

Supporting Information for

**A double-stranded DNA catalyzed strand displacement system  
for detection of small genetic variations**

*Na Liu<sup>[a]//</sup>, Xuzhe Zhang<sup>[a]//</sup>, Xiaofeng Tang<sup>[a]</sup>, Yizhou Liu, Donghui Huang<sup>\*[a]</sup> and  
Xianjin Xiao<sup>\*[a]</sup>*

<sup>a</sup> Institute of Reproductive Health, Tongji Medical College, Huazhong University of  
Science and Technology. Wuhan 430030, PR China.

//contributed equally to this work.

\* To whom correspondence should be addressed.

## Materials and methods

### 1. Materials

TE solution, proteinase K and PCR product purification kit were purchased from Tiangen Biotech Co. (Beijing, China). Taq polymerase and ThermoPol reaction buffer were purchased from NEB Co. (Beijing, China). DNA strands were synthesized and purified by HPLC (Sangon Co., China). The sequences of all the probes and targets that have been studied in this work are summarized in Table S1.

**Table S1.** Sequences used in this work

Stand name	Sequence (from 5' to 3') <sup>a</sup>
EGFR-L858R-FAM	TGGGCGGGCCAAACTGCTAACCTCTCTA-FAM
EGFR-L858R-BHQ	BHQ-TAGAGAGGTTAGCAGTTTGGCCCGCCCA AAATCTGTGATCCAT
EGFR-L858R-MT-12-10	GATCACAGATTTTGGGCGGGCCAAACTGCT
EGFR-L858R-MT-11-10	ATCACAGATTTTGGGCGGGCCAAACTGCT
EGFR-L858R-MT-11-9	ATCACAGATTTTGGGCGGGCCAAACTGC
EGFR-L858R-WT-12-10	GATCACAGATTTTGGGCTGGCCAAACTGCT
EGFR-L858R-WT-11-10	ATCACAGATTTTGGGCTGGCCAAACTGCT
EGFR-L858R-WT-11-9	ATCACAGATTTTGGGCTGGCCAAACTGC
EGFR-L858R-Blocker17	GCAGTTTGGCCAGCCCA
EGFR-L858R-Blocker16	CAGTTTGGCCAGCCCA
EGFR-L858R-Blocker15	AGTTTGGCCAGCCCA
EGFR-L858R-Blocker17+tail	GCAGTTTGGCCAGCCCATTTTTTTT
EGFR-L858R-Blocker15+tail	AGTTTGGCCAGCCCATTTTTTTT
EGFR-L858R-Invader1	TTGGGCGGGCCAAACTGCTAACCTCTCTA
EGFR-L858R-Invader2	TTTGGGCGGGCCAAACTGCTAACCTCTCTA
EGFR-L858R-Invader3	TTTTGGGCGGGCCAAACTGCTAACCTCTCTA
EGFR-L858R-Invader4	ATTTTGGGCGGGCCAAACTGCTAACCTCTCTA
EGFR-L858R-FP	GAACGTAAGTGGTAAAAACACCG
EGFR-L858R-RP	TCCTTCTGCATGGTATTCTTTCTC
KrasG13D-FAM	TGGTGACGTAGGCAAGAGTGAGCTACA-FAM

KrasG13D-BHQ	BHQ-TGTAGCTCACTCTTGCCTACGTCACCAGCTCCAACCTACCACA
KrasG13D-MT	GTAGTTGGAGCTGGTGACGTAGGCAAGAG
KrasG13D-WT	GTAGTTGGAGCTGGTGCCGTAGGCAAGAG
KrasG13D-Invader2	GCTGGTGACGTAGGCAAGAGTGAGCTACA
KrasG13D-Blocker17	TCTTGCCTACGCCACCA
ShareY-FAM	GTCTCGCCGACGGAG-FAM
ShareY-BHQ	BHQ-GGCCGTCACACGTGT
KrasG13D-Y-fam	TGGTGACGTAGGCAAGAGTGAGCTACAGTATCATAGCATAAGGCCGC
KrasG13D-Y-bhq	ACACGTGTGACGGCCTGTAGCTCACTCTTGCCTACGTCACCAGCTCCAACCTACCAC
KrasG13D-Y-MT	GTAGTTGGAGCTGGTGACGTAGGCAAGAG
KrasG13D-Y-WT	GTAGTTGGAGCTGGTGCCGTAGGCAAGAG
KrasG13D-Y-Invader0	TGGTGACGTAGGCAAGAGTGAGCTACA
KrasG13D-Y-Invader1	CTGGTGACGTAGGCAAGAGTGAGCTACA
KrasG13D-Y-Invader2	GCTGGTGACGTAGGCAAGAGTGAGCTACA
KrasG13D-Y-Blocker17	TCTTGCCTACGCCACCA
ShareX-FAM	AAGGGTGTGCTTATGCTATGATAC-FAM
ShareX-BHQ	BHQ-GTAGGAGAGAGATAGAGGTGG
KrasG13D-X-fam	TGGTGACGTAGGCAAGAGTGAGCTCATAGCATAAGCAACACCCTT
KrasG13D-X-bhq	CCACCTCTATCTCTCCGCTCACTCTTGCCTACGTCACCAGCTCCAACCTACCACA
KrasG13D-X-MT	GTAGTTGGAGCTGGTGACGTAGGCAAGAG
KrasG13D-X-WT	GTAGTTGGAGCTGGTGCCGTAGGCAAGAG
KrasG13D-X-Invader0	TGGTGACGTAGGCAAGAGTGAGCTACA
KrasG13D-X-Invader1	CTGGTGACGTAGGCAAGAGTGAGCTACA
KrasG13D-X-Invader2	GCTGGTGACGTAGGCAAGAGTGAGCTACA
KrasG13D-X-Blocker17	TCTTGCCTACGCCACCA
BRCars3765640-Y-fam	TGTATCATTCTAAAACCTGAGCTACAGTATCATAGCAT AAGGCCGC
BRCars3765640-Y-bhq	ACACGTGTGACGGCCTGTAGCTCAGGTTTTAGAATGATA CAAACCAAAGAACTAATG
BRCars3765640-Y-MT	TAGTTCTTTGTTTTGTATCATTCTAAAACC

BRCArs3765640-Y-WT	TAGTTCCTTTGGTTTGTATTATTCTAAAACC
BRCArs3765640-Y-Invader	TTTGTATCATTCTAAAACCTGAGCTACA
BRCArs3765640-Y-Blocker	GGTTTTAGAATAATACA
MTRRA66G-Y-fam	GAAATGTGTGAGCAAGCTGAGCTACAGTATCATAGCAT AAGGCCGC
MTRRA66G-Y-bhq	ACACGTGTGACGGCCTGTAGCTCAGCTTGCTCACACATT TCTTCTGCGATGGCCTTT
MTRRA66G-Y-MT	GGCCATCGCAGAAGAAATGTGTGAGCAAGC
MTRRA66G-Y-WT	GGCCATCGCAGAAGAAATATGTGAGCAAGC
MTRRA66G-Y-Invader	AAGAAATGTGTGAGCAAGCTGAGCTACA
MTRRA66G-Y-Blocker	GCTTGCTCACATATTTTC
MTHFRA1298C-Y-fam	TGAAGCAAGTGTCTTTGTGAGCTACAGTATCATAGCAT AAGGCCGC
MTHFRA1298C-Y-bhq	ACACGTGTGACGGCCTGTAGCTCACAAAGACACTTGCTTCACTGGTCAGCTCCTCCC
MTHFRA1298C-Y-MT	AGGAGCTGACCAGTGAAGCAAGTGTCTTTG
MTHFRA1298C-Y-WT	AGGAGCTGACCAGTGAAGAAAGTGTCTTTG
MTHFRA1298C-Y-Invader	AGTGAAGCAAGTGTCTTTGTGAGCTACA
MTHFRA1298C-Y-Blocker	CAAAGACACTTTCTTCA
MTHFRC677T-Y-fam	GGGAGTCGATTTTCATCATGAGCTACAGTATCATAGCATAAGGCCGC
MTHFRC677T-Y-bhq	ACACGTGTGACGGCCTGTAGCTCATGATGAAATCGACTCCCGCAGACACCTTCTCCT
MTHFRC677T-Y-MT	AGAAGGTGTCTGCGGGAGTCGATTTTCATCA
MTHFRC677T-Y-WT	AGAAGGTGTCTGCGGGAGCCGATTTTCATCA
MTHFRC677T-Y-Invader	GCGGGAGTCGATTTTCATCATGAGCTACA
MTHFRC677T-Y-Blocker	TGATGAAATCGGCTCCC
BRCArs1799949-Y-fam	GACAGTGATACTTTCCCTGAGCTACAGTATCATAGCAT AAGGCCGC
BRCArs1799949-Y-bhq	ACACGTGTGACGGCCTGTAGCTCAGGGAAAGTACTCTG TCATGTCTTTACTTGTC
BRCArs1799949-Y-MT	AAGTAAAAGACATGACAGTGATACTTTCCC
BRCArs1799949-Y-WT	AAGTAAAAGACATGACAGCGATACTTTCCC

BRCAs1799949-Y-Invader	ATGACAGTGATACTTTCCCTGAGCTACA
BRCAs1799949-Y-Blocker	GGGAAAGTATCGCTGTC
BRCAs16940-Y-fam	TTTCACTGGTACCTGGTTGAGCTACAGTATCATAGCATAAGGCCGC
BRCAs16940-Y-bhq	ACACGTGTGACGGCCTGTAGCTCAACCAGGTACCAGTGA AATACTGCTACTCTCTA
BRCAs16940-Y-MT	GAGAGTAGCAGTATTTCACTGGTACCTGGT
BRCAs16940-Y-WT	GAGAGTAGCAGTATTTCACTGGTACCTGGT
BRCAs16940-Y-Invader	TATTTCACTGGTACCTGGTTGAGCTACA
BRCAs16940-Y-Blocker	ACCAGGTACCAATGAAA
BRCAs80357234-Y-fam	GCCTATAAGAAAGTACGTGAGCTACAGTATCATAGCATAAGGCCGC
BRCAs80357234-Y-bhq	ACACGTGTGACGGCCTGTAGCTCACGACTTTCTTATA GGCTCCTGAAATTAATTG
BRCAs80357234-Y-MT	TTTAATTTGAGGAGCCTATAAGAAAGTACG
BRCAs80357234-Y-WT	TTTAATTTGAGGAGCCTACAAGAAAGTACG
BRCAs80357234-Y-Invader	GCCTATAAGAAAGTACGTGAGCTACA
BRCAs80357234-Y-Blocker	CGTACTTTCTTGTAGGC
MTRRA66G-FP	TCAGTTTCACTGTTACATGCCTTGAAGTG
MTRRA66G-RP	TCCACTGTAACGGCTCTAACCTTATCGGA
MTHFRA1298C-FP	CCCTCTGTCAGGAGTGTG
MTHFRA1298C-RP	CACTCCAGCATCACTCACT
BRCAs1799949-FP	CAAGAAGAGTAACAAGCCAAATG
BRCAs1799949-RP	TAAAAGAACCAGGTGCATTTGT
BRCAs16940-FP	GACCCCAAAGATTTTGCAAACTG
BRCAs16940-RP	CCAGTAACGAGATACTTTCCTGAGTG

<sup>a</sup>All the modification sites of interest were highlighted in red.

## 2. DNA sequence design

NCBI was used for consulting the corresponding SNP site sequence information. According to the probe structure, sequences were designed using NUPACK software to reduce secondary

structures and the interaction among the DNA strands.

### **3. DNA strand preparation and concentration quantification**

TE solution was used for dissolving the dry powder of DNA strands to the required concentration. UV spectrophotometer was used for measuring the concentration of the DNA strands at 260 nm.

### **4. Non-universal probe preparation**

The Dissociation strand and Template strand were mixed in a PCR tube at a ratio of 1.2: 1 with 1 × thermopol reaction buffer for the preparation of the probe to ensure the full bonding. After mixing, put the tube at 85 ° C for 1 minute, and then left at 55 ° C for 1 minute. Finally, the tube is transferred to 37 ° C, and kept at a constant temperature for 1 hour to several hours until use.

### **5. Universal probe preparation**

The Dissociation strand, Template strand, Invader strand, and BHQ strand were mixed in a PCR tube at a ratio of 1.2: 1:1:1 with 1 × thermopol reaction buffer (NEB) for the preparation of the probe to ensure the full bonding. After mixing, put the tube at 85 ° C for 5 minutes, and then left at 55 ° C for 5 minutes. Finally, the tube was transferred to 37 ° C, and kept at a constant temperature for 3 hours to several hours until use.

### **6. Two-step asymmetric PCR**

The first step was to perform conventional PCR with 1: 1 primer, and then took 1 ul of the PCR product of the first step, diluted it 100 times, and took 1 ul as the template for second step. The second step was to perform PCR with 10:1 primer and obtained a large number of target single strands.

### **7. Fluorescence signal detection**

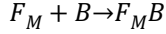
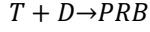
Add probes, target Fuel strands, Invader strands, Blocker strands to the ELISA Plate Strips. Centrifugated it with a palm centrifuge and quickly put the plate strips into the Bio-Tek ELISA instrument for fluorescence detection. Performed the fluorometric assay with an excitation/emission wavelengths of 485/582nm and gained value of 60. Fluorescence signal was collected once every 1 minute and lasted for hours.

## Modelling and theoretical calculations

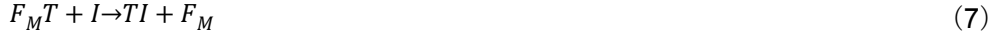
### 1. Demonstration of the chemical kinetic reaction process of dsCSD system.

We use initial letters to represent the ingredients in dsCSD system such as  $F_{M/W}$  stands for Fuel<sub>MT/WT</sub>,  $B$  stands for Blocker,  $I$  for Invader,  $D$  for Dissociation strand and  $T$  for Template strand. Take Fuel<sub>MT</sub> as example:

As we pre-treated and mixed Template strand and Dissociation strand, Fuel strand and Blocker for the preparation of the probe and the reactant system, based on the principle of base pairing, following reactions would occur respectively:



When all the ingredients were mixed and centrifugated in the ELISA Plate Strips, our dsCSD system would work as:



Considering Blocker strand was far more sufficient in the reaction system and they could easily bond with Fuel and Invader, reaction (1) and (4) would be finished in a short time, resulting in reaction (2), (5) and (7) being omitted in our modeling calculation. To summarize and simplify the above-mentioned reaction process, we could obtain a net reaction as following:



It was clear that the net reaction shared the same reactants and products with (6), thus we could take reaction (6) as the background. To summarize, we could view the whole dsCSD system as a reaction of Invader taking the place of Dissociation stand with  $F_M B$  and  $F_M T$  as catalyzers, that is the parataxis of reaction (3) and (8).

Similarly, for Fuel<sub>WT</sub>, we have:



With the net reaction:



Above all, in our modeling, the concentration of  $DB$  (noted as  $[DB]$ ) could stand for the fluorescence intensity. However, as we were eager to obtain the equation of  $[DB]$  versus time, the reaction rate constant of reaction (3), (8), (10) and (11) remained unknown. Therefore, it was obligatory for us to further detailing the reaction process to modeling.

For reaction (3), for example, it could be further detailed into the following 3 steps:





$F_M B \cdot PRB$  in equation (13) represents the intermediate state of the reactants. In this case, that is the accomplishment of the free bases in Invader pairing with Template strand before the occurrence of strand displacement. Equation (14) represents the process of strand displacement and equation (15) represents the dissociation of  $F_M T$  and  $DB$ . Similarly, for reaction (8):



For the same reason, we could extract reaction (10) and (11) into another 6 equations based on the above principles.

## 2. Theoretical calculation of dsCSD system model

Based on the principles of chemical kinetics, the corresponding ordinary differential equations for Fuel (MT) are:

$$\frac{d[DB]}{dt} = k_{3+} [F_M T \cdot DB] - k_{3-} [F_M T] \cdot [DB]$$

$$\frac{d[IB]}{dt} = -k_{4+} [F_M T] \cdot [IB] + k_{4-} [F_M T \cdot IB]$$

$$\frac{d[F_M B]}{dt} = -k_{1+} [F_M B] \cdot [PRB] + k_{1-} [F_M B \cdot PRB] + k_{6+} [TI \cdot F_M B] - k_{6-} [F_M B] \cdot [TI]$$

$$\frac{d[F_M T]}{dt} = k_{3+} [F_M T \cdot DB] - k_{3-} [F_M T] \cdot [DB] - k_{4+} [F_M T] \cdot [IB] + k_{4-} [F_M T \cdot IB]$$

$$\frac{d[PRB]}{dt} = -k_{1+} [F_M B] \cdot [PRB] + k_{1-} [F_M B \cdot PRB]$$

$$\frac{d[TI]}{dt} = k_{6+} [TI \cdot F_M B] - k_{6-} [F_M B] \cdot [TI]$$

$$\frac{d[F_M B \cdot PRB]}{dt} = k_{1+} [F_M B] \cdot [PRB] - k_{1-} [F_M B \cdot PRB] - k_{2+} [F_M B \cdot PRB] + k_{2-} [F_M T \cdot DB]$$

$$\frac{d[F_M T \cdot DB]}{dt} = k_{2+} [F_M B \cdot PRB] - k_{2-} [F_M T \cdot DB] - k_{3+} [F_M T \cdot DB] + k_{3-} [F_M T] \cdot [DB]$$

$$\frac{d[F_M T \cdot IB]}{dt} = k_{4+} [F_M T] \cdot [IB] - k_{4-} [F_M T \cdot IB] - k_{5+} [F_M T \cdot IB] + k_{5-} [TI \cdot F_M B]$$

$$\frac{d[TI \cdot F_M B]}{dt} = k_{5+} [F_M T \cdot IB] - k_{5-} [TI \cdot F_M B] - k_{6+} [TI \cdot F_M B] + k_{6-} [F_M B] \cdot [TI]$$

Similarly, we could obtain the corresponding ordinary differential equations for Fuel (WT).



### 3. Simulation and calculation of reaction rate constant

For reaction (13),  $F_{MB}$  and  $PRB$  are two different double-strand structures that formed during the preparation of the probe and the reaction platform. There exist several free and matched bases at the 3' end of the Template strand and the 5' end of Fuel strand.  $k_{1+}$  represents the reaction rate constant of bases binding. According to literature searching and simulation, the binding rate constant could be assumed to be  $3 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ . For the same reason,  $k_{3-}$ ,  $k_{4+}$  and  $k_{6-}$  for those who represents the binding of free bases, could be assumed to be  $3 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ . Their reverse rate constants could be calculated as

$$k_- = \frac{k_+}{K_{eq}} = k_+ e^{\Delta G_{rxn}^0 / RT}$$

where  $\Delta G_{rxn}^0$  denotes the standard free energy of the relevant reaction. By using

NUPACK to predict free energy of secondary structure for reactants and products,  $\Delta G_{rxn}^0$  for different reactions or sequences can be obtained. Thus, the reverse rate constants for the binding process can be calculated.

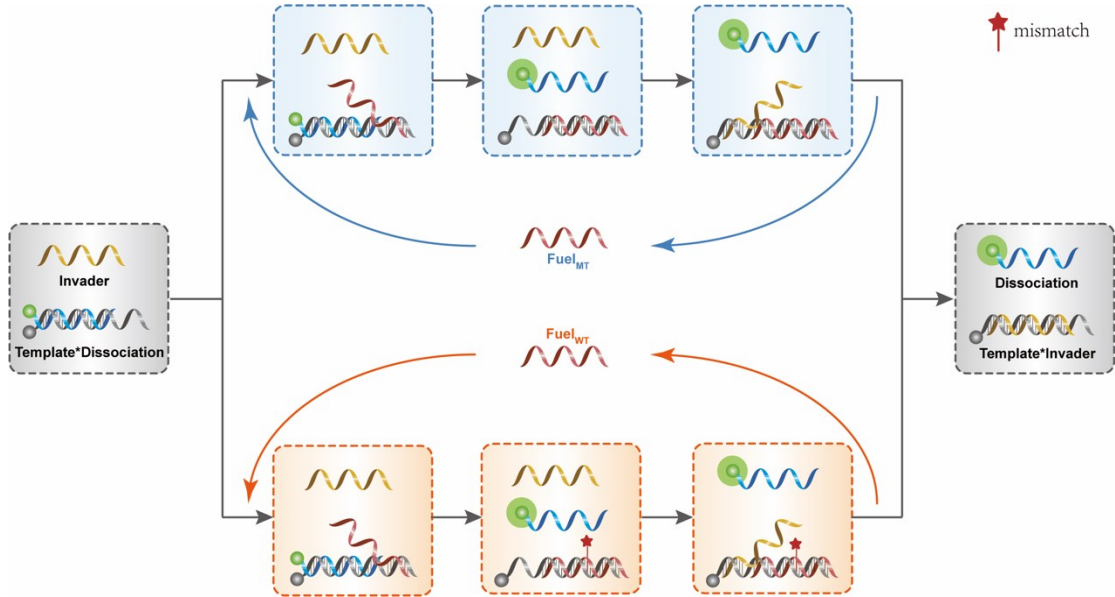
Equation (14) and (17) represent the process of strand displacement. Considering there is no relevant data on the rate constant of strand displacement, more experimental explorations and simulations are needed.

### 4. Simulation and calculation of model by MATLAB

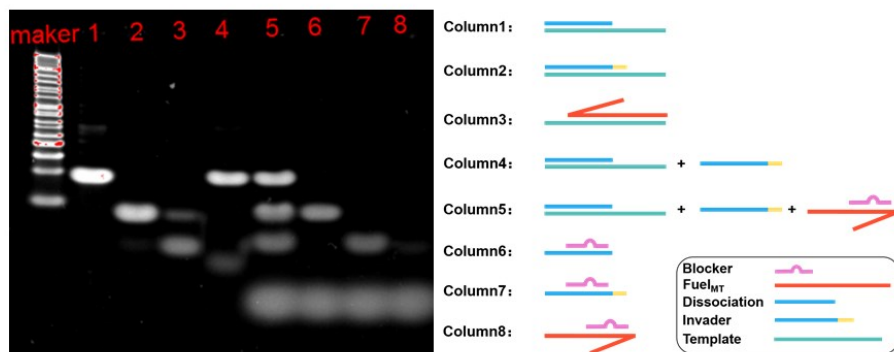
The ordinary differential equations mentioned above are simulated using MATLAB's stiff "ode23s" solver with the initial concentration of  $IB$ ,  $F_{M/WB}$  and  $PRB$  are 1uM, 100nM and 1uM respectively. By outputting the consequence of  $[DB]$  versus time in different condition, we could identify and evaluate their DF for the consultants of Invader optimization. Since the construction of our modeling is based on simulation and matching, it sure needs more perfection not only in the estimation of  $\Delta G$ , but also in more reveals of the chemical kinetic information of strand displacement process. We believe that with the in-depth study for kinetic process of strand displacement reaction, our model could serve as a powerful tool for the optimization of experiment conditions.

## Supplementary figures and tables

### 1. Principle of target DNA catalyzing DNA strand displacement reactions



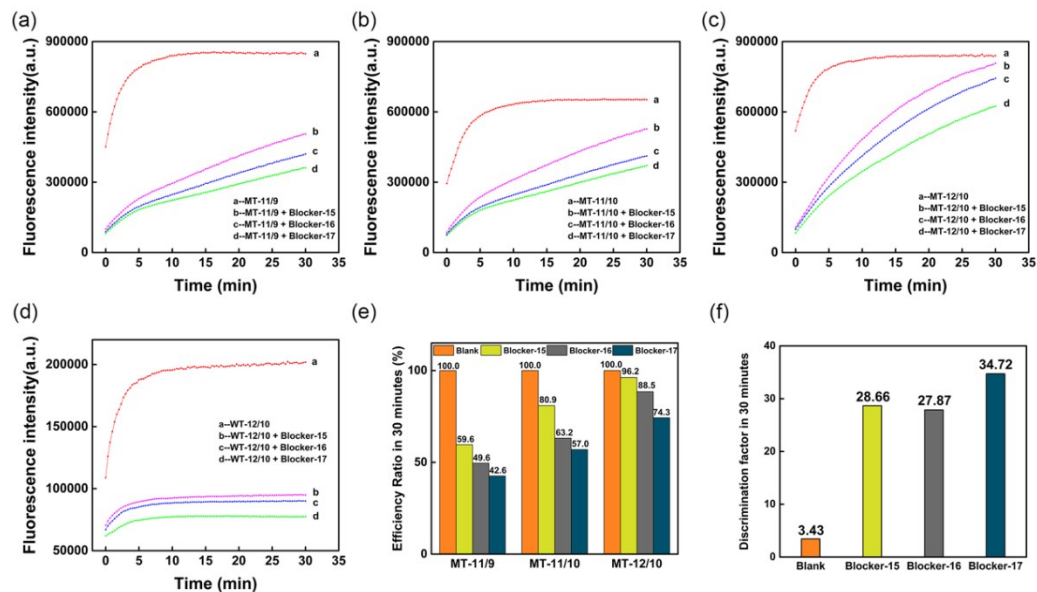
### 2. The feasibility of dsCSD system was identified by polyacrylamide gelelectrophoresis (PAGE).



e S2. Identification of the feasibility of dsCSD system by PAGE.

Figur

### 3. Optimization of dsCSD system targeting the detection of EGFR-L858R



Fig

re S3. Optimization of the toehold region for the target strand and the length of Blocker. (a) The target MT strands were set 11-9 with the length of Blocker from 15 to 17 nt. (b) The target MT strands were set 11-10 with the length of Blocker from 15 to 17 nt. (c) The target MT strands were set 12-10 with the length of Blocker from 15 to 17 nt. (d) The target WT strands were set 12-10 with the length of Blocker from 15 to 17 nt. (e) The efficiency ratio of dsCSD with different MT and Blocker in 30 minutes. (f) The discrimination factor of dsCSD with 12-10 target and different Blocker in 30 minutes.

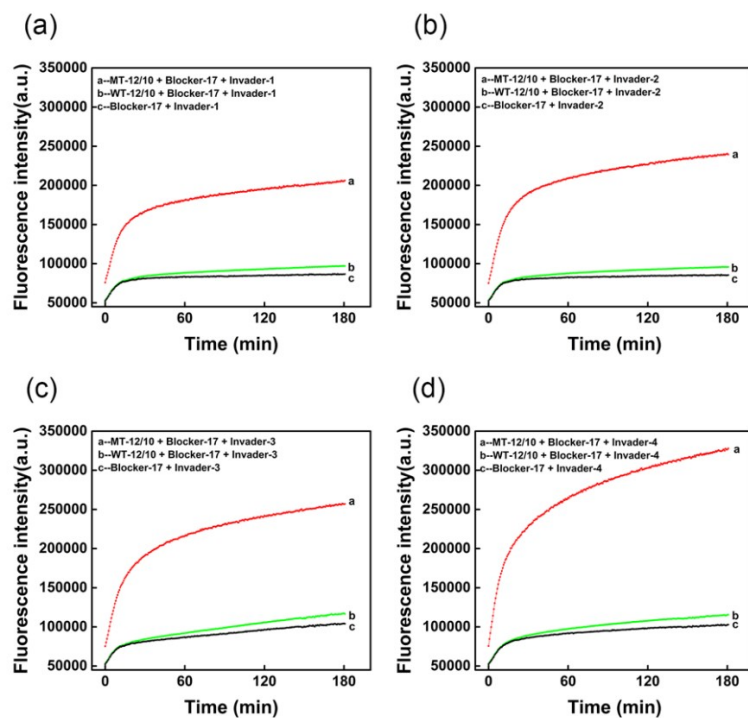


Figure S4. Optimization of the length for Invader strand. (a) Fuel (12-10) + BLK17 + Invader-1. (b) Fuel (12-10) + BLK17 + Invader-2. (c) Fuel (12-10) + BLK17 + Invader-3. (d) Fuel (12-10) + BLK17 + Invader-4.

**4. The detection limit of the dsCSD system targeting KrasG13D mutation in synthesized DNA sample**

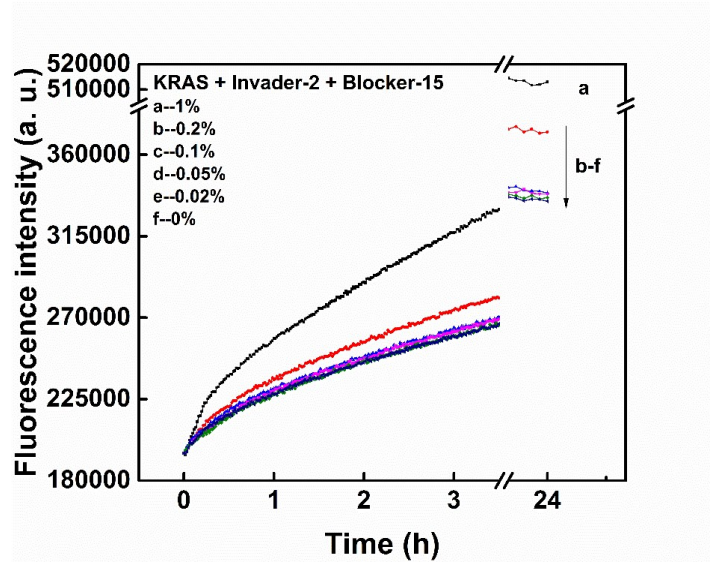


Figure S5. The detection limit of the dsCSD system targeting KrasG13D mutation in synthesized sample.

## 5. Optimization of universal dsCSD system

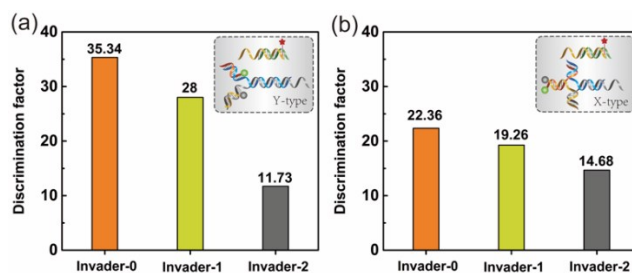


Figure S6. The discrimination factors (DF) of the universal dsCSD systems toward MT and WT with the length of invader chain ranging from 0 to 2nt. The concentration of invader chain was fixed at  $1\mu\text{M}$ . (a) Y-type universal dsCSD system. (b) X-type universal dsCSD system.

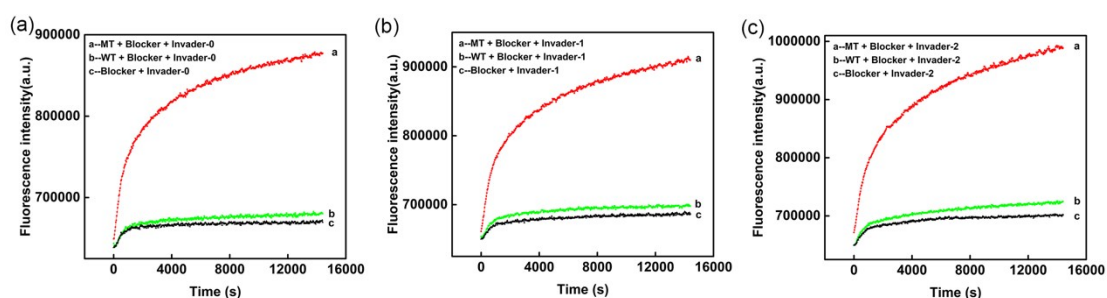


Figure S7. Fluorescence intensity responses of the Y-type universal dsCSD system in detection of the MT and WT with Invader-0/1/2. (a) Invader-0. (b) Invader-1. (c) Invader-2.

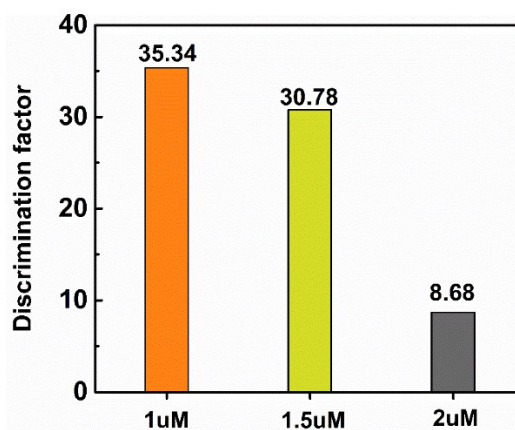


Figure S8. The discrimination factors (DF) of Y-type universal dsCSD system towards MT and WT with the concentration of invader chain ranging from 1 to  $2\mu\text{M}$ . The length of invader chain was fixed at 0nt.

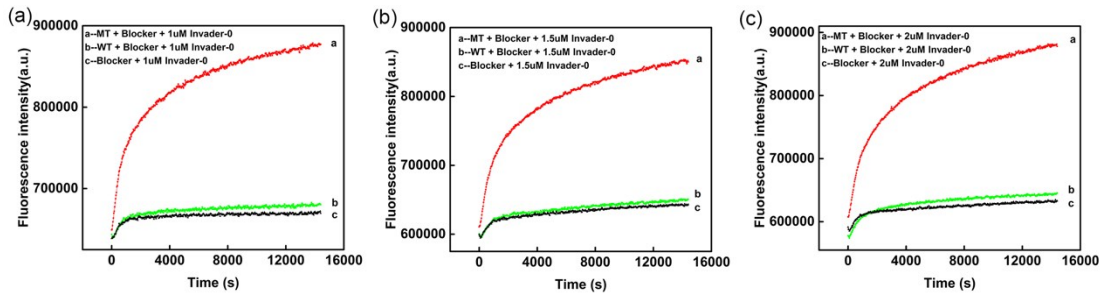


Figure S9. Fluorescence intensity responses of the Y-type universal dsCSD system in detection of the MT and WT with the concentration of invader chain ranging from 1 to 2 $\mu$ M. The length of invader chain was fixed at 0nt. (a) the concentration of invader chain was 1 $\mu$ M. (b) the concentration of invader chain was 1.5 $\mu$ M. (c) the concentration of invader chain was 2 $\mu$ M.

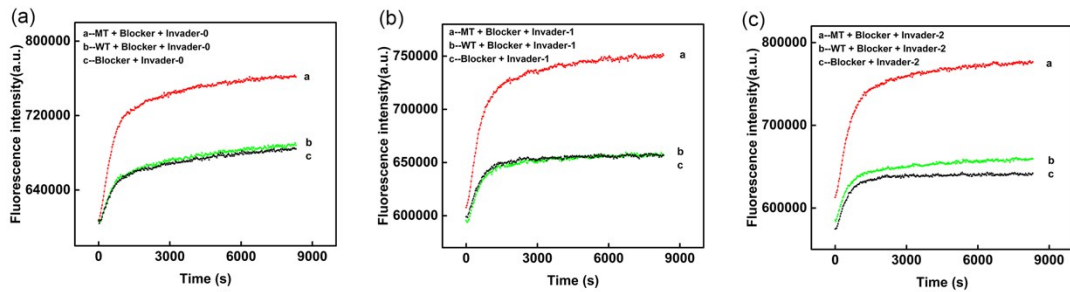


Figure S10. Fluorescence intensity responses of the X-type universal dsCSD system in detection of the MT and WT with the length of invader chain ranging from 0 to 2nt. The concentration of invader chain was fixed at 1 $\mu$ M. (a) the length of invader chain was 0nt. (b) the length of invader chain was 1nt. (c) the length of invader chain was 2nt.

6. The detection limit of Y-type universal dsCSD system targeting KrasG13D and BRCArs80357234 mutation in synthesized DNA sample

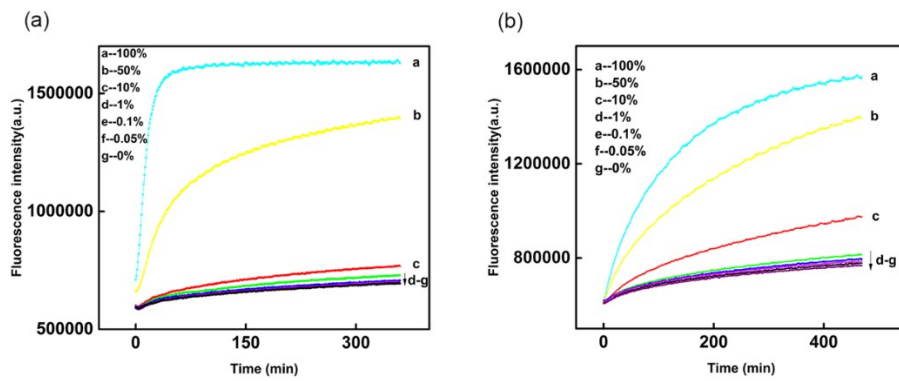


Figure S11. Fluorescence intensity responses of the dsCSD system in detection of the MT target with various abundances by using the Y-type universal probe. (a) Kras G13D. (b) BRCA rs80357234.

## 7. Genotyping using the Y-type universal dsCSD system to identify homozygous, heterozygous and wild type

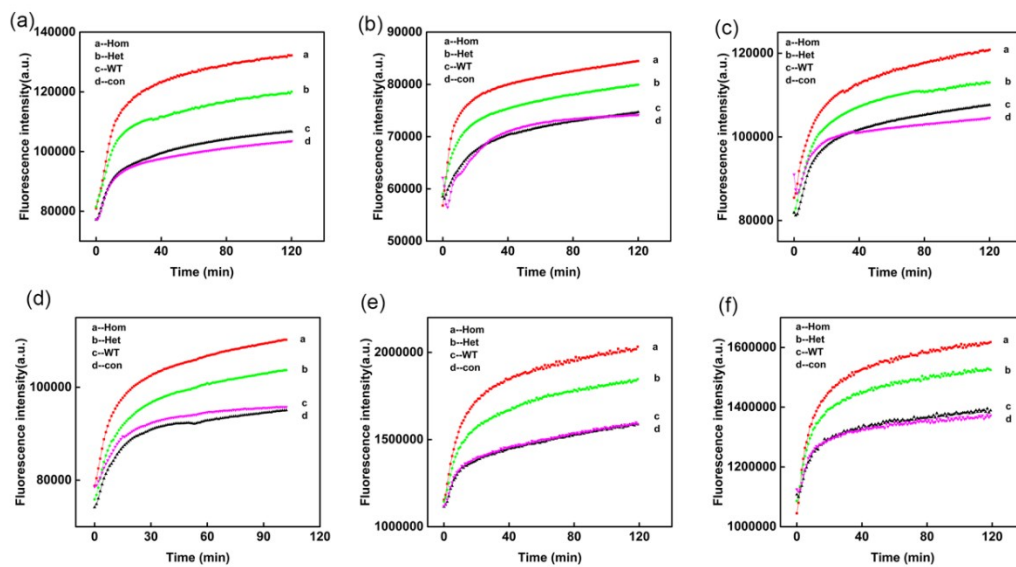


Figure S12. Genotyping using the Y-type universal dsCSD system. (a)BRCA rs1799949. (b)BRCA rs3765640. (c)BRCA rs16940. (d)MTRR A66G. (e)MTHFR A1298C. (f) MTHFR C677T.



8. The Sanger sequencing results of clinical samples used in this work

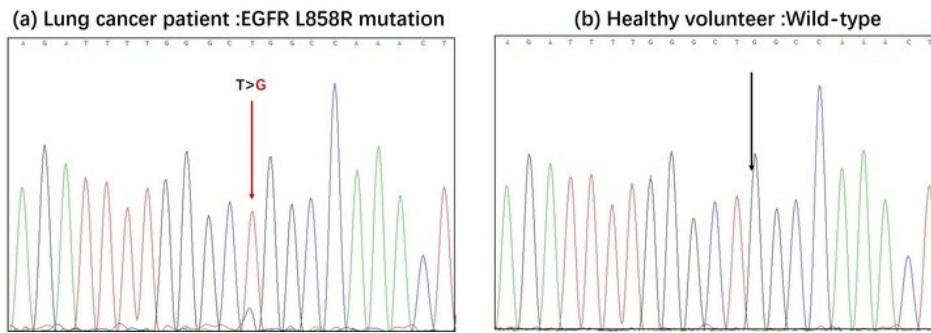


Figure S13. The sequencing results of the blood samples of (a) lung cancer patient and (b) healthy volunteer.

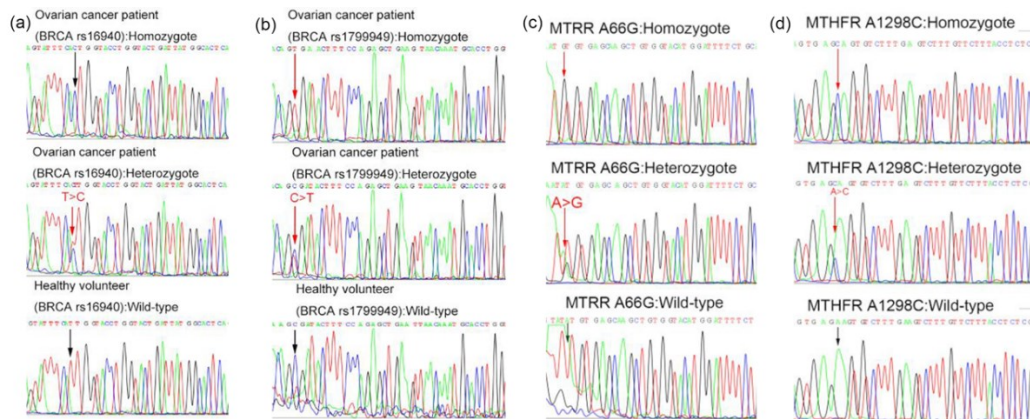


Figure S14. The sequencing results of the blood samples of ovarian cancer patients and healthy volunteers. (a)BRCA rs16940. (b)BRCA rs1799949. (c)MTRR A66G. (d)MTHFR A1298C.