Linear-double-stranded DNA (ldsDNA) based AND logic computation in mammalian cells

Weijun Su [1], Chunze Zhang [3], Shuai Li [2*]

1. School of Medicine, Nankai University, Tianjin 300071, China; 2. Department of Breast Cancer Pathology and Research Laboratory, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin 300060, China; 3. Department of Colorectal Surgery, Tianjin Union Medical Center, Tianjin 300121, China.

* To whom correspondence should be addressed. Tel: +86 22 23340123; Fax: +86 22 23340123; Email: shuaili@tmu.edu.cn & shuai.li2001@gmail.com.

Abstract

Synthetic biology employs engineering principles to redesign biological system for clinical or industrial purposes. The development and application of novel genetic devices for genetic circuits construction will facilitate the rapid development of synthetic biology. Here we demonstrate that mammalian cells could perform two- and three-input linear-double-stranded DNA (ldsDNA) based Boolean AND logic computation. Through hydrodynamic ldsDNA delivery, two-input ldsDNA-base AND-gate computation could be achieved *in vivo*. Inhibition of DNA-PKcs expression, a key enzyme in non-homologous end joining (NHEJ), could significantly downregulate the intensity of output signals from ldsDNA-based AND-gate. We further reveal that in mammalian cells ldsDNAs could undergo end processing and then perform AND-gate calculation to generate in-frame output proteins. Moreover, we show that ldsDNAs or plasmids with identical overlapping sequences could also serve as inputs of AND-gate computation. Our work establishes novel genetic devices and principles for genetic circuits construction, thus may open a new gate for the development of new disease targeting strategies and new protein genesis methodologies.

Key words: linear-double-stranded DNA based AND logic computation, synthetic biology, AND gate, linear-double-stranded DNA (ldsDNA), non-homologous end joining (NHEJ), homologous recombination (HR), artificial genetic circuit.

Introduction

Synthetic biology is a rapidly growing multidisciplinary branch of biology (1-3). Through applying engineering principles to biological systems, synthetic biology aims to design or redesign genetic devices, genes, genetic circuits and cells to accomplish sophisticated tasks in the areas of biotechnology and biomedicine (2,4-7). Several milestone achievements, such as artemisinin precursor production by engineered microbe (8,9), biofuel production using amino acid metabolism in *E. Coli* (10), creation of a bacterial cell with a synthetic genome (11,12), have already shown the revolutionary power of synthetic biology. Meanwhile, synthetic biology is now driving life science research into 'build life to understand it' era (3,4).

In addition to engineering microbes for industrial purposes, another major focus of synthetic biology is on building cell-based genetic circuits that implement artificial programs of gene expression (5,13,14). By assembling different genetic devices together, genetic circuits could control gene expression through various biochemical processes, including transcription, translation and post-translational processes (2,13). The recent advances in the design of genetic circuits to diagnose and target cells bring opportunities for effective disease treatment (14-17). However, building complex genetic circuits remains one of the great challenges in synthetic biology (13). The innovation and application of novel genetic devices could facilitate the development of new genetic circuit design strategies.

Like complex electronic circuits, genetic circuits accomplish bio-computation through the assembly of different logic gates. By performing Boolean calculations, cell-based AND logic gate convert information from two inputs into one output, which could facilitate diagnostic and therapeutic specificity (18-21). Over the past two decades, different low noise-signal-ratio AND logic gate design strategies have been developed to achieve delicate control of gene expression. For example, recently by utilizing split-intein protein-splicing strategy, transcriptional activation domain (TAD) fused programmable DNA binding proteins, like zinc finger proteins (ZFPs), transcriptional activator-like effectors (TALEs) and CRISPR-associated protein 9 (Cas9), were split into transcriptional mute parts, thus performing two- or three-input AND logic computation (22-24).

Here, we use linear-double-stranded DNA (ldsDNA, also PCR amplicon) to implement Boolean logic AND gates. We demonstrate that mammalian cells could conduct two- and three-input ldsDNA-based AND logic computation. Mouse hydrodynamic injection shows that ldsDNA-based AND calculation could be achieved *in vivo*. ldsDNA with one or two terminal nucleotide(s) addition, which leads to reading frame shift, could be end-processed to produce in-frame AND-gate output proteins. Moreover, we found ldsDNAs or plasmids with identical overlapping sequence could also perform AND-gate computation. Our research provides novel genetic devices and principles for synthetic genetic circuits design.

Materials and Methods

Cell culture

HEK293T cell line was maintained in Dulbecco's modified Eagle medium (DMEM) (Thermo Fisher) containing 10% fetal bovine serum (FBS) with 100 U/ml penicillin and 100 mg/ml streptomycin at 37 degree with 5% CO_2.

Plasmids

Renilla luciferase control reporter vector (pRL-CMV, Promega) and GFP expressing vector (pEGFP-C1, Clontech) are from lab storage. To get higher reporter gene expression level (SV40 promoter to CMV promoter), firefly-luciferase coding region was subcloned into pcDNA3.0 vector (Invitrogen) between Hind III and BamH I restriction sites (namely pcDNA3.0-firefly-luciferase). pMD-18T AT clone vector was bought from Takara.

ldsDNA synthesis

KOD DNA polymerase (Toyobo) was employed to amplify AND-gate ldsDNAs (PCR amplicons) taking pcDNA3.0-firefly-luciferase or pEGFP-C1 plasmids as templates. To remove plasmid templates and free dNTPs, PCR products underwent agarose electrophoresis and gel-purification (Universal DNA purification kit, Tiangen). For animal experiments, ldsDNAs were eluted in 0.9% NaCl solution. The amount of PCR products was determined by OD 260 absorption value.

Transfection of AND-gate genetic circuits

For dual luciferase assay, HEK293T cells were seeded into 24-well plates the day before transfection (60% - 70% confluency). Unless otherwise stated, 100 ng/well input amplicons or 400 ng/well input plasmids (for [1,1] states, 100 ng/well of each amplicon or 400 ng/well of each plasmid was used) and 10 ng/well pRL-CMV control vector were cotransfected into cells with Lipofectamine 2000 (Invitrogen) reagent following standard protocol. 48 h after transfection, dual reporter activities were measured by Dual Luciferase Reporter Assay System (Promega). For GFP experiments, 100 ng/well input amplicons or 400 ng/well input plasmids were transfected into 24-well seeded HEK293T cells (for [1,1] states, 100 ng/well of each amplicon or 400 ng/well of each plasmid was used). 48 h after transfection, cells were harvested for FACS or western blots.

Immunofluorescence staining

For immunofluorescence staining, specific antibodies against γ-H2AX (05-636, Millipore, 1:200 dilution), 53BP1 (NB100-304, Novus, 1:200 dilution) followed by Alexa Fluor 594 labeled-secondary antibodies (Invitrogen) were used for detection. Cells were counterstained with DAPI, mounted with anti-fade mounting medium, and photographs were taken under fluorescence microscope.

Flow cytometry analysis

HEK293T cells transfected with AND-gate genetic circuits were rinsed with PBS, then digested with 0.25% trypsin-EDTA. Cells were re-suspended by PBS supplemented with 2% FBS, then analyzed with BD FACSCalibur ™ (BD). Raw data was analyzed by FlowJo software.

Western blots

Proteins at the same amount were separated by SDS-PAGE and transferred onto PVDF membranes. After probing with anti-GFP (ab1218, abcam, 1:3000 dilution) or anti-β-actin antibody (sc-477778, Santa Cruz, 1:5000 dilution) and incubating with proper secondary antibody, antigen-antibody complex was visualized by chemiluminescent HRP substrate (Millipore).

RNA interfering and realtime qRT-PCR

siRNAs targeting DNA-PKcs (DNA-PKcs siRNA, CGU GUA UUA CAG AAG GAA ATT) (25), Rad51 (Rad51 siRNA, GAG CUU GAC AAA CUA CUU CTT) (26) and scramble control siRNA (Ctrl, UUC UCC GAA CGU GUC ACG UTT) were synthesized by GenePharma. siRNAs were transfected using Lipofectamine 2000 reagent following the protocol. Total RNA was prepared 48 h after transfection and qRT-PCR was employed to analyze the interfering effects. cDNA was generated by PrimeScript ™ RT reagent Kit with

gDNA Eraser (Takara). qPCR was performed using SYBP Premix Ex Taq ™ II (Tli RNaseH Plus, Takara). GAPDH mRNA levels were used for normalization. Forward (F) and reverse (R) primers used were as follows: GAPADH-F 5'- tcagtggtggacctgacctg -3', GAPDH-R 5'- tgctgtagccaaattcgttg -3'; DNA-PKcs-F 5'- gctgatctcttaaagcgggc -3', DNA-PKcs-R 5'- ctgtgtagcggatccaaacg -3'; Rad51-F 5'- tcacggttagagcagtgtgg -3', Rad51-R 5'- ttagctccttctttggcgca -3'.

Hydrodynamic injection of ldsDNAs

All of the *in vivo* mice experiments were approved by The Animal Ethical and Welfare Committee of Tianjin Medical University Cancer Institute and Hospital. 6-8-week-old syngeneic BALB/c female mice were bought from Beijing Vital River Laboratory Animal Technology Company. 2.5μg input ldsDNA (for [1,1] states, 2.5μg of each input ldsDNA was used) was diluted with 1.6ml normal saline at room temperature. Mice were anesthetized with 4% chloral hydrate and held immobile in a researcher's hand and then ldsDNA solution was injected into retro-orbital sinus by a 26-gauge needle in 12 seconds.

Bioluminescence imaging

Bioluminescence imaging (BLI) of live mice was performed at day 1, 2, 4, 7 and 14 after hydrodynamic injection of ldsDNAs. BLI of firefly luciferase was performed using the IVIS Lumina II system (Xenogen). After intraperitoneal injection of D-luciferin (Invitrogen,

150μg/g body weight), each mouse was imaged for 1-5 min. Bioluminescence signals were quantified in units of maximum photons per second (photons/s).

BLI of mice organs was performed at day 1 after hydrodynamic injection of ldsDNAs. Mice were injected with 150μg/g D-luciferin intraperitoneally and were euthanized 10 min later. Different organs (heart, liver, spleen, lung and kidneys) were collected immediately and placed into 6-well plates. BLI of organs was performed using the IVIS Lumina II system. Organs were imaged for 1–5 min. Bioluminescence signals were quantified in units of maximum photons per second (photons/s).

Statistics

All data are presented as means ± SD unless stated otherwise. Differences were assessed by two-tailed Student's t-test using GraphPad software. $p < 0.05$ was considered to be statistically significant.

Results

Two-input ldsDNA-based AND-gate computation in mammalian cells

A typical gene overexpression plasmid contains integrant sequence consisting of promoter, gene coding sequence (CDS) and ploy(A) signal (PAS). Here by PCR amplification, we split reporter gene expression cassette (promoter-CDS-PAS) into 2 or 3 linear-double-stranded

DNAs (ldsDNAs). Then these input ldsDNAs were introduced into mammalian cells to perform AND-gate calculation.

Taking pcDNA3.0-firefly-luciferase expression vector as PCR template, we first designed three different sets of two-input ldsDNA-based AND-gate genetic circuits (Figure 1A). Neither (state [0,0]), only one (states [0,1] or [1,0]), or both (state [1,1]) AND-gate genetic circuits were introduced into HEK293T cells by transfection, respectively. To balance transfection efficiency, CMV-Renilla-luciferase plasmid (pRL-CMV) was co-transfected. As shown in Figure 1B, all states [1,1] led to dramatic induction of relative luciferase activity (~16, 7, 26-thousands folds, respectively) compared with non-ldsDNA transfected cells (state [0,0]). Similar trend was observed when firefly luciferase activities were normalized to protein concentrations (Supplementary Figure S1). Moreover, there is a positive correlation between the amount of ldsDNA inputs and the intensity of output signals (Supplementary Figure S2).

We next designed a two-input ldsDNA-based AND-gate genetic circuit by amplifying CMV promoter fragment and GFP coding sequence-SV40 poly(A) signal fragment (GFP-SV40 PAS) from pEGFP-C1 vector, respectively (Figure 1E). Flow cytometry showed that the AND-gate computation led to 20.5% GFP positive ratio in transfected cells (state [1,1]) (Figure 1F and 1G). Moreover, western blots result also demonstrated the expression of GFP protein with introduction of state [1,1] (Figure 1H). To explore the ratio of input ldsDNAs

undergone right ligations, we used 5’ Cy5 modified primer to generate Cy5 labeled ldsDNA and transfected into HEK293T cells. Flow cytometry showed that about one quarter of Cy5 positive cells expressed GFP (Supplementary Figure S3).

In experimental plasmid ligation or cellular non-homologous double-strand end joining (NHEJ) process, covalent phosphodiester bond is formed between 5’ phosphate group from one dsDNA and 3’ hydroxyl group from another (27,28). As we know, regular synthesized PCR primers contain 5’ terminal hydroxyl group (not phosphate group), which leads to hydroxyl groups existing on both ends of PCR amplicons. Here by 5’ phosphated primers, we generated 5’ terminal phosphated ldsDNAs by PCR (taking pcDNA3.0-firefly-luciferase plasmid as template). However, unexpectedly, no significant enhancement of reporter activities was observed in 5’ phosphated ldsDNA-based AND computation compared with non-phosphated ldsDNAs (Figure 1B and Supplementary Figure S4).

Inhibition of DNA-PKcs attenuates output signals of ldsDNA-based AND calculation

As aforementioned ldsDNA end-relinkage process is similar to NHEJ pathway, we further explored whether inhibition of NHEJ pathway could attenuate AND-gate output signals. DNA-PKcs (DNA-dependent protein kinase, catalytic subunit), a pivotal Ku70/Ku80 interacting nuclear serine/threonine kinase, was selected to be targeted by RNA interfering (Figure 1C) (29). As shown in Figure 1D, the inhibition of DNA-PKcs significantly downregulates ldsDNA-based AND-gate output signals.

Introduction of ldsDNAs does not lead to DNA damage foci formation

Considering its potential applications in biological treatments, we decide to explore whether ldsDNA could lead to DNA damage foci formation in cells. After ldsDNA transfection, immunofluorescence staining of DNA damage foci markers, such as γ-H2AX and 53BP1, was performed in HEK293T cells, while 10Gy γ-radiated cells were used as positive control. As shown in Supplementary Figure S5, neither γ-H2AX nor 53BP1 foci appeared in ldsDNA transfected HEK293T cells.

ldsDNA-based two-input AND-gate calculation *in vivo*

To estimate whether ldsDNA-based AND-gate calculation could be achieved *in vivo*, we employed hydrodynamic injection as a tool for *in vivo* ldsDNA delivery (30). First, we confirmed the validity of this method by injecting CMV-Luc-BGH PAS ldsDNAs (full gene expression cassette) into mice retro-orbital sinus. With bioluminescence imaging (BLI), firefly luciferase signals were detected in liver region one day after hydrodynamic injection and diminished in 2 weeks (Supplementary Figure S6A-B). Postmortal BLI confirmed the expression of luciferase in liver but not other organs, which is in consistent with the results from other groups (Supplementary Figure S6C) (30). We next introduced only one (states [0,1], [1,0]) or both (state [1,1]) AND-gate genetic circuits into mice. State [1,1] (CMV-Luc[1-205] + Luc[206-1653]-BGH PAS), but not state [0,1] or [1,0], led to firefly

luciferase expression *in vivo* (Figure 2A-B). Correspondingly, significant luciferase signal was detected in liver under state [1,1] via postmortal BLI (Figure 2C-D).

Three-input ldsDNA-based AND-gate calculation in mammalian cells

We further explored the possibility to build up three-input ldsDNA-based AND logic computation circuits in mammalian cells. Three ldsDNAs (CMV promoter, GFP coding sequence and SV40 PAS) were amplified from pEGFP-C1 plasmid and transfected into HEK293T cells with different patterns of combination (Figure 3A). About 8% cells expressed GFP when all 3 ldsDNAs were introduced (state [1,1,1]) (Figure 3B and 3C). This implies that an increase in the number of inputs may cause a decrease in the intensity of output signals (two-input vs three-input, 20.5% vs 8.01%, Figure 1G and 3C).

AND computation of ldsDNAs supplemented with additional terminal nucleotide(s)

In NHEJ-mediated DNA damage repair, dsDNA ends undergo end processing in which nucleotide(s) insertion or deletion occurs to DNA terminals. Thus, we wondered whether introduced exogenous ldsDNAs were also subject to similar DNA end processing. By PCR amplification, we added one or two random nucleotide(s) to PCR amplicon terminals (Figure 4A). Theoretically, no in-frame GFP protein would be produced without DNA end processing. However, in practice, pEGFP-C1 vector-derived ldsDNAs supplemented with one or two additional terminal nucleotide(s) led to 10.76% or 6.77% GFP positive rates, respectively

(Figure 4B and 4C). These results demonstrate that ldsDNAs could undergo end nucleotide(s) deletion and thereby conduct AND logic calculation to produce in-frame proteins.

ldsDNAs with identical overlapping sequences could conduct AND logic computation

We further revealed that two ldsDNAs with identical overlapping sequences could also perform AND logic computation (Figure 5A and 5B). Dual luciferase assay showed that ldsDNAs with different length of identical overlapping sequence led to 56, 73, 77 or 63 thousand times increase of output signals compared with non-transfection cells (state [0,0]), respectively (Figure 5B), which is stronger than that by ldsDNA ends rejoining (Figure 1B and 5B). However, we found longer overlapping sequences do not necessarily lead to higher output signal activities (Figure 5B), while there is a positive correlation between the amount of ldsDNA inputs and the output signal intensity (Supplementary Figure S7). We further performed RNA interfering downregulation of Rad51, a protein in charge of homologous searching in homologous recombination (HR), and found Rad51 knockdown could attenuate output signals under state [1,1] (Figure 5C and 5D) (31). Additionally, flow cytometry results showed that identical sequence overlapping ldsDNAs (PCR amplified from pEGFP-C1 vector) could lead to GFP expression in 34.63% of transfected HEK293T cells (Figure 5E, 5F and 5G).

Plasmids with identical overlapping sequences could conduct AND logic computation

We then subcloned overlapping ldsDNAs into pMD-18T shuttle vector. Reporter assay showed that mammalian cells could conduct AND logic computation after being co-transfected with sequence-overlapping plasmids (Figure 6A). This was further verified with pMD-18T vectors with overlapping GFP fragments (Figure 6D-F). Different with the case of sequence-overlapping ldsDNAs, there is a significant positive correlation between the length of overlapping sequence and the output signal intensity under state [1,1] (Figure 6B). And unlike overlapping ldsDNAs, downregulation of Rad51 could not turndown state [1,1] signal from overlapping plasmids (Figure 6C). Both of the differences imply that the mechanisms underlying AND computation with overlapping ldsDNAs and overlapping plasmids would be distinct.

Discussion

Here, we present a proof-of-concept demonstration that ldsDNAs can be engineered for AND logic computation in mammalian cells. Luciferase reporter assay results showed that the ldsDNA-based AND-gate both states [1,1] could lead to over thousands folds of signal induction over neither state [0,0]. Flow cytometry and western blots revealed the induction of GFP expression by ldsDNA-based AND logic computation. This AND-gate calculation was also achieved in mice via ldsDNA hydrodynamic injection. We also found that ldsDNAs with 1N or 2N random nucleotide(s) end addition could conduct AND gate computation. Furthermore, ldsDNAs or plasmids with identical overlapping sequence could also form AND logic circuits in mammalian cells. Our study provides novel devices and principles for

synthetic genetic circuits design, which have great application potentials in future experimental and clinical researches.

One major focus of synthetic biology is the development of novel devices or principles for genetic circuits design. In eukaryotes, several principles relying on recombinases, DNA binding proteins and RNA-based devices, have been developed to conduct AND logic calculation (32). Here, we present a proof-of-concept ldsDNA device for synthetic circuits design. By splitting essential gene-expression cassette into different PCR fragments, our ldsDNA ends rejoining and identical overlapping sequence recombination genetic circuit design strategies could generate AND gate output signals with low noise-signal ratio. However, our new principle generates lower gene expression levels and lower GFP positive rates compared with intact gene-expression cassette (in PCR amplicon or gene expression vector, data not shown). This may be caused by the following reasons. First, as shown in Figure 1D, NHEJ machinery members (e.g. DNA-PKcs) participate in the generation of output signals. This implies that ldsDNAs have to undergo a rejoining process to get full length productive gene-expression cassette. Second, if not taking self-ligation into account, geometrically, only one-fourth ldsDNA ligation generates the right gene-expression cassette. To overcome it, two conceptual design strategies are proposed here: 1) addition of chemical groups to ldsDNA ends to block unwanted linkages; 2) using back-to-back promoter ldsDNAs as one input (Supplementary Figure S8).

Another goal of synthetic biology is to improve diagnostic or therapeutic targeting specificity of genetic circuits (19-21,33). To achieve this, genetic circuits could be designed to sense targeted-cell gene expression signature and then turn on cell specific gene expression. For example, Lior *et al* developed RNA-based AND-gate genetic circuits which could trigger tumor-specific immunomodulator. This sophisticated circuit assembles different devices, including tumor-specific promoters, miRNA sponge and miRNA-target regulation, to guarantee tumor-specific immunotherapy (34). Another way to improve targeting specificity is to optimize the delivery strategy. For example, Roybal *et al* engineered T cells to recognize combinations of antigens and thereby kill target tumor cells (20). Our novel principle can be utilized in improving genetic circuits targeting specificity. For example, combination of antibody coated nanoparticles, targeting different cell membrane protein antigens and containing ldsDNA-based AND-gate inputs can be applied to enhance the targeting specificity (Supplementary Figure S9).

In mammalian adaptive immunity system, highly diversified repertoires of antigen-receptor genes were generated by V(D)J recombination (35-37). Along with B and T cell maturation, variable (V), joining (J), and in some cases, diversity (D) gene fragments undergo nearly random recombination, which makes the major contribution to the diversity of antibodies/immunoglobulins (Igs) and T cell receptors (TCRs). Moreover, during DJ or VD junction formation, terminal deoxynucleotide transferase (TdT), Artemis nuclease, DNA polymerase and some NHEJ pathway members also make contributions to antigen-receptor diversity through introducing nucleotides insertion or deletion (indel) to the junctions (38-40).

In this study, we present the design of three-input ldsDNA-based AND genetic circuits which is similar to V(D)J recombination process (Figure 5). Furthermore, this ends-rejoining process could also conduct nucleotide(s) end processing and then produce in-frame right AND output signals (Figure 4). Inspired by the aforementioned characteristics of this novel genetic circuit, we propose a conceptual strategy to create highly diversified antibody repertoires. With ldsDNAs corresponding to V, D or J fragments, the AND logic computation may generate highly variable antibody library. Since D fragments only encode a few amino acids, they can be replaced by adding randomized nucleotides to the ends of V and J ldsDNAs (Supplementary Figure S10).

Furthermore, we showed another two patterns of AND-gate genetic circuits in mammalian cells, which are mediated by identical sequence overlapping ldsDNAs or plasmids (Figure 5 and 6). The overlapping ldsDNAs generated stronger both state [1,1] signals compared with end joining AND gate. No significant correlation between ldsDNA overlapping length and output signals was observed, while the plasmids with overlapping sequences apparently showed a positive correlation between overlapping length and output signals. Another different point between these two patterns is that Rad51 inhibition leads to attenuation of AND output signals mediated by overlapping ldsDNAs, but not overlapping plasmids. The mechanisms underlying the differences between these two types of sequence overlapping AND gate calculation need to be further explored.

In conclusion, we developed a novel ldsDNA-based AND logic genetic circuits design strategy. This AND-gate could be achieved both *in vitro* and *in vivo*. ldsDNA ends could undergo end chopping and generate the right in-frame AND gate output proteins. Our study provides new devices and principles for synthetic biology and opens a gate for the generation of functional proteins or antibodies.

Author Contributions

S.L. conceived and supervised the study. S.L. and W.S. designed and performed experiments. S.L. and W.S. analyzed data. W.S. and S.L. prepared figures. S.L. and W.S. wrote the paper. All authors discussed the results and reviewed the manuscript.

Acknowledgments

This work was supported by grants from National Natural Science Foundation of China (31400673 to S.L., 81402407 to W.S.); Tianjin Research Program of Application Foundation and Advanced Technology [14JCQNJC09800 to S.L., 15JCQNJC11700 to W.S.]; partially from Key Project of National Health and Family Planning Commission of Tianjin [2017057 to C.Z.].

Competing Interests

The authors have declared that no competing interest exists.

Reference

1. Andrianantoandro, E., Basu, S., Karig, D.K. and Weiss, R. (2006) Synthetic biology: new engineering rules for an emerging discipline. *Molecular systems biology*, **2**, 2006 0028.
2. Purnick, P.E. and Weiss, R. (2009) The second wave of synthetic biology: from modules to systems. *Nature reviews. Molecular cell biology*, **10**, 410-422.
3. Cameron, D.E., Bashor, C.J. and Collins, J.J. (2014) A brief history of synthetic biology. *Nature reviews. Microbiology*, **12**, 381-390.
4. Elowitz, M. and Lim, W.A. (2010) Build life to understand it. *Nature*, **468**, 889-890.
5. Bashor, C.J., Horwitz, A.A., Peisajovich, S.G. and Lim, W.A. (2010) Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. *Annual review of biophysics*, **39**, 515-537.
6. Wieland, M. and Fussenegger, M. (2012) Engineering molecular circuits using synthetic biology in mammalian cells. *Annual review of chemical and biomolecular engineering*, **3**, 209-234.
7. Smanski, M.J., Zhou, H., Claesen, J., Shen, B., Fischbach, M.A. and Voigt, C.A. (2016) Synthetic biology to access and expand nature's chemical diversity. *Nature reviews. Microbiology*, **14**, 135-149.
8. Martin, V.J., Pitera, D.J., Withers, S.T., Newman, J.D. and Keasling, J.D. (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nature biotechnology*, **21**, 796-802.
9. Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J. *et al.* (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, **440**, 940-943.
10. Atsumi, S., Hanai, T. and Liao, J.C. (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature*, **451**, 86-89.
11. Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.Y., Algire, M.A., Benders, G.A., Montague, M.G., Ma, L., Moodie, M.M. *et al.* (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, **329**, 52-56.
12. Hutchison, C.A., 3rd, Chuang, R.Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L. *et al.* (2016) Design and synthesis of a minimal bacterial genome. *Science*, **351**, aad6253.
13. Nielsen, A.A., Segall-Shapiro, T.H. and Voigt, C.A. (2013) Advances in genetic circuit design: novel biochemistries, deep part mining, and precision gene expression. *Current opinion in chemical biology*, **17**, 878-892.
14. Lienert, F., Lohmueller, J.J., Garg, A. and Silver, P.A. (2014) Synthetic biology in mammalian cells: next generation research tools and therapeutics. *Nature reviews. Molecular cell biology*, **15**, 95-107.
15. Weber, W. and Fussenegger, M. (2011) Emerging biomedical applications of synthetic biology. *Nature reviews. Genetics*, **13**, 21-35.
16. Ruder, W.C., Lu, T. and Collins, J.J. (2011) Synthetic biology moving into the clinic. *Science*, **333**, 1248-1252.

17. Bacchus, W., Aubel, D. and Fussenegger, M. (2013) Biomedically relevant circuit-design strategies in mammalian synthetic biology. *Molecular systems biology*, **9**, 691.
18. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. and Benenson, Y. (2011) Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science*, **333**, 1307-1311.
19. Liu, Y., Zeng, Y., Liu, L., Zhuang, C., Fu, X., Huang, W. and Cai, Z. (2014) Synthesizing AND gate genetic circuits based on CRISPR-Cas9 for identification of bladder cancer cells. *Nature communications*, **5**, 5393.
20. Roybal, K.T., Rupp, L.J., Morsut, L., Walker, W.J., McNally, K.A., Park, J.S. and Lim, W.A. (2016) Precision Tumor Recognition by T Cells With Combinatorial Antigen-Sensing Circuits. *Cell*, **164**, 770-779.
21. Morel, M., Shtrahman, R., Rotter, V., Nissim, L. and Bar-Ziv, R.H. (2016) Cellular heterogeneity mediates inherent sensitivity-specificity tradeoff in cancer targeting by synthetic circuits. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 8133-8138.
22. Lohmueller, J.J., Armel, T.Z. and Silver, P.A. (2012) A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. *Nucleic acids research*, **40**, 5180-5187.
23. Lienert, F., Torella, J.P., Chen, J.H., Norsworthy, M., Richardson, R.R. and Silver, P.A. (2013) Two- and three-input TALE-based AND logic computation in embryonic stem cells. *Nucleic acids research*, **41**, 9967-9975.
24. Ma, D., Peng, S. and Xie, Z. (2016) Integration and exchange of split dCas9 domains for transcriptional controls in mammalian cells. *Nature communications*, **7**, 13056.
25. Biehs, R., Steinlage, M., Barton, O., Juhasz, S., Kunzel, J., Spies, J., Shibata, A., Jeggo, P.A. and Lobrich, M. (2017) DNA Double-Strand Break Resection Occurs during Non-homologous End Joining in G1 but Is Distinct from Resection during Homologous Recombination. *Molecular cell*, **65**, 671-684 e675.
26. Follonier, C., Oehler, J., Herrador, R. and Lopes, M. (2013) Friedreich's ataxia-associated GAA repeats induce replication-fork reversal and unusual molecular junctions. *Nature structural & molecular biology*, **20**, 486-494.
27. Lehman, I.R. (1974) DNA ligase: structure, mechanism, and function. *Science*, **186**, 790-797.
28. Lindahl, T. and Barnes, D.E. (1992) Mammalian DNA ligases. *Annual review of biochemistry*, **61**, 251-281.
29. Davis, A.J., Chen, B.P. and Chen, D.J. (2014) DNA-PK: a dynamic enzyme in a versatile DSB repair pathway. *DNA repair*, **17**, 21-29.
30. Yan, S., Fu, Q., Zhou, Y., Wang, J., Liu, Y., Duan, X., Jia, S., Peng, J., Gao, B., Du, J. *et al.* (2012) High levels of gene expression in the hepatocytes of adult mice, neonatal mice and tree shrews via retro-orbital sinus hydrodynamic injections of naked plasmid DNA. *Journal of controlled release : official journal of the Controlled Release Society*, **161**, 763-771.

31. Baumann, P. and West, S.C. (1998) Role of the human RAD51 protein in homologous recombination and double-stranded-break repair. *Trends in biochemical sciences*, **23**, 247-251.
32. Brophy, J.A. and Voigt, C.A. (2014) Principles of genetic circuit design. *Nature methods*, **11**, 508-520.
33. Slomovic, S., Pardee, K. and Collins, J.J. (2015) Synthetic biology devices for in vitro and in vivo diagnostics. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 14429-14435.
34. Nissim, L., Wu, M.R., Pery, E., Binder-Nissim, A., Suzuki, H.I., Stupp, D., Wehrspaun, C., Tabach, Y., Sharp, P.A. and Lu, T.K. (2017) Synthetic RNA-Based Immunomodulatory Gene Circuits for Cancer Immunotherapy. *Cell*, **171**, 1138-1150 e1115.
35. Tonegawa, S. (1988) Nobel lecture in physiology or medicine--1987. Somatic generation of immune diversity. *In vitro cellular & developmental biology : journal of the Tissue Culture Association*, **24**, 253-265.
36. Weltman, J.K. (1988) The 1987 Nobel Prize for physiology or medicine awarded to molecular immunogeneticist Susumu Tonegawa. *Allergy proceedings : the official journal of regional and state allergy societies*, **9**, 575-576.
37. Schatz, D.G. (2004) V(D)J recombination. *Immunological reviews*, **200**, 5-11.
38. Komori, T., Okada, A., Stewart, V. and Alt, F.W. (1993) Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. *Science*, **261**, 1171-1175.
39. Nadel, B. and Feeney, A.J. (1997) Nucleotide deletion and P addition in V(D)J recombination: a determinant role of the coding-end sequence. *Molecular and cellular biology*, **17**, 3768-3778.
40. Ma, Y., Pannicke, U., Schwarz, K. and Lieber, M.R. (2002) Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell*, **108**, 781-794.

Figure Legends

Figure 1. Two-input ldsDNA-based AND computation *in vitro*. (A) Schematic of 3 sets of pcDNA3.0-firefly-luciferase vector derived two-input ldsDNAs-based AND-gate. 1st set: input 1 as CMV-Luc[1-205], input 2 as Luc[206-1653]-BGH poly(A) signal (PAS); 2nd set: input 1 as CMV-Luc[1-158], input 2 as Luc[159-1653]-BGH PAS; 3rd set: input 1 as CMV, input 2 as Luc[1-1653]-BGH PAS. (B) Indicated two-input ldsDNA-based genetic circuits were introduced into HEK293T cells for 48 h. Luciferase activity was normalized to co-transfected Renilla signals. (C) qRT-PCR analysis of DNA-PKcs siRNA effect. (D) Relative firefly luciferase output activities of indicated AND calculation (state [1,1]) in DNA-PKcs knockdown cells. (E) Schematic of pEGFP-C1 vector derived two-input ldsDNA-based AND-gate. Input 1 as CMV promoter, input 2 as GFP-SV40 PAS. (F) Indicated two-input ldsDNAs were transfected into HEK293T cells for 48 h, and GFP positive rates of states [0,0], [1,0], [0,1] and [1,1] were analyzed by FACS, respectively. (G) Statistics of FACS data in (F). (H) Western blots analysis of GFP expression under indicated states. Error bars indicate standard deviation from at least three biological replicates. $^{**}p < 0.01$.

Figure 2. Two-input ldsDNA-based AND computation *in vivo*. (A-B) Different sets of ldsDNA inputs were introduced into mice via hydrodynamic injection through retro-orbital sinus, and firefly luciferase bioluminescence imaging (BLI) were performed at indicated time points. Representative images of indicated time points were shown in (A). Signals were

quantified in units of maximum photons per second (photons/s), and the statistics was presented as means ± SEM in (B). (C-D) Mice were euthanized one day after hydrodynamic injections of ldsDNAs, and firefly luciferase BLI was performed with different organs. Representative images were shown in (C). Signals were quantified in units of units of maximum photons per second (photons/s), and the statistics was presented as means ± SEM in (D).

Figure 3. Three-input ldsDNA-based AND computation. (A) Schematic of three-input ldsDNA-based AND-gate. Input 1 as CMV promoter, input 2 as GFP CDS, input 3 as SV40-PAS. (B) Indicated three-input ldsDNAs were transfected into HEK293T cells for 48 h, and GFP positive rates were analyzed by FACS. (C) Statistics of FACS data in (B). Error bars indicate standard deviation from three biological replicates. **$p < 0.01$.

Figure 4. AND computation by ldsDNA with additional terminal nucleotide(s). (A) Schematic of two-input AND-gate computation using ldsDNAs with additional terminal nucleotide(s). Input 1 as CMV-GFP[1-204], input 2 as GFP[205-717]-SV40 PAS, N-GFP[205-717]-SV40 PAS or NN-GFP[205-717]-SV40 PAS. (B) Indicated two-input ldsDNAs with or without additional terminal random nucleotide(s) were transfected into HEK293T cells for 48 h, and GFP positive rates under indicated states were analyzed by FACS. (C) Statistics of FACS data in (B). Error bars indicate standard deviation from three biological replicates. **$p < 0.01$.

Figure 5. AND computation by ldsDNAs with identical overlapping sequences. (A) Schematic of 4 sets of pcDNA3.0-firefly-luciferase vector derived two-input AND-gate by ldsDNA with identical overlapping sequence. 1st set: input 1 as CMV-Luc[1-205], input 2 as Luc[159-1653]-BGH PAS; 2nd set: input 1 as CMV-Luc[1-254], input 2 as Luc[159-1653]-BGH PAS; 3rd set: input 1 as CMV-Luc[1-354], input 2 as Luc[159-1653]-BGH PAS; 4th set: input 1 as CMV-Luc[1-554], input 2 as Luc[159-1653]-BGH PAS. (B) Indicated two-input ldsDNA pairs were transfected into HEK293T cells for 48 h, and firefly luciferase output signals were measured and normalized to co-transfected Renilla signals. (C) qRT-PCR analysis of Rad51 siRNA knockdown efficiency. (D) Relative firefly luciferase output activities of indicated AND calculation (state [1,1]) in Rad51 knockdown cells. (E) Schematic of pEGFP-C1 vector derived two-input AND-gate by ldsDNAs with identical overlapping sequence. Input 1 as CMV-GFP[1-187], input 2 as GFP[1-717]-SV40 PAS. (F) Indicated ldsDNAs were transfected into HEK293T cells for 48 h, and GFP positive rates of states [0,0], [1,0], [0,1] and [1,1] were analyzed by FACS, respectively. (G) Statistics of FACS data in (F). Error bars indicate standard deviation from at least three biological replicates. **$p < 0.01$.

Figure 6. AND computation by plasmids with identical overlapping sequences. (A) Schematic of plasmids with identical overlapping sequences: pMD-18T-CMV-Luc 5' represents 3 plasmids in which CMV-Luc[1-205], CMV-Luc[1-254] and CMV-Luc[1-354]

fragment was subcloned into pMD-18T, respectively; and pMD-18T-Luc 3'-BGH represents pMD-18T plasmid with Luc[159-1653]-BGH PAS fragment. (B) Different combinations of plasmids with identical overlapping sequences were transfected into HEK293T cells as indicated, and firefly luciferase output signals under different states were measured and normalized to co-transfected Renilla signals. (C) Relative firefly luciferase activities of indicated AND calculation (state [1,1]) in Rad51 knockdown cells. (D) Schematic of plasmids with identical overlapping sequences: pMD-18T-CMV-GFP 5' represents 2 plasmids in which CMV-GFP[1-303] and CMV-GFP[1-503] fragment was subcloned into pMD-18T respectively; and pMD-18T-GFP 3'-SV40 represents pMD-18T plasmid with GFP[104-717]-SV40 PAS fragment. (E) Different combinations of plasmids with identical overlapping sequences were transfected into HEK293T cells as indicated, and GFP positive rates under different states were analyzed by FACS. (F) Statistics of FACS data in (E). Error bars indicate standard deviation from at least three biological replicates. **$p < 0.01$.

Supplementary Figure S1. Two-input ldsDNA-based AND-gate output signal normalized to protein concentrations. Indicated ldsDNAs were introduced into HEK293T cells via transfection. Firefly luciferase output signals of states [0,0], [1,0], [0,1] and [1,1] were measured and normalized to protein concentrations. The experiments were performed simultaneously with that in Figure 1B. Error bars indicate standard deviation from four biological replicates. **$p < 0.01$.

Supplementary Figure S2. Two-input AND computation with different amounts of ldsDNAs. Different amounts of ldsDNAs were transfected into HEK293T cells as indicated. Firefly luciferase output signals were measured and normalized to co-transfected Renilla signals. Error bars indicate standard deviation from four biological replicates. **$p < 0.01$.

Supplementary Figure S3. Two-input AND-gate computation with Cy5 labeled ldsDNA. (A) Cy5 labeled ldsDNA Cy5-CMV (Input 1) and non-labeled GFP-SV40 PAS (Input 2) were employed to perform AND-gate computation. Indicated ldsDNAs were transfected into HEK293T cells, and Cy5/GFP positive rates of states [0,0], [1,0], [0,1] and [1,1] were analyzed by FACS. (B) Statistics of GFP positive and negative percentage in Cy5+ cells. Error bars indicate standard deviation from three biological replicates. **$p < 0.01$.

Supplementary Figure S4. Two-input AND computation with 5' terminal phosphated ldsDNAs. (A) Schematic of 2 sets of two-input AND-gate 5' terminal phosphated ldsDNAs amplified from pcDNA3.0-firefly luciferase plasmid. 1st set: input 1-P as CMV-Luc[1-205]-P, P-input 2 as P-Luc[206-1653]-BGH PAS; 2nd set: input 1-P as CMV-Luc[1-158]-P, P-input 2 as P-Luc[159-1653]-BGH PAS. (B) Indicated 5' terminal phosphated ldsDNAs were transfected into HEK293T cells, and firefly luciferase output signals of states [0,0], [1,0], [0,1] and [1,1] were measured and normalized to co-transfected Renilla signals. Error bars indicate standard deviation from four biological replicates. **$p < 0.01$.

Supplementary Figure S5. Detection of DNA damage foci formation in cells transfected with ldsDNAs. HEK293T cells were transfected with ldsDNA-based AND-gate genetic circuits as indicated. 48h later, immunofluorescence staining of DNA damage foci markers, γ-H2AX and 53BP1, was performed. Cells undergone 10Gy γ-radiation served as positive controls. Representative images of γ-H2AX and 53BP1 staining were shown in (A) and (B), respectively.

Supplementary Figure S6. Expression of firefly luciferase *in vivo* via hydrodynamic injection of CMV-Luc-BGH PAS ldsDNA. (A-B) Bioluminescence imaging (BLI) of firefly luciferase in mice undergone CMV-Luc-BGH PAS ldsDNA hydrodynamic injection. Representative images at different time points were shown in (A). Signals were quantified in units of maximum photons per second (photons/s), and the statistics was shown in (B). (n=4 mice.) (C) Mice were euthanized one day after hydrodynamic injections, and firefly luciferase BLI was performed with different organs. Representative images were shown here.

Supplementary Figure S7. Two-input AND computation by different amounts of ldsDNAs with identical overlapping sequences. Different amounts of two-input ldsDNAs with identical overlapping sequences were transfected into HEK293T cells as indicated. Firefly luciferase output signals were measured and normalized to co-transfected Renilla signals. Error bars indicate standard deviation from four biological replicates. $^{**}p < 0.01$.

Supplementary Figure S8. Design of back-to-back promoter ldsDNAs for ligation efficiency improvement. (A) Using single-promoter ldsDNAs strategy, one-fourth of all ldsDNA ligations generates the right gene-expression cassette. (B) Using back-to-back promoter ldsDNA strategy, one half of all ldsDNA ligation generates the right gene-expression cassette.

Supplementary Figure S9. Design of ldsDNA-based AND-gate nanoparticle delivery strategy for targeting specificity improvement. Through the recognition of corresponding cell surface antigens, antibody-coated nanoparticles deliver different ldsDNA inputs into the targeted cells thereby achieving cell specific gene expression.

Supplementary Figure S10. Generation of highly diversified antibody library by ldsDNA-based AND-gate calculation. ldsDNAs corresponding to antibody V and D gene fragments are amplified. By ldsDNA end processing and fragments random combination, ldsDNA-based AND-gate computation could aid in generating highly diversified antibody library.

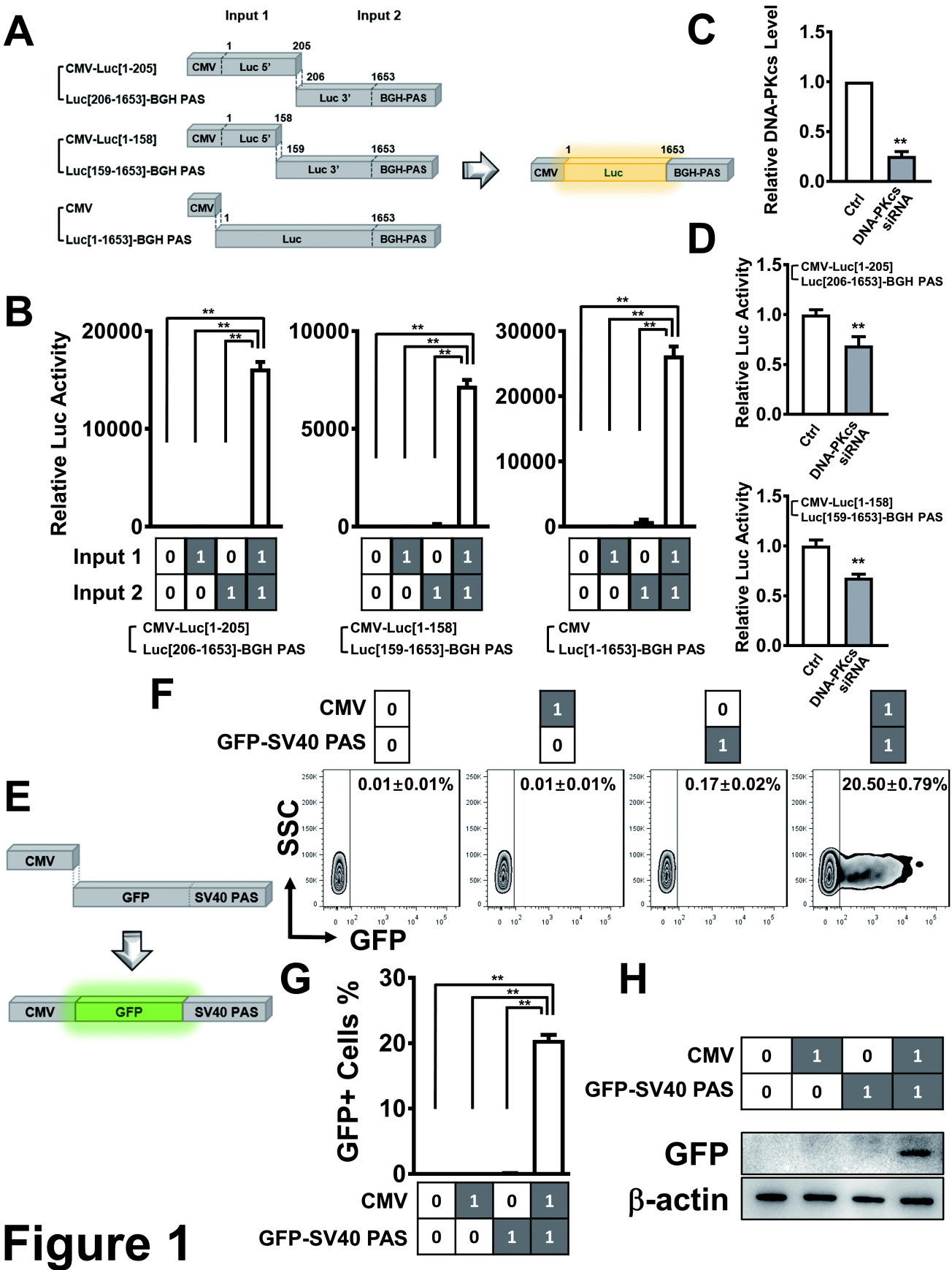
A
Input 1
Input 2
CMV-Luc[1-205]
Luc[206-1653]-BGH PAS
CMV-Luc[1-158]
Luc[159-1653]-BGH PAS
CMV
Luc[1-1653]-BGH PAS
CMV
Luc 5'
Luc 3'
Luc
BGH-PAS
B
Relative Luc Activity
Input 1
Input 2
C
Relative DNA-PKcs Level
Ctrl
DNA-PKcs siRNA
D
Relative Luc Activity
CMV-Luc[1-205]
Luc[206-1653]-BGH PAS
CMV-Luc[1-158]
Luc[159-1653]-BGH PAS
E
GFP
SV40 PAS
F
CMV
GFP-SV40 PAS
SSC
GFP
0.01±0.01%
0.01±0.01%
0.17±0.02%
20.50±0.79%
G
GFP+ Cells %
H
GFP
β-actin

Figure 1

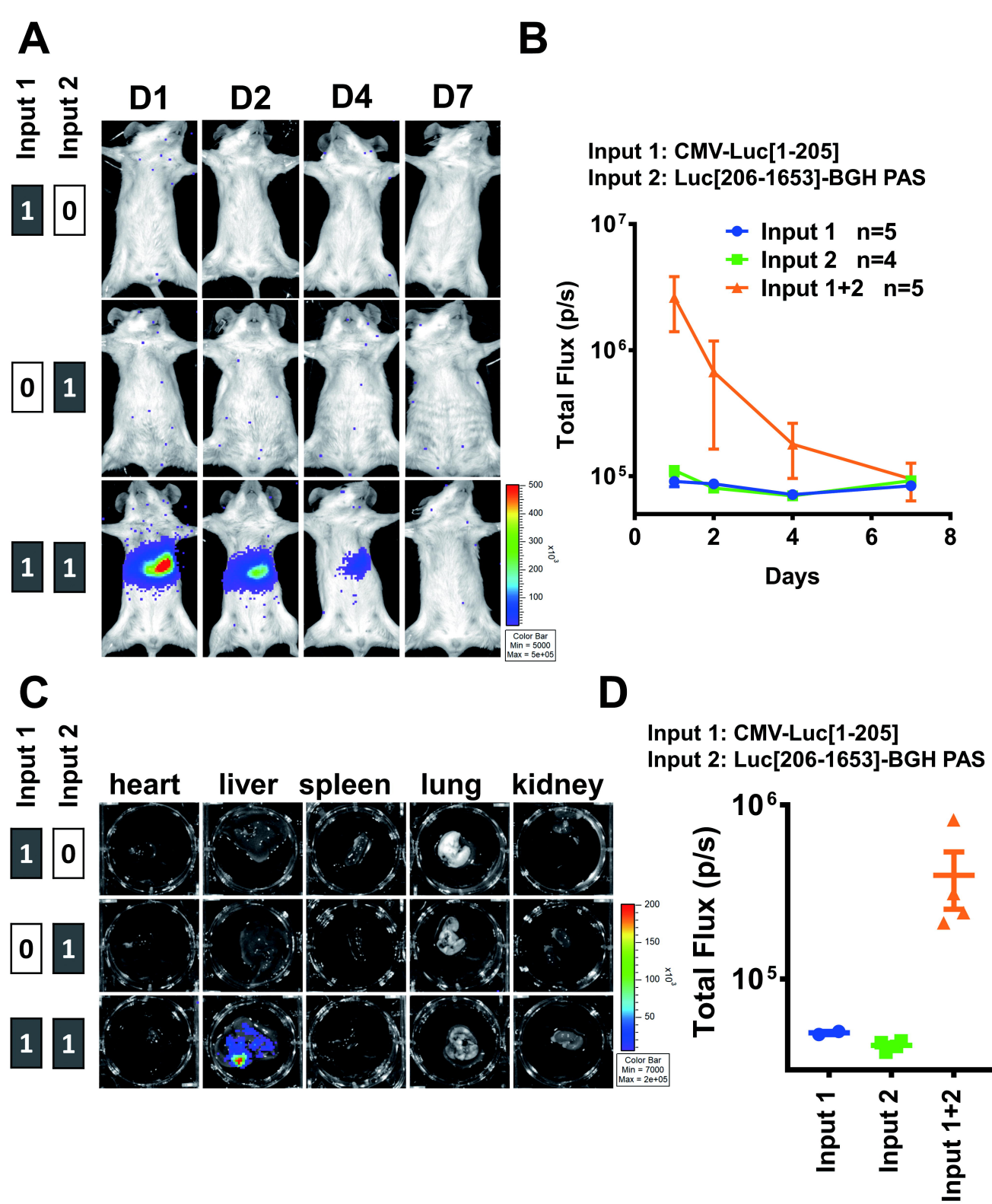
A
Input 1
Input 2
D1
D2
D4
D7
1
0
0
1
1
1
500
400
300
200
100
x10^3
Color Bar
Min = 5000
Max = 5e+05
B
Input 1: CMV-Luc[1-205]
Input 2: Luc[206-1653]-BGH PAS
Input 1 n=5
Input 2 n=4
Input 1+2 n=5
Total Flux (p/s)
10^7
10^6
10^5
0
2
4
6
8
Days
C
Input 1
Input 2
heart
liver
spleen
lung
kidney
1
0
0
1
1
1
200
150
100
50
x10^3
Color Bar
Min = 7000
Max = 2e+05
D
Input 1: CMV-Luc[1-205]
Input 2: Luc[206-1653]-BGH PAS
Total Flux (p/s)
10^6
10^5
Input 1
Input 2
Input 1+2

Figure 2

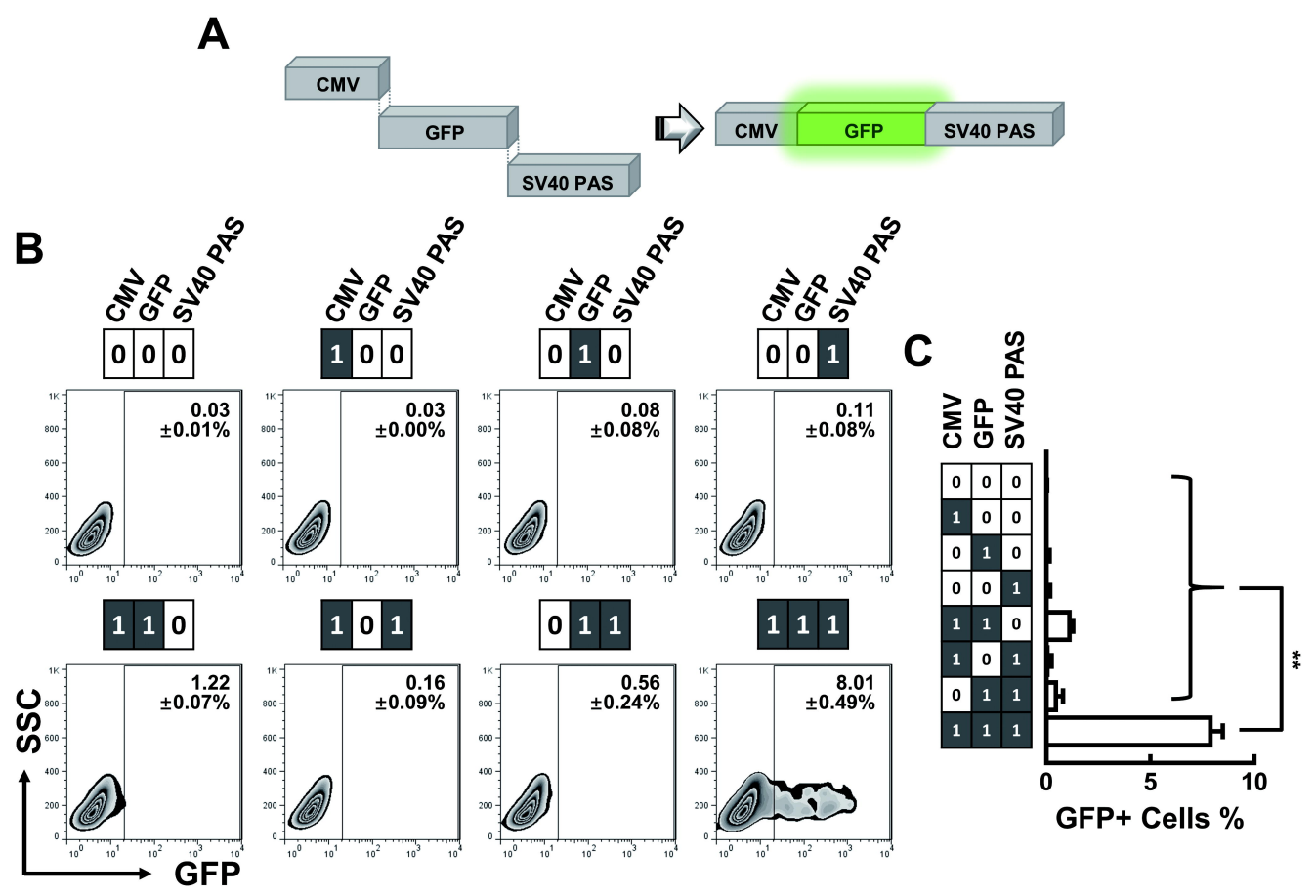
A
CMV
GFP
SV40 PAS
CMV
GFP
SV40 PAS
B
CMV GFP SV40 PAS
0 0 0
0.03 ±0.01%
1 0 0
0.03 ±0.00%
0 1 0
0.08 ±0.08%
0 0 1
0.11 ±0.08%
1 1 0
1.22 ±0.07%
1 0 1
0.16 ±0.09%
0 1 1
0.56 ±0.24%
1 1 1
8.01 ±0.49%
SSC
GFP
C
CMV GFP SV40 PAS
0 0 0
1 0 0
0 1 0
0 0 1
1 1 0
1 0 1
0 1 1
1 1 1
**
0
5
10
GFP+ Cells %

Figure 3

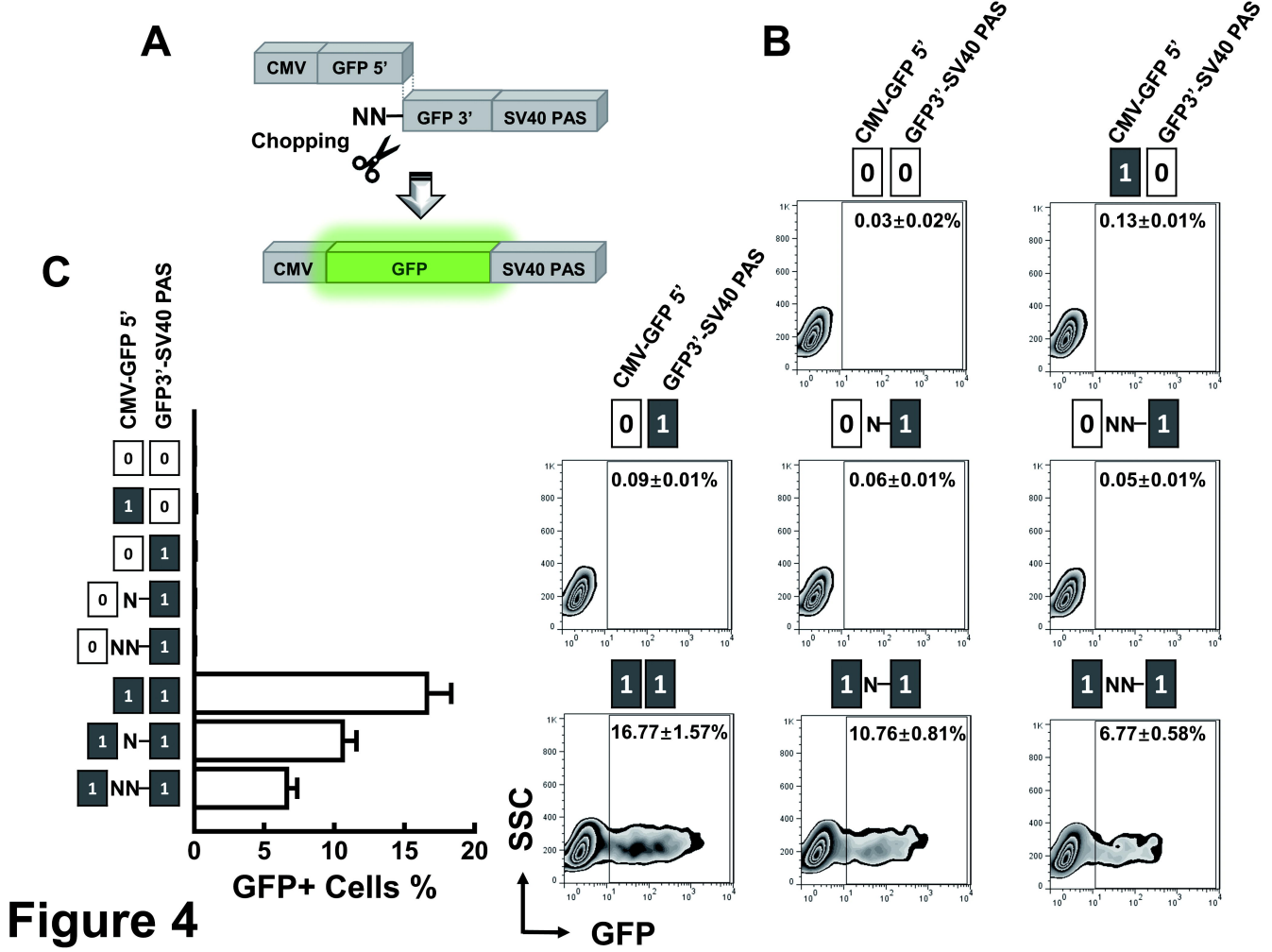
A
CMV
GFP 5'
NN
GFP 3'
SV40 PAS
Chopping
CMV
GFP
SV40 PAS
B
CMV-GFP 5'
GFP3'-SV40 PAS
0.03±0.02%
0.13±0.01%
0.09±0.01%
0.06±0.01%
0.05±0.01%
16.77±1.57%
10.76±0.81%
6.77±0.58%
SSC
GFP
C
CMV-GFP 5'
GFP3'-SV40 PAS
GFP+ Cells %
0
5
10
15
20

Figure 4

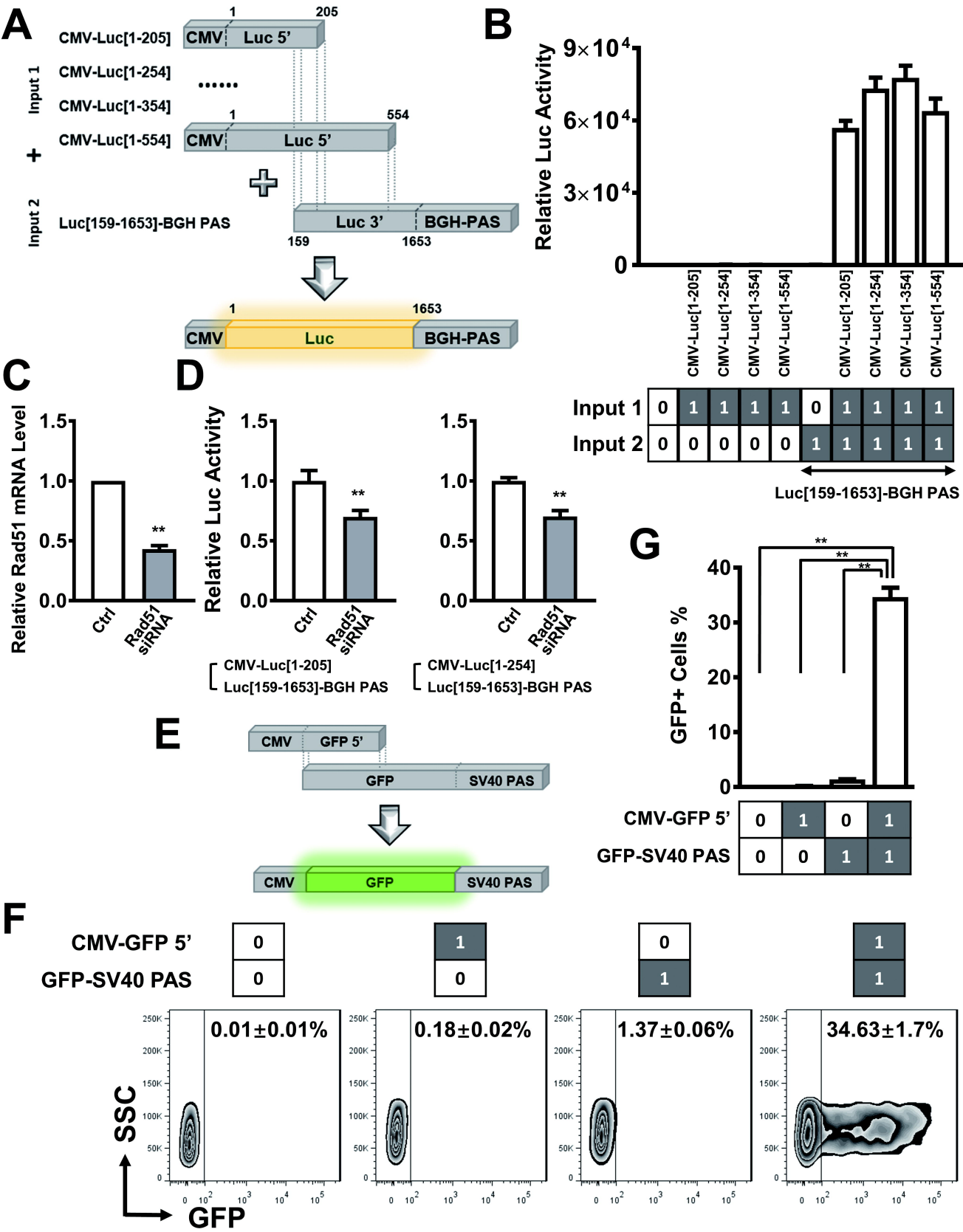
A
Input 1
CMV-Luc[1-205]
CMV-Luc[1-254]
CMV-Luc[1-354]
CMV-Luc[1-554]
+
Input 2
Luc[159-1653]-BGH PAS
CMV
Luc 5'
Luc 3'
BGH-PAS
Luc
1
205
554
159
1653
B
Relative Luc Activity
9×10^4
6×10^4
3×10^4
0
Input 1
Input 2
Luc[159-1653]-BGH PAS
C
Relative Rad51 mRNA Level
Ctrl
Rad51 siRNA
D
Relative Luc Activity
CMV-Luc[1-205]
Luc[159-1653]-BGH PAS
CMV-Luc[1-254]
Luc[159-1653]-BGH PAS
E
CMV
GFP 5'
GFP
SV40 PAS
G
GFP+ Cells %
CMV-GFP 5'
GFP-SV40 PAS
F
CMV-GFP 5'
GFP-SV40 PAS
0.01±0.01%
0.18±0.02%
1.37±0.06%
34.63±1.7%
SSC
GFP

Figure 5

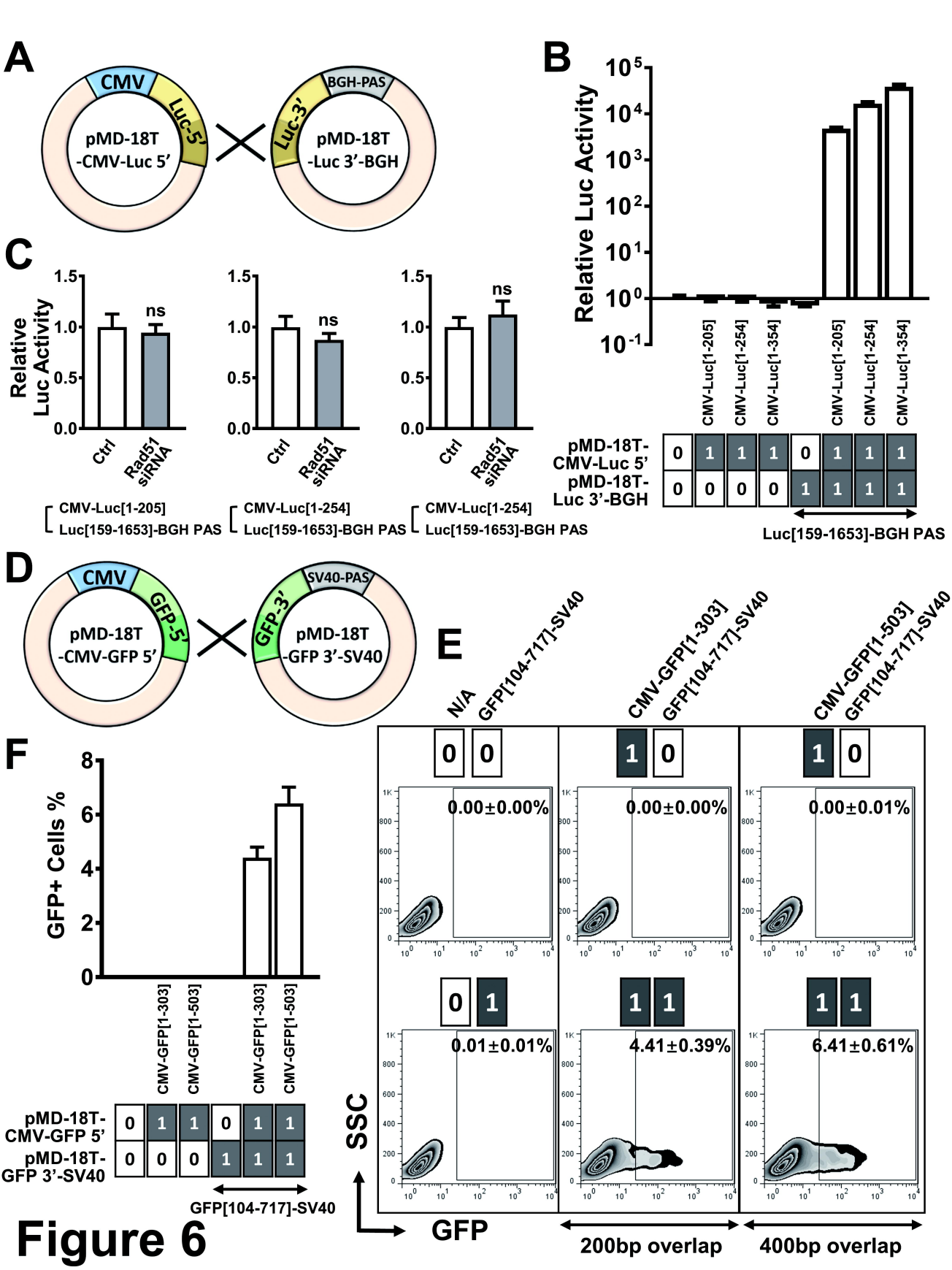
A
CMV
Luc-5'
pMD-18T
-CMV-Luc 5'
BGH-PAS
Luc-3'
pMD-18T
-Luc 3'-BGH
B
Relative Luc Activity
CMV-Luc[1-205]
CMV-Luc[1-254]
CMV-Luc[1-354]
pMD-18T-
CMV-Luc 5'
pMD-18T-
Luc 3'-BGH
Luc[159-1653]-BGH PAS
C
Relative
Luc Activity
ns
Ctrl
Rad51
siRNA
CMV-Luc[1-205]
Luc[159-1653]-BGH PAS
CMV-Luc[1-254]
Luc[159-1653]-BGH PAS
CMV-Luc[1-254]
Luc[159-1653]-BGH PAS
D
CMV
GFP-5'
pMD-18T
-CMV-GFP 5'
SV40-PAS
GFP-3'
pMD-18T
-GFP 3'-SV40
E
N/A
GFP[104-717]-SV40
CMV-GFP[1-303]
CMV-GFP[1-503]
0.00±0.00%
0.00±0.00%
0.00±0.01%
0.01±0.01%
4.41±0.39%
6.41±0.61%
SSC
GFP
200bp overlap
400bp overlap
F
GFP+ Cells %
CMV-GFP[1-303]
CMV-GFP[1-503]
pMD-18T-
CMV-GFP 5'
pMD-18T-
GFP 3'-SV40
GFP[104-717]-SV40

Figure 6