Electronic Supplementary Information

Specific multiplexed detection of mRNA splice variants based on size-coding DNA probes and universal PCR amplification

Yuting Jia, Honghong Wang*, Hui Wang, Fangfang Wang, Kejian Gao and Zhengping Li*

Beijing Key Laboratory for Bioengineering and Sensing Technology; School of Chemistry and Biological Engineering, University of Science and Technology Beijing, 30 Xueyuan Road, Haidian District, Beijing 100083, China.

Email: wanghh@ustb.edu.cn; lzpbd@ustb.edu.cn.

Content List

1. Table S1. The sequences of the oligonucleotides used in this work

Name	Sequence (5′-3′ direction)
hTERT-Δ(7, 8)	GUCCGCAAGGCCUUCAAGAGCCACGUCCUACGUCCAGUGCCAGGGGAU
hTERT-Δ(11)	CACCCACGCGAAAACCUUCCUCAGCUAUGCCCGGACCUCCAUCAGAGC
BRCA1-Δ(9, 10)	AAAGACGUCUGUCUACAUUGAAUUGGCUGCUUGUGAAUUUUCUGAGACGG
BRCA2-Δ(18, 4-)	GGGUGCUUCUUCAACUAAAAUACAGAUAUGAUACGGAAAUUGAUAGAAGC
TP53-Δ(4)	AAACAUUUUCAGACCUAUGGAAACUACUUCCUGAAAACAACGUUCUGUCC
FGFR1-Δ(4, 6-)	GAAACAGAUAACACCAAACCAAACCCCGUAGCUCCAUAUUGGACAUCCCC
$PA_{\Delta(7, 8)}$	CCATCTCATCCCTGCGTGTCTGGCACTGGACGTAGGAC
$PB_{\Delta(7, 8)}$	CCGCTTTCCTCTCTATGGGCCAAGGCCTTCAAGAGCCAC
$PA_{\Delta(11)}$	CCATCTCATCCCTGCGTGTCTGATGGAGGTCCGGGCATAG
$PB_{\Delta(11)}$	CCGCTTTCCTCTCTATGGGCAGTCACGCGAAAACCTTCCTCAG
$PA_{\Delta(9, 10)}$	CCATCTCATCCCTGCGTGTCGTCTCAGAAAATTCACAAGCAGC
$PB_{\Delta(9, 10)}$	CCGCTTTCCTCTCTATGGGCAAGACGTCTGTCTACATTGAATTG
$PA_{\Delta(18, 4-)}$	CCATCTCATCCCTGCGTGTCATCGCTTCTATCAATTTCCGTATCATAT
$PB_{\Delta(18, 4-)}$	CCGCTTTCCTCTCTATGGGCGGGTGCTTCTTCAACTAAAATACAG
$PA_{\Delta(4)}$	CCATCTCATCCCTGCGTGTCATCGCTAGGACAGAACGTTGTTTTCAGGAAGT
$PB_{\Delta(4)}$	CCGCTTTCCTCTCTATGGGCTATAAACATTTTCAGACCTATGGAAACT
$PA_{\Delta(4, 6-)}$	CCATCTCATCCCTGCGTGTCATGTACGTAGGGGATGTCCAATATGGAGCTACGG
$PB_{\Delta(4, 6-)}$	CCGCTTTCCTCTCTATGGGCCATGACGCATGAAACAGATAACACCAAACCAAA CC
UP1	CCGCTTTCCTCTCTATGGGC
UP2	FITC-CCATCTCATCCCTGCGTGTC
Internal standard poly(A)	FAM-AAAAAAAAAAAAAAAAAAA-PO_4

2. Validation of amplification mechanism of mRNA splice variants assay by using polyacrylamide gel electrophoresis (PAGE)

To further support the amplification mechanism of the proposed mRNA splice variant assay, the hybridizations between oligonucleotide strands and PCR amplification products are characterized by non-denaturing PAGE. As shown in Fig. S1, *BRCA1*-Δ(9, 10) (Lane 2) and $PA_{\Delta(9,\ 10)}$ (Lane 3) respectively show a defined band corresponding to 50 bp and 43 bp. In the presence *BRCA1*-Δ(9, 10) target, the hybridization products between *BRCA1*-Δ(9, 10) and $PA_{\Delta(9,\ 10)}$ (Lane 4) and their reverse transcription reaction products (Lane 5) respectively show a clear band, where the reverse transcriptional reaction products (ss-cDNA) and the commercial synthetic cDNA have the same electrophoretic position, indicating that the *BRCA1*-Δ(9, 10) target is reversely transcribed to form cDNA. Then, the $PB_{\Delta(9,\ 10)}$ hybridized with ss-cDNA and performed 12 cyclic amplification reactions, resulting in a clear double-stranded cDNA product band (Lane 9). After 26 PCR cycles' amplification, the PCR amplified product is shown in Lane 12. By contrast, the Blank control without *BRCA1*-Δ(9, 10) does not produce any observable band of amplification products except for the unreacted UP1 and UP2 (Lane 13). These results demonstrate that the proposed method is feasible for mRNA splice variants assay.

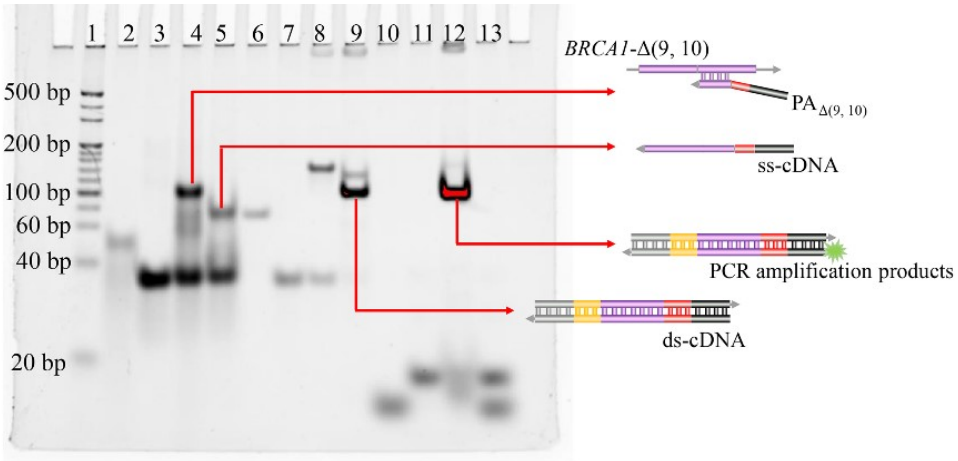

Fig. S1. Electrophoregram of non-denaturing PAGE analysis. Lane 1: double-stranded DNA markers; Lane 2: *BRCA1*-Δ(9, 10) (10 μM); Lane 3: $PA_{\Delta(9,\ 10)}$ (20 μM); Lane 4: the hybridization products of *BRCA1*-Δ(9, 10) + $PA_{\Delta(9,\ 10)}$; Lane 5: the reverse transcription product (ss-cDNA); Lane 6: the commercial synthetic cDNA (1 μM); lane 7: $PB_{\Delta(9,\ 10)}$ (2 μM); Lane 8: the hybridization products of the commercial synthetic cDNA and $PB_{\Delta(9,\ 10)}$; Lane 9: the double-stranded cDNA (ds-cDNA); Lane 10: UP1 (10 μM); Lane 11: UP2 (10 μM); Lane 12: the PCR amplification products for *BRCA1*-Δ(9, 10) assay; Lane 13: the PCR amplification product of Blank.

3. Optimization of the concentration of PA and PB

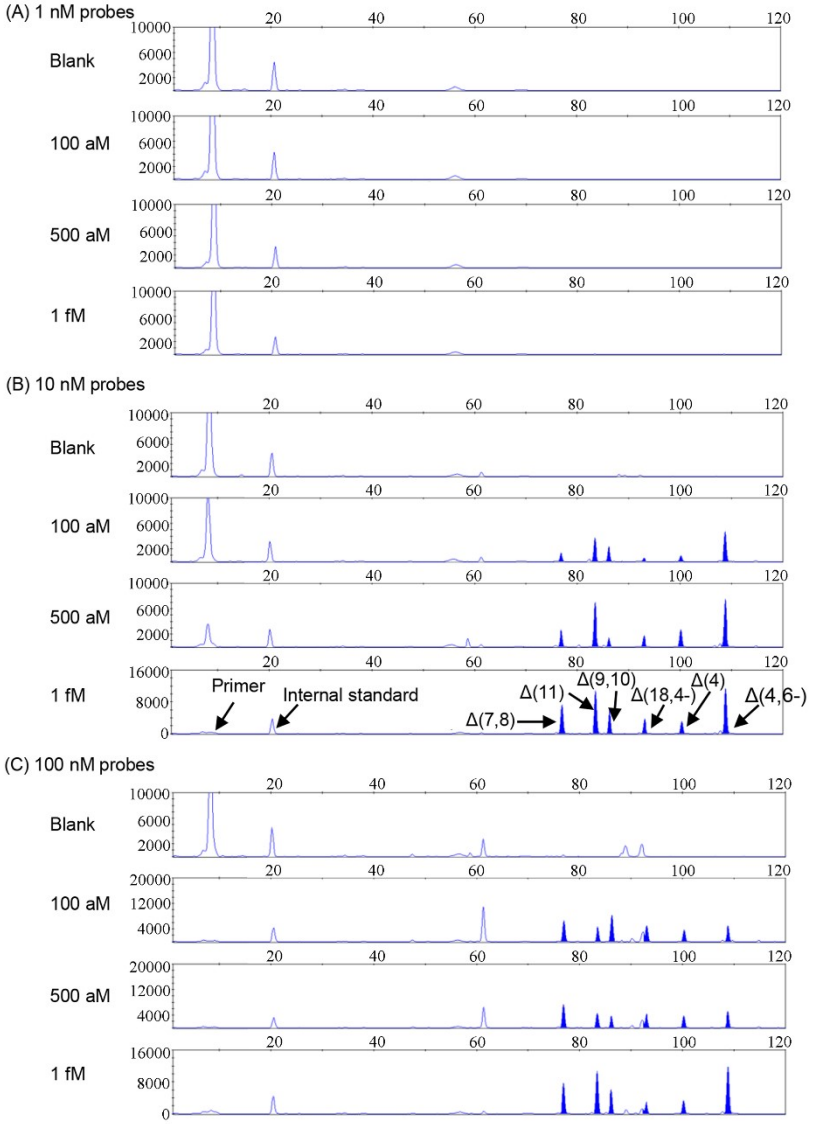

Fig. S2. The effect of the dosage of PA and PB on the mRNA splice variants assay based on size-coding DNA probes and universal PCR amplification. The dosage of PA and PB is 1 nM (A), 10 nM (B), and 100 nM (C), respectively.

The concentration of PA and PB will affect the efficiency of cDNA synthesis. To evaluate the influence of the dosage of probes, 1 nM, 10 nM, and 100 nM PA and PB were investigated by detecting the Blank and 6 types of mRNA splice variants including *hTERT*-Δ(7, 8), *hTERT*-Δ(11), *BRCA1*-Δ(9, 10), *BRCA2*-Δ(18, 4-), *TP53*-Δ(4), *FGFR1*-Δ(4, 6-) with the proposed assay. The concentration of each mRNA splicing variant was in the order of 100 aM, 500 aM, and 1 fM, respectively. As depicted in Fig. S2, when the concentration of the probe was 1 nM (Fig. S2A), the splice variants do not produce a corresponding product peak after PCR amplification, indicating that

the concentration of the probe was too low to effectively generate cDNA for PCR amplification. In sharp contrast, when the concentration of the probe was 10 nM (Fig. S2B), six product peaks were observed for each splice variant, and no heterologous peaks were found. One can see from Fig. S2C, when the probe concentration increases to 100 nM, there were heterologous peaks at approximately 90 bp. This may be due to excessive probes leading to severe nonspecific hybridization causing nonspecific amplification. Therefore, considering the detection sensitivity and interference of non-specific amplification, 10 nM probe was selected for the multiple mRNA splice variants assay.

4. Optimization of the cycle numbers of PCR

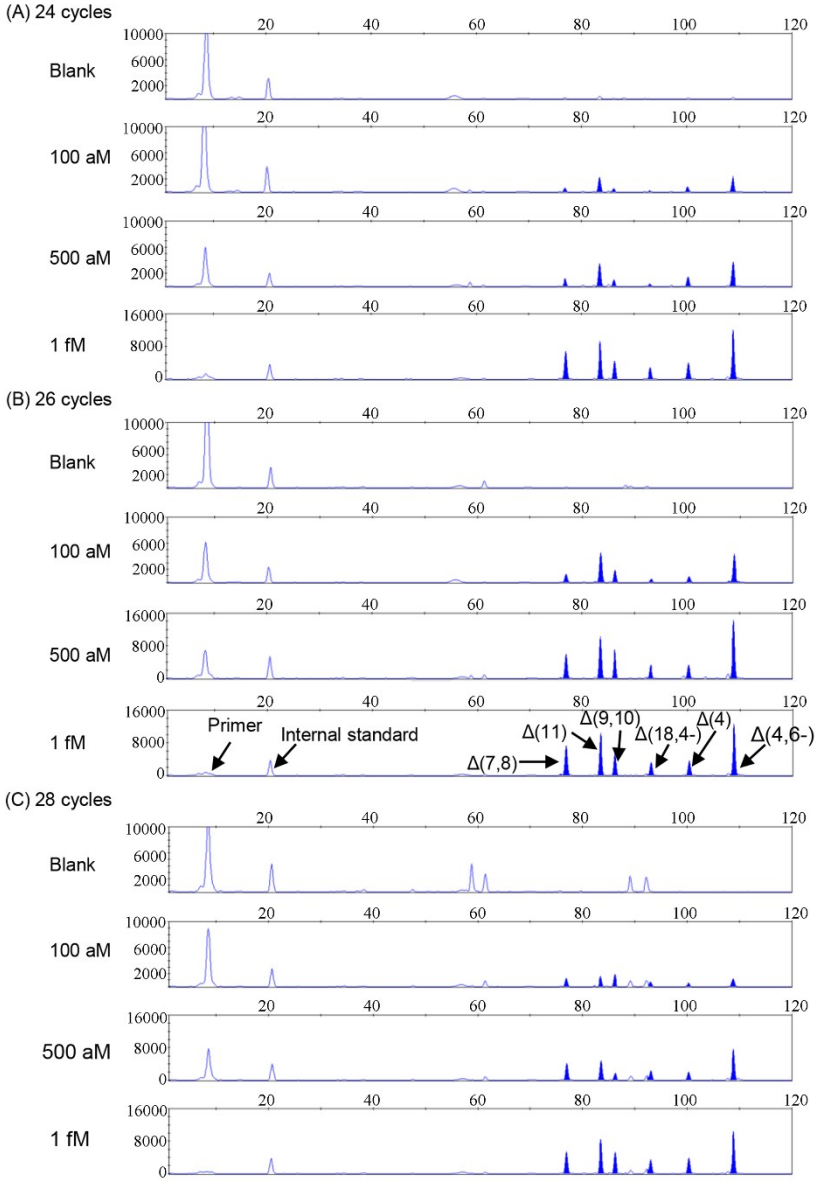

Fig. S3. The effect of the cycle number of PCR on the mRNA splice variants assay based on size-coding DNA probes and universal PCR amplification. The cycle number of PCR is 24 (A), 26 (B), and 28 (C), respectively.

Nextly, the cycle number of PCR was further optimized by detecting Blank and 6 types of mRNA splice variants including *hTERT*-Δ(7, 8), *hTERT*-Δ(11), *BRCA1*-Δ(9, 10), *BRCA2*-Δ(18, 4-), *TP53*-Δ(4), *FGFR1*-Δ(4, 6-) with the proposed assay. The concentration of each mRNA splicing variant was in order of 100 aM, 500 aM, and 1 fM, respectively. As shown in Fig. S3, when the cycle number of PCR was 24 (Fig. S3A), the product peak of the 100 aM splice variants cannot be observed, indicating that the PCR reaction has not reached the plateau stage of amplification. When the cycle number of PCR was 26 (Fig. S3B), the product peaks of splice variants as low as 100 aM can be observed. As the cycle number of PCR increases to 28 (Fig. S3C), an obvious impurity peak was observed at 90 bp, indicating that non-specific amplification has occurred, which can interfere with the detection of target splice variants. Considering the detection sensitivity and interference from non-specific amplification, 26 cycles were selected as the optimal reaction conditions.

5. Table S2. The correlation equations and corresponding correlation coefficients for multiple detections of mRNA splice variants

Splice variant	The correlation equation and corresponding correlation coefficient (100 aM - 5 fM)	
hTERT-Δ(7, 8)	$A=15.70+0.94\lg C_{\Delta(7, 8)}$	$R^2=0.9913$
hTERT-Δ(11)	$A=20.83+1.26\lg C_{\Delta(11)}$	$R^2=0.9953$
BRCA1-Δ(9, 10)	$A=12.26+0.73\lg C_{\Delta(9, 10)}$	$R^2=0.9992$
BRCA2-Δ(18, 4-)	$A=8.55+0.52\lg C_{\Delta(18, 4-)}$	$R^2=0.9766$
TP53-Δ(4)	$A=10.20+0.63\lg C_{\Delta(4)}$	$R^2=0.9590$
FGFR1-Δ(4, 6-)	$A=24.41+1.44\lg C_{\Delta(4, 6-)}$	$R^2=0.9931$

6. The repeatability of the proposed strategy for multiple mRNA splice variants assay

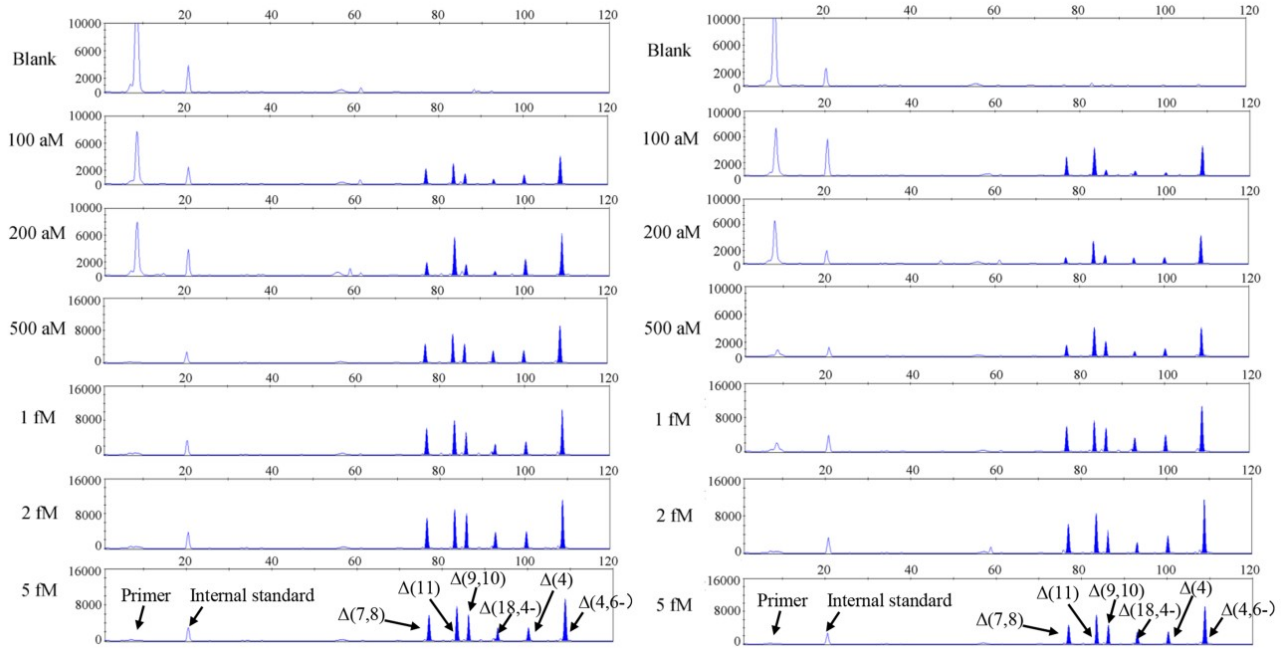

Fig. S4. The repeatability of the proposed strategy for multiple mRNA splice variants assay.

7. Table S3. Comparison of the proposed strategy with previously reported methods for mRNA splicing variants assay

Target	Background Signal	Detection Limit	Detection Throughput	Experimental Time	Reference
CD45 mRNA splicing variants	+	Single-cell analysis	high throughput	overnight	ACS. Cent. Sci., 2018, 4(6): 680-687.
JM and CYT mRNA splicing variants	+	Semi-quantitative analysis	high throughput	4 h	Mol. Neurobiol., 2018, 55(7): 6169-6181.
FGFR mRNA splicing variants	-	100 aM	2 kinds	3 h	RSC Adv., 2020, 10(11): 6271-6276.
endogenous genes mRNA splicing variants	+	single cell analysis	high throughput	overnight	Nucleic Acids Res. 2022, 50(22): e130
Alternative splicing analysis in total RNA	-	Single-cell analysis	high throughput	overnight	Anal. Chem., 2022, 94(36): 12342-12351.
hTERT, BRCA1, BRCA2, TP53, FGFR mRNA splicing variants, et al.	-	100 aM	high throughput	4 h	This work