Supporting Information

Sensitive discrimination of stable mismatched base pairs by abasic site modified fluorescent probe and lambda exonuclease

Tongbo Wu, Xianjin Xiao, Feidan Gu, and Meiping Zhao*

Beijing National Laboratory for Molecular Sciences, MOE Key Laboratory of Bioorganic Chemistry and Molecular Engineering, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

Experimental Section

Materials

Lambda exonuclease (λ exo) was purchased from Thermo Fisher Scientific Inc. (MA, USA). Uracil-DNA glycosylase (UDG), Exonuclease I (exo I), Lambda Exonuclease Buffer (67 mM Glycine-KOH (pH 9.4 @ 25°C), 2.5 mM $MgCl_2$, 50 μg/ml BSA) and ThermoPolReaction Buffer (20 mM Tris -HCl, 10 mM KCl，10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$ and 0.1% Triton X-100, pH 8.8 @ 25°C) were purchased from New England Biolabs (MA, USA). Taq DNA polymerase and dNTPs were purchased from Tiangen Biotech Co. (Beijing, China). DNA strands were synthesized and purified by HPLC (Sangon Biotech Co., China). The sequences of all the probes and targets that have been studied in this work are summarized in Table S1. DNase/RNase free deionized water purchased from Tiangen Biotech Co. (Beijing, China) was used with all the experiment.

Detection of target DNA strands (Fig. 1B, Table 1, Fig. S1, Fig. S2, Fig. S3, Table S2 and Table S3)

To a 200 μL PCR tube, 37 μL (for column 3 in Table 1, Fig. S3 and Table S3) or 39 μL (for the other tables and figures) of water, 5 μL of 10× ThermolPol Reaction Buffer, 2 μL of probe (10 pmol) and 1 μL of target ssDNA (5 pmol) were added and mixed well. For column 3 in Table 1, Fig. S3, Fig. S4 and Table S3, additional 2 μL of blocker (20 pmol) was added. The solution was heated to 85°C and then gradually cooled down to 37°C. Then 3 μL of λ exo (0.83 U for Fig. 1B, Table S2 and Fig. S1 and 5 U for the other figures and tables) was added and the detection was performed at 37°C on a Rotor-Gene Q 5plex HRM Instrument (QIAGEN, Hilden, Germany) with gain level of 7 (for Fig. 1B, Table S2 and Fig. S1) or 8 (for the other figures and tables). Fluorescence intensity was measured once a cycle (5 s per cycle) for 240 cycles. The excitation and emission wavelengths were set to 470 nm and 510 nm, respectively. The rate of fluorescence increase was determined by the slope of the linear portion of the time curve.

Detection of low abundance point mutations (Fig. 2, Fig. S4, Fig. S7 and Fig. S8)

To a 200 μL PCR tube, 37 μL of water, 5 μL of 10× ThermolPol Reaction Buffer, 2 μL of probe (10 pmol), 1 μL of mixed target ssDNA (total amount 5 pmol) and 2 μL of blocker (20 pmol) were added and mixed well. After the heating-annealing procedure same as above, 3 μL of λ exo (5 U) was added and the detection was performed at 37°C in the same manner as described above (gain=10 for Fig. 2, Fig. S7 and Fig. S8, gain=8 for Fig. S4).

Determination of the sensitivity of the method for *JAK2*V617F mutation detection (Fig. S5)

To a 200 μL PCR tube, 37 μL of water, 5 μL of 10× ThermolPolReaction Buffer, 2 μL of probe (10 pmol), 2 μL of blocker (20 pmol), and 1 μL of target ssDNA (1 pmol, 100 fmol, 10 fmol and 1fmol) or 1 μL of water as the blank were added and mixed well. After the heating-annealing procedure same as above, 3 μL of λ exo (5 U) was added and the detection was performed at 37 ^{o}C in the same manner as described above (gain=10).

Detection of low abundance *JAK2*V617F mutation in PCR amplicons (Fig. S6)

To a 200 μL PCR tube, 26.5 μL of water, 5 μL of 10× ThermolPol Reaction Buffer, 4 μL of dNTPs (10 nmol), 1 μL of forward primers (20 pmol), 1 μL of reverse primers (20 pmol), 1 μL of mixed template ssDNA (total amount 1 fmol, 32 pg), 0.5 μL of Taq (1.25 U) were added and mixed well and performed PCR procedure (94°C 30 s, 60°C 30 s, 72°C 20 s, 25 cycles) on a Rotor-Gene Q 5plex HRM Instrument (QIAGEN, Hilden, Germany). 1 μL of exo I (5 U) was added to the amplicons remove the unreacted primers, followed by inactivation at 85 ^{o}C for 10 min. Then 3 μL of λ exo (5 U) was added to digest the strand containing 5'-PO_4 in the duplex products for 20 min at 37 ^{o}C. After inactivation of λ exo at 85°C for 10 min, 2 μL of probe (10 pmol) and 2 μL of blocker (20 pmol) were added and mixed well. After the same heating-annealing procedure as above, 3 μL of λ exo (5 U) was added and the detection was performed at 37 ^{o}C in the same manner as described above (gain=10).

Table S1. DNA sequences used in this work.

Name	Sequences（5'->3'）
Sequences of the probes	
P-2A-3FAM-4C	PO_4-TAT(-FAM)CCACAGACACATACTCCA-BHQ1
P-2U-3FAM-4T	PO_4-TUT(-FAM)TCACAGACACATACTCCA-BHQ1
P-2AP-3FAM-4T	PO_4-TΔT(-FAM)TCACAGACACATACTCCA-BHQ1[a]
P-2U-3FAM-4A	PO_4-TUT(-FAM)ACACAGACACATACTCCA-BHQ1
P-2□-3FAM-4C	PO_4-T□T(-FAM)CCACAGACACATACTCCA-BHQ1[b]
P-2□-3FAM-4T	PO_4-T□T(-FAM)TCACAGACACATACTCCA-BHQ1
P-2□-3FAM-4A	PO_4-T□T(-FAM)ACACAGACACATACTCCA-BHQ1
P-2□-4C-15FAM	PO_4-T□TCCACAGACACAT(-FAM)ACTCCATAATTTAA-BHQ1
P-9□-10FAM-11C	PO_4-TCTCCACA□T(-FAM)CACATACTCCA-BHQ1
Tested target sequences	
Tar(21)-2V-4G (V=G,C,A,T)	GTTTTAAATTATGGAGTATGTGTCTGTGGA**V**ACGAGAGTAAG[c]
Tar(21)-2V-4C	GTTTTAAATTATGGAGTATGTGTCTGTG**C**A**V**ACGAGAGTAAG
Tar(21)-9V-11G	GTTTTAAATTATGGAGTATGTGA**V**TGTGGAGACGAGAGTAAG
Tar(21)-9V-11C	GTTTTAAATTATGGAGTATGT**C**A**V**TGTGGAGACGAGAGTAAG
Tar(21)-2G-4V	GTTTTAAATTATGGAGTATGTGTCTGTG**V**AGACGAGAGTAAG
Tar(16)-2G-4V	TATGTGTCTGTG**V**AGACGAGAGTAAG
Tar(15)-2G-4V	ATGTGTCTGTG**V**AGACGAGAGTAAG
Tar(14)-2G-4V	TGTGTCTGTG**V**AGACGAGAGTAAG
Tar(13)-2G-4V	GTGTCTGTG**V**AGACGAGAGTAAG
Tested blocker sequences for the targets	
Blo(15)-2C-4V	TCT**V**CACAGACACAT
Blo(14)-2C-4V	TCT**V**CACAGACACA
Blo(13)-2C-4V	TCT**V**CACAGACAC
Blo(12)-2C-4V	TCT**V**CACAGACA
Blo(11)-2C-4V	TCT**V**CACAGAC
Sequences used for *JAK2*V617F detection	
P-*JAK2*(16)	PO_4-T□T(-FAM)TCCTGTGGAGACG-BHQ1
Tar-*JAK2*-mutant	AGTTTTACTTACTCTCGTCTCCACAGAAA**C**ATACTCCATAA
Tar-*JAK2*-wild	AGTTTTACTTACTCTCGTCTCCACAGA**C**A**C**ATACTCCATAA
Blo-*JAK2*(16)	TGTGCCTGTGGAGACG
Blo-*JAK2*(15)	TGTGCCTGTGGAGAC
Blo-*JAK2*(14)	TGTGCCTGTGGAGA
Blo-*JAK2*(13)	TGTGCCTGTGGAG
Blo-*JAK2*(12)	TGTGCCTGTGGA
Forward-primer	PO_4-GCAGCAAGTATGATGAGCAA
Reverse-primer	GGCATTAGAAAGCCTGTAGTT
PCR-*JAK2*-Wild	GCAGCAAGTATGATGAGCAAGCTTTCTCACAAGCATTTGGTTT TAAATTATGGAGTATGT**G**TCTGTGGAGACGAGAGTAAGTAAAA CTACAGGCTTTCTAATGCC
PCR-*JAK2*V617F-	GCAGCAAGTATGATGAGCAAGCTTTCTCACAAGCATTTGGTTT TAAATTATGGAGTATGT**T**TCTGTGGAGACGAGAGTAAGTAAAA

mutant	CTACAGGCTTTCTAATGCC
Sequences used for Fig. S7	
P-*JAK*2(15)	PO_4-T□T(-FAM)TCCTGTGGAGACG-BHQ1
Tar-*JAK*2-mutant	AGTTTTACTTACTCTCGTCTCCACAGAAA**C**ATACTCCATAA
Tar-*JAK*2-wild-4G	AGTTTTACTTACTCTCGTCTCCACAGA**G**A**C**ATACTCCATAA
Blo-*JAK*2-4C(13)	TGTCCCTGTGGAG
Sequences used for Fig. S8	
P-*KRAS*(15)	PO_4-G□T(-FAM)CCTGGCGTAGGC
Tar-*KRAS*-mutant	CGTCAAGGCACTCTTGCCTACGCCACGA**G**CTCCAACTACC
Tar-*KRAS*-wild	CGTCAAGGCACTCTTGCCTACGCCAC**C**A**G**CTCCAACTACC
Blo-*KRAS*(13)	GATCCTGGCGTAG

[a] Δ stands for the natural abasic site generated by UDG.

[b] □ stands for the tetrahydrofuran abasic site.

[c] V in the sequence and the name of the targets and blockers represents G, C, A or T.

[d] The bases shown in bold and underlined in the tested targets indicate that they are mismatched with the opposite bases in the probe.

[e] The point mutation in the sequences of wild-type and V617F mutant type of *JAK2* is indicated in red.

Table S2. The discrimination factors of C:C type mismatched probe/target duplex by using different probes[a].

probe	**The base at -I position in the target**			
	G	C	A	T
P-2A-3FAM-4C	2.3	2.5	2.6	N/A[b]
P-2□-3FAM-4C	85	161	107	107
P-2□-4C-15FAM	3.0	1.4	1.4	1.5
P-9□-10FAM-11C	120	246	248	181

[a] The two single-nucleotide different target sequences for probes P-2A-3FAM-4C, P-2□-3FAM-4C and P-2□-4C-15FAM were Tar(21)-2V-4G and Tar(21)-2V-4C. The two targets for porbe P-9□-10FAM-11C were Tar(21)-9V-11G and Tar(21)-9V-11C.

[b] When the base at Position -I is T, it forms a matched basepair in the probe/target duplex for P-2A-3FAM-4C. So it is not applicable to our Position -I mismatch modification strategy.

Table S3. The discrimination factors (DF) for duplexes containing different types of mismatched base pairs at Position +I by using the abasic site modified probe in the absence and presence of different blocker strands.

Probe[a]	Type of mismatched base pair (N:M)	DF with different blockers				
		w/o[b]	Blo(11)-2C-4V	Blo(12)-2C-4V	Blo(13)-2C-4V[c]	Blo(14)-2C-4V
P-2□-3FAM-4C	C:C	427	580	**902**	*	*
	C:A	279	394	**530**	*	*
	C:T	237	311	**499**	*	*
P-2□-3FAM-4T	T:C	50	78	141	**562**	*
	T:T	43	54	66	**603**	*
	T:G	5.8	8.3	15	140	**447**
P-2□-3FAM-4A	A:A	46	45	54	**158**	*
	A:C	26	35	45	**179**	*
	A:G	16	22	28	88	**238**

[a] The target strands for probe P-2□-3FAM-4C are Tar(14)-2G-4V. The target strands for probe P-2□-3FAM-4T and P-2□-3FAM-4A are Tar(15)-2G-4V. V represents G, C, A or T. For target strands, V is matched with the base at Position +I in the mutant-type sequence. While for the blocker strands, V is matched with the base at Position +I in the wild -type sequence.

[b] DFs are obtained without the addition of blocker strands.

[c] * indicates that the blocker strands are too long, which significantly affected the hybridization between the probe and the MT.

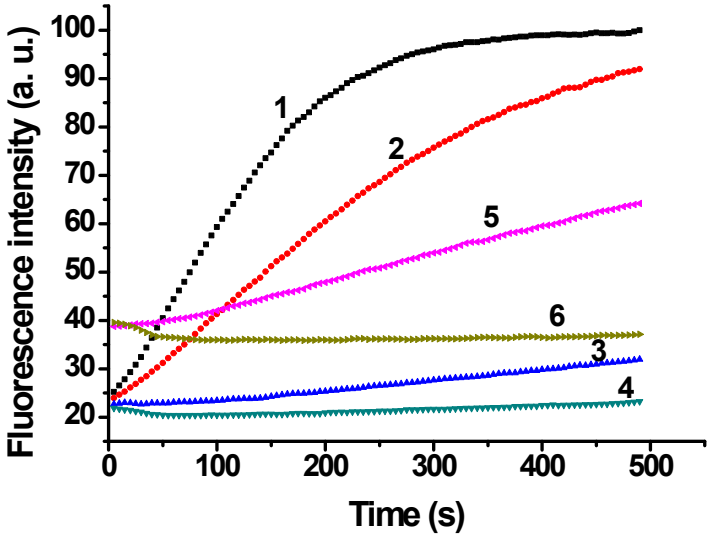

Fig. S1 The fluorescence intensity responses of probe P-2U-3FAM-4A (curve 1 and 2), P-2AP-3FAM-4A (curve 3 and 4), and P-2□-3FAM-4A (curve 5 and 6) for detection of targets Tar(21)-2G-4T (curve 1, 3 and 5) and Tar(21)-2G-4T (curve 2, 4 and 6). The sequences of P-2U-3FAM-4A and P-2□-3FAM-4A are listed in Table S1. P-2AP-3FAM-4A was obtained by treatment of P-2U-3FAM-4A with UDG.

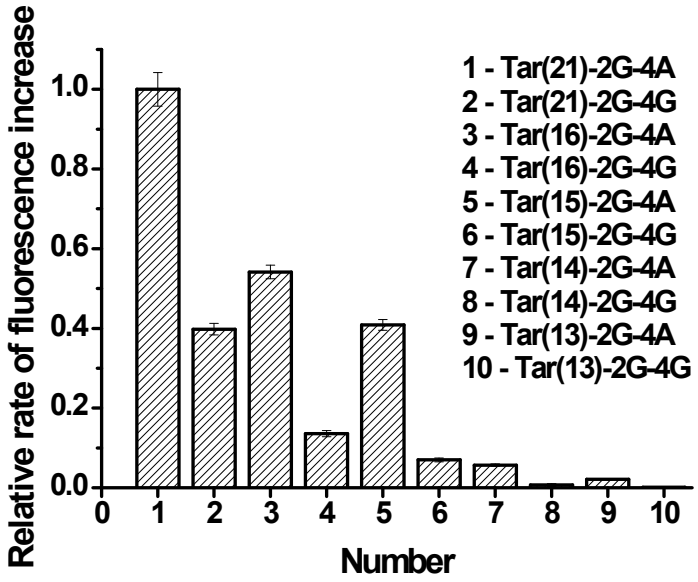

Fig. S2 Optimization of the length of the complementary region between P-2□-3FAM-4T and different target strands. The sequences of the targets are listed in Table S1. The rate of fluorescence increase of the P-2□-3FAM-4T/Tar(21)-2G-4A duplex was set to be 1.0. The DFs for T:G-mismatch at Position +I in P-2□-3FAM-4T/target duplexes with different length of complementary region were found to be 2.5 (21-nt), 4.0 (16-nt), 5.8(15-nt), 8.4(14-nt), 18.6(13-nt), respectively.

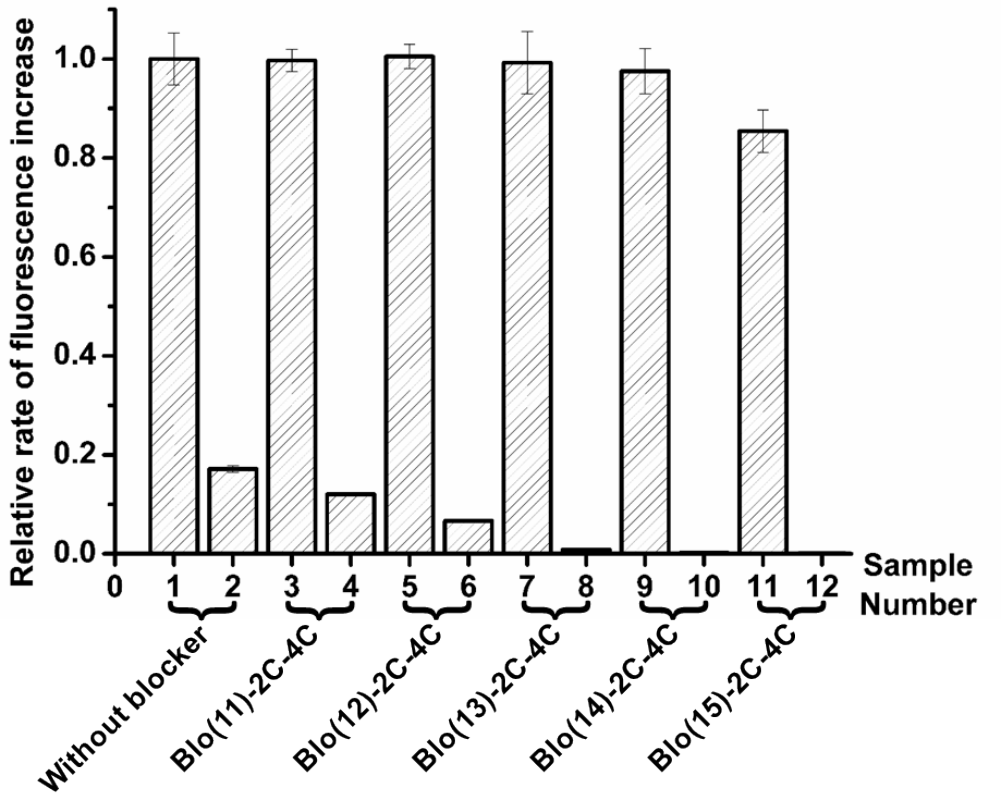

Fig. S3 Optimization of the length of the blocker strands with P-2□-3FAM-4T as the probe, Tar(15)-2G-4A as the target strand for samples 1, 3, 5, 7, 9, 11 and Tar(15)-2G-4G as the target for samples 2, 4, 6, 8, 10, 12. The sequences of the probe, targets and blockers are listed in Table S1. The amount of blocker is 20 pmol, which is two times of the probe and four times of the target. The rate of fluorescence increase of the P-2□-3FAM-4T/Tar(15)-2G-4A duplex without the addition of blocker (sample Number 1) was set to be 1.0. The resultant DFs are summarized in Table S3.

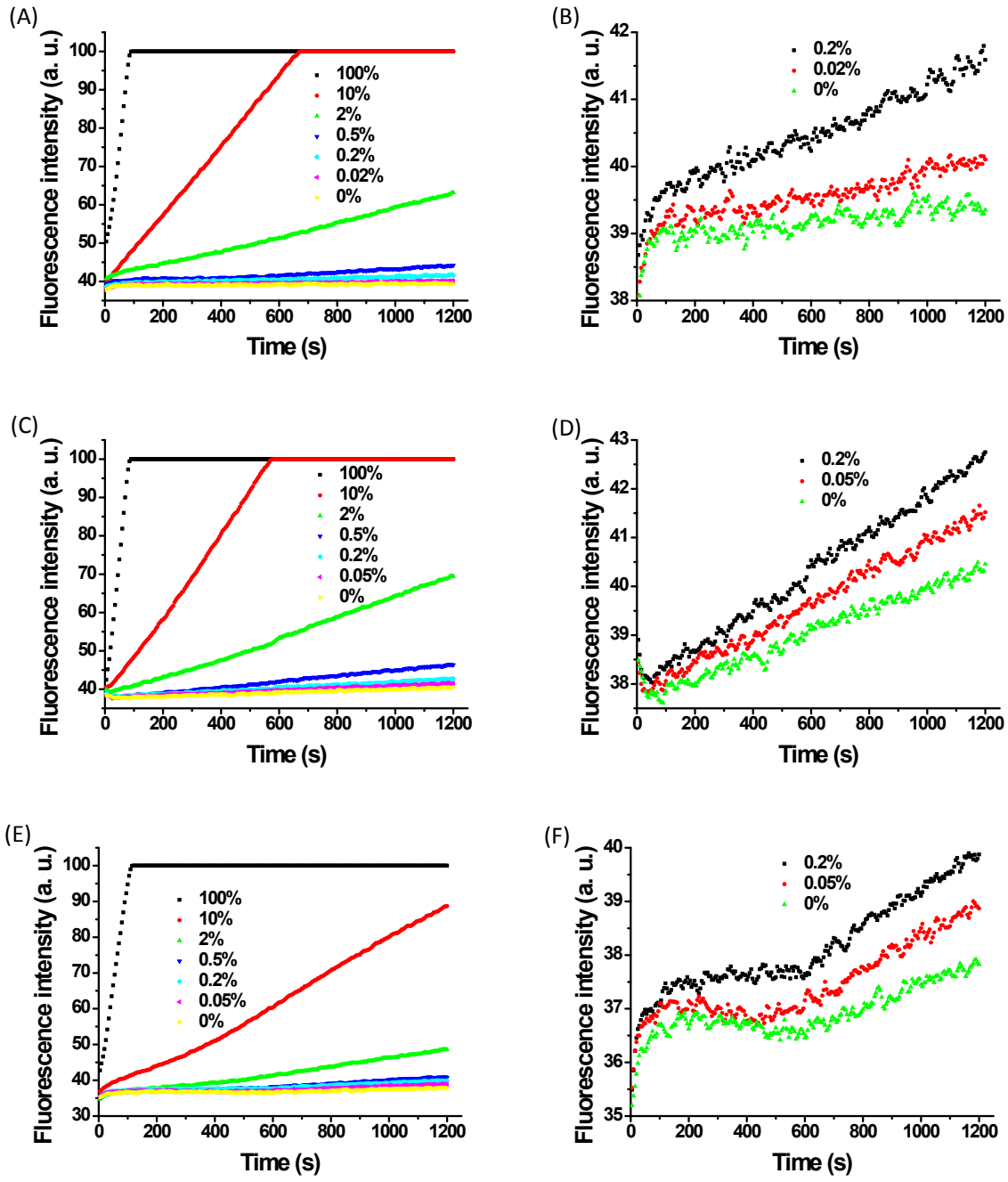

Fig. S4 Fluorescence intensity responses of P-2□-3FAM-4C (A and C) and P-2□-3FAM-4T (E) in the detection of mutations at different abundances. B, D and F show the curves for low-abundance samples (from 0% to 0.2%) in enlarged scale. 100% means the tested strands are all mutant type (Tar(14)-2G-4G for A, B, C and D, Tar(15)-2G-4A for E and F). 0% means the tested strands are all wild type (Tar(14)-2G-4C for A and B, Tar(14)-2G-4T for C and D, Tar(15)-2G-4G for E and F). The blocker used are Blo(12)-2C-4G for A and B, Blo(12)-2C-4A for C and D, and Blo(14)-2C-4C for E and F, respectively. The sequences of the probes, targets and blockers are all listed in Table S1.

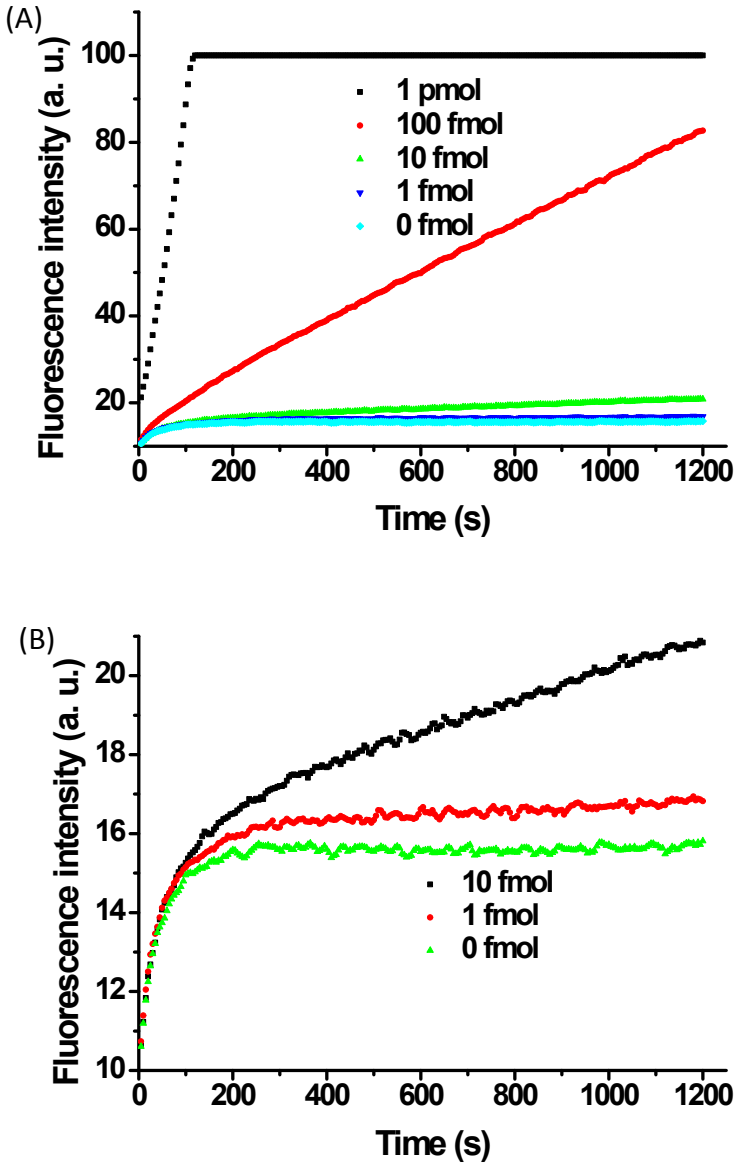

Fig. S5 Determination of the assay sensitivity by using P-*JAK2*(16) as the probe, Tar-*JAK2*-mutant as the target and Blo-*JAK2*(14) as the blocker. B show the curves for 0 fmol, 1 fmol and 10 fmol in enlarged scale. The sequences of the probe and target are listed in Table S1. The limit of detection is 1 fmol.

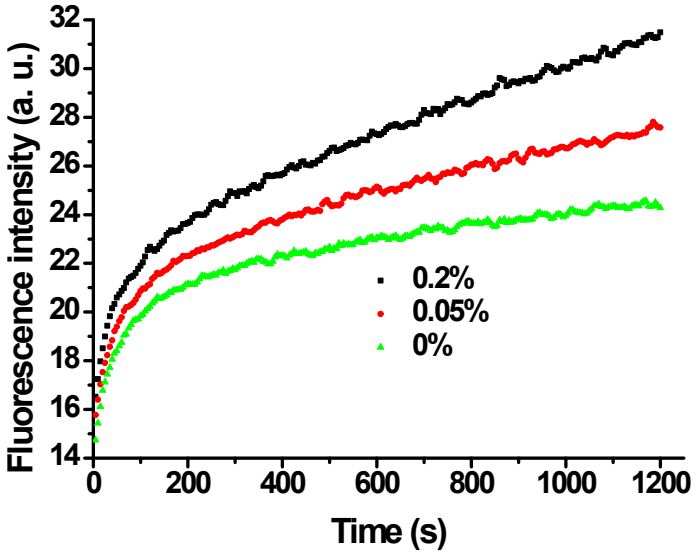

Fig. S6 Fluorescence intensity responses of P-*JAK2*(16) in the detection of low-abundance mutations 0.2% and 0.05% after PCR procedure. The mutant-type target is PCR-*JAK2*V617F-mutant and the wild-type strand is PCR-*JAK2*-wild. The blocker used is Blo-*JAK2*(14).

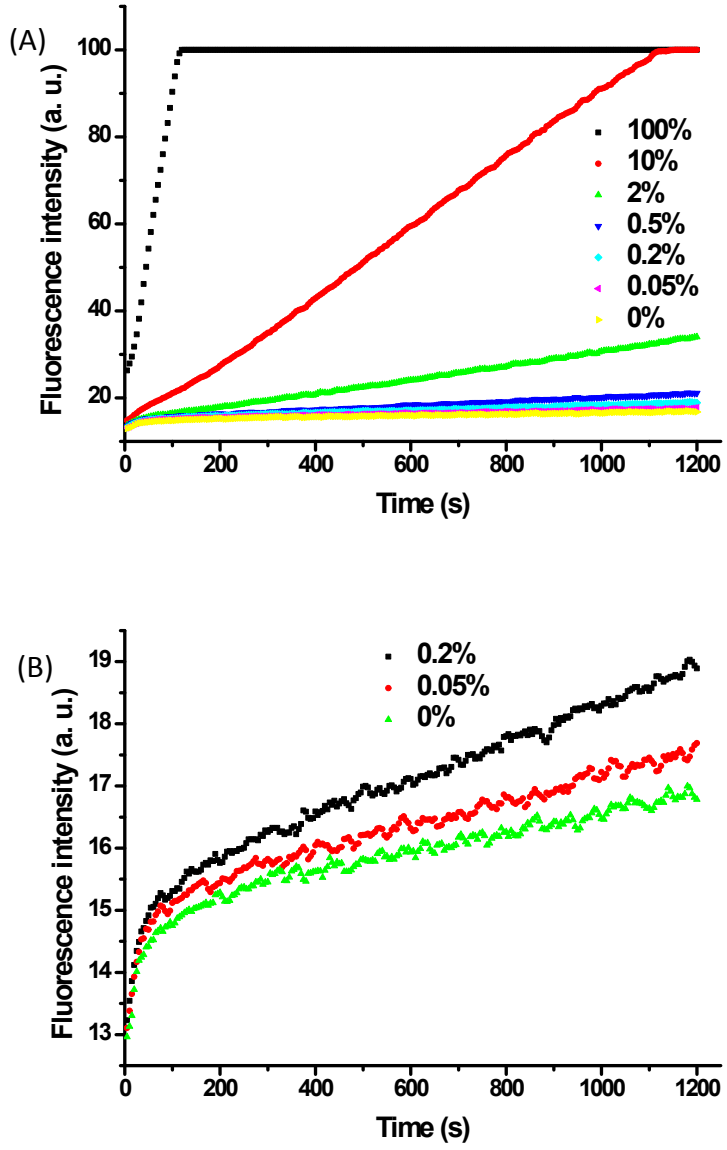

Fig. S7 (A) Fluorescence intensity responses of P-*JAK*2(15) in the detection of mutant type (Tar-*JAK*2-mutant) at different abundances (from 0% to 100%, T:G mismatch type). (B) The curves for low-abundance samples (0%, 0.05% and 0.2%) in enlarged scale. 100% means the tested strands are all mutant type (Tar-*JAK*2-mutant). 0% means the tested strands are all wild type (Tar-*JAK*2-wild-4G). The blocker used is Blo-*JAK*2-4C(13). The sequences of the probe, blocker and targets are listed in Table S1.

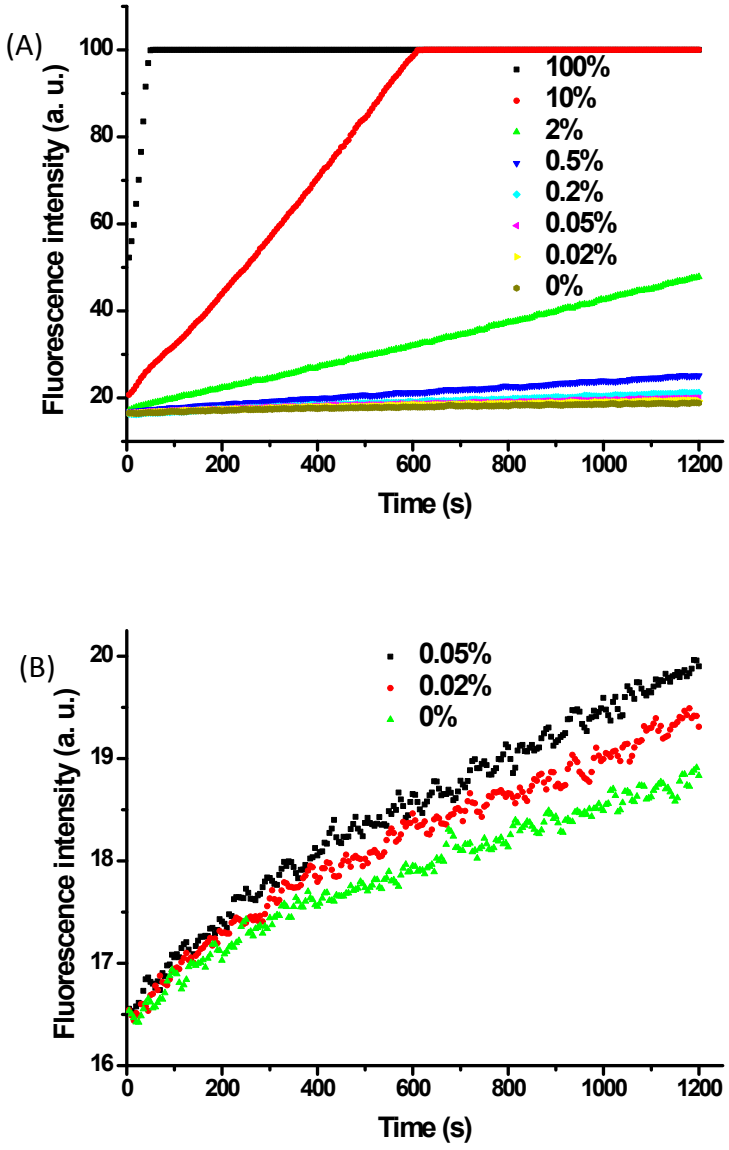

Fig. S8 (A) Fluorescence intensity responses of P-*KRAS*(15) in the detection of mutant type (Tar-*KRAS*-mutant) at different abundances (from 0% to 100%, C:C mismatch type). (B) The curves for low-abundance samples (0%, 0.02% and 0.05%) in enlarged scale. 100% means the tested strands are all mutant type (Tar-*KRAS*-mutant). 0% means the tested strands are all wild type (Tar-*KRAS*-wild). The blocker used is Blo-*KRAS*(13). The sequences of the probe, blocker and targets are listed in Table S1.

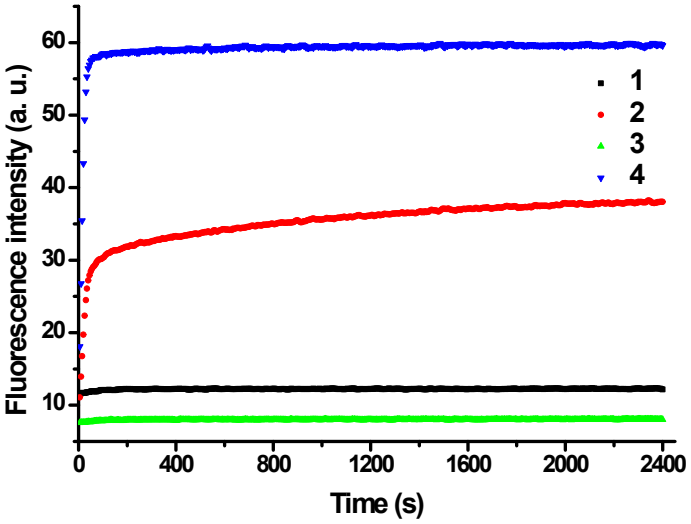

Fig. S9 Test of the stability of the THF-modified probe in ThermoPol Reaction Buffer used in this study. Curve 1 and Curve 3 show the fluorescence intensity responses of the probe (5'-3': FAM-TATCTGCAC□AFATFCACCT(-BHQ1) CATAAT) itself in ThermoPol Reaction Buffer (Curve 1) and Lambda Exonuclease Buffer (Curve 3), respectively. For comparison, Curve 2 and Curve 4 showed the fluorescence intensity responses of the same probe in the presence of T5 exonuclease in the two buffers (ThermoPol Reaction Buffer: Curve 2; Lambda Exonuclease Buffer: Curve 4), respectively. The results confirmed that the probe was stable in both buffers for at least 40 min, which was much longer than the assay time.