB, Gao Y, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotechnol*. 2009; 27(4):361-368.

86. Hellman A, Chess A. Gene Body-Specific Methylation on the Active X Chromosome. *Science*. 2007; 315(5815):1141-1143.

87. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from

epigenomics. *Nat Rev Genet*. 2008; 9(6):465-476.

88. Shenker N, Flanagan JM. Intragenic DNA methylation: implications of this epigenetic mechanism for cancer research. *Br J Cancer*. 2012; 106(2):248-253.

89. Istrail S, Davidson EH. Logic functions of the genomic cis-regulatory code. *Proc Natl Acad Sci*. 2005; 102(14):4954-4959.

90. Pan Y, Tsai C-J, Ma B, Nussinov R. Mechanisms of transcription factor selectivity. *Trends Genet*. 2010; 26(2):75-83.

91. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 1997; 389(6648):251-260.

92. Li G, Reinberg D. Chromatin higher-order structures and gene regulation. *Curr Opin Genet Dev*. 2011; 21(2):175-186.

93. Woodcock CL, Ghosh RP. Chromatin Higher-order Structure and Dynamics. *Cold Spring Harb Perspect Biol*. 2010; 2(5):a000596-a000596.

94. Mizzen CA, Allis CD. Linking histone acetylation to transcriptional regulation. *Cell Mol Life Sci*. 1998; 54(1):6-20.

95. Rea S, Eisenhaber F, O'Carroll D, et al. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*. 2000; 406(6796):593-599.

96. Felsenfeld G, Groudine M. Controlling the double helix. *Nature*. 2003; 421(6921):448-453.

97. Struhl K. Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev*. 1998; 12(5):599-606.

98. Nathan D, Sterner DE, Berger SL. Histone modifications: Now summoning sumoylation. *Proc Natl Acad Sci*. 2003; 100(23):13118-13120.

99. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev*. 2002; 16(1):6-21.

100. Hendrich B, Bird A. Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins. *Mol Cell Biol*. 1998; 18(11):6538-6547.

101. Lehnertz B, Ueda Y, Derijck AAHA, et al. Suv39h-Mediated Histone H3 Lysine 9

Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin. *Curr Biol*. 2003; 13(14):1192-1200.

102. Chang Y, Sun L, Kokura K, et al. MPP8 mediates the interactions between DNA methyltransferase Dnmt3a and H3K9 methyltransferase GLP/G9a. *Nat Commun*. 2011; 2(1):533.

103. Fuks F. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res*. 2003; 31(9):2305-2312.

104. Li H, Rauch T, Chen ZX, Szabó PE, Riggs AD, Pfeifer GP. The Histone Methyltransferase SETDB1 and the DNA Methyltransferase DNMT3A Interact Directly and Localize to Promoters Silenced in Cancer Cells. *J Biol Chem*. 2006; 281(28):19489-19500.

105. Bochtler M, Kolano A, Xu G-L. DNA demethylation pathways: Additional players and regulators. *Bioessays*. 2017; 39(1):1-13.

106. Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol*. 2010; 11(9):607-620.

107. Yang YA, Zhao JC, Fong K, et al. FOXA1 potentiates lineage-specific enhancer activation through modulating TET1 expression and function. *Nucleic Acids Res*. 2016; 44(17):8153-8164.

108. Tanaka T, Izawa K, Maniwa Y, et al. ETV2-TET1/TET2 Complexes Induce Endothelial Cell-Specific Robo4 Expression via Promoter Demethylation. *Sci Rep*. 2018; 8(1):5653.

109. Zhong J, Li X, Cai W, et al. TET1 modulates H4K16 acetylation by controlling auto-acetylation of hMOF to affect gene regulation and DNA repair function. *Nucleic Acids Res*. 2017; 45(2):672-684.

110. Zhang R, Erler J, Langowski J. Histone Acetylation Regulates Chromatin Accessibility: Role of H4K16 in Inter-nucleosome Interaction. *Biophys J*. 2017; 112(3):450-459.

111. Deplus R, Delatte B, Schwinn MK, et al. TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J*. 2013; 32(5):645-655.

112. Okitsu CY, Hsieh CL. DNA Methylation Dictates Histone H3K4 Methylation. *Mol Cell Biol*. 2007; 27(7):2746-2757.

113. Roeder R. The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci*. 1996; 21(9):327-335.

114. Buratowski S, Hahn S, Guarente L, Sharp PA. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell*. 1989; 56(4):549-561.

115. Boeger H, Bushnell DA, Davis R, et al. Structural basis of eukaryotic gene transcription. *FEBS Lett*. 2005; 579(4):899-903.

116. Goodrich JA, Cutler G, Tjian R. Contacts in Context: Promoter Specificity and Macromolecular Interactions in Transcription. *Cell*. 1996; 84(6):825-830.

117. Edmunds JW, Mahadevan LC, Clayton AL. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J*. 2008; 27(2):406-420.

118. McDaniel SL, Strahl BD. Shaping the cellular landscape with Set2/SETD2 methylation. *Cell Mol Life Sci*. 2017; 74(18):3317-3334.

119. Lorincz MC, Dickerson DR, Schmitt M, Groudine M. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol*. 2004; 11(11):1068-1075.

120. Dhayalan A, Rajavelu A, Rathert P, et al. The Dnmt3a PWWP Domain Reads Histone 3 Lysine 36 Trimethylation and Guides DNA Methylation. *J Biol Chem*. 2010; 285(34):26114-26120.

121. Bai L, Morozov A V. Gene regulation by nucleosome positioning. *Trends Genet*. 2010; 26(11):476-483.

122. Schones DE, Cui K, Cuddapah S, et al. Dynamic Regulation of Nucleosome Positioning in the Human Genome. *Cell*. 2008; 132(5):887-898.

123. Deal RB, Henikoff JG, Henikoff S. Genome-Wide Kinetics of Nucleosome Turnover Determined by Metabolic Labeling of Histones. *Science*. 2010; 328(5982):1161-1164.

124. Hartley PD, Madhani HD. Mechanisms that Specify Promoter Nucleosome Location and Identity. *Cell*. 2009; 137(3):445-458.

125. Kaplan N, Moore IK, Fondufe-Mittendorf Y, et al. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*. 2009; 458(7236):362-366.

126. Mao C, Brown CR, Griesenbeck J, Boeger H. Occlusion of Regulatory Sequences by Promoter Nucleosomes In Vivo. Tora L, ed. *PLoS One*. 2011; 6(3):e17521.

127. Kaplan N, Moore I, Fondufe-Mittendorf Y, et al. Nucleosome sequence preferences influence in vivo nucleosome organization. *Nat Struct Mol Biol*. 2010; 17(8):918-920.

128. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev*. 2011; 25(21):2227-2241.

129. Magnani L, Eeckhoute J, Lupien M. Pioneer factors: directing transcriptional regulators within the chromatin environment. *Trends Genet*. 2011; 27(11):465-474.

130. Tolkunov D, Zawadzki KA, Singer C, Elfving N, Morozov A V., Broach JR. Chromatin remodelers clear nucleosomes from intrinsically unfavorable sites to establish nucleosome-depleted regions at promoters. Chang F, ed. *Mol Biol Cell*. 2011; 22(12):2106-2118.

131. Wang X, Bai L, Bryant GO, Ptashne M. Nucleosomes and the accessibility problem. *Trends Genet*. 2011; 27(12):487-492.

132. Rose NR, Klose RJ. Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim Biophys Acta - Gene Regul Mech*. 2014; 1839(12):1362-1372.

133. Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*. 2013; 152(5):1173-1183.

134. Xu X, Tao Y, Gao X, et al. A CRISPR-based approach for targeted DNA demethylation. *Cell Discov*. 2016; 2(1):16009.

135. Lei Y, Zhang X, Su J, et al. Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. *Nat Commun*. 2017; 8(1):16026.

136. Vojta A, Dobrinić P, Tadić V, et al. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res*. 2016; 44(12):5615-5628.

137. Liu XS, Wu H, Ji X, et al. Editing DNA Methylation in the Mammalian Genome. *Cell*. 2016; 167(1):233-247.e17.

138. Kearns NA, Pham H, Tabak B, et al. Functional annotation of native enhancers with a Cas9–histone demethylase fusion. *Nat Methods*. 2015; 12(5):401-403.

139. Hilton IB, D'Ippolito AM, Vockley CM, et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol*. 2015; 33(5):510-517.

140. Kwon DY, Zhao YT, Lamonica JM, Zhou Z. Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC. *Nat Commun*. 2017; 8(1):15315.

141. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA–guided activation of endogenous human genes. *Nat Methods*. 2013; 10(10):977-979.

142. Chavez A, Scheiman J, Vora S, et al. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods*. 2015; 12(4):326-328.

143. Thakore PI, D'Ippolito AM, Song L, et al. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods*. 2015; 12(12):1143-1149.

144. Gilbert LA, Larson MH, Morsut L, et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*. 2013; 154(2):442-451.

145. Morgan SL, Mariano NC, Bermudez A, et al. Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nat Commun*. 2017; 8:15993.

146. Hao N, Shearwin KE, Dodd IB. Programmable DNA looping using engineered bivalent dCas9 complexes. *Nat Commun*. 2017; 8(1):1628.

147. Chen B, Gilbert LA, Cimini BA, et al. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell*. 2013; 155(7):1479-1491.

148. Ma H, Naseri A, Reyes-Gutierrez P, Wolfe SA, Zhang S, Pederson T. Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc Natl Acad Sci*. 2015; 112(10):3002-3007.

149. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016; 533(7603):420-424.

150. Gaudelli NM, Komor AC, Rees HA, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. 2017; 551(7681):464-471.

151. McDonald JI, Celik H, Rois LE, et al. Reprogrammable CRISPR/Cas9-based system for

inducing site-specific DNA methylation. *Biol Open*. 2016; 5(6):866-874.

152. Suetake I, Shinozaki F, Miyagawa J, Takeshima H, Tajima S. DNMT3L Stimulates the DNA Methylation Activity of Dnmt3a and Dnmt3b through a Direct Interaction. *J Biol Chem*. 2004; 279(26):27816-27823.

153. Stepper P, Kungulovski G, Jurkowska RZ, et al. Efficient targeted DNA methylation with chimeric dCas9–Dnmt3a–Dnmt3L methyltransferase. *Nucleic Acids Res*. 2017; 45(4):1703-1713.

154. Choudhury SR, Cui Y, Lubecka K, Stefanska B, Irudayaraj J. CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter. *Oncotarget*. 2016; 7(29):46545-46556.

155. Liu XS, Wu H, Krzisch M, et al. Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. *Cell*. 2018; 172(5):979-992.

156. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc*. 2013; 8(11):2180-2196.

157. Witzgall R, O'Leary E, Leaf A, Onaldi D, Bonventre J V. The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc Natl Acad Sci*. 1994; 91(10):4514-4518.

158. Margolin JF, Friedman JR, Meyer WKH, Vissing H, Thiesen HJ, Rauscher FJ. Kruppel-associated boxes are potent transcriptional repression domains. *Proc Natl Acad Sci*. 1994; 91(10):4509-4513.

159. Groner AC, Meylan S, Ciuffi A, et al. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. Madhani HD, ed. *PLoS Genet*. 2010; 6(3):e1000869.

160. Pengue G, Lania L. Kruppel-associated box-mediated repression of RNA polymerase II promoters is influenced by the arrangement of basal promoter elements. *Proc Natl Acad Sci*. 1996; 93(3):1015-1020.

161. Amabile A, Migliara A, Capasso P, et al. Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell*. 2016; 167(1):219-232.e14.

162. Zalatan JG, Lee ME, Almeida R, et al. Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. *Cell*. 2015; 160(1-2):339-350.

163. Gilbert LA, Horlbeck MA, Adamson B, et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*. 2014; 159(3):647-661.

164. O'Geen H, Ren C, Nicolet CM, et al. dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic Acids Res*. 2017; 45(17):9901-9916.

165. Wysocka J, Herr W. The herpes simplex virus VP16-induced complex: The makings of a regulatory switch. *Trends Biochem Sci*. 2003; 28(6):294-304.

166. Perez-Pinera P, Kocak DD, Vockley CM, et al. RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nat Methods*. 2013; 10(10):973-976.

167. Mali P, Aach J, Stranges PB, et al. Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*. 2013; 31(9):833-838.

168. Cheng AW, Wang H, Yang H, et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res*. 2013; 23(10):1163-1171.

169. Mukherjee SP, Behar M, Birnbaum HA, Hoffmann A, Wright PE, Ghosh G. Analysis of the RelA:CBP/p300 Interaction Reveals Its Involvement in NF-κB-Driven Transcription. Saccani S, ed. *PLoS Biol*. 2013; 11(9):e1001647.

170. Hardwick JM, Tse L, Applegren N, Nicholas J, Veliuona MA. The Epstein-Barr virus R transactivator (Rta) contains a complex, potent activation domain with properties different from those of VP16. *J Virol*. 1992; 66(9):5500-5508.

171. Burkhart BA, Hebbar PB, Trotter KW, Archer TK. Chromatin-dependent E1A Activity Modulates NF-κB RelA-mediated Repression of Glucocorticoid Receptor-dependent Transcription. *J Biol Chem*. 2005; 280(8):6349-6358.

172. O'shea JM, Perkins ND. Regulation of the RelA (p65) transactivation domain. *Biochem Soc Trans*. 2008; 36(4):603-608.

173. Blair WS, Bogerd HP, Madore SJ, Cullen BR. Mutational analysis of the transcription

activation domain of RelA: identification of a highly synergistic minimal acidic activation module. *Mol Cell Biol*. 1994; 14(11):7226-7234.

174. Xu X, Qi LS. A CRISPR–dCas Toolbox for Genetic Engineering and Synthetic Biology. *J Mol Biol*. 2019; 431(1):34-47.

175. Lavender P, Kelly A, Hendy E, McErlean P. CRISPR-based reagents to study the influence of the epigenome on gene expression. *Clin Exp Immunol*. 2018; 194(1):9-16.

176. Limsirichai P, Gaj T, Schaffer D V. CRISPR-mediated Activation of Latent HIV-1 Expression. *Mol Ther*. 2016; 24(3):499-507.

177. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol*. 2013; 31(9):839-843.

178. Fu Y, Foden JA, Khayter C, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*. 2013; 31(9):822-826.

179. Cho SW, Kim S, Kim Y, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*. 2014; 24(1):132-141.

180. Lessard S, Francioli L, Alfoldi J, et al. Human genetic variation alters CRISPR-Cas9 on- and off-targeting specificity at therapeutically implicated loci. *Proc Natl Acad Sci*. 2017; 114(52):E11257-E11266.

181. Hsu PD, Scott DA, Weinstein JA, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013; 31(9):827-832.

182. Tsai SQ, Zheng Z, Nguyen NT, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol*. 2015; 33(2):187-197.

183. Zhang Y, Heidrich N, Ampattu BJ, et al. Processing-Independent CRISPR RNAs Limit Natural Transformation in Neisseria meningitidis. *Mol Cell*. 2013; 50(4):488-503.

184. Ran FA, Hsu PD, Lin C-Y, et al. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell*. 2013; 154(6):1380-1389.

185. Lin Y, Cradick TJ, Brown MT, et al. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res*. 2014; 42(11):7473-7485.

186. Naeem M, Majeed S, Hoque MZ, Ahmad I. Latest Developed Strategies to Minimize the Off-Target Effects in CRISPR-Cas-Mediated Genome Editing. *Cells*. 2020; 9(7):1608.

187. Zhang XH, Tee LY, Wang XG, Huang QS, Yang SH. Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol Ther - Nucleic Acids*. 2015; 4(11):e264.

188. Huaman MA, Fiske CT, Jones TF, et al. Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity. 2015; 143(5):951-959.

189. Chapman JE, Gillum D, Kiani S. Approaches to reduce CRISPR off-target effects for safer genome editing. *Appl Biosaf*. 2017; 22(1):7-13.

190. Cebrian-Serrano A, Davies B. CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm Genome*. 2017; 28(7-8):247-261.

191. O'Geen H, Henry IM, Bhakta MS, Meckler JF, Segal DJ. A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. *Nucleic Acids Res*. 2015; 43(6):3389-3404.

192. Kuscu C, Arslan S, Singh R, Thorpe J, Adli M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol*. 2014; 32(7):677-683.

193. Lin L, Liu Y, Xu F, et al. Genome-wide determination of on-target and off-target characteristics for RNA-guided DNA methylation by dCas9 methyltransferases. *Gigascience*. 2018; 7(3):1-19.

194. Galonska C, Charlton J, Mattei AL, et al. Genome-wide tracking of dCas9-methyltransferase footprints. *Nat Commun*. 2018; 9(1):597.

195. Pflueger C, Tan D, Swain T, et al. A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. *Genome Res*. 2018; 28(8):1193-1206.

196. Josipović G, Tadić V, Klasić M, et al. Antagonistic and synergistic epigenetic modulation using orthologous CRISPR/dCas9-based modular system. *Nucleic Acids Res*. 2019; 47(18):9637-9657.

197. Kiani S, Chavez A, Tuttle M, et al. Cas9 gRNA engineering for genome editing, activation and repression. *Nat Methods*. 2015; 12(11):1051-1054.

198. Tabebordbar M, Zhu K, Cheng JKW, et al. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science*. 2016; 351(6271):407-411.

199. Nelson CE, Hakim CH, Ousterout DG, et al. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science*. 2016; 351(6271):403-407.

200. Ran FA, Cong L, Yan WX, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*. 2015; 520(7546):186-191.

201. Thakore PI, Kwon JB, Nelson CE, et al. RNA-guided transcriptional silencing in vivo with S. aureus CRISPR-Cas9 repressors. *Nat Commun*. 2018; 9(1):1674.

202. Gao Y, Xiong X, Wong S, Charles EJ, Lim WA, Qi LS. Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nat Methods*. 2016; 13(12):1043-1049.

203. Bernstein BE, Stamatoyannopoulos JA, Costello JF, et al. The NIH roadmap epigenomics mapping consortium. *Nat Biotechnol*. 2010; 28(10):1045-1048.

204. Fernández JM, de la Torre V, Richardson D, et al. The BLUEPRINT Data Analysis Portal. *Cell Syst*. 2016; 3(5):491-495.e5.

205. Feingold EA, Good PJ, Guyer MS, et al. The ENCODE (ENCyclopedia of DNA Elements) Project. *Science*. 2004; 306(5696):636-640.

206. Bujold D, Morais DA de L, Gauthier C, et al. The International Human Epigenome Consortium Data Portal. *Cell Syst*. 2016; 3(5):496-499.

207. Chen X, Zaro JL, Shen WC. Fusion protein linkers: Property, design and functionality. *Adv Drug Deliv Rev*. 2013; 65(10):1357-1369.

208. Szymczak AL, Vignali DAA. Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin Biol Ther*. 2005; 5(5):627-638.

209. Dean DA. Import of Plasmid DNA into the Nucleus Is Sequence Specific. *Exp Cell Res*. 1997; 230(2):293-302.

210. Young JL, Benoit JN, Dean DA. Effect of a DNA nuclear targeting sequence on gene

transfer and expression of plasmids in the intact vasculature. *Gene Ther*. 2003; 10(17):1465-1470.

211. Blomberg P, Eskandarpour M, Xia S, Sylvén C, Islam KB. Electroporation in combination with a plasmid vector containing SV40 enhancer elements results in increased and persistent gene expression in mouse muscle. *Biochem Biophys Res Commun*. 2002; 298(4):505-510.

212. Li S, MacLaughlin FC, Fewell JG, et al. Muscle-specific enhancement of gene expression by incorporation of SV40 enhancer in the expression plasmid. *Gene Ther*. 2001; 8(6):494-497.

213. Josipović G, Zoldoš V, Vojta A. Active fusions of Cas9 orthologs. *J Biotechnol*. 2019; 301:18-23.

214. Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013; 339(6121):823-826.

215. Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*. 2014; 513(7519):569-573.

216. Chakraborty S, Ji H, Kabadi AM, Gersbach CA, Christoforou N, Leong KW. A CRISPR/Cas9-Based System for Reprogramming Cell Lineage Specification. *Stem Cell Reports*. 2014; 3(6):940-947.

217. Gisler S, Gonçalves JP, Akhtar W, et al. Multiplexed Cas9 targeting reveals genomic location effects and gRNA-based staggered breaks influencing mutation efficiency. *Nat Commun*. 2019; 10(1):1598.

218. Xu H, Xiao T, Chen CH, et al. Sequence determinants of improved CRISPR sgRNA design. *Genome Res*. 2015; 25(8):1147-1157.

219. Mandal PK, Ferreira LMR, Collins R, et al. Efficient Ablation of Genes in Human Hematopoietic Stem and Effector Cells using CRISPR/Cas9. *Cell Stem Cell*. 2014; 15(5):643-652.

220. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science*. 2014; 343(6166):80-84.

221. Moreno-Mateos MA, Vejnar CE, Beaudoin JD, et al. CRISPRscan: designing highly

efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods*. 2015; 12(10):982-988.

222. Doench JG, Hartenian E, Graham DB, et al. Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. *Nat Biotechnol*. 2014; 32(12):1262-1267.

223. Chen X, Rinsma M, Janssen JM, Liu J, Maggio I, Gonçalves MAFV. Probing the impact of chromatin conformation on genome editing tools. *Nucleic Acids Res*. 2016; 44(13):6482-6492.

224. Horlbeck MA, Witkowsky LB, Guglielmi B, et al. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife*. 2016; 5:e12677.

225. Hanzawa N, Hashimoto K, Yuan X, et al. Targeted DNA demethylation of the Fgf21 promoter by CRISPR/dCas9-mediated epigenome editing. *Sci Rep*. 2020; 10(1):5181.

226. Balboa D, Weltner J, Eurola S, Trokovic R, Wartiovaara K, Otonkoski T. Conditionally Stabilized dCas9 Activator for Controlling Gene Expression in Human Cell Reprogramming and Differentiation. *Stem Cell Reports*. 2015; 5(3):448-459.

227. Carleton JB, Berrett KC, Gertz J. Multiplex Enhancer Interference Reveals Collaborative Control of Gene Regulation by Estrogen Receptor α-Bound Enhancers. *Cell Syst*. 2017; 5(4):333-344.e5.

228. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell*. 2014; 159(3):635-646.

229. Konermann S, Brigham MD, Trevino AE, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 2015; 517(7536):583-588.

230. Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods*. 2013; 10(11):1116-1121.

231. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*. 2009; 155(3):733-740.

232. Nishimasu H, Cong L, Yan WX, et al. Crystal Structure of Staphylococcus aureus Cas9.

Cell. 2015; 162(5):1113-1126.

233. Boettcher M, Tian R, Blau JA, et al. Dual gene activation and knockout screen reveals directional dependencies in genetic networks. *Nat Biotechnol*. 2018; 36(2):170-178.

234. Lauc G, Huffman JE, Pučić M, et al. Loci Associated with N-Glycosylation of Human Immunoglobulin G Show Pleiotropy with Autoimmune Diseases and Haematological Cancers. Gibson G, ed. *PLoS Genet*. 2013; 9(1):e1003225.

235. Igarashi K, Ochiai K, Itoh-Nakadai A, Muto A. Orchestration of plasma cell differentiation by Bach2 and its gene regulatory network. *Immunol Rev*. 2014; 261(1):116-125.

236. Lauc G, Essafi A, Huffman JE, et al. Genomics Meets Glycomics—The First GWAS Study of Human N-Glycome Identifies HNF1α as a Master Regulator of Plasma Protein Fucosylation. Gibson G, ed. *PLoS Genet*. 2010; 6(12):e1001256.

237. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012; 13(7):484-492.

238. Kungulovski G, Nunna S, Thomas M, Zanger UM, Reinhardt R, Jeltsch A. Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. *Epigenetics Chromatin*. 2015; 8(1):12.

239. Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-γuided platform for sequence-specific control of gene expression. *Cell*. 2013; 152(5):1173-1183.

240. Zoldoš V, Horvat T, Novokmet M, et al. Epigenetic silencing of HNF1A associates with changes in the composition of the human plasma N -glycome. *Epigenetics*. 2012; 7(2):164-172.

241. Maeder ML, Angstman JF, Richardson ME, et al. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat Biotechnol*. 2013; 31(12):1137-1142.

242. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci*. 2002; 99(6):3740-3745.

243. Larsen F, Gundersen G, Lopez R, Prydz H. CpG islands as gene markers in the human genome. *Genomics*. 1992; 13(4):1095-1107.

244. Miller JL, Grant PA. The role of DNA methylation and histone modifications in transcriptional regulation in humans. *Subcell Biochem*. 2013; 61:289-317.

245. Yang X, Ferguson AT, Nass SJ, et al. Transcriptional activation of estrogen receptor alpha in human breast cancer cells by histone deacetylase inhibition. *Cancer Res*. 2000; 60(24):6890-6894.

246. Bovenzi V, Momparler R. Antineoplastic action of 5-aza-2′-deoxycytidine and histone deacetylase inhibitor and their effect on the expression of retinoic acid receptor β and estrogen receptor α genes in breast carcinoma cells. *Cancer Chemother Pharmacol*. 2001; 48(1):71-76.

247. Pruitt K, Zinn RL, Ohm JE, et al. Inhibition of SIRT1 Reactivates Silenced Cancer Genes without Loss of Promoter DNA Hypermethylation. Reik W, ed. *PLoS Genet*. 2006; 2(3):e40.

248. Cano-Rodriguez D, Gjaltema RAF, Jilderda LJ, et al. Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. *Nat Commun*. 2016; 7(1):12284.

249. Kim J-M, Kim K, Schmidt T, et al. Cooperation between SMYD3 and PC4 drives a distinct transcriptional program in cancer cells. *Nucleic Acids Res*. 2015; 43(18):8868-8883.

250. Kohler RS, Anugraham M, López MN, et al. Epigenetic activation of MGAT3 and corresponding bisecting GlcNAc shortens the survival of cancer patients. *Oncotarget*. 2016; 7(32):51674-51686.

251. Halmai JANM, Deng P, Gonzalez CE, et al. Artificial escape from XCI by DNA methylation editing of the CDKL5 gene. *Nucleic Acids Res*. 2020; 48(5):2372-2387.

252. Verkuijl SA, Rots MG. The influence of eukaryotic chromatin state on CRISPR–Cas9 editing efficiencies. *Curr Opin Biotechnol*. 2019; 55:68-73.

253. Baumann V, Wiesbeck M, Breunig CT, et al. Targeted removal of epigenetic barriers during transcriptional reprogramming. *Nat Commun*. 2019; 10(1):2119.

254. Morita S, Horii T, Kimura M, Hatada I. Synergistic Upregulation of Target Genes by TET1 and VP64 in the dCas9–SunTag Platform. *Int J Mol Sci*. 2020; 21(5):1574.

255. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res*. 2013; 41(20):9584-9592.

256. Haysom G. Target specificity of the CRISPR-Cas9 system. *Work Pap 33*. 2016; 2:24.

257. Jinek M, Jiang F, Taylor DW, et al. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science*. 2014; 343(6176):1247997-1247997.

258. Nishimasu H, Ran FA, Hsu PD, et al. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell*. 2014; 156(5):935-949.

259. Varriale A, Bernardi G. Distribution of DNA methylation, CpGs, and CpG islands in human isochores. *Genomics*. 2010; 95(1):25-28.

260. Mugal CF, Ellegren H. Substitution rate variation at human CpG sites correlates with non-CpG divergence, methylation level and GC content. *Genome Biol*. 2011; 12(6):R58.

261. Chuang TJ, Chen FC, Chen YZ. Position-dependent correlations between DNA methylation and the evolutionary rates of mammalian coding exons. *Proc Natl Acad Sci*. 2012; 109(39):15841-15846.

262. Trbojević Akmačić I, Ventham NT, Theodoratou E, et al. Inflammatory bowel disease associates with proinflammatory potential of the immunoglobulin G glycome. *Inflamm Bowel Dis*. 2015; 21(6):1237-1247.

263. Theodoratou E, Thaçi K, Agakov F, et al. Glycosylation of plasma IgG in colorectal cancer prognosis. *Sci Rep*. 2016; 6(1):28098.

264. Tanaka T, Yoneyama T, Noro D, et al. Aberrant N-Glycosylation Profile of Serum Immunoglobulins is a Diagnostic Biomarker of Urothelial Carcinomas. *Int J Mol Sci*. 2017; 18(12):2632.

265. Zhang D, Chen B, Wang Y, et al. Disease-specific IgG Fc N-glycosylation as personalized biomarkers to differentiate gastric cancer from benign gastric diseases. *Sci Rep*. 2016; 6(1):25957.

266. Jennewein MF, Alter G. The Immunoregulatory Roles of Antibody Glycosylation. *Trends Immunol*. 2017; 38(5):358-372.

267. Goulabchand R, Vincent T, Batteux F, Eliaou J, Guilpain P. Impact of autoantibody

glycosylation in autoimmune diseases. *Autoimmun Rev*. 2014; 13(7):742-750.

6 ŽIVOTOPIS

Goran Josipović rođen je 23. prosinca 1992. godine u Zagrebu. Završio je diplomski studij molekularne biologije na Prirodoslovno-matematičkom fakultetu Sveučilišta u Zagrebu 2016. godine. Diplomski rad „Ciljana metilacija DNA promotora gena *MGAT3* korištenjem alata CRISPR-Cas9“ napravio je pod voditeljstvom izv. prof. dr. sc. Aleksandra Vojta i neposrednim voditeljstvom dr. sc. Paule Dobrinić na Zavodu za molekularnu biologiju Biološkog odsjeka PMF-a. Od siječnja 2017. godine zaposlen je kao suradnik na projektu Horizon 2020 „Systems medicine approach to chronic inflammatory disease“ (SYSCID) na istom zavodu, u grupi za epigenetiku prof. dr. sc. Vlatke Zoldoš. Sudjelovao je u različitim projektima grupe za epigenetiku – IRI „Nova generacija visokoprotočnih gliko-servisa“, ICGEB „Comprehensive toolbox for epigenetic modulation of gene expression“ i HRZZ „Epigenetička regulacija glikozilacije imunoglobulina G“ pod vodstvom prof. dr. sc. Vlatke Zoldoš.

Doktorski studij biologije upisao je 2017. godine na Biološkom odsjeka PMF-a. Koautor je na dva izvorna znanstvena rada i na jednom preglednom radu objavljenima u časopisima s međunarodnom recenzijom, a sudjelovao je i na međunarodnim znanstvenim skupovima s posterskim priopćenjima. Dobitnik je nagrade Hrvatskog društva za biokemiju i molekularnu biologiju (HDBMB) i nagrade dekana Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu.

Od akademske godine 2016./2017. sudjeluje u izvođenju nastave na Zavodu za molekularnu biologiju Biološkog odsjeka PMF-a u sklopu kolegija Genetika i Laboratorijska stručna praksa.