Supplementary Information for

Metformin selectively inhibits metastatic colorectal cancer with *KRAS* mutation by intracellular accumulation through silencing MATE1

Jinye Xie[1,2#], Liangping Xia[3#], Wei Xiang[1,2#], Wenzhuo He[3#], HaoFan Yin[2], Fang Wang[4], Tianxiao Gao[3], Weiwei Qi[1,2], Zhonghan Yang[1,2], Xia Yang[1,2,5*], Ti Zhou[1,2*], Guoquan Gao[1,2,6*]

[1] Program of Molecular Medicine, Affiliated Guangzhou Women and Children's Hospital, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou, China
[2] Department of Biochemistry, Zhongshan School of Medicine, SunYat-sen University, Guangzhou, China
[3] Sun Yat-sen University Cancer Center; State Key Laboratory of Oncology in South China; Collaborative Innovation Center for Cancer Medicine, Guangzhou, China
[4] Department of Molecular Diagnostics, Sun Yat-sen University Cancer Center, Guangzhou, China
[5] Guangdong Provincial Key Laboratory of Brain Function and Disease, Sun Yat-sen University, Guangzhou, China
[6] Guangdong Engineering & Technology Research Center for Gene Manipulation and Biomacromolecular Products, Sun Yat-sen University, Guangzhou, China
# Jinye Xie, Liangping Xia, Wei Xiang, and Wenzhuo He contributed equally to this study
Correspondence to:
Guoquan Gao, Email: gaogq@mail.sysu.edu.cn
Ti Zhou, Email: zhouti2@mail.sysu.edu.cn
Xia Yang, E-mail: yangxia@mail.sysu.edu.cn

This PDF file includes:

MATERIALS AND METHODS

Medical records and tissue collection

4751 patients were diagnosed as mCRC at Sun Yat-sen University Cancer Center (SYUCC) between 2004 and 2016, 2378 patients with complete past medical history were enrolled. Among them, 325 patients were diagnosed as mCRC with type 2 mellitus diabetes (T2DM). Firstly, we excluded 7 patients who were lost to follow-up and 36 patients whose anti-diabetic drug usage information was missing. Then, according to the hypoglycemic mechanism, 282 patients were divided into insulin or insulin secretagogues group (n=119), who took insulin or/and sulfonylureas or glinides; non-insulin secretagogues group (n=22), using α-glucosidase inhibitors or thiazolidinediones; metformin only group (n=50), who had about 750-1000mg metformin once a day; metformin and other hypoglycemic agents group (n=59), who took about 500mg metformin combined with insulin, glinides, or α-glucosidase inhibitors per day; and those without anti-diabetic therapy group (n=32), who did not control the blood glucose level before diagnosed as mCRC (*SI Appendix*, Fig.S1). There was no difference in blood glucose between those antidiabetic medication groups. The distribution of gender, age, BMI, primary tumor site, pathological grading, metastatic site, and *KRAS* exon 2 genotyping were shown in *SI Appendix*, Table. S1. Specimens were obtained from patients who had undergone CRC resection and with tumor grade categorized as well-differentiated, moderately differentiated, poorly or undifferentiated by three pathologists. All patients' informed consent has been obtained before surgery, and the use of medical records and histological sections has also been approved by the ethics committee in SYUCC.

Establishment of xenograft models and animal studies

For the patient-derived xenograft model (PDX), colorectal cancer specimens were obtained from patients undergoing resections of primary disease at SYUCC. All patients' informed consent has been obtained before surgery, and the samples used for PDX establishment were procured with the approval of the ethics committee in SYUCC. The fresh specimens were transferred in DMEM/F12 media (Corning, R10-092-CV) supplemented with 10% fetal bovine serum (FBS, Gibco, #10100-147) and 1% Penicillin-Streptomycin (Gibco, #15140-122) within 4 hours after surgery. The PDX model mice were established and passaged by IDMO Co., Ltd, (Beijing, China, approval No.111023500000091). Briefly, a portion of specimens was excised for genotyping, immunohistochemistry, and DNA methylation detection. The rest portions were divided into fragments <3mm^{3}for subcutaneous implantation into male 6- to 8-week-old NPI (NOD-Prkdc^{scid}-Il2rg^{em1IDMO}) mice as described. Tumor-implanted mice were monitored for xenograft growth for up to 300-500 mm^3 following implantation. Then the xenografts from the NPI mice were passaged into the next-generation mice. The patient-derived xenografts were maintained in NPI mice for no more than 4 passages.

In our animal studies, two PDX models (*SI Appendix*, Table. S6) were selected based on *KRAS* genotypes to explore the respective *in vivo* consequences of metformin used in *KRAS*-wildtype and *KRAS*-mutation groups. Briefly, these two groups PDX xenografts of 2-passage were cut into several 2-3 mm^3 pieces; then, the tumor fragments were implanted subcutaneously into twelve BALB/c nude mice per group. The male 4- to 5-week-old and 16- to 18-gram-weight BALB/c nude (CAnN.Cg-Foxn1nu/Crl) mice were purchased from Vital River (Beijing, China, approval No.11400700310365). When the tumor volumes reached 30-50 mm^3, mice with *KRAS*^{WT} and *KRAS*^{G12D} xenograft were randomized into 4

groups: *KRAS*^{WT} with vehicle or metformin treatment groups, *KRAS*^{G12D} with vehicle or metformin treatment groups.

For cell line-derived xenograft model (CDX) experiments, we first established the *KRAS*^{G13D} SW48 cell line by CRISPR/Cas9; and sh-MATE1-SW48 cell line, which knocks down MATE1 in SW48 using shRNA lentiviral transduction; LV-MATE1- *KRAS*^{G13D} SW48 cell line which overexpresses MATE1 in *KRAS*^{G13D} SW48 by lentivirus infection. Then 1×10^6 cells suspended in Matrigel [100μL high concentration Matrigel (Corning, #354248) mixed with 100μL PBS] were injected subcutaneously into BALB/c nude mice. The CDX mice were randomly divided into a metformin treatment group and its control group for each cell line model.

According to the Reagan-Shaw method for dose translation [human equivalent dose = animal dose (mg/kg) ×(3/37)] and 4-7 mL daily water intake per mice, the mice were treated with metformin dissolved in drinking water (1mg/mL, approximately 200mg/kg body weight for mice and 1000mg/60kg bodyweight for human) for 30 days. The drinking water with the drug was changed every 3 days. Subcutaneous tumor volume [(length × width^3)/2] and body weights were also measured every 3 days. All procedures are related to animal feeding, treatment, and welfare were conducted in accordance with the Institutional Animal Care and Use Committee of Sun Yat-sen University.

Immunohistochemistry

Tissues from patients and PDX models were fixed with 4% paraformaldehyde overnight, embedded in paraffin and cut into 3μm sections. Haematoxylin and eosin (H&E) staining and immunohistochemical staining were conducted following standard protocol. Briefly, the sections were deparaffinized, rehydrated in a gradient of ethanol and pretreated with 0.01M citrate buffer (pH 6.0) through the high-pressure method. Then the sections were immersed in 3% H_2O_2 for 20 minutes to quench endogenous peroxidase, following with goat serum to block non-specific background staining site. Next, the following primary antibodies were used: Ki67 (AB9260) from Merck Millipore, MATE1 (ab104016), DNMT1 (ab19905), TET1 (ab191698) and TET2 (ab94580) from Abcam. After incubating with primary antibodies at 4℃ overnight, the sections were washed 3 times in PBS with Tween-20 (PBST, pH7.2), following incubation with HRP-conjugated secondary antibody for 40 minutes at room temperature and then stained with DAB+ chromogenic substrates. Sections were photographed through a slide scanner (Axio Scan. Z1, ZEISS) and the integral optical density (IOD), the area of cancer cells or the number of nuclear with positive stained were calculated by ImageJ software.

Construction of plasmids and lentivirus

For sgRNA cloning, the CRISPR/Cas9 vector PX459 (Addgene, #62988) was digested with BbsI (ThermoFisher, FD1014) and ligated with BbsI compatible annealed oligos. The sgRNA targeting the upstream sequence of *KRAS* exon2 (5'-GCATTTTTCTTAAGCGTCGATGG-3') was designed using Optimized CRISPR Design. The homologous fragments using for introducing the point mutation were amplified by overlapping PCR and assembled into T vector pGM-T (TianGen, VT202) using In-Fusion technology (Clontech, #639636). Next, the donor vector using for homology-directed repair (HDR) was generated with the homologous fragment (amplified by overlapping PCR with attB1 at the ends) and pDONR 211 (Invitrogen, #12536017) by BP reaction. The used primers mention above were listed in *SI Appendix*, Table. S7. For shRNA lentivirus production, the shRNAs were designed to target *KRAS* sequence (5'-

ATGGTCCTAGTAGGAAATAAA-3' and 5'-TTGATGATGCCTTCTATACAT-3') and then the shRNAs DNA fragments were cloned into GV248 (hU6-MCS-Ubiquitin-EGFP-IRES-puromycin) and packaged as lentivirus by GeneChem Co., Ltd, (Shanghai, China).

Cell culture and gene edit

The human CRC cell lines (SW48, CaCO2, LoVo, HCT-116, HT-29) were obtained from the American Type Culture Collection (Manassas, VA, USA). Cell lines were authenticated by Cellcook Biotech Co., Ltd, (Guangzhou, China). *KRAS*^{G13D} SW48 was established by an improved CRISPR/Cas9-mediated precise genetic modification by using 1μM non-homologous end joining (NHEJ) inhibitor Scr7 (Selleck, S7742)[1]. For shRNA experiments, cells were infected with shRNA lentivirus at 20 MOI followed by drug selection (1μg/mL puromycin for 1 week). For RNAi experiments, SW48 were transfected with HiPerFect reagent (QIAGEN, #301705) using siRNA molecules (RayBio, Guangzhou, China) at a final concentration of 20nM. The siRNA sequences used for MATE1 RNAi experiments are as follows: si-MATE1-1 sense, 5'-GUCUUUCAAGCUGUGUGUUdTdT-3' and anti-sense, 3'-dTdTCAGAAAGUUCGACACACAA-5'; si-MATE1-3 sense, 5'-GCUUCAUAAGCUCCGUGUUdTdT-3' and anti-sense, 3'-dTdTCGAAGUAUUCGAGGCACAA-5'. All cells were cultured in the recommended medium (Gibco) and supplemented with 10% FBS and 1% Penicillin-Streptomycin at 37℃ in a humidified incubator containing 5% CO_2.

Cell viability assay

The viability of CRC cells was determined through Cell Counting Kit-8 (CCK-8) assay (Dojindo, CK04), following the manufacturer's instructions. Briefly, 8,000 CRC cells were seeded in 100μL of growth medium per well in 96-well plates for 12 hours. Then cells were treated with 2.5, 5, 10, 20mM of metformin (Sigma, D150959), when necessary with or without Lansoprazole (Selleck, S1354) or Azacitidine (Selleck, S1782) for 48 hours (the total volume was up to 200μL). Cell viability was measured by adding 20μL of CCK-8 to each well. After 2 hours of incubation at 37℃ in a humidified incubator containing 5% CO_2, the OD value was determined by absorbance at 450 nm using the Sunrise microplate reader (TECAN, Mäannedorf, Switzerland).

Clonogenic assay

Cells were seeded at 5,000 per well in 6-well plates and allowed to adhere overnight, then cultured in the absence or presence of metformin as indicated above in complete media for 14 days. The media with or without drug was replaced every 2 days. In the end, cells were fixed with 4% formaldehyde, stained with 0.1% crystal violet, and photographed using GS-800 Calibrated Imaging Densitometer (Bio-Rad). The colonies were counted by ImageJ software.

Cell cycle analysis

CRC cells were cultured in 6-well plates and treated with the indicated treatments for 24 hours. Cells were harvested and fixed with 70% ethanol overnight at -20℃. Then the cells were washed twice to remove the ethanol and incubated with PI/RNase Staining Buffer (BD Biosciences, #558662) for 15 minutes at room temperature. The cells were subjected to flow cytometry analysis using CytoFLEX S (Beckman Coulter), and the cell cycle was analyzed by FlowJo software.

qRT-PCR

Total RNA was extracted using TRIZOL reagent (Invitrogen, #15596026), and

complementary DNA was obtained using the PrimeScript reverse transcription reagent (TaKaRa, RR036A) following the manufacturer's instructions. Real-time PCR was performed using SYBR Green PCR Mix (TakaRa, RR420A) on the CFX96 Touch PCR system (BioRad). β-actin was used as an internal normalization control. The normalized fold change of gene mRNA levels was calculated using the $2^{-\Delta\Delta Ct}$. The PCR primer sequences are listed in *SI Appendix*, Table. S8.

Western Blotting

The total protein of tumors sample and cultured cells were extracted using SDS lysis buffer (Beyotime, P0013G) supplemented with 1mM phenylmethanesulfonyl fluoride (Beyotime, ST506), and protein concentration were determined by Bicinchoninic Acid (BCA) method according to the manufacturer's protocol (KeyGen, KGP902). Proteins were separated by SDS-PAGE and transferred to 0.45μm polyvinyl difluoride (PVDF) membranes (Millipore) according to standard immunoblotting protocols. After blocking with 5% of defatted milk in TBST (20 mM of Tris-HCl pH 7.4, 500 mM of NaCl and 0.1% of Tween-20) for 1 h at room temperature, the membranes were incubated with the following primary antibodies: p-AMPKα (Thr172) (#4188),AMPK (#2532), p-ERK1/2 (Thr202/Tyr204) (#9101), ERK1/2 (#9102), p-AKT (T308) (#13038), AKT (#4691), p-mTOR (Ser2448) (#5536), mTOR (#2983), p-4E-BP1 (Thr37/46) (#236B4), 4E-BP1 (#9452) from Cell Signaling Technology; p-RB (Ser807/811) (AF1135), RB (AF1564), Cyclin D1 (AF1183), CDK4 (AF2515), CDK6 (AF2536) from Beyotime Co., Ltd.; MATE1 (ab104016), DNMT1 (ab19905) from Abcam. After incubation at 4℃ overnight, membranes were probed with HRP-conjugated anti-rabbit IgG (Cell Signaling Tech, #7074) or anti-mouse IgG (Sigma-Aldrich, AP308P), then developed by ECL substrate (Merck Millipore) and visualized using the Bio-Rad ChemiDoc Touch Imaging System.

Detection of metformin concentration

The concentration of metformin in plasma, intracellular of cell lines and xenografts were determined by liquid chromatography-mass spectrometry (LC-MS). Briefly, for the detection of metformin in mice plasma, 10 μL of plasma per mice was mixed with 190 μL MeCN. Before quantifying the metformin levels in xenografts, each tumor was cut into 10-50 mg and lysed under ultrasonication with 200 μL water, then mixed with 800 μL MeCN; while cultured cells in a 6 cm^2 dish were harvested with 100-200 μL water, following with ultrasonication and mixed with 300-600 μL MeCN. These mixtures were centrifuged at 14,000 rpm for 6.5 minutes. The supernatant was collected and analyzed using LC-MS by Servicebio Co., Ltd. (Wuhan, China) under the method as described[2].

Bisulfite sequencing PCR

Tumors or cells DNA was extracted using a Genomic DNA Extraction Kit (TaKaRa, T9765), and then modified by a DNA Bisulfite Conversation Kit (TIANGEN, DP215). The bisulfite-treated DNA was amplified using an EpiTaq HS PCR assay (TaKaRa, R110A) according to the manufacturer's protocol. Five independent clones were sequenced from each DNA sample, and the amount of 5'-mC was analyzed by a tool from the QUMA website. The used primers sequences were listed in *SI Appendix*, Table. S7.

Statistical analysis

For the clinical retrospective study, proportional hazards (PH) assumptions were examined using the Kolmogorov-Smirnov test and the Cramer von Mises test, and $P>0.05$ represents the interaction between metformin use and event (death or progression) is not time-dependent. Moreover, then a hierarchical proportional hazards regression analysis was

performed to estimate the adjusted hazard ratio (HR) and 95% confidence interval (CI) for the association of metformin use with overall survival (OS) and progression-free survival (PFS), after stratifying by each characteristic. The HR, 95% CI, and a p-value of interaction between metformin use and other characteristics were calculated using EmpowerStats based on R programme. The survival of OS and PFS was measured using the Kaplan-Meier method. Significance was set at $P<0.05$.

The other experimental data were expressed as mean±SEM, while the statistical significance was determined by one-way ANOVA, unpaired two-tailed Student's t-test, or Pearson's Chi-square test using GraphPad Prism 7 or IBM SPSS Statistics 21 software, according to the data type. P value < 0.05 indicated a statistically significant difference.

Funding This study was supported by the National Nature Science Foundation of China, grant numbers: 81471033, 81572342, 81600641, 81770808, 81370945, 81570871, and 81570764, 81872165,81272641 and 81572409; the National Key SciTech Special Project of China, grant numbers: 2013ZX09102-053 and 2015GKS-355; the Key Project of Nature Science Foundation of Guangdong Province, China, grant numbers: 2015A030311043 and 2016A030311035; the Guangdong Natural Science Fund, grant numbers: 2014A020212023, 2014A030313073, 2015A030313029, 2015A030313103, and 2015A030313010; the Guangdong Science Technology Project, grant numbers: 2017A020215075 and 2015B090903063; Initiate Research Funds for the Central Universities of China (Youth Program), grant numbers: 14ykpy05, and 16ykpy24; the Key Sci-tech Research Project of Guangzhou Municipality, China, grant numbers: 201508020033, 201707010084, 201803010017, and 201807010069; the Pearl River Nova Program of Guangzhou Municipality, China, grant number: 201610010186; and the 2017 Milstein Medical Asian American Partnership Foundation Research Project Award in Translational Medicine; Science and Technology Program of Guangzhou, China (1563000305).

Supplementary Figure 1

A

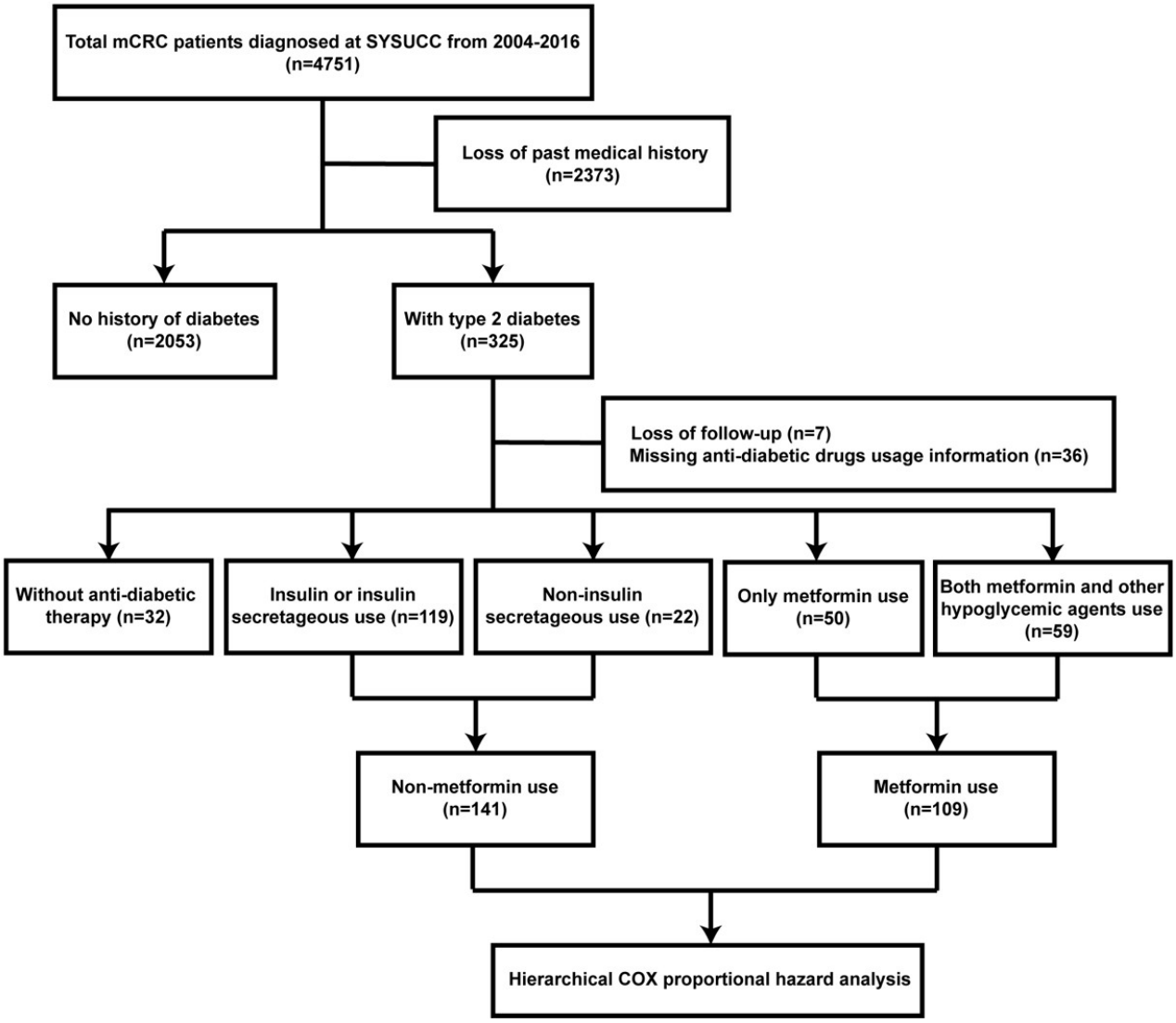

Fig. 1. Flowchart for patient inclusion and exclusion in the clinical retrospective study. mCRC, metastatic colorectal cancer. The final distribution is shown in Supplementary Table 1.

Supplementary Figure 2

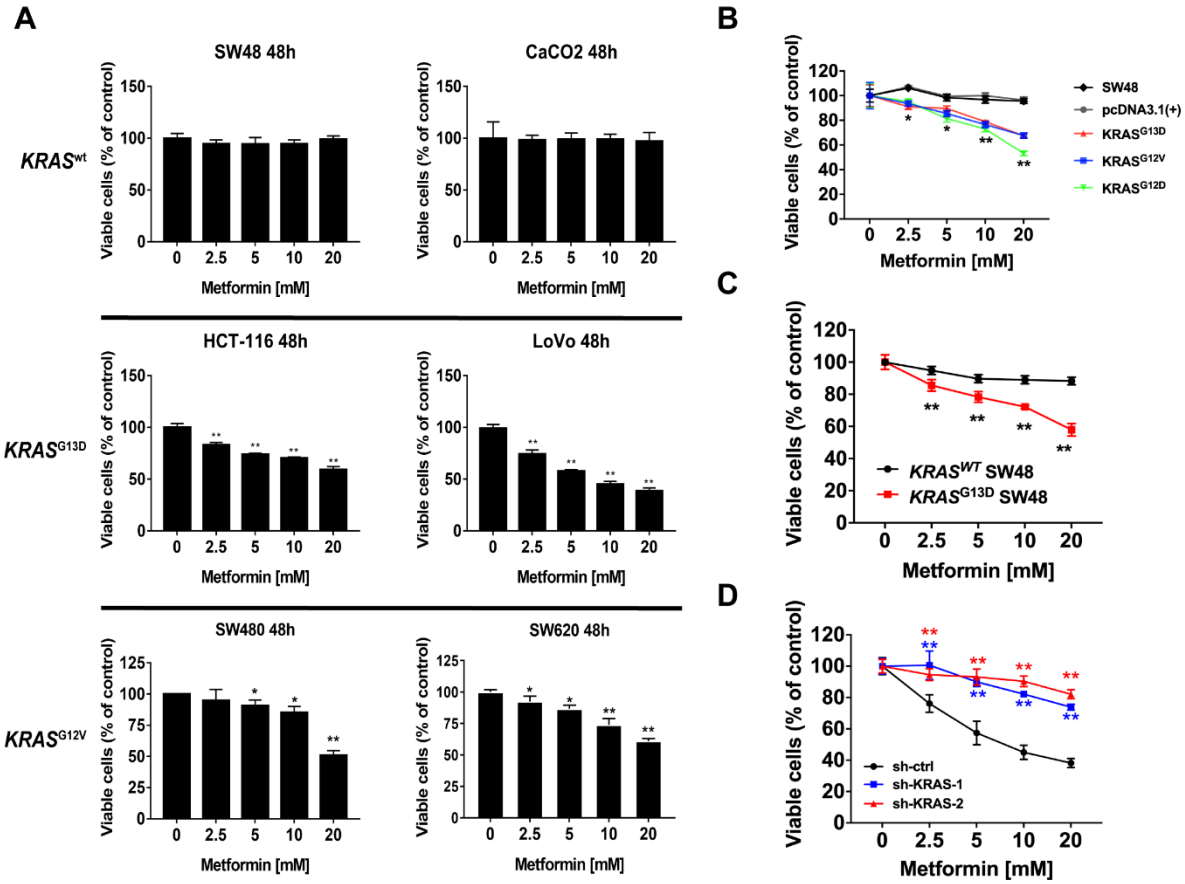

Fig. 2. KRAS mutation enhances the anti-tumor activity of metformin in the CRC cells. A, 48-hour cell viability of *KRAS*^{WT} CRC cell lines SW48 and CaCO2, *KRAS*^{G13D} CRC cell lines HCT-116 and LoVo, *KRAS*^{G12V} CRC cell lines SW480 and SW620 were detected by CCK-8 after treatment with 0, 2.5, 5, 10, 20 mM metformin. Data were shown as mean±SEM (n=5). **P*<0.05, ***P*<0.01 was compared with 0 mM metformin. **B**, *KRAS*^{WT} CRC cell line SW48 was transducted with *KRAS*^{G12D}, *KRAS*^{G12V}, and *KRAS*^{G13D} recombinant plasmid, and 48-hour cell viability was determined after treatment with metformin. Data were shown as mean±SEM (n=5). **P*<0.05, ***P*<0.01 was compared with SW48-pcDNA3.1(+) group. **C**, *KRAS*^{G13D} SW48 established by Crispr/Cas9 system and 48-hour cell viability were determined after treatment with metformin. Data were shown as mean±SEM (n=5). ***P*<0.01 was compared with the *KRAS*WT SW48 group. **D,** LoVo was infected by shRNA lentivirus and treated with incremental metformin. Data were shown as mean±SEM (n=5). ***P*<0.01 was compared with LoVo sh-ctrl group. All P value was determined by two-way ANOVA.

Supplementary Figure 3

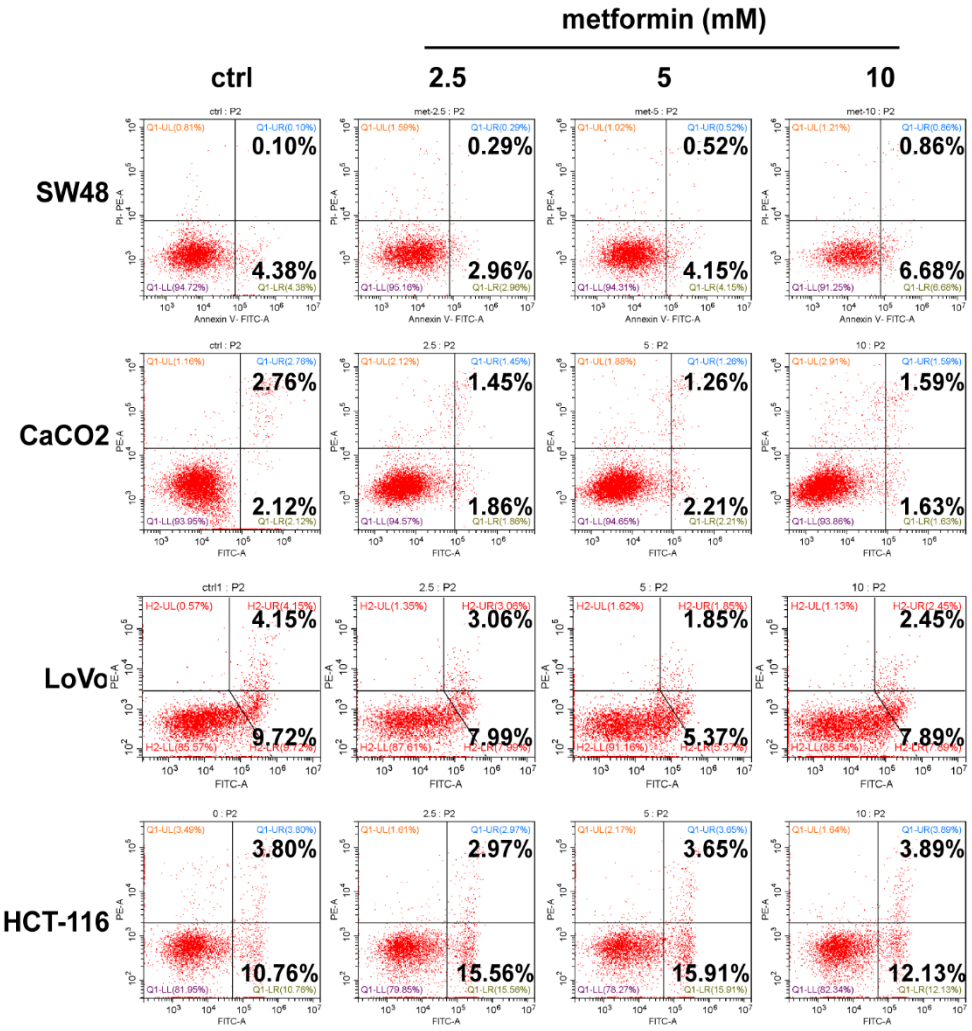

Fig. 3 . The effect of metformin on the apoptosis of colon cancer cells was detected using Annexin V/PI double staining. Results showed that 2.5mM, 5mM, 10mM metformin did not promote apoptosis of KRAS wild-type colon cancer SW48,CaCO2 and KRAS mutant colon cancer cell LoVo, HCT116.

Supplementary Figure 4

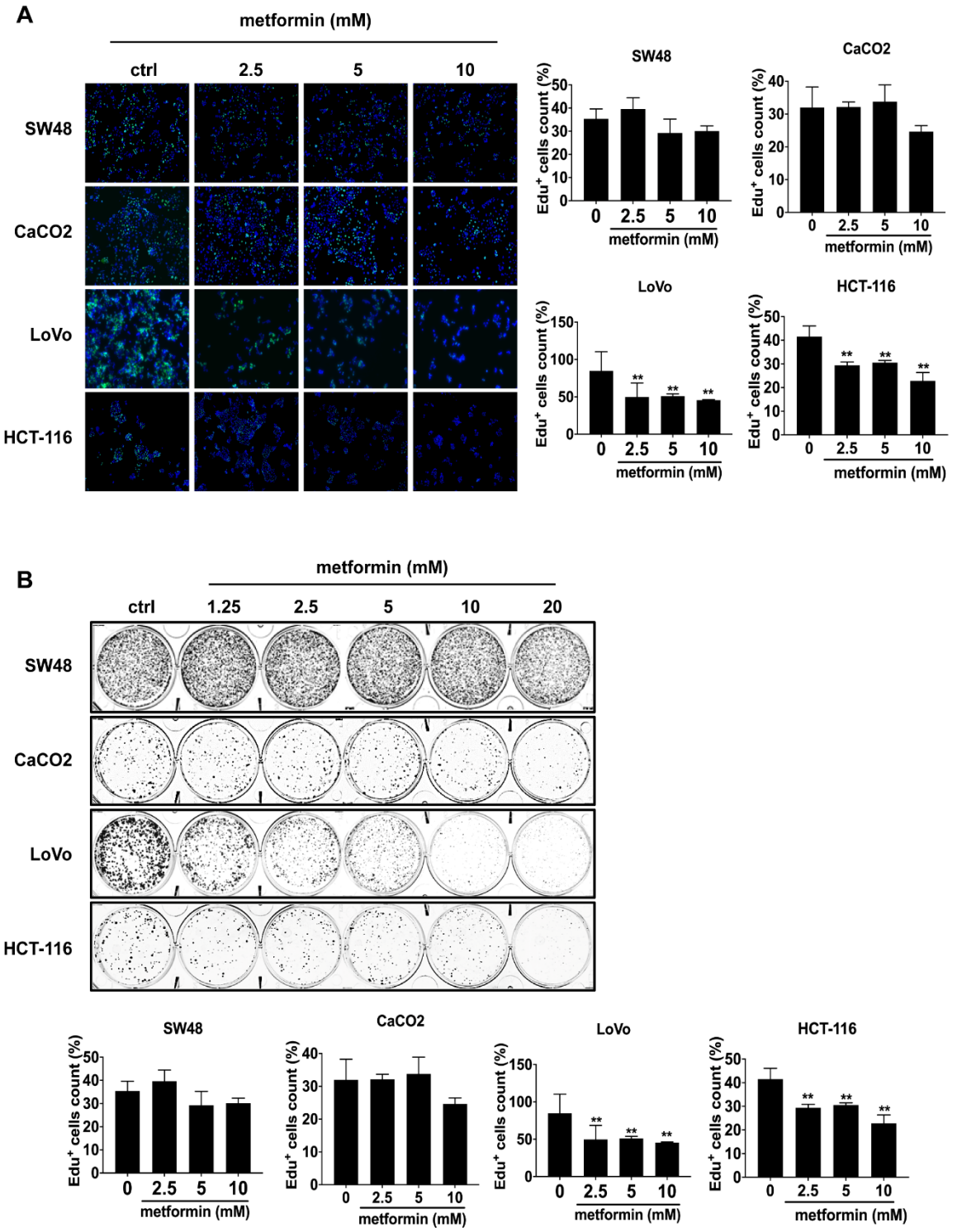

Fig. 4. Metformin inhibits the growth of *KRAS*-mutation CRC cells in a dose-dependent manner. **A,** Representative images of proliferative cells labeled with green EdU staining, and all nucleus with blue DAPI staining were detected after treatment with 0, 1.25, 2.5, 5, 10, 20 mM metformin for 24 hrs, and the proportion of EdU^+ cells were shown (n=3). **B,** Representative images of colony formation of SW48, CaCO2 and HCT-116 and LoVo was photographed after treatment with 0, 1.25, 2.5, 5, 10, 20 mM metformin for 14 days, and the numbers of colonies were calculated relative to the respective control group (right; n=3). All P value was determined by two-way ANOVA.

Supplementary Figure 5

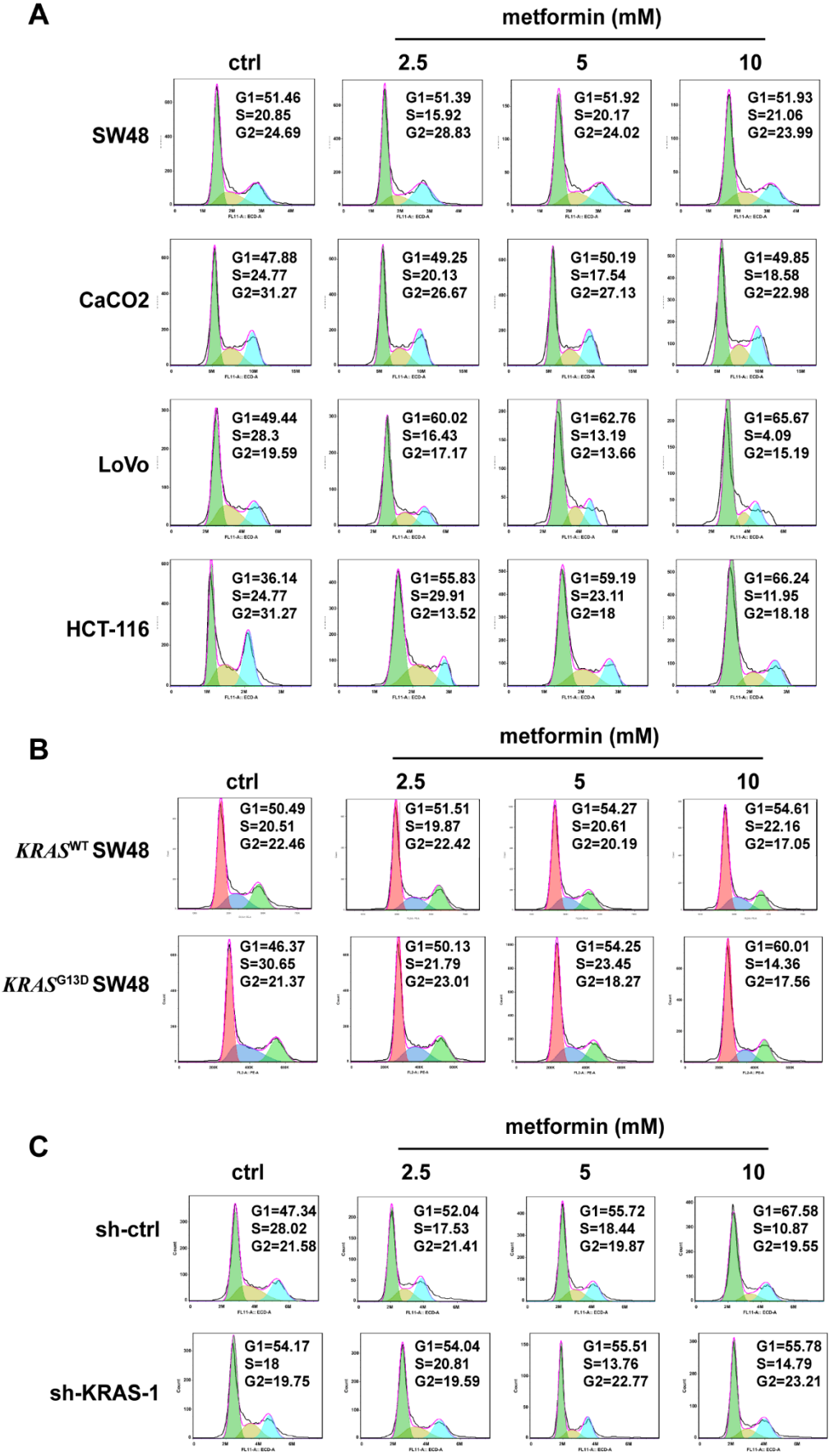

Fig. 5. Metformin inhibits the growth of KRAS-mutation CRC cells through a G0/G1 arrest in dose-dependent manners. Representative results showing the percentage of cells in G1, S and G2 phase in *KRAS*^{WT} CRC cell lines SW48 and CaCO2, *KRAS*^{G13D} CRC cell lines LoVo and HCT-116 (**A**), *KRAS*^{G13D} SW48 established by Crispr/Cas9 system (**B**), LoVo infected by shRNA lentivirus (**C**) were detected after treatment with 0, 2.5, 5, 10 mM metformin for 24 hrs (n=3).

Supplementary Figure 6

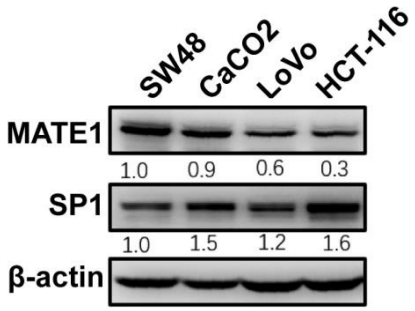

Fig. 6. The SP1 expression is not associated with MATE1 in CRC cells. Immunoblot analysis of SP1 and MATE1 in SW48, CaCO2, LoVo and HCT116 cells.

Supplementary Figure 7

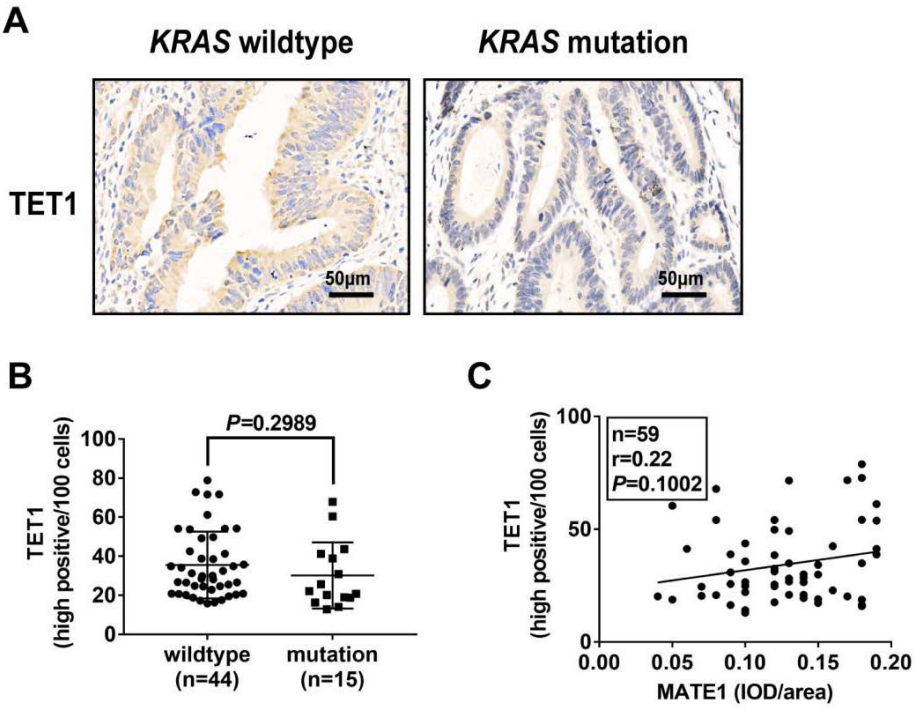

Fig. 7. The TET1 expression is not associated with *KRAS* mutation in mCRC patients tumor section. A, Representative images of TET1 immunohistochemistry on cross-sections from patients of mCRC with T2DM. **B,** The cells proportion with high positive TET1 expression in the nucleus of *KRAS* wildtype and mutation group were present. **C,** The association between IOD of MATE1 (shown in **Fig. 4D**) and cells proportion with high positive TET1 expression were determined by Pearson's correlation analysis.

Supplementary Figure 8

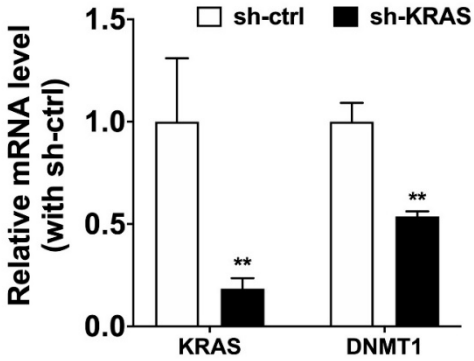

Fig. 8. The mRNA level of KRAS and DNMT1 in shKRAS LoVo cells. Data were shown as mean±SEM (n=3). **P<0.01 was determined by unpaired two-tailed Student's t-test.

Supplementary Figure 9

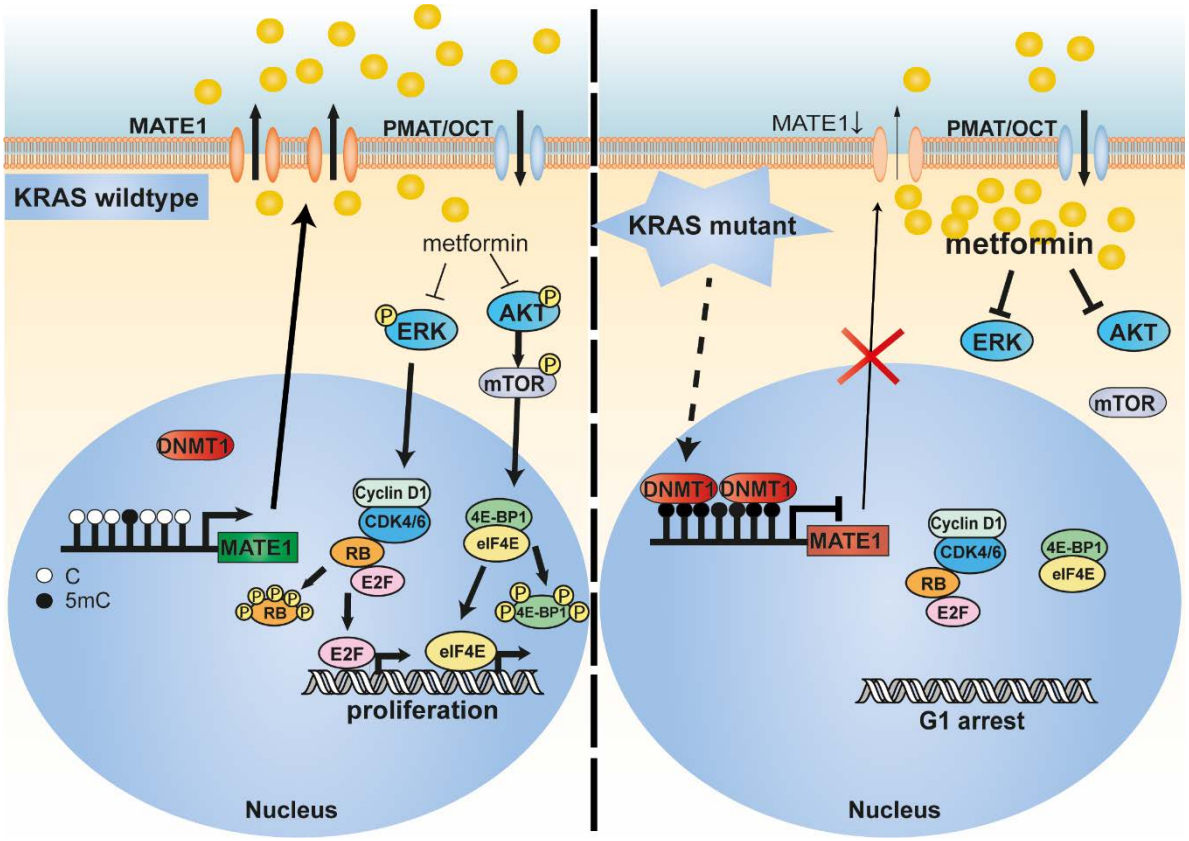

Fig. 9. A schematic of the regulation of metformin sensitivity in *KRAS*- wildtype and *KRAS*-mutation CRC cells. Aberrant KRAS signaling downregulates the MATE1 expression by hypermethylation in the CpG island of MATE1 promoter through breaking the balance of DNMT1, resulting in the more metformin accumulation in CRC cells, which further mediates the sensitivity of CRC cells to the anti-tumor effect of metformin.

Supplementary Figure 10

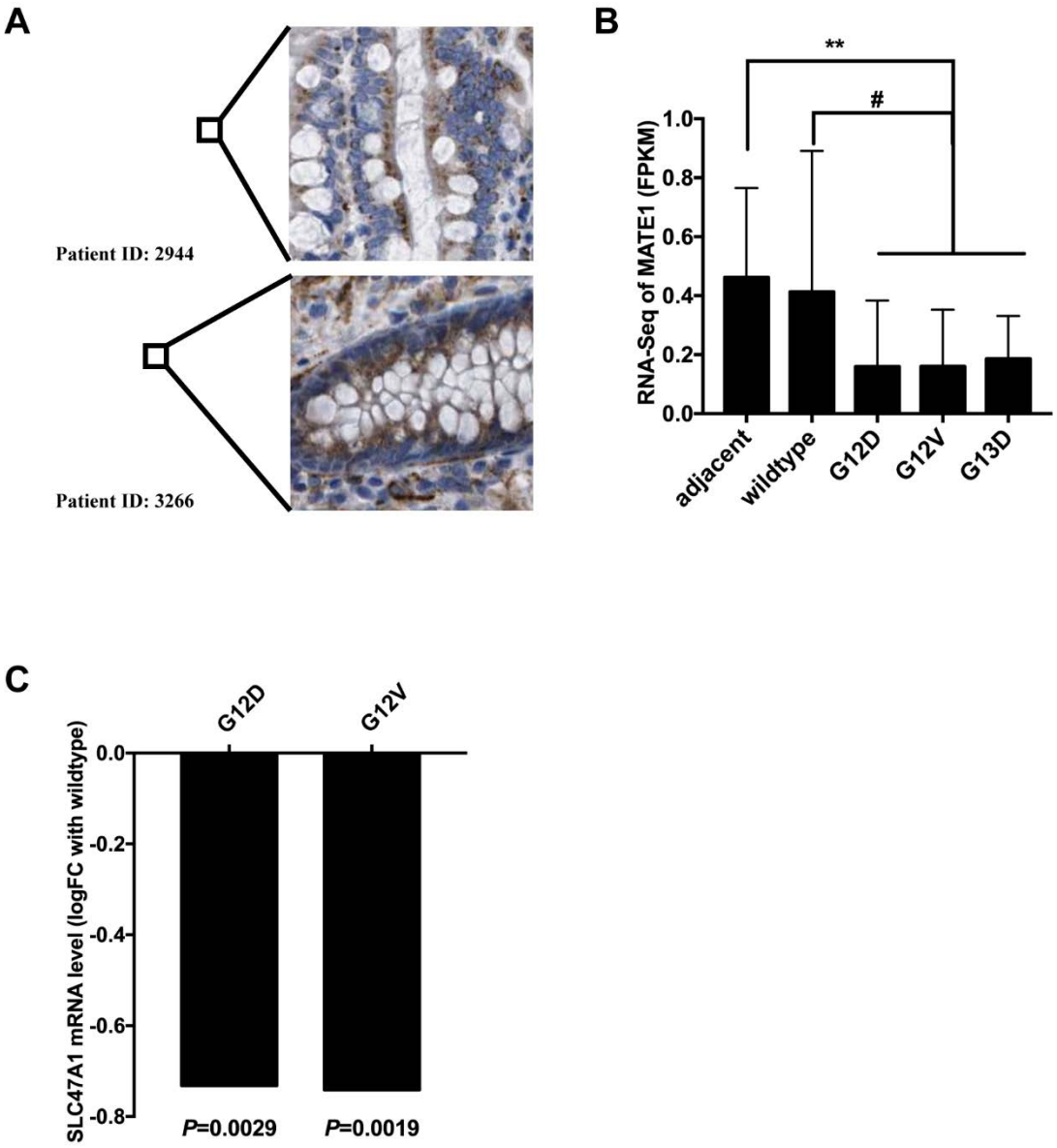

Fig. 10. MATE1 protein expression in normal tissue and mRNA level in wildtype and *KRAS*-mutation cancer tissue from TCGA-COAD and Colo741 cell line transfected by *KRAS*-G12D or *KRAS*-G12V from GEO database. A, Immunostaining of MATE1 in normal colon tissue. **B,** RNA analysis of MATE1 in adjacent and wildtype and *KRAS*-G12D, G12V, G13D cancer tissue from TCGA. **C,** RNA analysis of MATE1 in wildtype and Colo741 cell line transfected by *KRAS*-G12D or *KRAS*-G12V from the GEO database. P value was determined by unpaired two-tailed Student's t-test.

Supplementary Figure 11

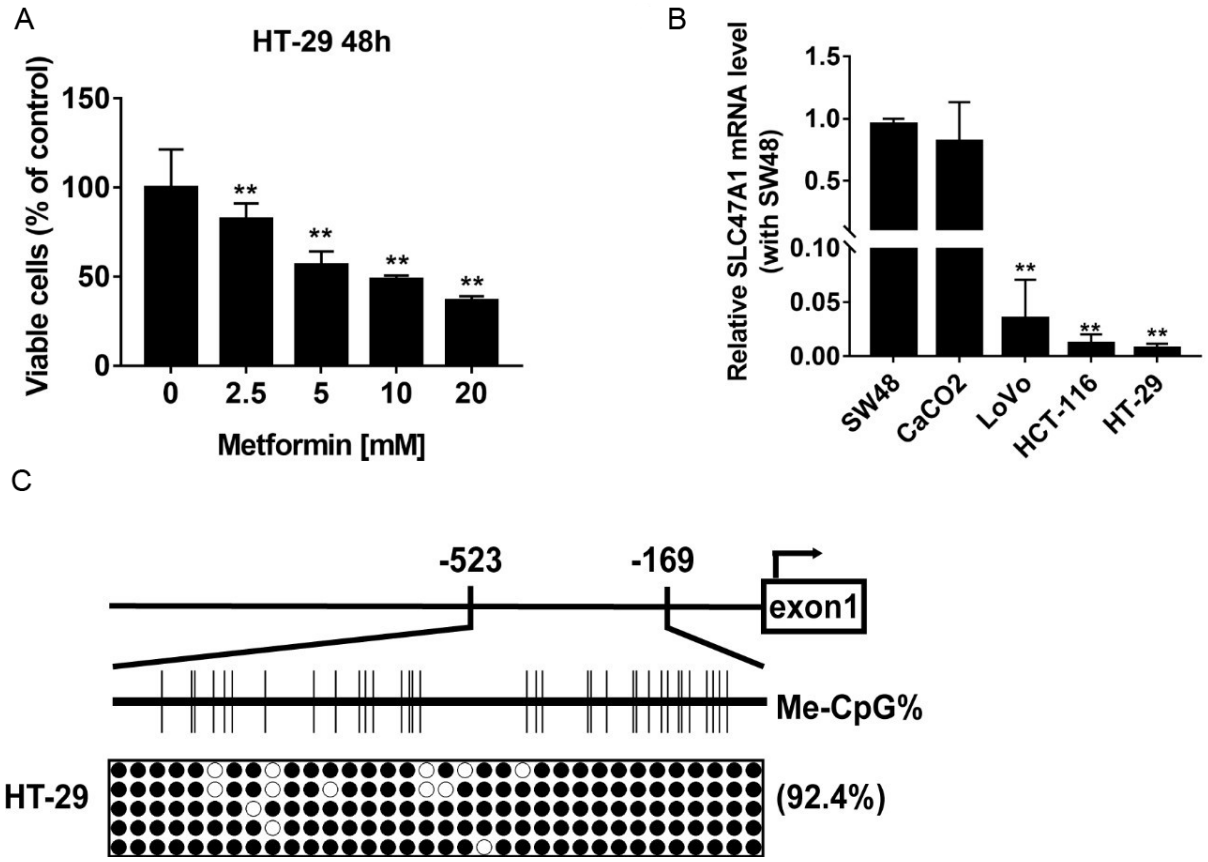

Fig. 11. Metformin inhibits *BRAF* mutation CRC cell line HT-29 by downregulating the MATE1 expression through hypermethylation on the CpG island. A, 48-hour growth of *BRAF*^{V600E} CRC cell line HT-29 was determined after treatment with 2.5, 5, 10, 20 mM or without metformin (n=5). Data were shown as mean±SEM. **P*<0.05, ***P*<0.01. **B,** Transcriptional levels of *SLC47A1* (MATE1) in *KRAS*^{WT} CRC cell lines SW48 and CaCO2, *KRAS*^{G13D} CRC cell lines HCT-116 and LoVo, and *BRAF*^{V600E} CRC cell line HT-29 was determined by qRT-PCR (n=3). Data were shown as mean±SEM. *P* value was determined by unpaired two-tailed Student's t-test with SW48 as the control, ***P*<0.01. Each analysis was replicated 3 times. **C,** BSP analysis of MATE1 promoter in HT-29.

Table S1. Baseline characteristics of mCRC patients with T2DM, according to hypoglycemic agents use, SYSUCC 2004-2016

Characteristics	Total No. (%)	Hypoglycemic agents use	
		Nonusers No. (%)	Users No. (%)
Sex			
Male	208 (73.76)	25 (78.13)	183 (73.20)
Female	74 (26.24)	7 (21.88)	67 (26.80)
Age at diagnosis, years			
< 60	93 (32.98)	9 (28.13)	84 (33.60)
≥ 60	189 (67.02)	23 (71.88)	166 (66.40)
BMI (kg/m^2)			
Lean (<18.5)	16 (5.67)	1 (3.13)	15 (6.00)
Normal (18.5~23.0)	154 (54.61)	15 (46.88)	154 (54.61)
Pre- & obese (>23.0)	112 (39.72)	16 (50.00)	112 (39.72)
Primary site			
Right colon	62 (21.99)	8 (25.00)	54 (21.60)
Left colon	109 (38.65)	8 (25.00)	101 (40.40)
Rectum	111 (39.36)	16 (50.00)	95 (38.00)
Pathological grading			
Well differentiated	7 (2.53)	0 (0.00)	7 (2.86)
Moderately differentiated	177 (63.90)	20 (62.50)	157 (64.08)
Poorly & un-differentiated	93 (33.57)	12 (37.50)	81 (33.06)
Unknown	5	0	5
Metastatic site			
Liver	122 (43.26)	12 (37.50)	110 (44.00)
Other organs	43 (15.25)	4 (12.50)	39 (15.60)
Distant lymph node	28 (9.93)	6 (18.75)	22 (8.80)
Multiple metastatic sites	54 (19.15)	3 (9.38)	51 (20.40)
Peritoneum	35 (12.41)	7 (21.88)	28 (11.20)
***KRAS* genotype**			
Wildtype	127 (60.19)	10 (52.63)	117 (60.94)
Mutation	84 (39.81)	9 (47.37)	75 (39.06)
Unknown	71	13	58

Abbreviations: mCRC, metastatic colorectal cancer; T2DM, type 2 diabetes mellitus; BMI, body mass index.

Table S2. Association between different hypoglycemic agents uses and OS and PFS, SYSUCC 2004-2016.

Hypoglycemic agents use	OS (n=282)		PFS (n=208)	
	Total (events)	HR (95% CI)[c]	Total (events)	HR (95% CI)[c]
Nonusers	32 (25)	Reference	26 (21)	Reference
Any-hypoglycemic agents use	250 (150)	0.547 (0.327, 0.913)	182 (157)	1.031 (0.585, 1.818)
Insulin or/and insulin secretagogues[a]	119 (72)	0.647 (0.374, 1.118)	91 (82)	1.129 (0.599, 2.129)
Non-insulin secretagogues[b]	22 (15)	0.875 (0.397, 1.926)	12 (8)	0.669 (0.209, 2.142)
Metformin-only	50 (25)	0.356 (0.183, 0.693)	35 (29)	0.630 (0.310, 1.280)
Both metformin and other hypoglycemic agents	59 (38)	0.491 (0.269, 0.894)	44 (38)	0.811 (0.406, 1.621)

Abbreviations: OS, overall survival; PFS, progression-free survival; SYSUCC, Sun Yat-sen University Cancer Center; HR, hazard ratio.

Notes:

a. Insulin secretagogues include sulfonylureas (e.g. Gliclazide, Glipizide, Glimepiride) and glinides (e.g. Repaglinide, Nateglinide).

b. Non-insulin secretagogues include α-glucosidase inhibitors (e.g. Acarbose, Voglibose, Miglitol) and Thiazolidinediones (e.g. Rosiglitazone, Pioglitazone).

c. HR (95%CI) was stratified by age at diagnosis (<60, ≥ 60 years) and adjusted for sex, body mass index, primary tumor site, metastatic site, pathological grading, KRAS genotype.

Table S3. List of the probability distribution of mCRC patients characteristics with metformin use or other hypoglycemic agents use.

Characteristics	Total No. (%)	Groups		*P* value
		Metformin use No. (%)	Other hypoglycemic agents use No. (%)	
Sex				0.511
Male	133 (73.89)	64 (76.19)	69 (71.88)	
Female	47 (26.11)	20 (23.81)	27 (28.12)	
Age at diagnosis, years				0.495
< 60	66 (36.67)	33 (39.29)	33 (34.38)	
≥ 60	114 (63.33)	51 (60.71)	63 (65.63)	
BMI (kg/m^2) [a]				0.544
Normal (18.5~23.0)	105 (58.33)	51 (60.71)	54 (56.25)	
Pre- & obese (>23.0)	75 (41.67)	33 (39.29)	42 (43.75)	
Primary site				0.069
Right colon	43 (23.89)	15 (17.86)	28 (29.17)	
Left colon	78 (43.33)	35 (41.67)	43 (44.79)	
Rectum	59 (32.78)	34 (40.48)	25 (26.04)	
Metastatic site				0.133
Liver	81 (45.00)	35 (41.67)	46 (47.92)	
Other organs	31 (17.22)	20 (23.81)	11 (11.46)	
Distant lymph node	12 (6.67)	7 (8.33)	5 (5.21)	
Multiple metastatic sites	34 (18.89)	15 (17.86)	19 (19.79)	
Peritoneum	22 (12.22)	7 (8.339)	15 (15.63)	
Pathological grading				0.658
Well differentiated	4 (2.22)	1 (1.20)	3 (3.12)	
Moderately differentiated	112 (62.22)	52 (61.90)	60 (62.50)	
Poorly & un-differentiated	64 (35.56)	31 (36.90)	33 (34.38)	
***KRAS* genotype**				0.237
Wildtype	109 (60.56)	47 (55.95)	62 (64.58)	
Mutation	71 (39.44)	37 (44.05)	34 (35.42)	

Abbreviations: mCRC, metastatic colorectal cancer; T2DM, type 2 diabetes mellitus; BMI, body mass index.

Note: a. 10 lean patients (BMI<18.5) were all other hypoglycemic agents' user, so this characteristic is inappropriate for hierarchical COX proportional hazards analysis.

Table S4. Association between the duration of metformin use before mCRC diagnosis and OS and PFS, SYSUCC 2004-2016.

Pre-duration of metformin use (years)	OS			PFS		
	Total (events) [a]	HR (95% CI) [b]	P_{trend}	Total (events) [c]	HR (95% CI) [b]	P_{trend}
***KRAS*-wildtype**						
Non-metformin user	69 (43)	Reference	0.084	50 (44)	Reference	0.266
≤3	10 (6)	1.409 (0.552, 3.596)		6 (6)	1.091 (0.400, 2.976)	
4-7	22 (15)	0.926 (0.483, 1.775)		16 (12)	0.816 (0.392, 1.699)	
>8	16 (11)	0.716 (0.338, 1.541)		13 (13)	1.131 (0.567, 2.258)	
***KRAS*-mutation**						
Non-metformin user	38 (25)	Reference	0.002	26 (23)	Reference	0.016
≤3	11 (8)	0.883 (0.333, 2.341)		10 (10)	0.998 (0.424, 2.348)	
4-7	14 (6)	0.211 (0.071, 0.630)		9 (7)	0.318 (0.122, 0.834)	
>8	12 (5)	0.038 (0.007, 0.219)		10 (7)	0.161 (0.049, 0.523)	

Abbreviations: OS, overall survival; PFS, progression-free survival; SYSUCC, Sun Yat-sen University Cancer Center; HR, hazard ratio; P_{trend}, the P value for trend analysis.

Notes: a. The number of death cases after diagnosis as mCRC.

b. HR (95%CI) was stratified by age at diagnosis (<60, ≥ 60 years) and adjusted for sex, body mass index, primary tumor site, metastatic site, pathological grading of CRC.

c. The number of dead or advanced cases during first-line chemotherapy.

Table S5. Association between metformin use and OS and PFS in different *KRAS* genotype, SYSUCC 2004-2016.

KRAS genotype	OS			PFS		
	Total (events) [a]	HR (95% CI) [b]	$P_{difference}$	Total (events) [c]	HR (95% CI) [b]	$P_{difference}$
Wildtype	47 (32)	Reference		35 (32)	Reference	
Mutation	37 (19)	0.362 (0.184, 0.715)		28 (23)	0.511 (0.264, 0.990)	
G12D	16 (9)	0.450 (0.195, 0.969)	0.450	12 (11)	0.440 (0.209, 0.928)	0.520
G12V	4 (1)	—		4 (4)	—	
G13D	8 (5)	0.251 (0.086, 0.735)		8 (7)	0.227 (0.052, 0.990)	
Other	9 (4)	—		4 (1)	—	

Abbreviations: OS, overall survival; PFS, progression-free survival; SYSUCC, Sun Yat-sen University Cancer Center; HR, hazard ratio; $P_{difference}$, the *P* value for difference among G12D, G12V, G13D, and other codon 12 mutations.

Notes: a. The number of death cases after diagnosis as mCRC with metformin use.

b. HR (95%CI) was stratified by age at diagnosis (<60, ≥ 60 years) and adjusted for sex, body mass index, primary tumor site, metastatic site, pathological grading of CRC.

c. The number of dead or advanced cases in mCRC patients with metformin use during first-line chemotherapy.

"—" indicates the variable is excluded because of <5 observations in this category.

Table S6. Clinical and pathological features of two colon cancer patients used for the PDX models establishment.

ID	Gender	Age at diagnosis, years	Primary site	Type	Pathological grading	*KRAS* genotype	Prior chemotherapy
374469	Male	48	Right colon	MUC	Moderate	G12D	No
386650	Male	68	Left colon	ADC	Poor	Wildtype	No

Abbreviations: PDX, patient-derived xenograft model; MUC, mucinous adenocarcinoma; ADC, adenocarcinoma.

Table S7. List of primers and their application.

Primer	Sequecnce (5’ to 3’)	Application
hKRAS gRNA-F-oligo	CACCGCATTTTTCTTAAGCGTCGA	sgRNA cloning
hKRAS gRNA-R-oligo	AAACTCGACGCTTAAGAAAAATGC	
	ATTCGTATCCAAGATTAAACATTCATATTATTTC	
hKRAS-Left arm-F	AAAAAATTTGAAAC	Point mutation introducing
hKRAS-TGG mut-Left arm-R	tCATCGACGCTTAAGAAAAATGCATAAATGCT	
	GAATTCGCAGCTACtAGGAGTTTGTAAATGAAGT	
hKRAS-TGG mut-Right arm-F	ACAGTT	Donor vector cloning
hKRAS-Right arm-R	TAGTGTATCCAAGTATGCCCTCAGAACTTGCCTC	
	AGCGCA	
hKRAS-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAA	
	CATTCATATTATTTCAAAAAATTTGAAAC	
hKRAS-attB1	GGGGACCACTTTGTACAAGAAAGCTGGGTTGCC	
	CTCAGAACTTGCCTCAGCGCA	
MATE1-promoter-F	GGTGTACCCAACTCAATCTGCACAGCAAATG	BSP sequence cloning
MATE1-promoter-R	CAGTGCAGCCTGCTCTGCAGTCCC	

Notes: underlined characters were used as the homologous arms for In-fusion recombination; double underlined characters were used as the homologous arms for BP reaction.

Table S8. List of primers for qRT-PCR.

Primer	Sequence (5’ to 3’)
homo-SLC29A4-F	AGTACCCAGGGACCTCCATC
homo-SLC29A4-R	GTGTGCAGGGTCAGTCTCTC
homo-SLC22A1-F	CCCCTCATTTTGTTTGCGGT
homo-SLC22A1-R	TTTCTCCCAAGGTTCTCGGC
homo-SLC22A2-F	AATCTCTACCCGCCTCCCTT
homo-SLC22A2-R	CACAGAGCTCGTGAACCAGT
homo-SLC22A3-F	CCCACTCCACCATCGTCAG
homo-SLC22A3-R	ATCCTGCCATACCTGTCTGC
homo-SLC47A1-F	GGAGTGATGGGTCTGTGGTC
homo-SLC47A1-R	ACTCCGAGGCACGTTGTTTA
homo-SLC47A2-F	TGAGATCGGGAGCTTCCTCA
homo-SLC47A2-R	GAGCCCCAAGGGAATCATGT

SI REFERENCE:

[1] Chu VT, Weber T, Wefers B, et al. (2015) Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol 33(5): 543-548. 10.1038/nbt.3198

[2] Dowling RJ, Lam S, Bassi C, et al. (2016) Metformin Pharmacokinetics in Mouse Tumors: Implications for Human Therapy. Cell Metab 23(4): 567-568. 10.1016/j.cmet.2016.03.006