Linear-double-stranded DNA (ldsDNA) Based AND Logic Computation in Mammalian Cells

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

1. 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;

2. School of Medicine, Nankai University, Tianjin 300071, China;

3. Department of Colorectal Surgery, Tianjin Union Medical Center, Tianjin 300121, China.

Contribute equally.

* 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 systems for biomedical or industrial purposes. Innovation and application of ingenious biological parts for genetic circuit construction will greatly facilitate and expedite the development of synthetic biology. We here demonstrate that linear-double-stranded DNA (ldsDNA) can perform two- or three input Boolean AND gate logic computation in mammalian cells. Bioluminescence living imaging reveals the feasibility of ldsDNA-based AND-gate computation *in vivo*. Inhibition of DNA-PKcs, a pivotal enzyme in non-homologous end joining (NHEJ), could significantly attenuate output signals from ldsDNA-based AND calculation. We further show that ldsDNA with terminal additional random nucleotide(s) could undergo end nucleotide(s) deletion and generate in-frame protein via AND logic computation. Additionally, ldsDNAs or plasmids with identical overlapping sequence could also serve as inputs for AND-gate calculation. Our work establishes ldsDNA as innovative biological parts for building low noise-signal ratio AND gate circuits with great application potential in targeting therapy and in protein diversity generation methodology.

Keywords

synthetic biology, artificial genetic circuit, AND gate, linear-double-stranded DNA (ldsDNA), non-homologous end joining (NHEJ).

Introduction

One major focus of synthetic biology is on designing genetic circuits for biomedical or industrial purposes (*1-3*). To date, a series of biological parts of different underlying composition (DNA, RNA, or protein) have been developed to function as basic building blocks of genetic circuits. DNA is commonly used as biological parts, working as either regulator or adaptor in genetic circuits (*4*). For instance, synthetic transcriptional factors could bind to specific DNA sequence and turn on gene expression (*5-9*). DNA recombinases could recognize recombination DNA sequence and catalyze recombination, thereby accomplishing sophisticated bio-computation (*10, 11*). Innovation of new DNA-based biological parts would greatly propel genetic circuit design.

Similar to complex electronic circuits, genetic circuits accomplish biocomputing via integrating different logic gates into networks. Among these logic gates, AND gate converts two (or more) input signals into one output signal by means of Boolean calculation, which can be exploited to enhance diagnostic or therapeutic specificity in biomedical studies (*12-15*). Over the past two decades, various AND logic gate design strategies have been developed to achieve delicate control of gene expression. These strategies could be divided into two subtypes: ‘circuit in place’ and ‘bio-parts self-assembly'. For commonly used ‘circuit in place’ strategy, inputs signal integration is executed in the presence of all computation modules. For ‘bio-parts self-assembly’ AND gate strategy, TURE output signal could be generated only when all inputs (or their products) are assembled together as one output-signal-generating molecule (*16-18*).

Here, we apply linear-double-stranded DNA (ldsDNA, or ‘PCR amplicon’) as innovative biological parts to implement Boolean logic AND gates. Via splitting essential gene expression cassette into

signal-mute ldsDNA, this strategy achieves AND gate calculation in mammalian cells with a low noise-signal ratio both *in vitro* and *in vivo*. Additionally, ldsDNA with one or two additional terminal nucleotide(s), which leads to reading frame shift, could undergo end-processing and produce in-frame AND-gate output proteins. Moreover, ldsDNAs or plasmids with identical overlapping sequence are also able to perform AND-gate computation. Our present study provides novel biological parts and principles for synthetic genetic circuits design, which may have great application potential in biomedical engineering fields.

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 1% penicillin-streptomycin solution 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 ldsDNA (PCR amplicons) taking pcDNA3.0-firefly-luciferase or pEGFP-C1 plasmid as template. To remove plasmid templates and free dNTPs, PCR products underwent agarose electrophoresis and gel-purification (Universal DNA purification kit, Tiangen). For animal experiments, ldsDNA was 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). Protein concentrations were determined with Pierce BCA Protein Assay Kit (Thermo Fisher) according to manufacturer's instruction. Firefly luciferase activities (Relative luminescence unit, RLU) were normalized either to Renilla luciferase activities (RLU) or to protein concentrations (μg/μl).

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 with confocal laser scanning microscope (Olympus).

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) (*19*), Rad51 (Rad51 siRNA, GAG CUU GAC AAA CUA CUU CTT) (*20*) 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: GAPDH-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 ldsDNA

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 ldsDNA. 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 ldsDNA. 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 (for mammalian cells) 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). Raw firefly luciferase relative luminescence unit (RLU) values were present in Supplementary Figure S2. And

there is a positive correlation between the amount of ldsDNA inputs and the intensity of output signals (Supplementary Figure S3).

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).

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 dsDNA (*21, 22*). 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 ldsDNA 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 ldsDNA (Figure 1B and Supplementary Figure S4).

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

Since 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) (*23*). As shown in

Figure 1D, the inhibition of DNA-PKcs significantly downregulates ldsDNA-based AND-gate output signals.

Introduction of ldsDNA does not lead to DNA damage foci formation

Considering its potential applications in biological treatments, we investigated 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 formed 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 (*24*). First, we confirmed the validity of this method by rapidly injecting a large volume of CMV-Luc-BGH PAS ldsDNA solution (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) (*24*). 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 ldsDNA supplemented with terminal additional 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 ldsDNA was 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 ldsDNA 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 ldsDNA 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 ldsDNA pairs 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 VS 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) (*25*). 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). These differences imply that the mechanisms underlying AND computation with overlapping ldsDNAs and overlapping plasmids would be different, and need further exploration.

Discussion

Here by splitting core gene expression cassette into signal-mute ldsDNA inputs, we present a proof-of-concept demonstration that ldsDNA can be utilized for AND logic computation in mammalian cells. In our study, both firefly luciferase and GFP expression could be generated by ldsDNA-based AND computation *in vitro*. Furthermore, bioluminescence living imaging of firefly luciferase also demonstrated the feasibility of AND-gate calculation *in vivo*. We also revealed that ldsDNA with terminal additional random nucleotide(s) could undergo end nucleotide(s) deletion and generate in-frame protein via AND logic calculation. Moreover, ldsDNAs or plasmids with identical overlapping sequence could also form AND logic circuits in mammalian cells. This study provides novel biological parts and principles for synthetic genetic circuits design, which have great application potentials in future experimental and clinical researches.

Ingenious biological parts, well-characterized principles and in-depth understanding of circuit composition all give rise to the development of genetic circuit design. Principles based on various biological parts, e.g. recombinases, DNA binding proteins or RNA-based elements have been developed to conduct AND logic calculation (*26*). Our study, for the first time, suggested the utilization of ldsDNA as novel biological parts in synthetic circuit design, whose straightforwardness and accessibility would greatly promote future broad applications. Considering that double strand DNA (dsDNA) is universally used as storage medium of genetic information in biological world, our design leads genetic circuit construction back to the essence. Moreover, via splitting essential gene-expression cassette, in our circuit design individual input only produces signal equivalent to background noise, thus achieving an excellent signal-to-noise ratio. Additionally, with this

split/combination strategy, strong or weak promoters can be conveniently assembled onto the 5' end of gene CDS module accordingly, working as rockets to propel transcription of downstream genes.

Though with these advantages, this ldsDNA-based circuit design has room for improvement nonetheless. In the current version, luciferase reporter signal levels generated by our principle are still lower than that from ldsDNA with intact gene-expression cassette (Supplementary Figure S8). One limiting factor to efficiency is the intermediary process. As shown in our results, NHEJ machinery members exemplified by DNA-PKcs participate in output signal generation, implying that ldsDNA undergoes a rejoining process prior to the formation of full length productive gene-expression cassette.

One major application of synthetic Boolean logic gate in biomedical research is to discriminate cell types or disease-specific conditions, thus improve diagnostic or therapeutic specificity (*12-14, 27*). To achieve the goal, one approach is to indiscriminately introduce genetic circuits into cells and specifically turn on gene expression in cells with particular signature. For instance, Lior *et al* developed RNA-based AND-gate genetic circuits assembling different biological parts, including tumor-specific promoters, miRNA sponge and miRNA-target regulation, which specifically output immunomodulator in tumor cells (*28*). Another approach is to sense cell membrane antigens and take them as AND-gate inputs to active gene expression, e.g. Roybal *et al* engineered T cells to recognize antigen combination and thereby kill targeted tumor cells (*14*). Taking the latter route, our novel principle can be applied to improve targeting specificity of both diagnosis and therapy. By way of illustration, as shown in Supplementary Figure S9, nanoparticles containing different ldsDNA-based

AND-gate inputs are designed to target different cell surface antigens with antibodies immobilized on their surfaces, thus achieving specific gene expression output in targeted cells.

Another potential application of ldsDNA-based AND logic calculation is in the construction of highly diversified antibody repertoires (Supplementary Figure S10). Along with B and T cell maturation, variable (V), diversity (D) and joining (J) gene fragments undergo nearly random recombination, making major contributions to the diversity of antibodies/immunoglobulins (Igs) and T cell receptors (TCRs) (*29-31*). Moreover, during VD or DJ junction formation, indel (nucleotides insertion or deletion) introduced by terminal deoxynucleotide transferase (TdT), Artemis nuclease, DNA polymerase or NHEJ pathway members also contributes to antigen-receptor diversity (*32-34*). As shown in our results, three-input ldsDNA-based AND genetic circuits similar to V(D)J recombination process are feasible in mammalian cells (Figure 5), and ldsDNA could undergo end nucleotide(s) processing in the re-joining procedure (Figure 4). According to these features of ldsDNA-based genetic circuit, in our assumption, ldsDNAs corresponding to V, D or J gene fragments are synthesized, and their AND logic computation could 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).

In conclusion, here we developed a novel ldsDNA-based AND logic genetic circuit design strategy. This AND-gate strategy was achieved both *in vitro* and *in vivo* in our study. It was also clarified that ldsDNA with terminal additional nucleotide(s) could undergo end chopping and output right in-frame

AND gate signals. Our study provides brand new biological parts and principles for synthetic biology and opens a door for the generation of functional proteins or antibodies.

Acknowledgements

This work was supported by National Natural Science Foundation of China [31870860, 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.]; and Key Project of National Health and Family Planning Commission of Tianjin [2017057 to C.Z.]. The authors have declared that no competing interest exists.

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.

Competing Interests

The authors have declared that no competing interest exists.

References

1. C. J. Bashor, A. A. Horwitz, S. G. Peisajovich, W. A. Lim, Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. *Annu Rev Biophys* **39**, 515 (2010).
2. F. Lienert, J. J. Lohmueller, A. Garg, P. A. Silver, Synthetic biology in mammalian cells: next generation research tools and therapeutics. *Nat Rev Mol Cell Biol* **15**, 95 (Feb, 2014).

3. A. A. Nielsen, T. H. Segall-Shapiro, C. A. Voigt, Advances in genetic circuit design: novel biochemistries, deep part mining, and precision gene expression. *Curr Opin Chem Biol* **17**, 878 (Dec, 2013).
4. Y. H. Wang, K. Y. Wei, C. D. Smolke, Synthetic biology: advancing the design of diverse genetic systems. *Annu Rev Chem Biomol Eng* **4**, 69 (2013).
5. A. J. Bogdanove, D. F. Voytas, TAL effectors: customizable proteins for DNA targeting. *Science* **333**, 1843 (Sep 30, 2011).
6. B. Gonzalez *et al.*, Modular system for the construction of zinc-finger libraries and proteins. *Nat Protoc* **5**, 791 (Apr, 2010).
7. A. S. Khalil *et al.*, A synthetic biology framework for programming eukaryotic transcription functions. *Cell* **150**, 647 (Aug 3, 2012).
8. S. Konermann *et al.*, Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583 (Jan 29, 2015).
9. F. Zhang *et al.*, Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* **29**, 149 (Feb, 2011).
10. P. Siuti, J. Yazbek, T. K. Lu, Synthetic circuits integrating logic and memory in living cells. *Nat Biotechnol* **31**, 448 (May, 2013).
11. B. H. Weinberg *et al.*, Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nat Biotechnol* **35**, 453 (May, 2017).
12. Y. Liu *et al.*, Synthesizing AND gate genetic circuits based on CRISPR-Cas9 for identification of bladder cancer cells. *Nat Commun* **5**, 5393 (Nov 6, 2014).
13. M. Morel, R. Shtrahman, V. Rotter, L. Nissim, R. H. Bar-Ziv, Cellular heterogeneity mediates inherent sensitivity-specificity tradeoff in cancer targeting by synthetic circuits. *Proc Natl Acad Sci U S A* **113**, 8133 (Jul 19, 2016).
14. K. T. Roybal *et al.*, Precision Tumor Recognition by T Cells With Combinatorial Antigen-Sensing Circuits. *Cell* **164**, 770 (Feb 11, 2016).
15. Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, Y. Benenson, Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science* **333**, 1307 (Sep 2, 2011).
16. F. Lienert *et al.*, Two- and three-input TALE-based AND logic computation in embryonic stem cells. *Nucleic Acids Res* **41**, 9967 (Nov, 2013).
17. J. J. Lohmueller, T. Z. Armel, P. A. Silver, A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. *Nucleic Acids Res* **40**, 5180 (Jun, 2012).
18. D. Ma, S. Peng, Z. Xie, Integration and exchange of split dCas9 domains for transcriptional controls in mammalian cells. *Nat Commun* **7**, 13056 (Oct 3, 2016).
19. R. Biehs *et al.*, DNA Double-Strand Break Resection Occurs during Non-homologous End Joining in G1 but Is Distinct from Resection during Homologous Recombination. *Mol Cell* **65**, 671 (Feb 16, 2017).
20. C. Follonier, J. Oehler, R. Herrador, M. Lopes, Friedreich's ataxia-associated GAA repeats induce replication-fork reversal and unusual molecular junctions. *Nat Struct Mol Biol* **20**, 486 (Apr, 2013).
21. I. R. Lehman, DNA ligase: structure, mechanism, and function. *Science* **186**, 790 (Nov 29, 1974).
22. T. Lindahl, D. E. Barnes, Mammalian DNA ligases. *Annu Rev Biochem* **61**, 251 (1992).

23. A. J. Davis, B. P. Chen, D. J. Chen, DNA-PK: a dynamic enzyme in a versatile DSB repair pathway. *DNA Repair (Amst)* **17**, 21 (May, 2014).
24. S. Yan *et al.*, 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. *J Control Release* **161**, 763 (Aug 10, 2012).
25. P. Baumann, S. C. West, Role of the human RAD51 protein in homologous recombination and double-stranded-break repair. *Trends Biochem Sci* **23**, 247 (Jul, 1998).
26. J. A. Brophy, C. A. Voigt, Principles of genetic circuit design. *Nat Methods* **11**, 508 (May, 2014).
27. S. Slomovic, K. Pardee, J. J. Collins, Synthetic biology devices for in vitro and in vivo diagnostics. *Proc Natl Acad Sci U S A* **112**, 14429 (Nov 24, 2015).
28. L. Nissim *et al.*, Synthetic RNA-Based Immunomodulatory Gene Circuits for Cancer Immunotherapy. *Cell* **171**, 1138 (Nov 16, 2017).
29. D. G. Schatz, V(D)J recombination. *Immunol Rev* **200**, 5 (Aug, 2004).
30. S. Tonegawa, Nobel lecture in physiology or medicine--1987. Somatic generation of immune diversity. *In Vitro Cell Dev Biol* **24**, 253 (Apr, 1988).
31. J. K. Weltman, The 1987 Nobel Prize for physiology or medicine awarded to molecular immunogeneticist Susumu Tonegawa. *Allergy Proc* **9**, 575 (Sep-Oct, 1988).
32. T. Komori, A. Okada, V. Stewart, F. W. Alt, Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. *Science* **261**, 1171 (Aug 27, 1993).
33. Y. Ma, U. Pannicke, K. Schwarz, M. R. Lieber, 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 (Mar 22, 2002).
34. B. Nadel, A. J. Feeney, Nucleotide deletion and P addition in V(D)J recombination: a determinant role of the coding-end sequence. *Mol Cell Biol* **17**, 3768 (Jul, 1997).

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 ldsDNA-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 ldsDNA was 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) was 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 ldsDNA, and firefly luciferase BLI was performed with different organs. Representative images were shown in (C). Signals were quantified in 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 ldsDNA was 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 random terminal additional nucleotide(s). (A) Schematic of two-input AND-gate computation using ldsDNA with terminal additional 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 terminal additional 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. The experiments were performed simultaneously with that in Figure 1B. (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 ldsDNA was 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 Legends

Supplementary Figure S1. Two-input ldsDNA-based AND-gate output signal normalized to protein concentrations. Indicated ldsDNA was 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 ldsDNA-based AND-gate output signal shown by raw RLU (relative luminescence unit). Indicated ldsDNA was introduced into HEK293T cells via transfection. Firefly luciferase output signals of states [0,0], [1,0], [0,1] and [1,1] were measured and raw RLU values were present. Error bars indicate standard deviation from four biological replicates. **$p < 0.01$.

Supplementary Figure S3. Two-input AND computation with different amount of ldsDNA. Different amount of ldsDNA was 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 S4. Two-input AND computation with 5' terminal phosphated ldsDNA. (A) Schematic of 2 sets of two-input AND-gate 5' terminal phosphated ldsDNA 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 ldsDNA was 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. The experiments were performed simultaneously with that in Figure 1B. 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 ldsDNA. 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 amount of ldsDNAs with identical overlapping sequences. Different amount 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. Comparison of output signals between ldsDNA-based two-input AND computation and ldsDNA with intact gene expression cassette. Indicated ldsDNAs were introduced into HEK293T cells via transfection. Firefly luciferase output signals from different AND-gate states [1,1] or ldsDNA with intact gene expression cassette (full, CMV-luc[1-1653]-BGH PAS) were measured and the RLU values were shown. Error bars indicate standard deviation from four biological replicates. $^{**}p < 0.01$.

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.

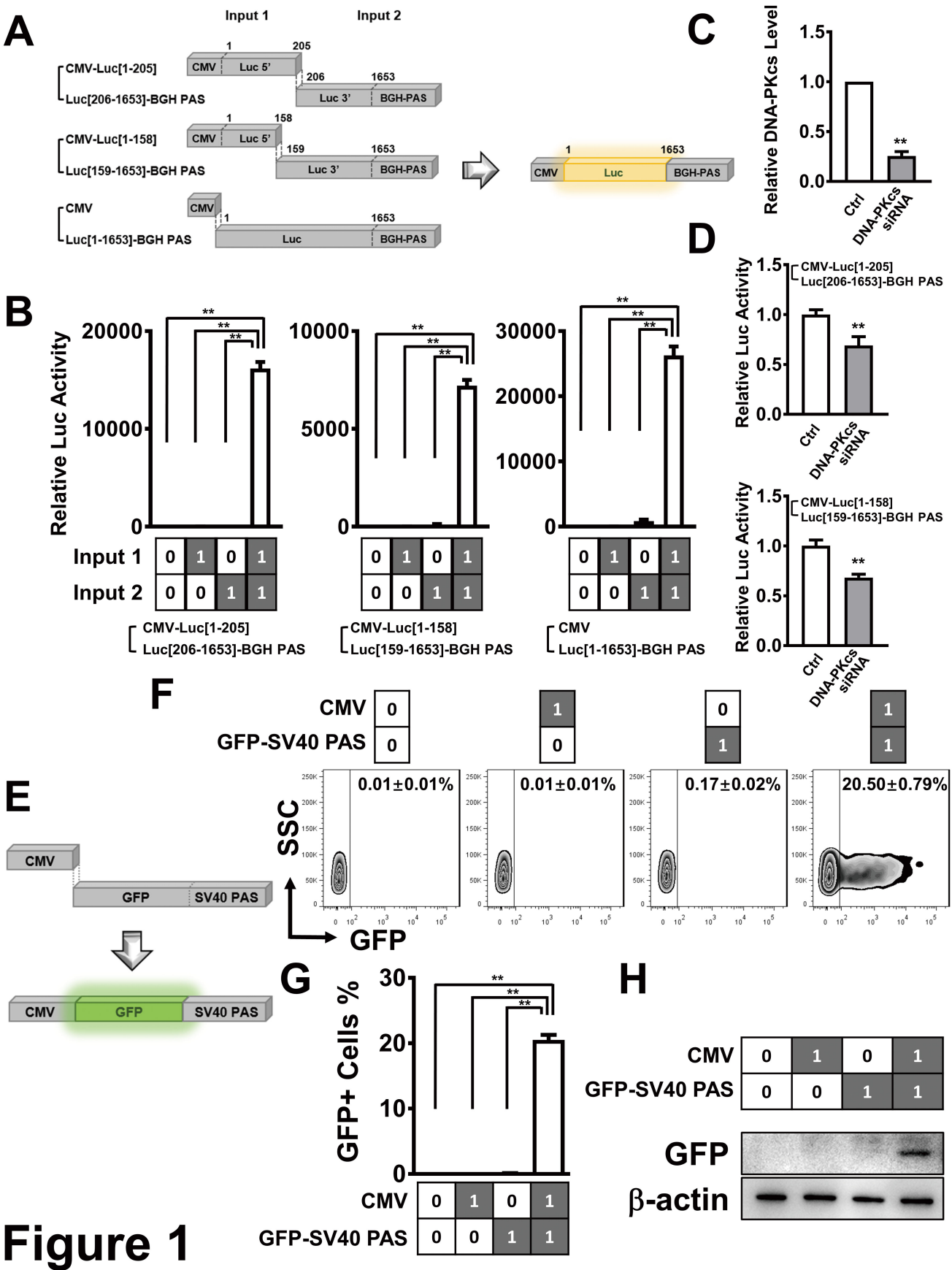

A
Input 1
Input 2
CMV-Luc[1-205]
Luc[206-1653]-BGH PAS
CMV-Luc[1-158]
Luc[159-1653]-BGH PAS
CMV
Luc[1-1653]-BGH PAS
CMV
Luc 5'
Luc 3'
BGH-PAS
Luc
B
Relative Luc Activity
Input 1
Input 2
C
Relative DNA-PKcs Level
Ctrl
DNA-PKcs siRNA
D
Relative Luc Activity
CMV-Luc[1-205]
Luc[206-1653]-BGH PAS
CMV-Luc[1-158]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

Figure 1

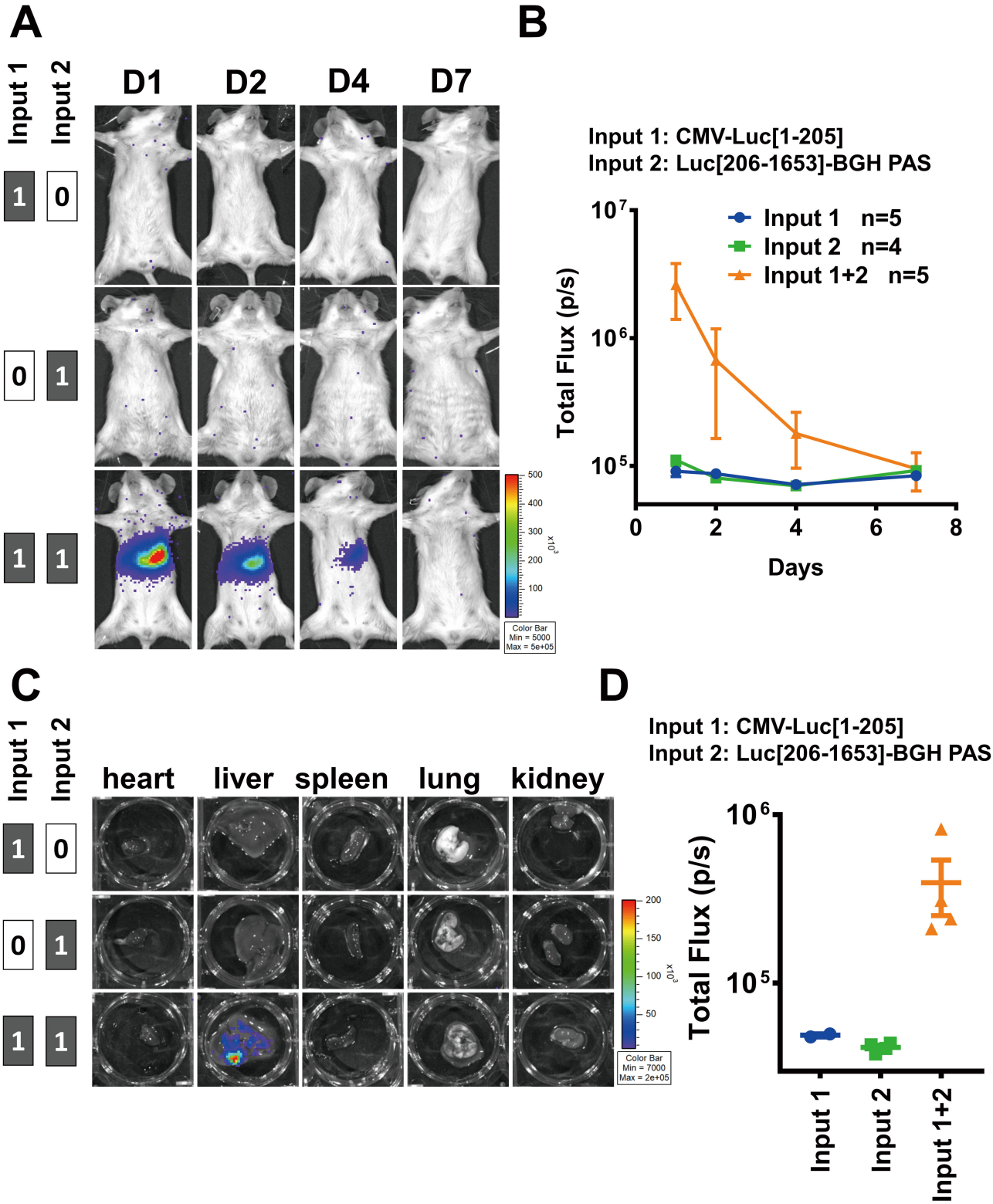

A
Input 1
Input 2
D1
D2
D4
D7
1
0
0
1
1
1
500
400
300
200
100
x10^3
Color Bar
Min = 5000
Max = 5e+05
B
Input 1: CMV-Luc[1-205]
Input 2: Luc[206-1653]-BGH PAS
Input 1 n=5
Input 2 n=4
Input 1+2 n=5
Total Flux (p/s)
10^7
10^6
10^5
0
2
4
6
8
Days
C
Input 1
Input 2
heart
liver
spleen
lung
kidney
1
0
0
1
1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

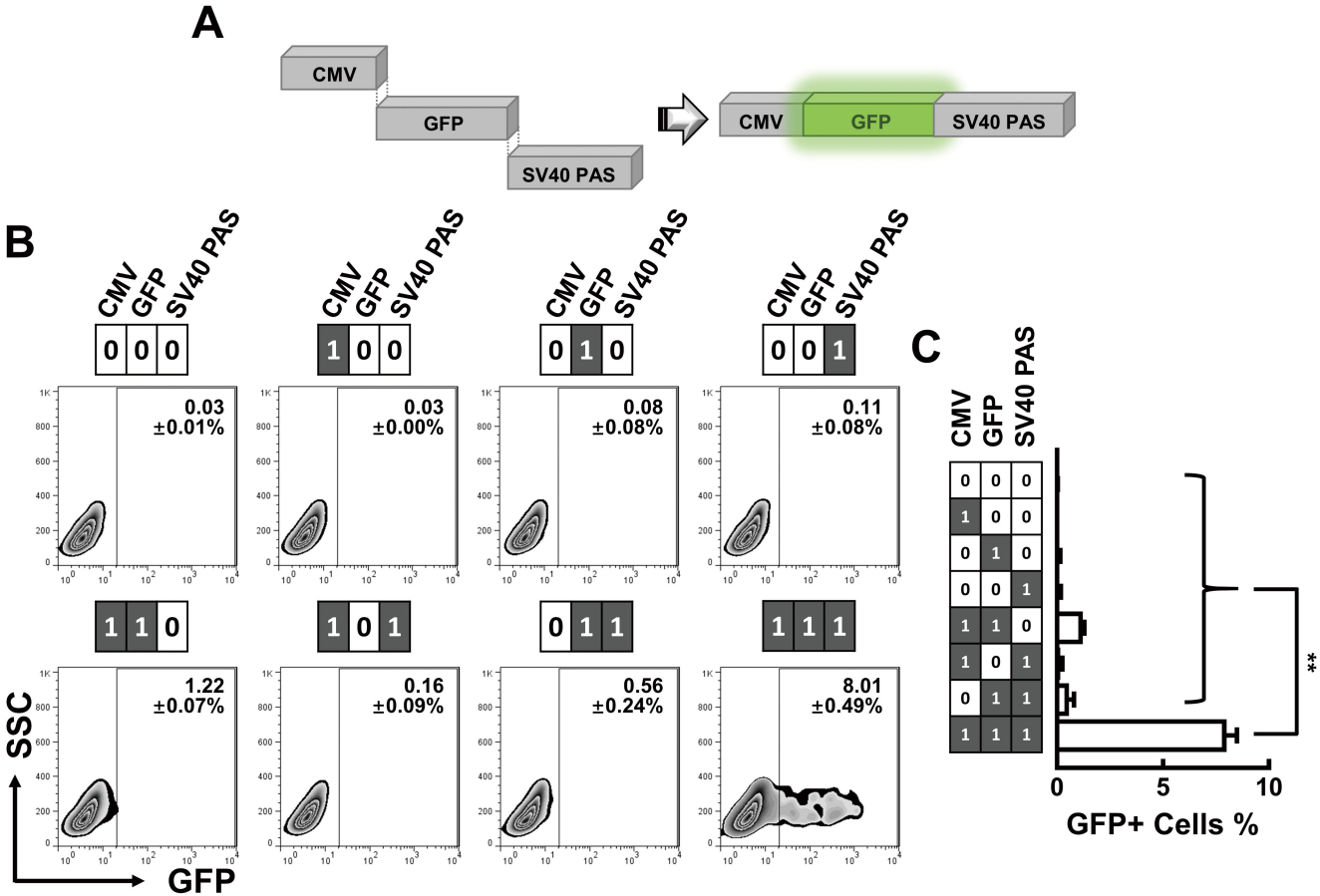

A
CMV
GFP
SV40 PAS
CMV
GFP
SV40 PAS
B
CMV GFP SV40 PAS
0 0 0
0.03 ±0.01%
1 0 0
0.03 ±0.00%
0 1 0
0.08 ±0.08%
0 0 1
0.11 ±0.08%
1 1 0
1.22 ±0.07%
1 0 1
0.16 ±0.09%
0 1 1
0.56 ±0.24%
1 1 1
8.01 ±0.49%
SSC
GFP
C
CMV GFP SV40 PAS
0 0 0
1 0 0
0 1 0
0 0 1
1 1 0
1 0 1
0 1 1
1 1 1
**
0
5
10
GFP+ Cells %

Figure 3

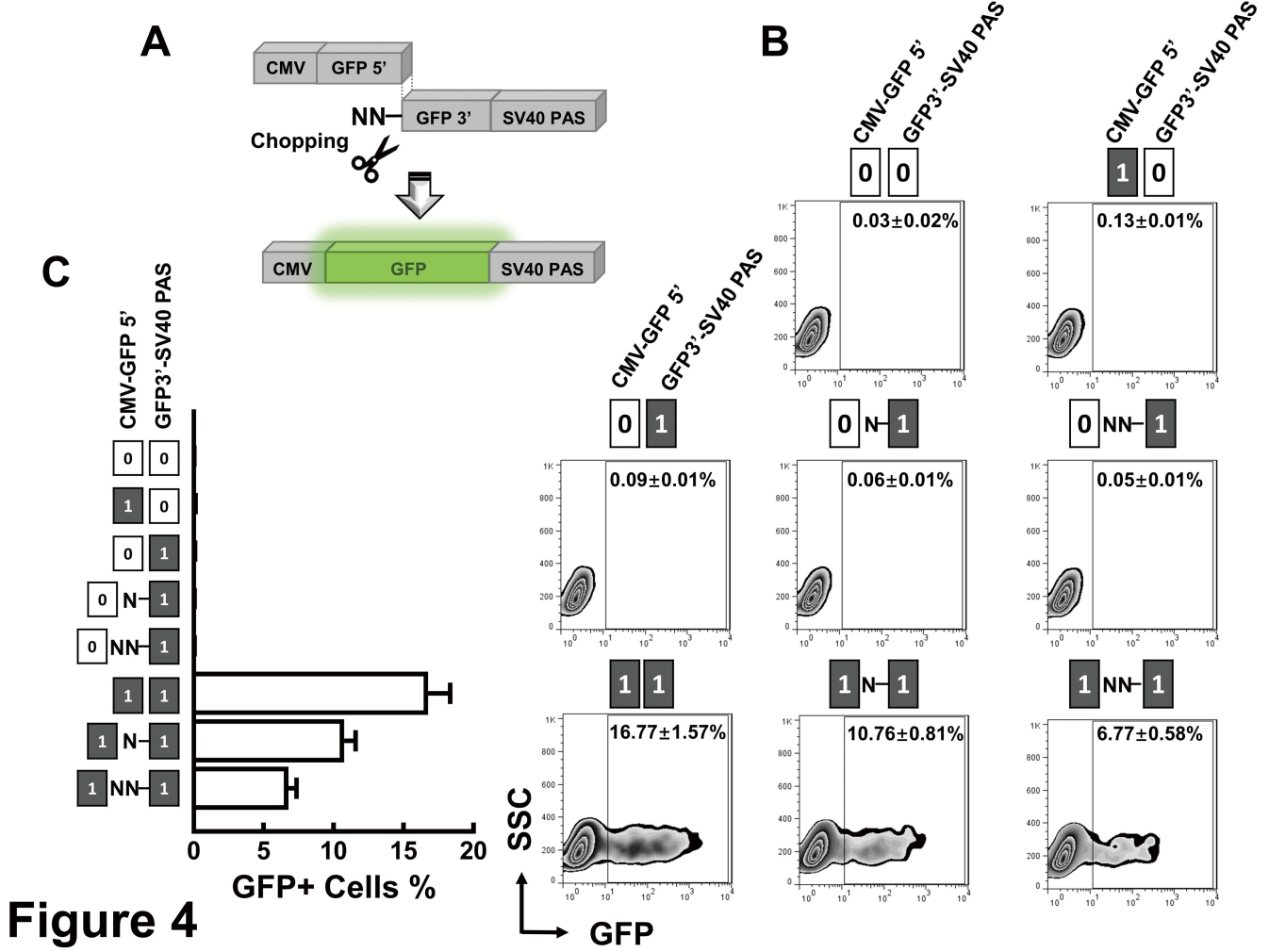

A
CMV
GFP 5'
NN
GFP 3'
SV40 PAS
Chopping
CMV
GFP
SV40 PAS
B
CMV-GFP 5'
GFP3'-SV40 PAS
0.03±0.02%
0.13±0.01%
0.09±0.01%
0.06±0.01%
0.05±0.01%
16.77±1.57%
10.76±0.81%
6.77±0.58%
SSC
GFP
C
CMV-GFP 5'
GFP3'-SV40 PAS
GFP+ Cells %
0
5
10
15
20

Figure 4

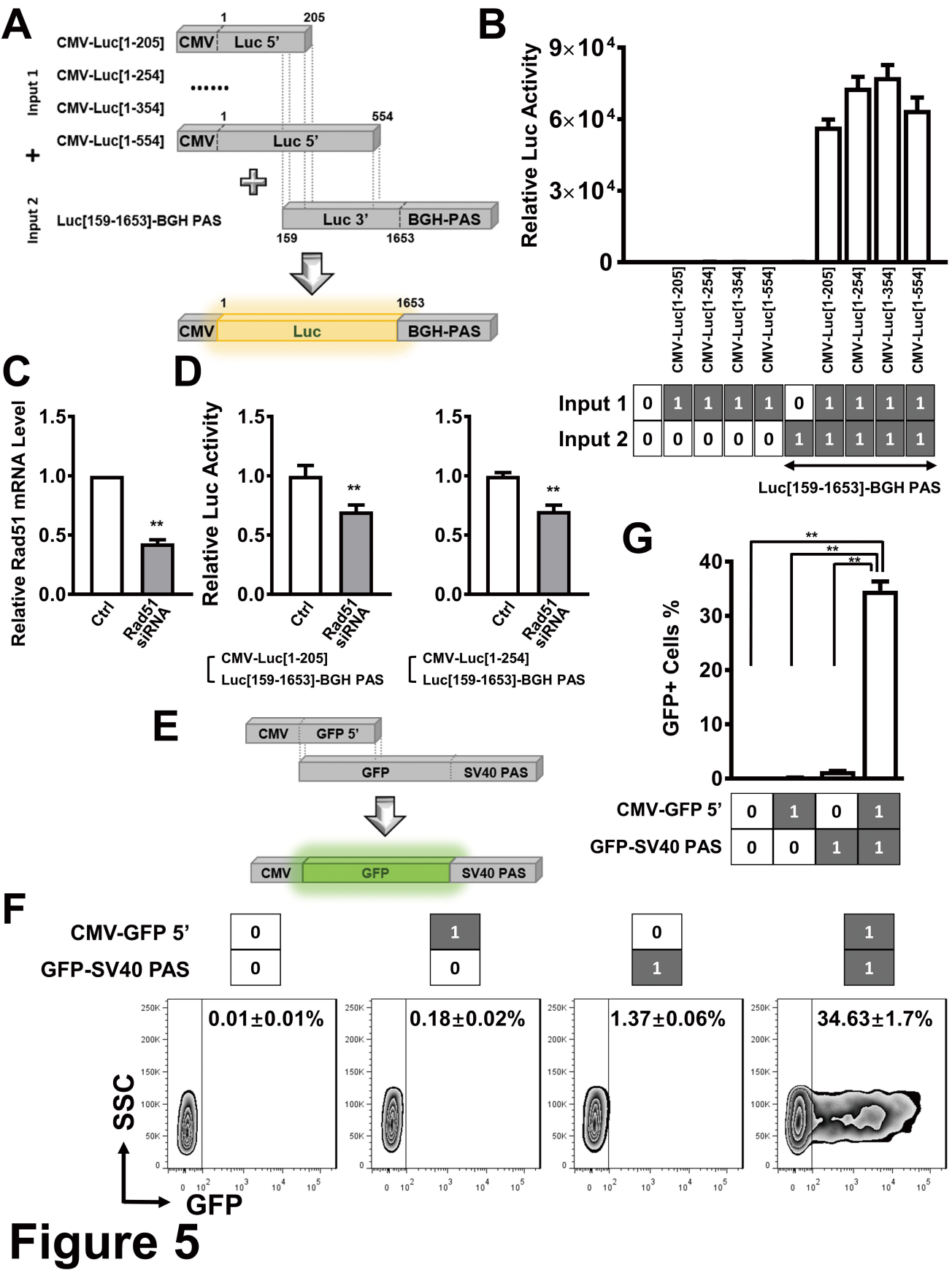

A
Input 1
CMV-Luc[1-205]
CMV-Luc[1-254]
CMV-Luc[1-354]
CMV-Luc[1-554]
+
Input 2
Luc[159-1653]-BGH PAS
CMV
Luc 5'
Luc 3'
BGH-PAS
Luc
1
205
554
159
1653
B
Relative Luc Activity
9×10^4
6×10^4
3×10^4
0
Input 1
Input 2
Luc[159-1653]-BGH PAS
C
Relative Rad51 mRNA Level
Ctrl
Rad51 siRNA
D
Relative Luc Activity
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

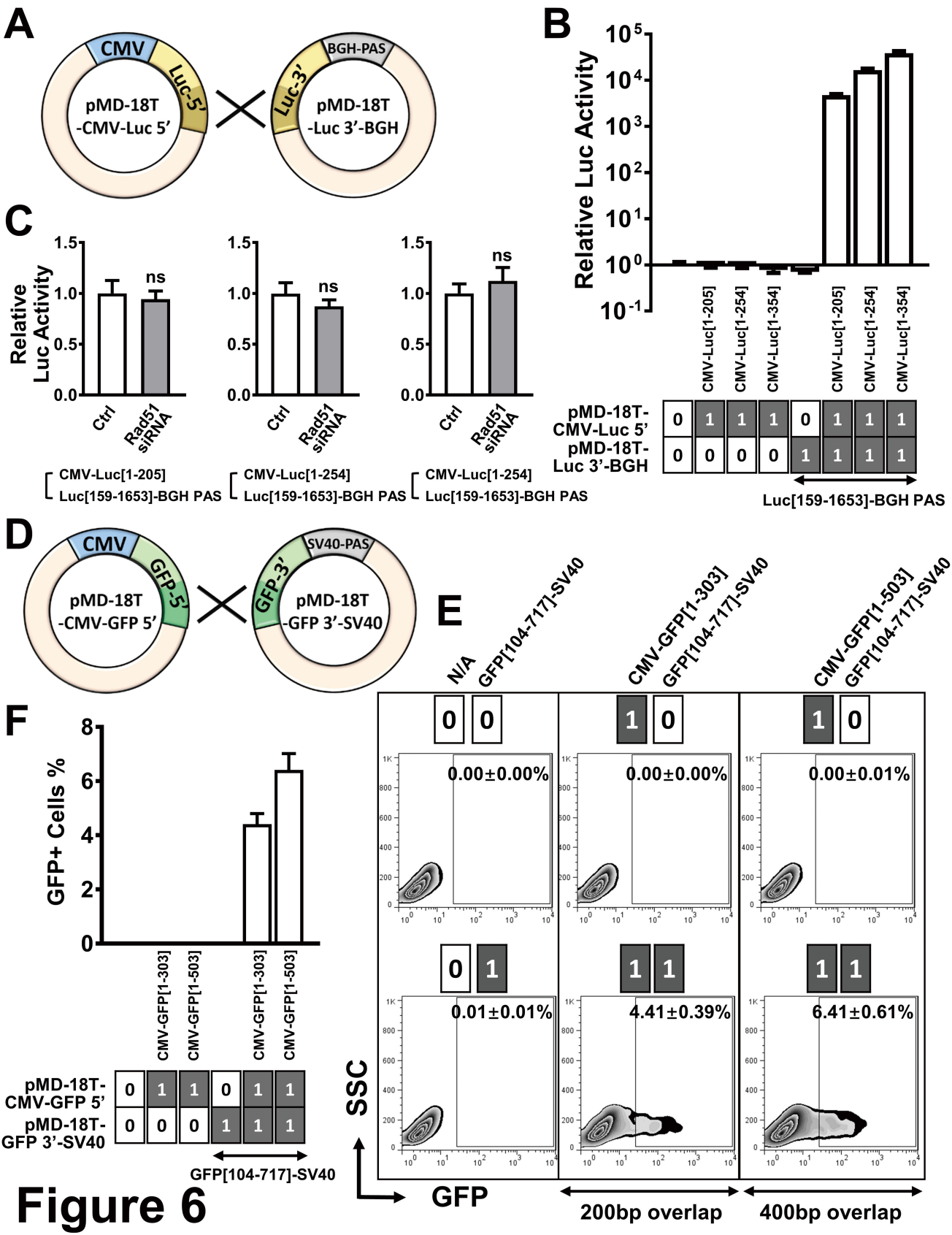
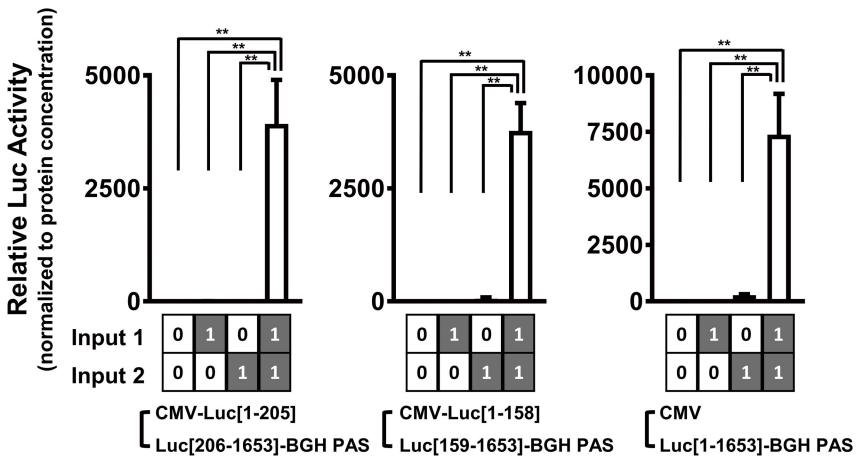
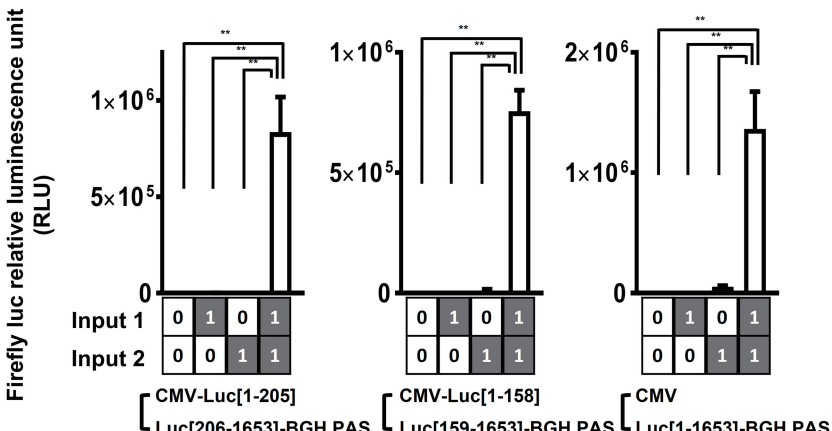

A
CMV
Luc-5'
pMD-18T
-CMV-Luc 5'
BGH-PAS
Luc-3'
pMD-18T
-Luc 3'-BGH
B
Relative Luc Activity
CMV-Luc[1-205]
CMV-Luc[1-254]
CMV-Luc[1-354]
pMD-18T-
CMV-Luc 5'
pMD-18T-
Luc 3'-BGH
Luc[159-1653]-BGH PAS
C
Relative
Luc Activity
ns
Ctrl
Rad51
siRNA
CMV-Luc[1-205]
Luc[159-1653]-BGH PAS
CMV-Luc[1-254]
Luc[159-1653]-BGH 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

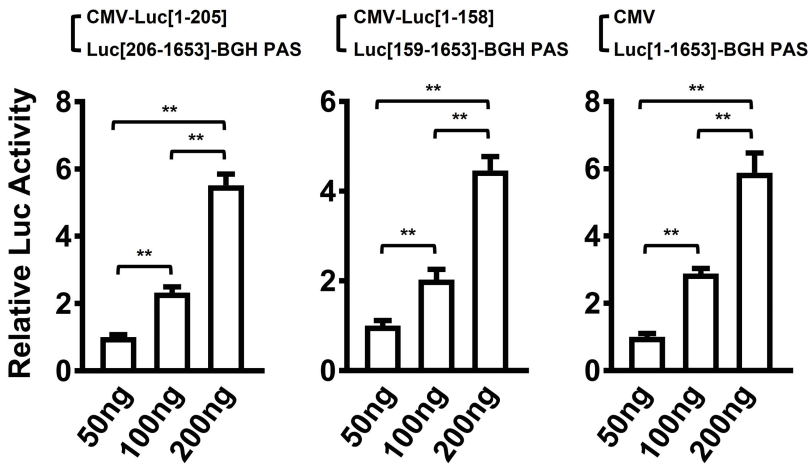
Figure 6



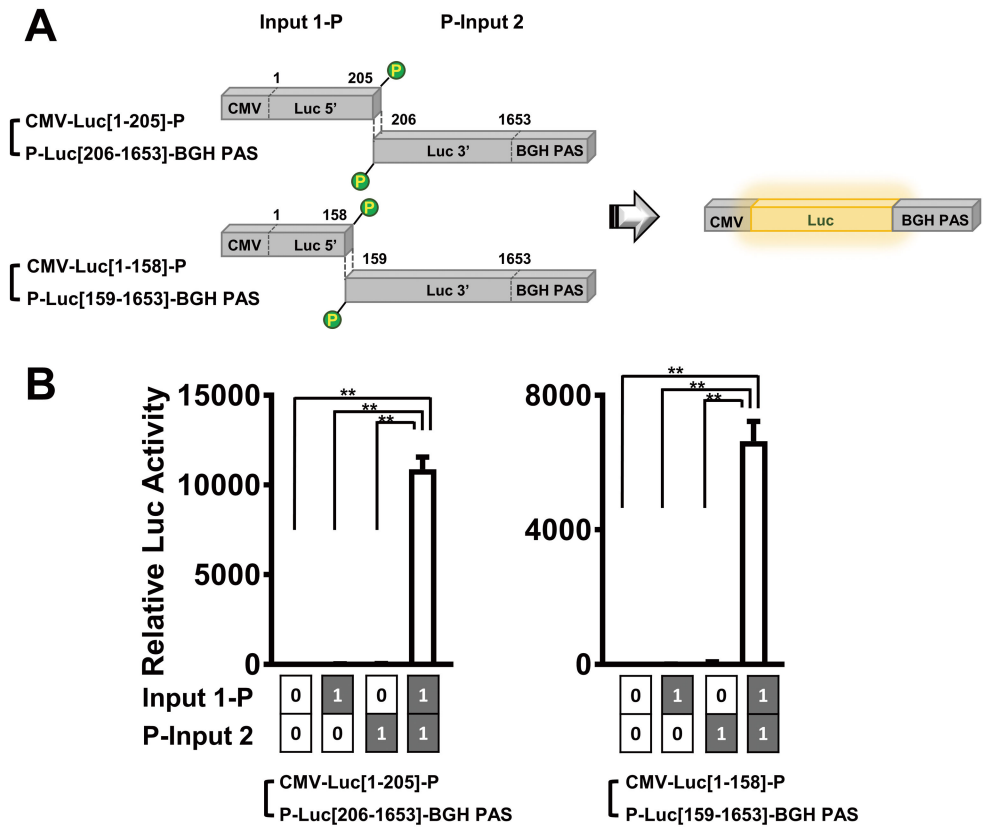
Supplementary Figure S1



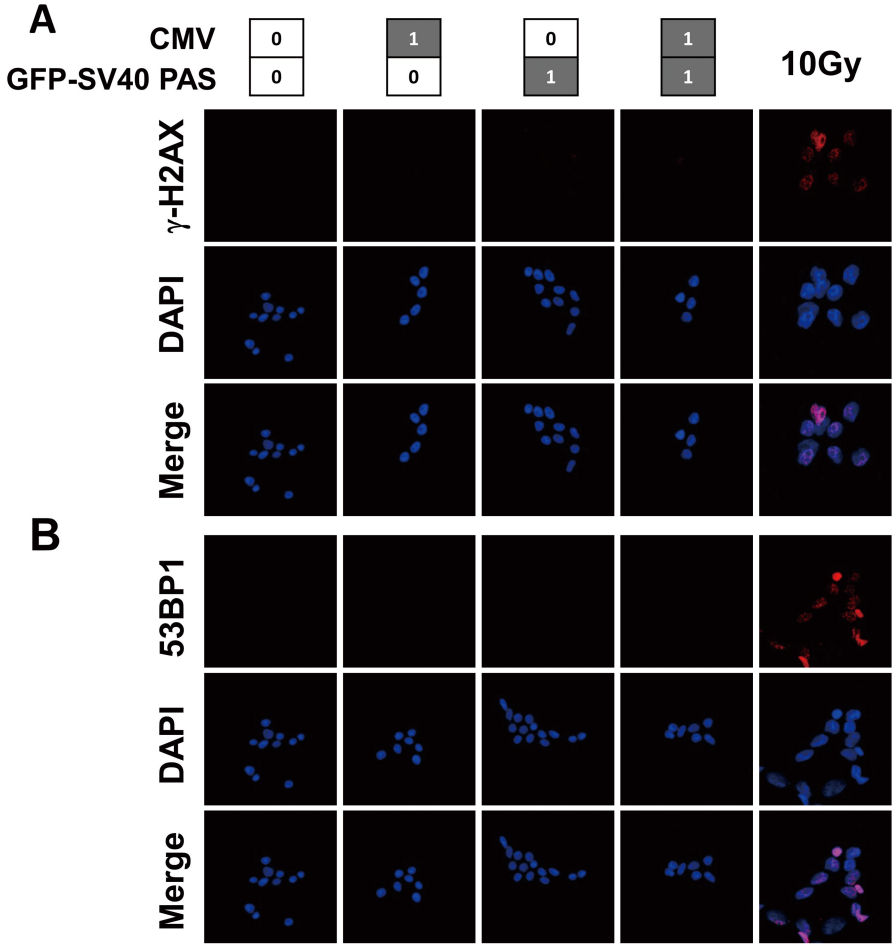
Supplementary Figure S2



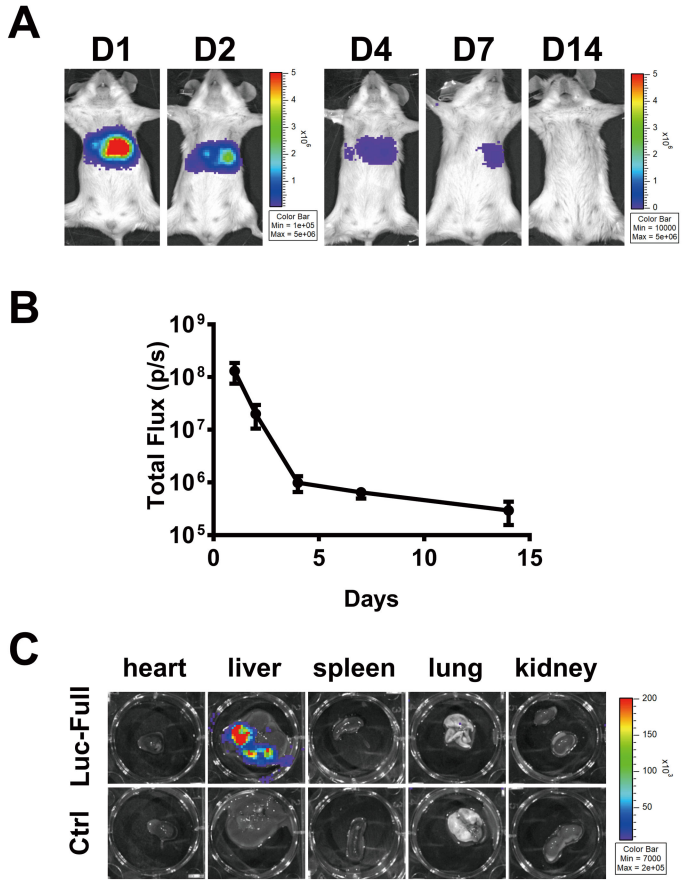
Supplementary Figure S3



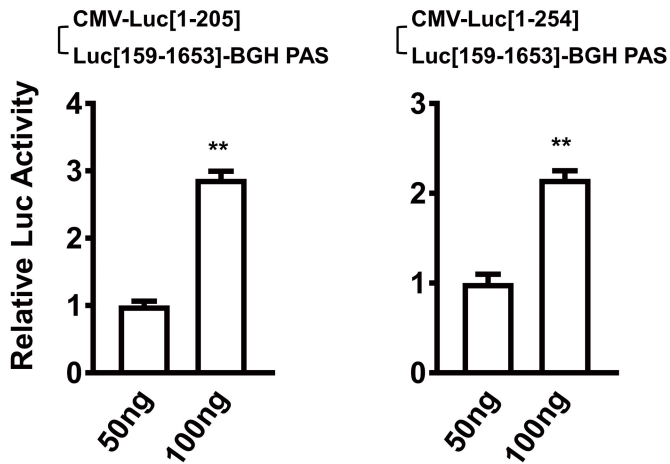
Supplementary Figure S4



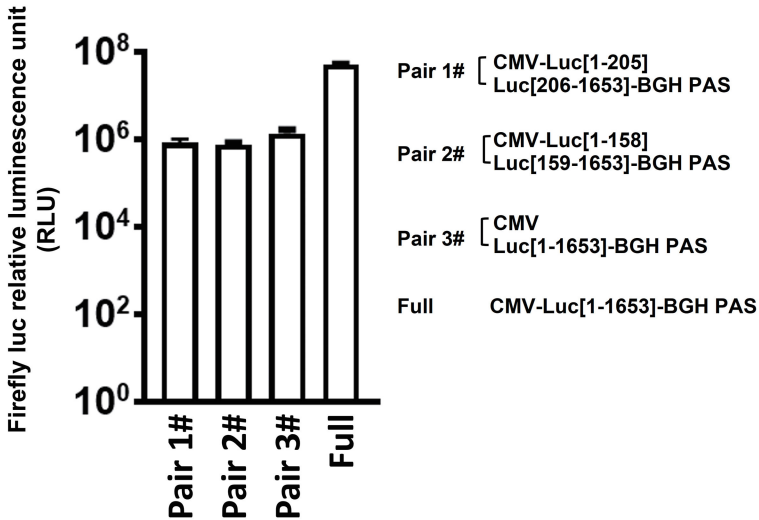
Supplementary Figure S5



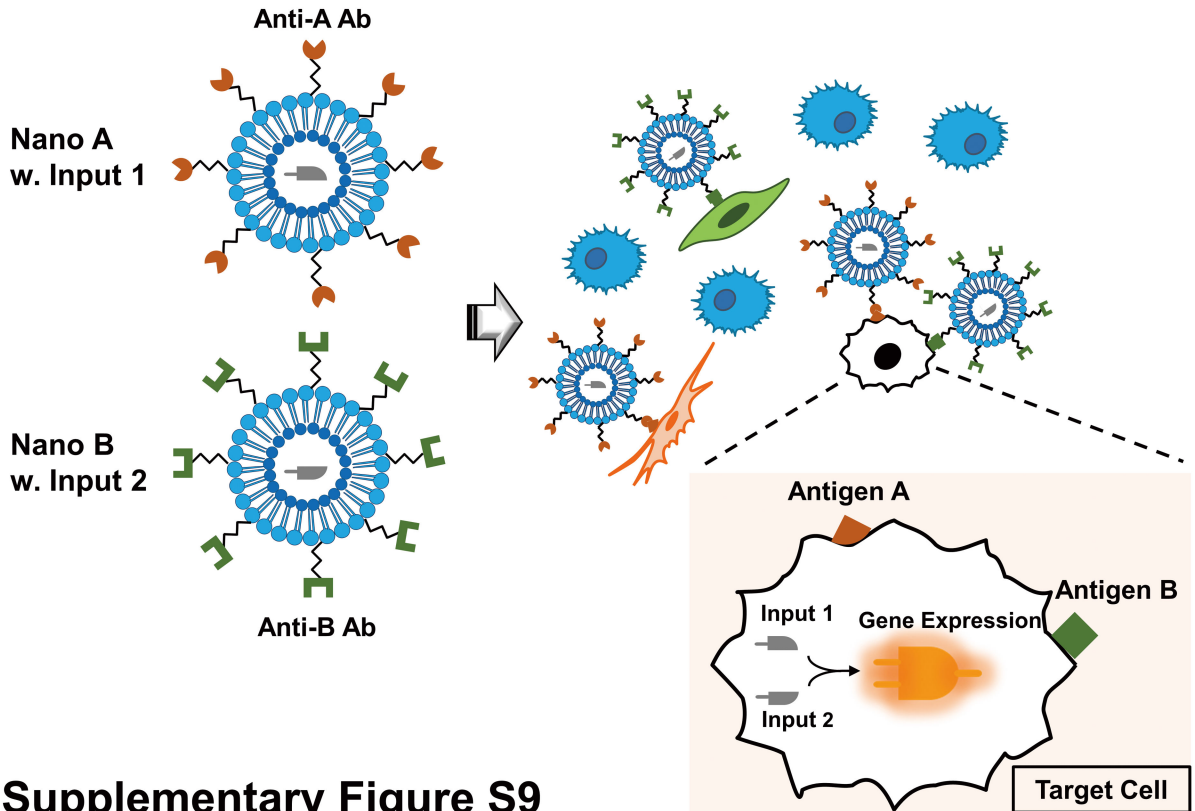
Supplementary Figure S6



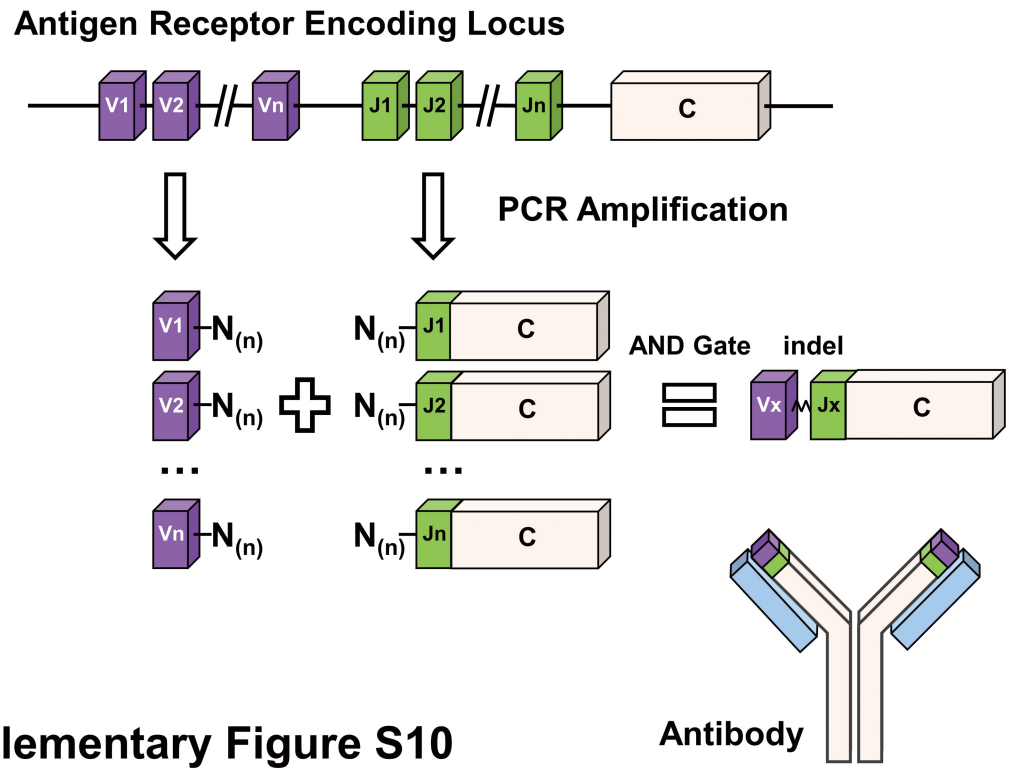
Supplementary Figure S7



Supplementary Figure S8



Supplementary Figure S9



Supplementary Figure S10