Supporting Information

An Ultrasensitive Nucleic Acid Detection Platform Based on Phosphorothioate-DNA

Hang Gao‡, Yuting Shuai‡, Yiyi Zheng, Anan Xu, Zhaoxi Han, Dini Ma, Zixin Deng, Xinyi He*, and Guang Liu*

State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic & Developmental Sciences, School of Life Sciences & Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

*Corresponding authors: Xinyi He (E-mail: xyhe@sjtu.edu.cn), Guang Liu (E-mail: jyayann@sjtu.edu.cn)

This file includes:

Supporting Experimental Methods

Supporting Tables S1 to S5

Supporting Figures S1 to S8

Supporting Experimental Methods

Protein expression and purification

The recombinant protein SBD_{Mmo}+Fluc-N was cloned into the pMAL-C2 plasmid. Fluc-C + SBD_{Mmo} was cloned into pET28a plasmid.[1] The recombinant protein expression vectors were transformed into *Escherichia coli* strain BL21 (DE3). For protein expression. The BL21 (DE3) strains, which carried the plasmid coding for the target protein, were cultured at 37 °C overnight in Luria-Bertani (LB) medium supplemented with the appropriate antibiotics. The overnight culture was inoculated into 1000 mL of the same medium at a 1:100 dilution and incubated at 37 °C until the optical density at 600 nm (OD_{600}) reached 0.6–0.8. Isopropyl β-D-thiogalactoside was added at a final concentration of 0.2 mM, and the culture was incubated at 17 °C for 18 hours to induce protein expression. For protein purification, the cells were collected by centrifugation and resuspended in Ni-NTA column binding buffer (25 mM Tris-HCl, pH 7.2, 300 mM NaCl). The cell pellet was lysed using a homogenizer (JN-Mini Pro) and then centrifuged at 10,000 × g for 60 minutes. The supernatant was collected and applied to a 2 mL Ni-NTA column. After washing with 80 mL of washing buffer (25 mM Tris-HCl, pH 7.2, 300 mM NaCl, 20 mM imidazole), the protein was eluted with 5 mL of elution buffer (25 mM Tris-HCl, pH 7.2, 300 mM NaCl, 500 mM imidazole). The MBP-tagged protein was cleaved overnight at 4 °C using TEV protease (this step was skipped for proteins without MBP tags) and then desalted with a PD-10 desalting column using desalting buffer (25 mM Tris-HCl, pH 7.2, 50 mM NaCl). The protein product was further purified using a HiTrap Heparin HP affinity chromatography column, equilibrated with desalting buffer, and eluted using a linear NaCl gradient (50-1000 mM) with an ÄKTA FPLC system (Cytiva). The purified protein was analyzed by 10% SDS-PAGE, and the protein concentration was determined using the Bradford Protein Assay Kit (Sangon).

Nucleic acid extraction of clinical samples

Rapid Nucleic Acid Release Reagent Kit was used to extract nucleic acids from clinical samples. For each 50 μL sample, 8 μL LD-1 and 2 μL LD-2 were added. After the sample and the release agent were added, they were mixed by vortex. After mixing, it was incubated at 40 °C for 5 min. After centrifugation, 5 μL of supernatant was taken as a template and directly used for amplification reaction.

Droplet Digital PCR (ddPCR)

The ddPCR mixture was prepared in PCR tubes by adding 10.5 μL of 2× EvaGreen

Supermix (Bio-Rad), 1.05 μL of 4 μM forward and reverse primers, 1 μL of diluted DNA, and 7.4 μL of deionized water. For each ddPCR reaction, approximately 15,000 nanoliter-sized water-in-oil droplets were generated using a QX200 Automated Droplet Generator with DG8 Cartridges (186-4007, Bio-Rad) according to the manufacturer's instructions. The resulting droplet mixture was transferred to a 96-well PCR plate, and gene amplification was performed using the following thermal cycling program: 95 °C for 5 min; 40 cycles of 95 °C for 30 s and 59.1 °C for 60 s; 4 °C for 5 min; 90 °C for 5 min; followed by cooling to 4 °C. The amplified products were analyzed using the QX200 Droplet Reader (Bio-Rad), and the target gene concentration was quantified using QuantaSoft™ software (Bio-Rad, version 1.7.4.019).

Table S1. Sequence information of oligonucleotides used for probe optimization and sensitivity investigation of SENSOR 2.0 (Figures 1, 2, and 3, and Supporting Figure S2, S3, S4 and S5).

Name	Sequence (5′ to 3′)
DNA15_S	AGCTTGAGGCCATTG
DNA15_P	CAATG$_{PS}$GCCTCAAGCT
RNA15_S	*AGCUUGAGGCCAUUG*
RNA15_P	*CAAUG$_{PS}$GCCUCAAGCU*
LNA15_P	*C+A+A+T+G$_{PS}$+G+C+C+TCAAGCU*
2OM15_P	*CmAmAmUmG$_{PS}$mGmCmCmUCAAGCU*
2ER15_P	*Ci2MOErAi2MOErAi2MOErTi2MOErG$_{PS}$i2MOErGi2MOErCi2MOErCi2MOErTCAAGCU*
FNA15_P	*Ci2FAi2FAi2FUi2FG$_{PS}$i2FGi2FCi2FCi2FUCAAGCU*
TMP-TB	CTCGGTGGATCAGCTTGTACCAGGCCTTGGCGAA
TB-S1	CGCCA$_{PS}$A$_{PS}$G$_{PS}$G$_{PS}$C$_{PS}$CTGGTA
TB-S2	CAAGC$_{PS}$T$_{PS}$G$_{PS}$A$_{PS}$T$_{PS}$CCACCG
TB-ddPCR-F	ATGAGCGGATCACGCGTCGCTGGCTGGAACAC
TB-ddPCR-R	CGACGAGGTCGTGGCTGACCGCAGGGACCG
TB-RPA-F	AGCGGATCACGCGTCGCTGGCTGGAACAC
TB-T7-RPA-R	ATTCGATAATACGACTCACTATAGGGCGACGAGGTCGTGGCTGACCGCAGGGACCG
TB-temp-DNA	CTCGGTGGATCAGCTTGTACCAGGCCTTGGCGAA
TB-temp-RNA	*CUCGGUGGAUCAGCUUGUACCAGGCCUUGGCGAA*
TB-temp-F	GTGCCCGAGCAACACCCACCCATTACAG
TB-temp-R	TCAGCGCACGTCGAACCTGTCGAGGTT
ORF1ab-ddPCR-F	CCATCAACTCAATATGAGTATGGTACTGAAG
ORF1ab-ddPCR-R	TATCATCATCTAACCAATCTTCTTCTTGCT
ORF1ab-RPA-F	CCATCAACTCAATATGAGTATGG
ORF1ab-T7-RPA-R	CCGGTAATACGACTCACTATAGGGTCATCTAACCAATCTTCTTCTTGCT
ORF1ab-S1	GAATTT$_{PS}$G$_{PS}$G$_{PS}$T$_{PS}$G$_{PS}$CCACT
ORF1ab-S2	TCTGCT$_{PS}$G$_{PS}$C$_{PS}$T$_{PS}$C$_{PS}$TTCAA

The RNA nucleotides are marked with italic. $_{PS}$ represents the PT-modification, + represents LNA modified nucleotides, m represents 2'-MeO modified nucleotides, i2MOEr represents 2'-MOE modified nucleotides, and i2FA represents 2'-F modified nucleotides. The T7 promoter is marked with underline.

Table S2. Sequence information of oligonucleotides used for SNV detection (Figure 4, 5 and 6).

Name	Sequence (5′ to 3′)
G944C-P1	GTATGGCACCG$_{PS}$GAAC
G944C-P2-1	CGGTAAGG$_{PS}$ACGCGATCACC
G944C-P2-2	CGGTAAGG$_{PS}$ACGCGATCAC
G944C-P2-3	CGGTAAGG$_{PS}$ACGCGATCA
G944C-P3-B3-1	AC$_{PS}$CCCGATCGAGGTCGTATGGACGAACACCC
G944C-P3-B3-2	CAC$_{PS}$CCCGATCGAGGTCGTATGGACGAACACCC
G944C-P3-B4-1	AC$_{PS}$CCCGTTCGAGGTCGTATGGACGAACACCC
G944C-P3-B4-2	CAC$_{PS}$CCCGTTCGAGGTCGTATGGACGAACACCC
G944C-P3-B4-3	CCAC$_{PS}$CCCGTTCGAGGTCGTATGGACGAACACCC
G944C-P3-B5-1	CAC$_{PS}$CCCGTACGAGGTCGTATGGACGAACACCC
G944C-P3-B5-2	CCAC$_{PS}$CCCGTACGAGGTCGTATGGACGAACACCCC
KatG-G944C	CGGGGUGUUCGUCCAUACGACCUCGAUGCCGGUGGUGAUCGC GUCCUUACCGGUUCCGGUGCCAUACG
KatG-WT	CGGGGUGUUCGUCCAUACGACCUCGAUGCCGCUGGUGAUCGC GUCCUUACCGGUUCCGGUGCCAUACG
A798C-P1	GCTACTTATATG$_{PS}$AAA
A798C-P2	GAAATTAAA$_{PS}$CTGAGAAGA
A798C-P3	GAC$_{PS}$CCACCTGATATTGATCCAAATCGTGATTT
ERG11-A798C	UAAAUCACGAUUUGGAUCAAUAUCACCACGGUCUCUUCUCA GUUUAAUUUCUUUCAUAUAAGUAGCA
ERG11-WT	UAAAUCACGAUUUGGAUCAAUAUCACCACGUUCUCUUCUCA GUUUAAUUUCUUUCAUAUAAGUAGCA
ERG11-F	TTTTTGACCGTTCATTTGCTCA
ERG11-R	CCGGTAATACGACTCACTATAGGGCAGAAGTATGTTGACCACCC
G944C-s1-P1	G$_{PS}$T$_{PS}$A$_{PS}$T$_{PS}$G$_{PS}$GCACCGGAACC
G944C-s1-P2	GGT$_{PS}$A$_{PS}$A$_{PS}$G$_{PS}$G$_{PS}$ACGCGATCA
G944C-s1-P3	CCAC$_{PS}$CGGCATCGAGGT

$_{PS}$ represent the PT-modification. The SNV sites are labeled with red, and the Bubble regions are labeled with blue.

Table S3. Strains and plasmids used in this study.

Plasmid	Source
pET28a-SBD_{Mmo}	[2]
pMAL-C2_(SBD_{Mmo}+Fluc-N)	[1]
pET28a_(Fluc-C+SBD_{Mmo})	[1]
E. coli TOP10 Competent Cells	Invitrogen
E. coli BL21(DE3) Competent Cells	Novagen

Table S4. Types of clinical samples

Number	Sex	Age	Specimen type
1	Male	68	sputum
2	Male	49	sputum
3	Male	71	sputum
4	Female	77	urine
5	Male	67	synovial fluid
6	Male	69	urine
7	Female	82	urine
8	Female	77	sputum
9	Male	74	sputum
10	Male	75	peritoneal drainage fluid
11	Male	87	urine
12	Male	85	sputum
13	Male	73	lavage fluids
14	Male	72	urine
15	Male	83	urine
16	Male	85	pus
17	Female	54	urine
18	Male	55	sputum
19	Male	72	urine
20	Male	73	sputum
21	Male	76	sputum
22	Female	77	urine
23	Male	58	urine
24	Male	49	urine
25	Male	68	gastrointestinal mucosal
26	Male	68	sputum
27	Male	72	sputum
28	Male	68	urine
29	Female	63	wound drainage
30	Male	72	sputum
31	Female	88	bile
32	Male	67	sputum
33	Female	70	urine
34	Female	79	feces
35	Male	87	sputum
36	Female	80	lavage fluids
37	Male	73	sputum
38	Male	39	bile
39	Male	69	urine
40	Male	82	sputum

41	Female	90	sputum
42	Female	81	sputum
43	Female	75	sputum
44	Male	71	sputum
46	Male	70	sputum
47	Female	63	urine
48	Female	77	urine
49	Female	56	purulent drainage
50	Male	82	lavage fluids
51	Male	69	sputum
52	Male	72	nasopharyngeal secretion
53	Female	69	sputum
54	Male	76	sputum
55	Male	74	sputum
56	Male	74	urine
57	Male	78	urine

Table S5. Cost comparison of SENSOR 2.0 and CRISPR

	SHERLOCK[3](CRISPR)		SENSOR2.0	
RPA	Same (430 USD/1,000 reactions)			
T7/rNTPs	Same (16 USD/1,000 reactions)			
Protein	Same (both custom purification, 4 USD/1,000 reactions)			
crRNA/Probe	One 67 nt RNA	200 USD /5 nmol /10,000 reactions	Two 15 nt DNA probes with 5 PTs each	10 USD /5 nmol /10,000 reactions
Reporter	One 5 nt RNA with 6-FAM and BHQ-1	800 USD /25 nmol /10,000 reactions	Steady-Glo® Luciferase Assay System	850 USD /100 ml /10,000 reactions
RNAse inhibitors	Murine inhibitors	147/1,000 reactions	Not required	0
Cost for 1,000 reactions (20 μL volume)	697 USD		537 USD	

Supporting Figures

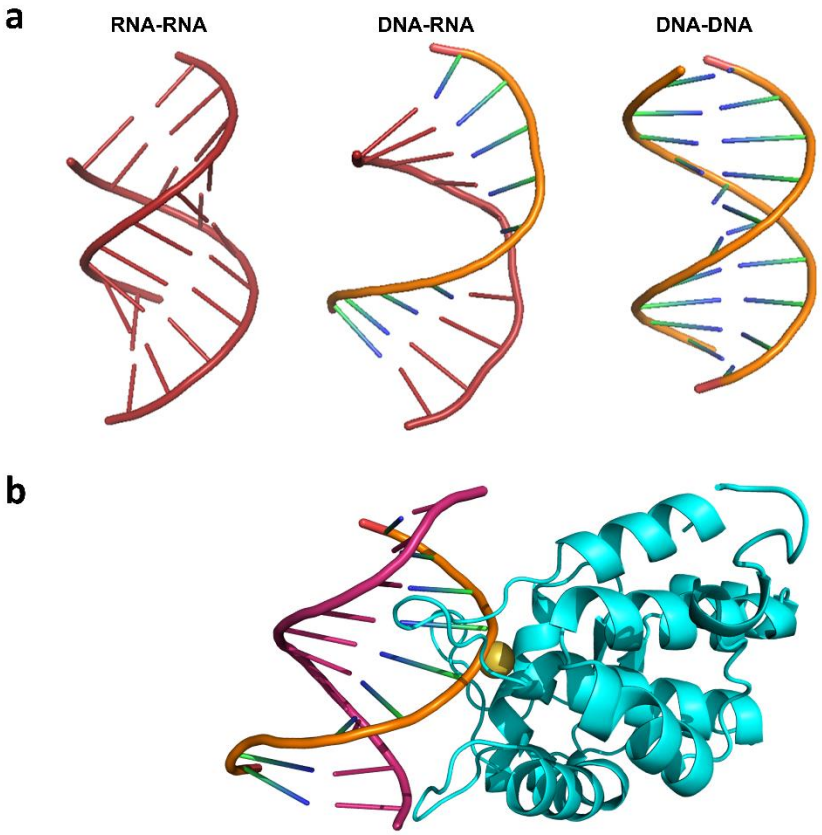


Figure S1. (a) Structure of ideal dsRNA (left panel), DNA/RNA duplex (middle panel, PDB:404d), and ideal dsDNA (right panel)**.** (b) Structure of SBD_{Mmo} binding with PT-DNA/RNA duplex predicted by AlphaFold 3.0. The purple chain represents RNA, and the orange chain represents PT-ssDNA probe. The sulfur atom is marked as a yellow ball.

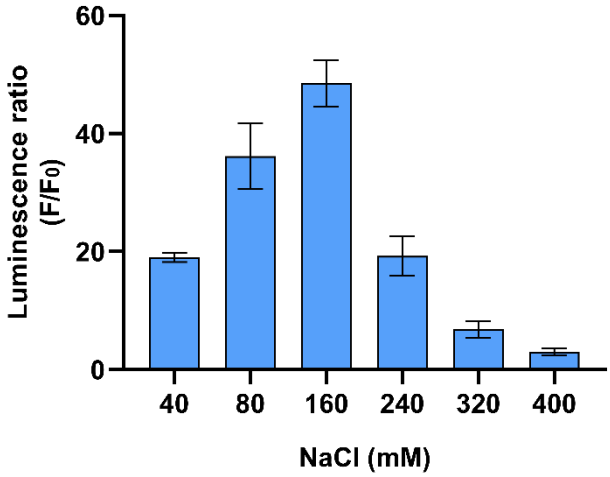


Figure S2. Optimization of salt concentration for SENSOR 2.0. The detection system was evaluated at pH 8.8 and a protein ratio of 1:1. The optimal salt concentration was found to be 160 mM. F and F_0 represent the luminescence intensity at 10 minutes in the sample and control groups, respectively. Error bars represent the standard deviation of three replicates.

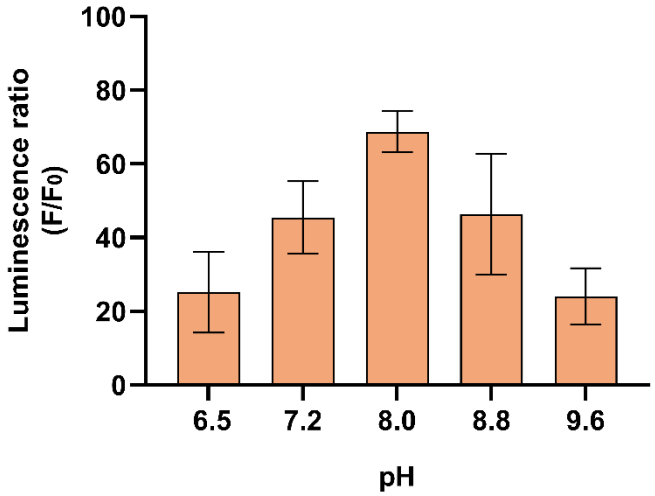


Figure S3. Optimization of pH for SENSOR 2.0. The detection system was evaluated at a NaCl concentration of 160 mM and a protein ratio of 1:1. F and F_0 represent the luminescence intensity at 10 minutes in the sample and control groups, respectively. Error bars represent the standard deviation of three replicates.

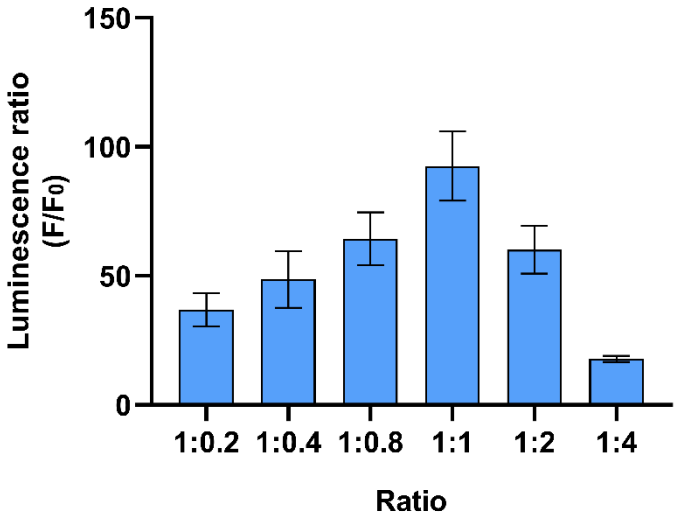


Figure S4. Optimization of protein ratio for SENSOR 2.0. The detection system was evaluated at pH 8.0 and a NaCl concentration of 160 mM. F and F_0 represent the luminescence intensity at 10 minutes in the sample and control groups, respectively. Error bars represent the standard deviation of three replicates.

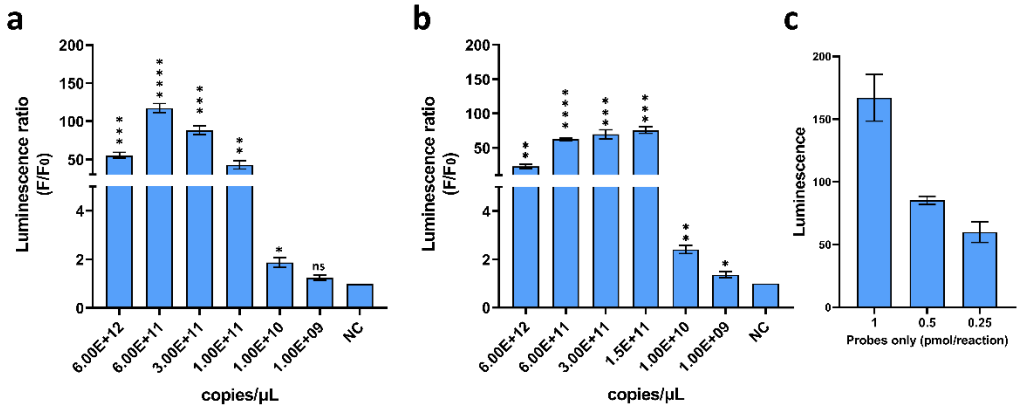


Figure S5. Optimization of probe concentration. (a) Detections of serially diluted substrates with 0.1 pmol probes. (b) Detections of serially diluted substrates with 0.25 pmol probes. (c) Background luminescence of different probe concentrations. Statistical significance was determined using a two-tailed student's t-test (*P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001). Error bars indicate the standard deviation of at least three replicates.

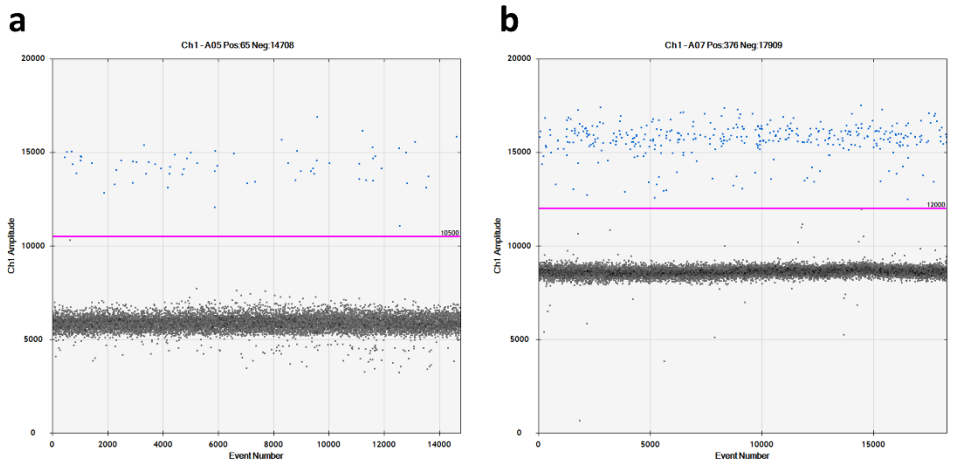


Figure S6. Representative 1D plots of ddPCR reactions. The ordinate scales indicate fluorescent amplitude. The pink line is the threshold, above which are positive droplets (blue) containing at least one copy of target DNA and below which are negative droplets (gray) without any target DNA. (a) The copy number of the *ORF1ab* gene DNA was determined to be 5.6×10^7 copies/μL. (b) The copy number of *KatG* gene DNA was determined to be 3.9×10^9 copies/μL. The copy number was determined by averaging the results of three replicates of ddPCR reactions using diluted samples.

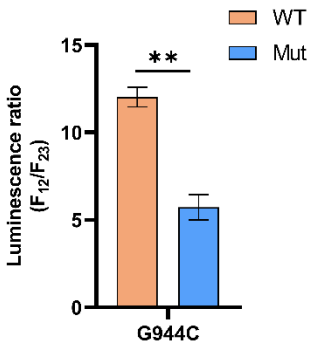


Figure S7. SNV detection for *M. tuberculosis* (G944C of KatG) using SENSOR 1.0. Statistical significance was determined using a two-tailed student's t-test (*P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001). Error bars indicate the standard deviation of at least three replicates.

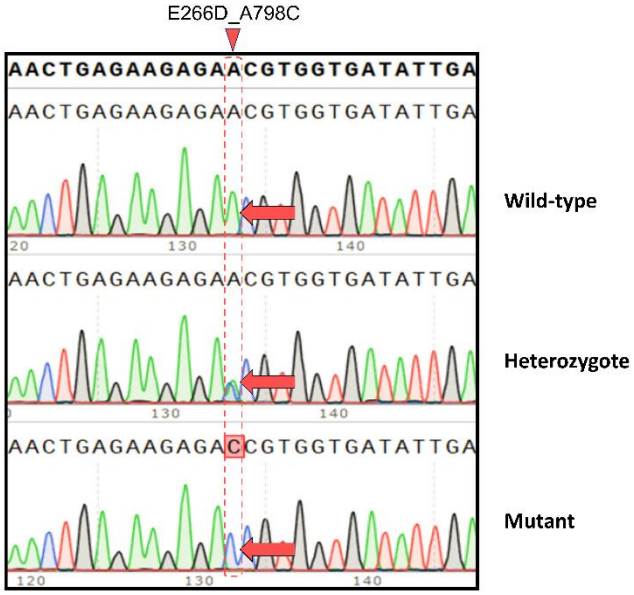


Figure S8. Representative Sanger sequencing results of *C. albicans* clinical samples. The SNV sites in the *ERG11* gene were indicated by red arrows. Adenine denotes the wild-type allele, Cytosine denotes the mutant allele, and the observation of a double peak indicates heterozygosity.

REFERENCES

1. Shuai, Y.; Ju, Y.; Li, Y.; Ma, D.; Jiang, L.; Zhang, J.; Tan, G.-Y.; Liu, X.; Wang, S.; Zhang, L.; Liu, G., A rapid nucleic acid detection platform based on phosphorothioate-DNA and sulfur binding domain. *Synth Syst Biotechnol* **2023,** *8* (2), 213-219.
2. Liu, G.; Fu, W.; Zhang, Z.; He, Y.; Yu, H.; Wang, Y.; Wang, X.; Zhao, Y. L.; Deng, Z.; Wu, G.; He, X., Structural basis for the recognition of sulfur in phosphorothioated DNA. *Nat Commun* **2018,** *9* (1), 4689.
3. Gootenberg, J. S.; Abudayyeh, O. O.; Lee, J. W.; Essletzbichler, P.; Dy, A. J.; Joung, J.; Verdine, V.; Donghia, N.; Daringer, N. M.; Freije, C. A.; Myhrvold, C.; Bhattacharyya, R. P.; Livny, J.; Regev, A.; Koonin, E. V.; Hung, D. T.; Sabeti, P. C.; Collins, J. J.; Zhang, F., Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* **2017,** *356* (6336), 438-442.