Supporting Information

SpliceRCA: *in situ* single-cell analysis of mRNA splicing variants

Xiaojun Ren,[a,b] Ruijie Deng,[a] Kaixiang Zhang,[a] Yupeng Sun,[a] Xucong Teng,[a] Jinghong Li*[a]

a. Department of Chemistry, Key Laboratory of Bioorganic Phosphorus Chemistry & Chemical Biology, Tsinghua University, Beijing 100084, China
b. School of Chemistry and Chemical Engineering, Beijing Institute of Technology, Beijing, 100081, P. R. China

* To whom correspondence should be addressed. Email: jhli@mail.tsinghua.edu.cn

Table of Contents

SI Materials and methods

Materials. T4 DNA ligase, phi29 DNA polymerase, T4 polynucleotide kinase, RiboLock RNase Inhibitor, RevertAid First Strand cDNA Synthesis Kit and SYBR select master mix were purchased from Thermo Fisher Scientific (Waltham, USA). Diethy pyrocarbonate (DEPC), formamide, Tween-20 and Triton-X100 were obtained from Sigma-Aldrich (St. Louis, USA). The 20×SSC buffer (pH 7.4), salmon sperm DNA and 4% paraformaldehyde in PBS buffer were bought from Beijing Solarbio Science & Technology Co., Ltd. (Beijing, China). Deoxyribonucleotides mixtures (dNTPs) were purchased from Beijing DingGuo Biotechnology Co., Ltd. (Beijing, China). TransZol and TransScript one-step gDNA removel and cDNA synthesis were bought from Transgen biotech Co., Ltd. (Beijing, China). All of the solution and deionized water used were treated with DEPC and autoclaved to be protected from RNase degradation.

***In vitro* detection of RNA splicing variants by SpliceRCA.** The ligation reaction was conducted in a 20 μL reaction mixture containing 1×T4 ligation buffer [40 mM Tris–HCl, 10 mM DTT, 10 mM $MgCl_2$, 0.5 mM ATP (pH 7.8 at 25 °C)], 2 μL of the splice-junction anchored padlock probe (1 μM), 5 μL of target RNA and 5 U of T4 ligase at 25°C for 1 h. Then 0.5 μL phi29 DNA polymerase, 5 μL dNTPs and 1× isothermal amplification buffer [20 mM Tris– HCl, 50 mM KCl, 10 mM $(NH4)_2SO4$, 2 mM $MgSO_4$, 0.1% Tween 20 (pH 8.8 @ 25°C)] were added to the above reaction mixture. The RCA reaction was performed at 37 °C for 2 h and terminated by heating to 65 °C for 10 min. Then The RCA products were mixed with 100×Sybr Green I to be the final solution with a 1×Sybr Green I. 30 μL of the solution was transferred to a cuvette and the fluorescent spectra were measured on a EnVision Multilabel Plate Readers (PerkinElmer, USA).The excitation wavelength was 488 nm, and the emission wavelengths were 520 nm.

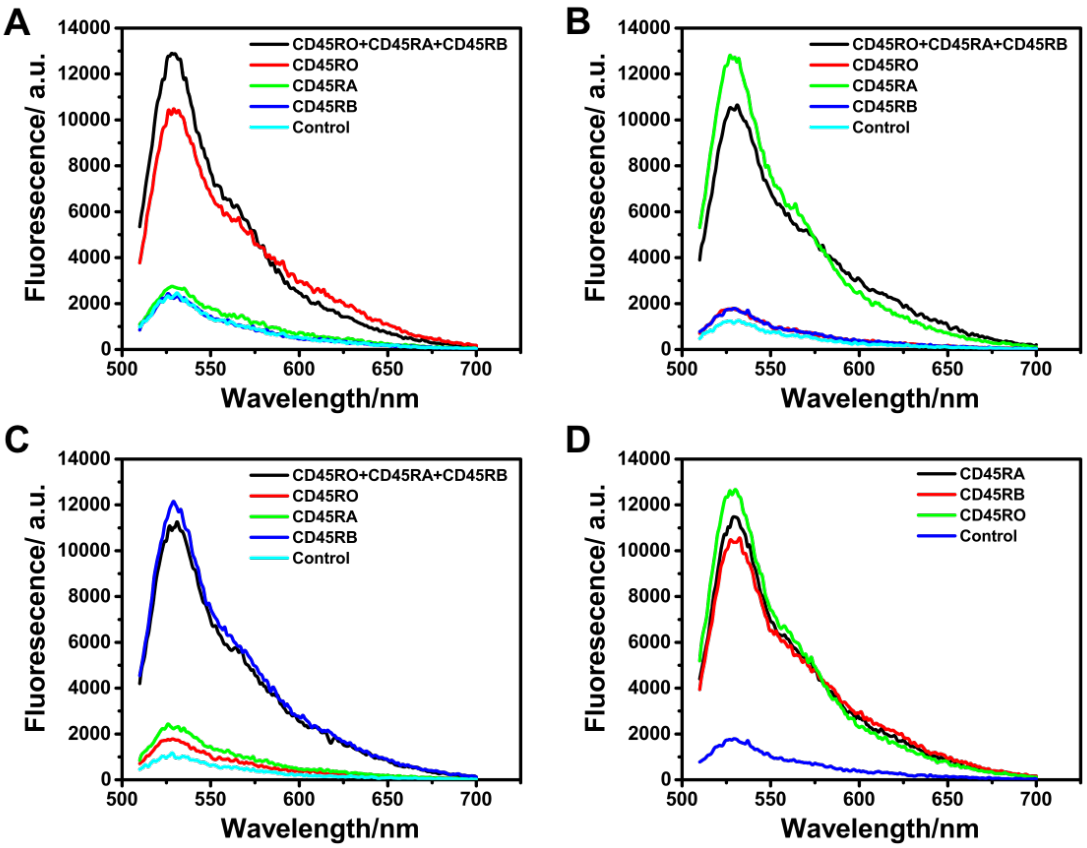

Figure S1. Evaluation specificity of padlock probes by SpliceRCA *in vitro.* Fluorescence intensity of amplification products triggered by splice-junction padlock probes S1 (**A**), S2 (**B**) and S3 (**C**) respectively in the presence of three targets DNA (black), and dependent presence of CD45RO (red), CD45RA (green), CD45RB (dark blue) (S1, S2, S3, padlock probes targeting isoforms CD45RO, CD45RA, CD45RB, respectively). Control means that in the absence of three target DNA (light blue). **D)** Fluorescence intensity of amplification products by three padlock probes mixture in the presence of CD45RO (green), CD45RA (black), CD45RB (red) and the absence of three target RNA (dark blue).

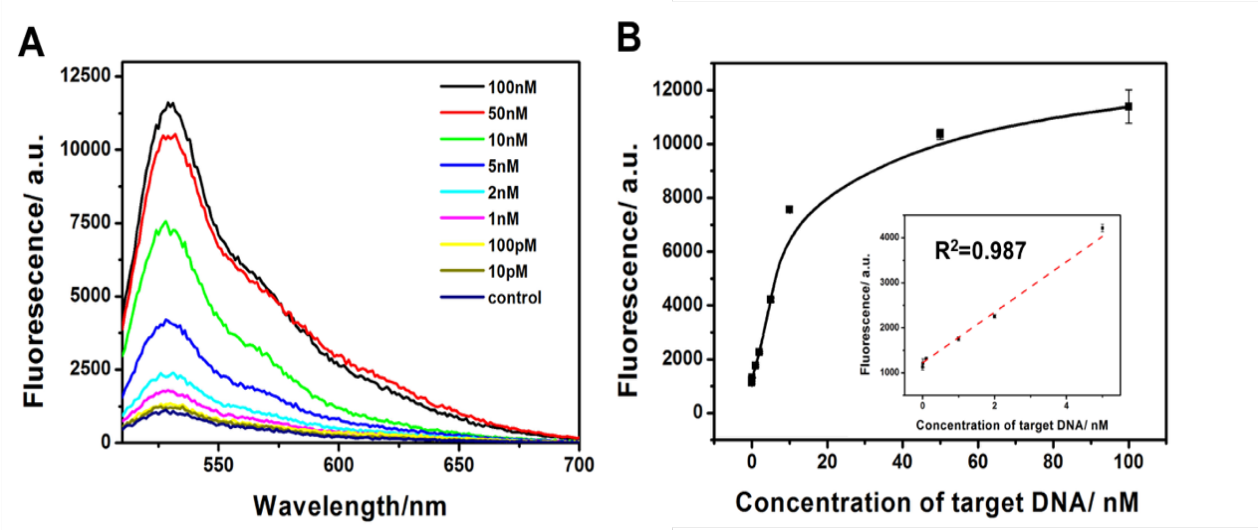

Figure S2. SpliceRCA for quantification of the target sequence *in vitro*. (A) Fluorescence spectral responses to target splice junction sequence of varying concentrations ranging from 0 to 100 nM *in viro*. (B) Dependence of the fluorescence intensity at 520 nm on the target concentration. Inset shows a linear range from 10 pM to 5 nM. Error bars are based on triplicate experiments. The concentrations of padlock probes, phi29 polymerase were 10 nM and 0.25 U/μL.

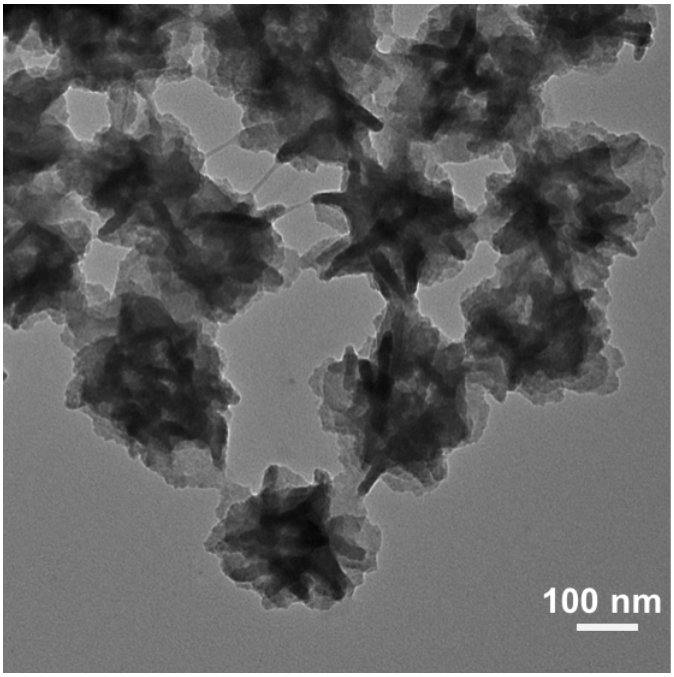

Figure S3. Transmitted electron microscopy (TEM) image of RCA amplicons.

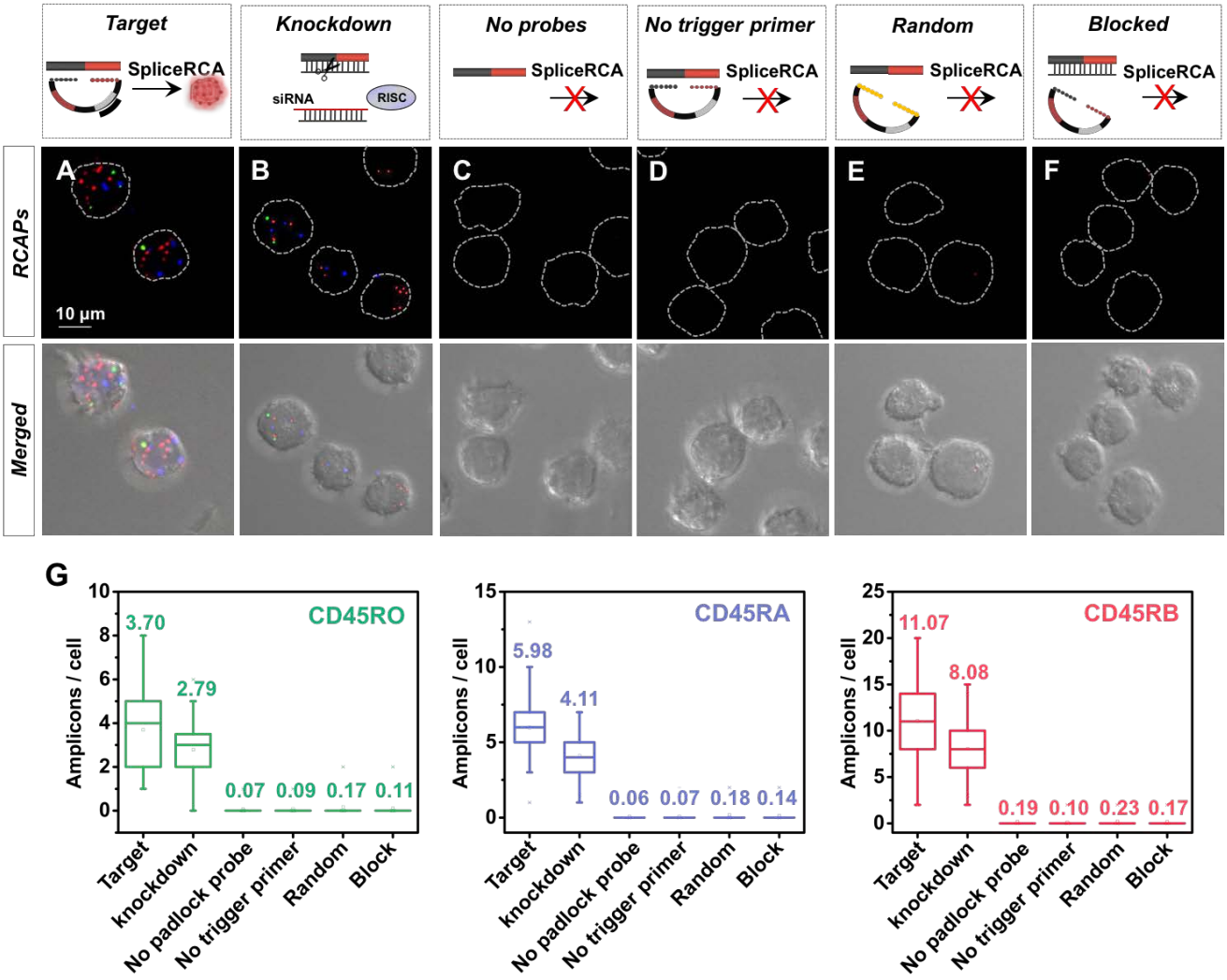

Figure S4. Evaluation of SpliceRCA for target mRNA splicing isoforms imaging in jurkat cells. The fluorescence images and DIC merged images of CD45 RNA isoforms in jurkat T cells imaged by SpliceRCA under different conditions: A) using target padlock probes, B) after knockdowning CD45 gene, C) without padlock probes, D) without the trigger primer, E) using random padlock probes and F) after blocking the target sites. The green spots represent RCA amplicons hybridized with Alexa488-labeled detection probes, the blue spots represent RCA amplicons hybridized with Cy5-labeled detection probes, the red spots represent RCA amplicons hybridized with Alexa555-labeled detection probes, and the outline of jurkat T cells is marked in a gray dot line. Scale bars: 10 μm; G) Quantification of the average numbers of RCA amplicons per cell detected under above conditons.

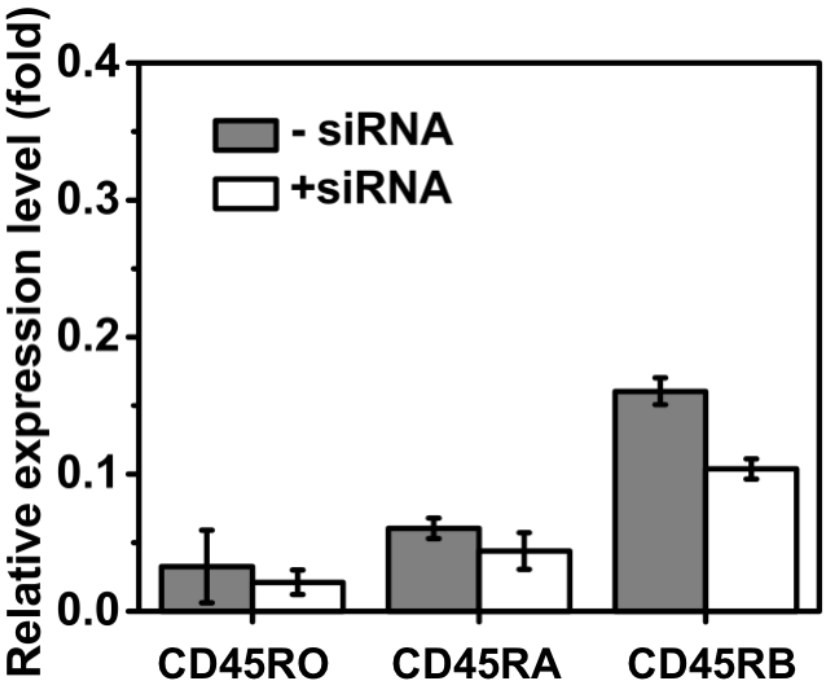

Figure S5. RT-qPCR analysis of CD45 splice isoforms upon siRNA knockdown. The histogram of relative expression level for isoforms $CD_{45}RO$, $CD_{45}RA$ and CD45RB. Error bars are based on triplicate experiments.

Multiplex *in situ* imaging of BRCA1 mRNA splicing variants in single cells by SpliceRCA.

To demonstrate general applicability of our method in RNA splicing study, we further applied SpliceRCA to image splicing isoforms of BRCA1, breast cancer susceptibility gene 1,[1-2] and the alternative splicing of BRCA1 is closely related with the transformation of malignant breast cancer[3-4]. Three critical mRNA splicing isoforms of BRCA1 were chose, including $BRCA1_{wt}$ (all exons retained), Δ(9,10,11q) (exon 9, exon 10 and the last 333 nt of exon 11 deleted), Δ(11q) (the last 333 nt of exon 11 deleted) (Figure S6 A). From Figure S6 B, the generated superbright dots amplified from target splice isoform could be clearly distinguished from the background inside cell. No bright spot was observed when random padlock probes were used. As shown in Figure S6 C, The average numbers of amplicons for $BRCA1_{wt}$, Δ(9,10,11q) and Δ (11q) were 8.02, 4.20, 5.29 per cell, repectively. As a validation, we performed RT-qPCR assay for expression comparison. The results of SpliceRCA are in good accordance with the RT-qPCR results in general (Table S5). Therefore, SpliceRCA technology can accurately quantify the splice variants of different genes in single-cell level.

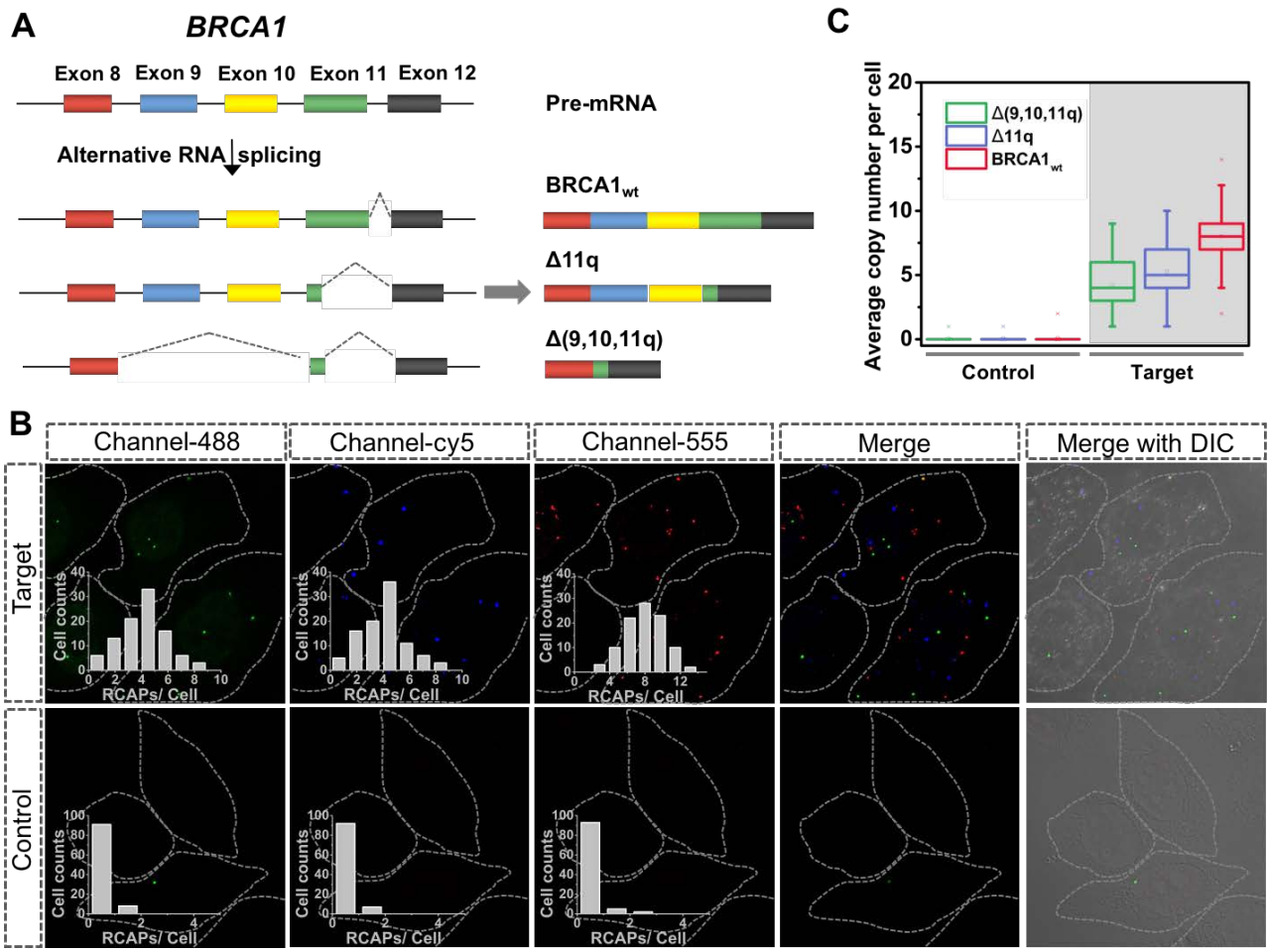

Figure S6. Simultaneous imaging of CD45 splicing variants in single jurkat T cells. A) Alternative splicing patterns of BRCA1; B) The fluorescent image of BRCA1 isoforms visualized by SpliceRCA in MCF-7 cells. Inset: frequency histogram of RCA amplicons per cell detected (cell number>100). The green spots represent RCA amplicons of Δ(9,10,11q), the blue spots represent RCA amplicons of Δ(11q), the red spots represent RCA amplicons of $BRCA1_{wt}$.The outline of MCF-7 cell is marked with a gray dot line. Scale bars: 10 μm; C) Quantification of amplicons per cell for isoforms Δ(9,10,11q), Δ(11q) and $BRCA1_{wt}$ in MCF-7 cells (cell number>100). MCF-7 cell imaged by SpliceRCA with random padlock probes were used as control.

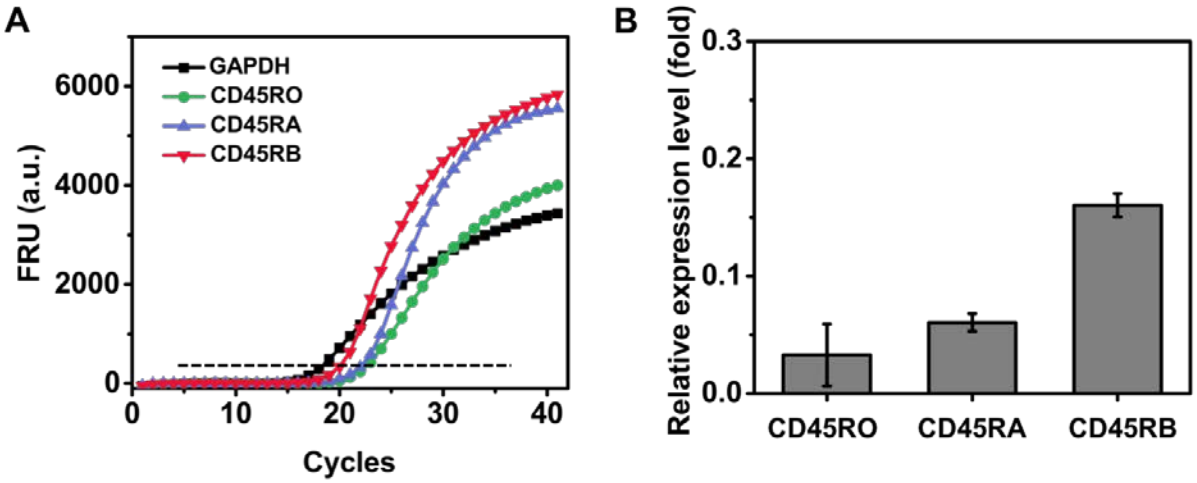

Figure S7. Expression analysis of CD45 splice isoforms in jurkat T cells by RT-qPCR. (A) Real-time fluorescence curves in RT-qPCR analysis. The black dot horizontal line represents the threshold line. (B) The histogram of relative expression level for isoforms CD45RO, CD45RA and CD45RB by RT-qPCR. Error bars are based on triplicate experiments.

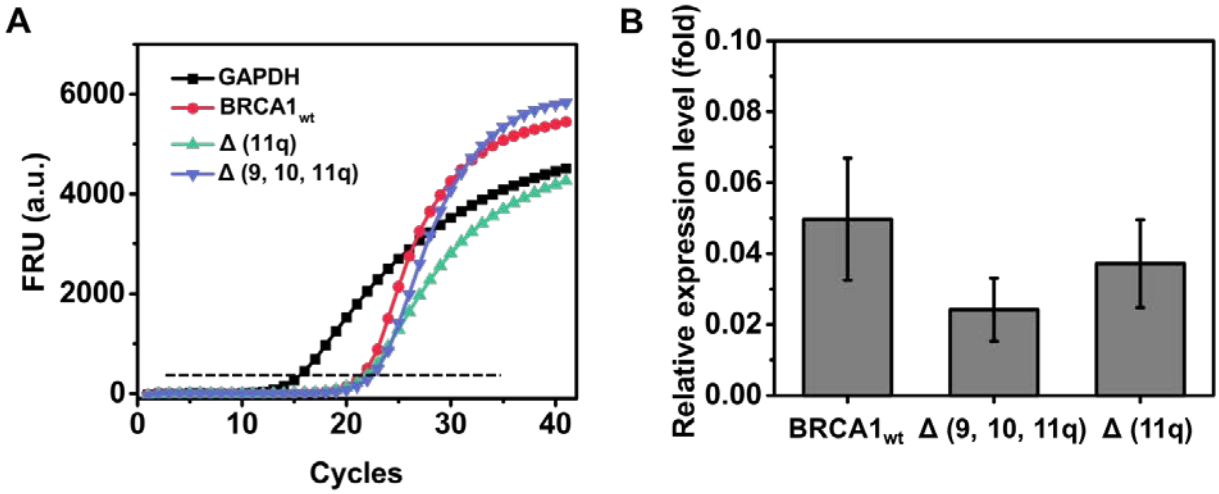

Figure S8. Expression analysis of BRCA1 splice isoforms in MCF-7 cells by RT-qPCR. (A) Real-time fluorescence curves in RT-qPCR analysis. The black dot horizontal line represents the threshold line. (B) The histogram of relative expression level for isoforms $BRCA1_{wt}$, Δ(9, 10,11q) and Δ(11q) by RT-qPCR. Error bars are based on triplicate experiments.

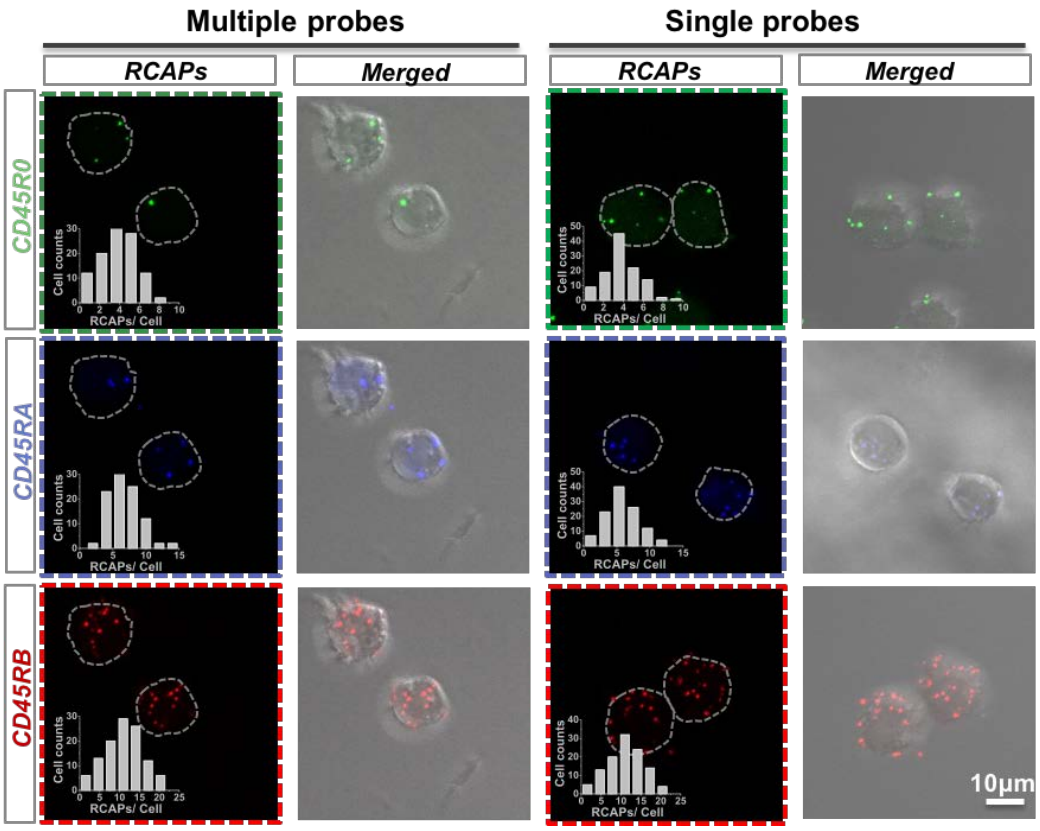

Figure S9. Comparison of SpliceRCA results in simultaneous detection with separate detection. Fluorescence image of CD45 splice isoforms by SpliceRCA with simultaneous detection and separate detection, respectively. Insets: Frequency histogram of amplicons in the cells. The outline of jurkat T cell is marked with a gray dot line. Scale bars:10 μm.

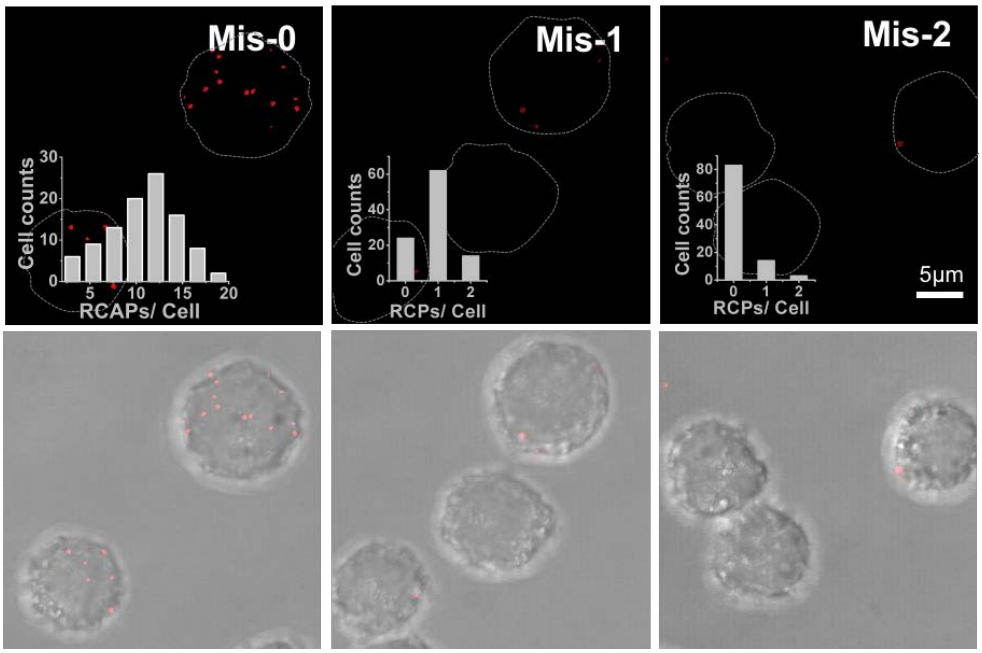

Figure S10. Demonstration of the specificity of SpliceRCA for splice isoforms imaging in jurkat T cells. Single cell imaging of CD45RB in jurkat T cells was performed by using perfectly matched (Mis-0; A), single mismatched (Mis-1; B), and double mismatched (Mis-2; C) splice-junction anchored padlock probe S2. Scale bars: 5 μm. Insets: Frequency histogram of RCAPs in the cells.

The detection efficiency of SpliceRCA and possible improvement.

The *in situ* detection efficiency of mRNA splice isoforms-initiated RCA was estimated to be 10%- 20% on the basis of a comparison to RT-qPCR data (Table S3). There are many factors that influence the detection efficiency of SpliceRCA. The first effect is relatively low efficiency of liagtion process using T4 DNA ligase. A recently discovered ligase, Splint R, can efficiently catalyze the ligation of the padlock probe by a RNA template,[5] enabling efficiently detecting mRNA without reverse transcription (the detection efficiency is over 20 %).[6-7] The second factor is the amplification bias of RCA. The secondary structures in target mRNA splice isoforms and padlock probes, or the association with splicing factors might hinder the hybridization between the target mRNA splice isoforms and the padlock probes.[8] To reduce the amplification bias, the padlock probe should be designed with none or minor secondary structure. And the structure prediction of target mRNA isoforms maybe helpful for avoiding the steric hindrance. Another factor may decrease the detection efficiency is the imaging process. Due to the fluorescence images are usually obtained from combing z-sliced images by MIP, the 3D distributed amplicons were coalesced in the flattening process, making the z-axis segregate amplicons unresolvable. Furthermore, some amplicons with relatively dim fluorescence may be lost in the process of setting the intensity threshold to differentiate a single amplicon from the background signal.

smFISH and *in situ* RCA method are powerful single-molecule RNA imaging methods. Subject to the spatial resolution of fluorescence imaging, it is hard to resolve very closely located mRNAs. It's the common problem with single-molecule RNA fluorescence imaging methods.[9-10] Recently, new methods using super-resolution imaging[11] or correlation decoding[12] have improved the ability to image condensed mRNAs. RCA amplicons are generally large with diameters of ~300 nm to provide efficient detection by fluorescence imaging. The formation of hundreds of such RCA amplicons per cell causes the signals to coalesce, limiting the maximum number for digital quantification of target molecules.[13] Another solution is reducing the size of RCA amplicons by control the time of amplification.

In conclusion, there's much room to improve the detection efficiency of RCA based imaging method. And the effect of underestimate now could hardly be resolved as the lack of methods for absolute quantification for single-cell splice isoform. Nevertheless, at present, SplicRCA is with relatively high detection efficiency (10%-20%) for detecion of splice isoforms. Most importantly, SpliceRCA can target the mRNA splice isoforms with short exon and discriminate the highly similar sequence with single-nucleotide resolutions, which not available in other method. The quantitative information of expression profile and the spatial pattern of mRNA

isoforms in single cells could be acquired, and these single-cell information is fairly useful for studying the gene function and regulatory network.

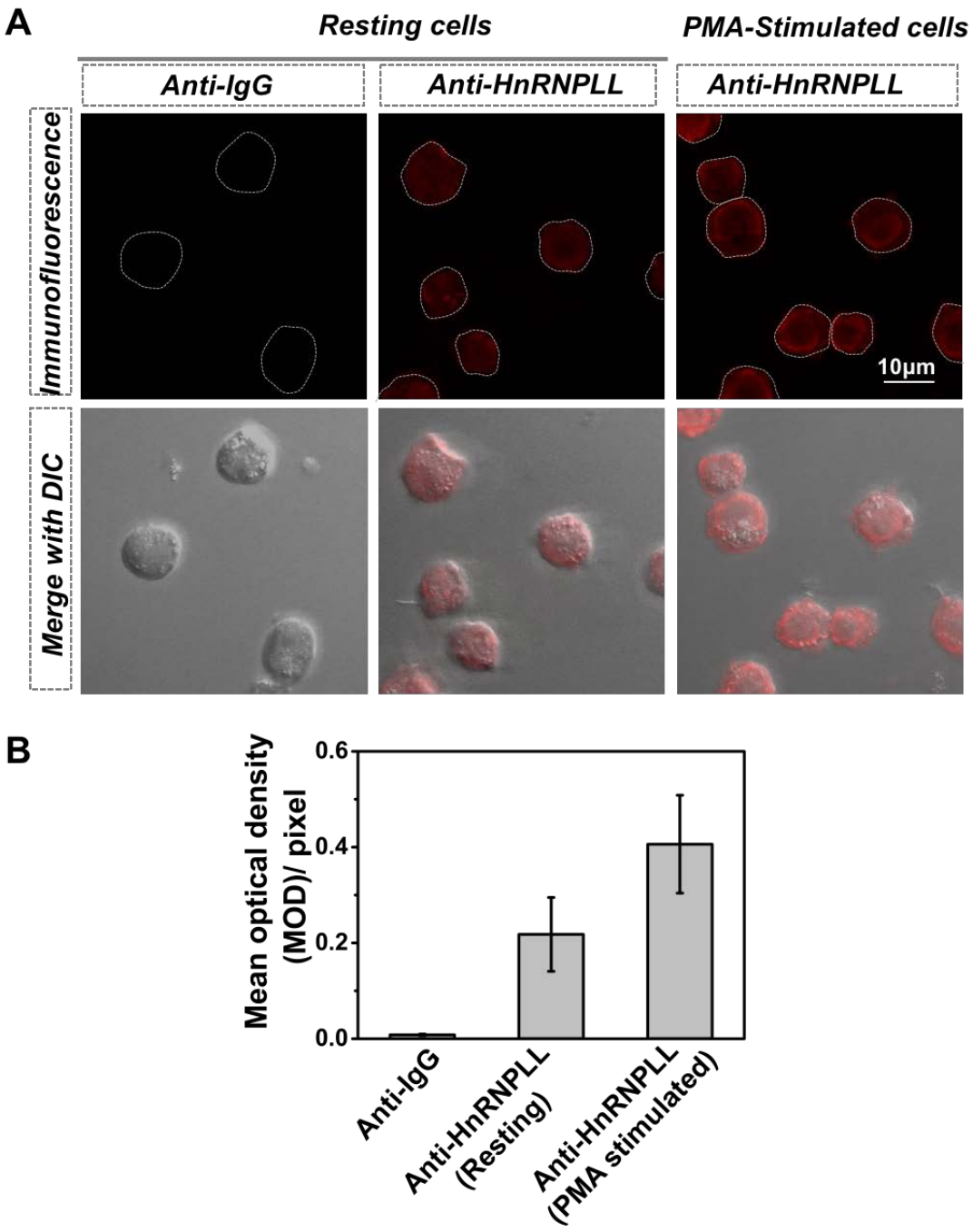

Figure S11. Immuno-fluorescence imaging of HnRNPLL in Jurlat T cells. A) The fluorescence images of HnRNPLL[14] in resting and PMA stimulated jurkat T cells. Immuo-fluorescence imaging of IgG in resting cells was performed as control. Scale bars: 10 μm. The cell nuclei are shown in blue. B) Mean optical density (MOD) of the samples in figure S7 A (cell number > 50), The MOD of anti-HnRNPLL shows a skewed increase in PMA stimulated T cells indicating that jurkat T cells were efficiently stimulated.

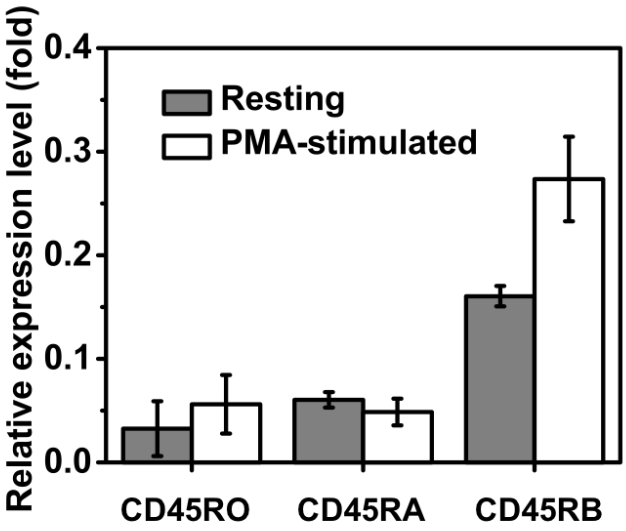

Figure S12. Expression analysis of CD45 splice isoforms upon T cell activation by RT-qPCR. The histogram of relative expression level for isoforms CD45RO, CD45RA and CD45RB before and after PMA stimulation by RT-qPCR. Error bars are based on triplicate experiments.

Table S1. Oligonucleotide sequences

Name	Sequences (5'-3')	Description
T-CD45RO	CCCAACACCTTCCCCCACTGATGCCTACCTTAATGCCTCT	Used as the templates for transcription to produce the target RNA sequence for *in vitro* testing
T-CD45RA	CCCAACACCTTCCCCCACTGGATTGACTACAGCAAAGATG	
T-CD45RB	CCCAACACCTTCCCCCACTGGTGTTTCATCAGTACAGACG	
S-CD45RO	CAGTGGGGGAAGGTGGTTGGATTGTCTGTACTACAACTATACAACTATACAACATACTACCTCAAGGAGCATTAAGGTAGGCAT	Padlock probes for *in situ* imaging splice isoform CD45RO, CD45RA, CD45RB and GAPDH. The bases marked in red indicate the mismatched bases between the padlock probe and target sequence.
S-CD45RA	CAGTGGGGGAAGGTGGTTGGATTGTCTGTACTACAACTATACTGCGTCTATTTTCTGGAGCCATCACTTTTGCTGTAGTCAATC	
S-CD45RB	CAGTGGGGGAAGGTGGTTGGATTGTCTGTACTACAACTATACCTCAATTCTGCTACTGTACTACCGTTCGTACTGATGAAACAC	
S-CD45RB-mut1	AAGTGGGGGAAGGTGGTTGGATTGTCTGTACTACAACTATACCTCAATTCTGCTACTGTACTACCGTTCGTACTGATGAAACAC	
S-CD45RB-mut2	AAGTGGGGGAAGGTGGTTGGATTGTCTGTACTACAACTATACCTCAATTCTGCTACTGTACTACCGTTCGTACTGATGAAACAA	
S-GAPDH	CATGACAAGGTGCGGTTTTTAGTGCGAACTACTACTCTCTCTTTTTAGTGCGAACTACTACTCTCTCTTTTTTACTTTATTGATGGTA	
Random-1	ATGCATCCGTTATTAGTTGGATTGTCTGTACTACAACTATACAACTATACAACATACTACGGAGCCTCTACTGTCTTACGTT	Padlock probes with random sequence in recognition region
Random-2	ATGCATCCGTTATTAGTTGGATTGTCTGTACTACAACTATACTGCGTCTATTTTCTGGAGCCATCACTTCCTACTGTCTTACGTT	

Random-3	ATGCATCCGTTATTAGTTGGGTTGGATTGTCTGTACTACAACTATACCTCAATTCTGCTACTGTACTACCGTTCTCTACTGTCTTACGTT	
S-BRCA1$_{wt}$	CAAGTCTTCCAATTCAAACTATACAACATACTACCTCACCAACTATACAACATACTACCTCAGGTGTTTGTATTTGCAGT	Padlock probes for *in situ* imaging splice isoform BRCA1$_{wt}$, Δ(9, 10, 11q) and Δ(11q).
S-Δ(9, 10, 11q)	CCAATTCAATGTAGACTGCGTCTATTTTCTGGAGCCATCATTGCGTCTATTTTCTGGAGCCATAGAAAATTCACAAGCAG	
S-Δ(11q)	CCTGATACTTTTCTGGCTCAATTCTGCTACTGTACTACTTTTTCTCAATTCTGCTACTGTACTACCAGATGCTGCTTCAC	
Block-CD45RO	AGAGGCATTAAGGTAGGCATCAGTGGGGGAAGGTGTTGGG	Probes masking the target splice junction sequence
Block-CD45RA	CATCTTTGCTGTAGTCAATCCAGTGGGGGAAGGTGTTGGG	
Block-CD45RB	CGTCTGTACTGATGAAACACCAGTGGGGGAAGGTGTTGGG	
CD45-homo-218	CCAUGUAUUUGUGGCUUAATT	siRNA used for knockdown of CD45 gene
CD45-homo-2500	GCAGAAUACUGGCCGUCAATT	
CD45-homo-1783	GCUGGAAAUACUCUGGUUATT	
DP-488	Alexa488-AACTATACAACATACTACCTCA	Fluorophore-labeled detection probes for visualizing RCA amplicons
DP-Cy5	Cy5-TGCGTCTATTTTCTGGAGCCAT	
DP-555	Alexa555-CTCAATTCTGCTACTGTACTAC	
Trigger primer (CD45RO, CD45RA, CD45RB)	GTATAGTTGTAGTACAGACAAT	Used as the trigger primer for RCA
Trigger primer (GAPDH)	AGAGAGAGTAGTAGTTCGCACT	Used as the trigger primer for RCA

[a]The color marked in the sequence of padlock probe indicates the modules corresponding to module Rx (green), P (gray), T (purple) and Ry (blue).

Table S2. Primers for RT-qPCR

Target name	Primer name	Sequences (5'-3')
CD45RO	CD45RO-F	GAAATTGTTCCTCGTCTGAT
	CD45RO-R	AGGTAGGCATCAGTGGGGGA
CD45RA	CD45RA-F	TCCCCCACTGGATTGACTAC
	CD45RA-R	GATGAAACACCTGTGGTATT
CD45RB	CD45RB-F	GAAATTGTTCCTCGTCTGAT
	CD45RB-R	GATGAAACACCAGTGGGGGA
GAPDH	GAPDH-F	TATGACAACAGCCTCAAGAT
	GAPDH-R	AGTCCTTCCACGATACCA
$BRCA1_{wt}$	$BRCA1_{wt}$-F	ATTGGAAGACTTGACTGCAAATACAAAC
	$BRCA1_{wt}$-R	CACACCCAGATGCTGCTTCAC
Δ(11q)	Δ(11q)-F	TTACAAATCACCCCTCAAGGAACC
	Δ(11q)-R	TGCTGCTTCACCCTGATACTTT
Δ(9,10,11q)	Δ(9,10,11q)-F	CTACATTGAATTGGCTGCTTGTGA
	Δ(9,10,11q)-R	CACACCCAGATGCTGCTTCAC

Table S3. The average copy number of CD45 splice isoforms profiled by SpliceRCA after knockdown of CD45 gene.

Splice isoform	SpliceRCA (copy number/ cell, - siRNA)	SpliceRCA (copy number/ cell, + siRNA)	Decrease percent (%)
CD45RO	3.70	2.79	24.59
CD45RA	5.98	4.11	31.27
CD45RB	11.07	8.08	27.01

Table S4. The average copy number of CD45 splice isoforms profiled by RT-qPCR after knockdown of CD45 gene.

Splice isoform	- siRNA		+ siRNA		Decrease percent (%)
	Ct	RT-qPCR (copy number/ cell)	Ct	RT-qPCR (copy number/ cell)	
CD45RO	23.18	21.44	22.58	14.19	29.15
CD45RA	22.29	39.74	21.86	29.51	25.73
CD45RB	20.88	105.60	20.28	69.81	33.89

Table S5. The average copy number of mRNA splicing variants in single cells profiled by SpliceRCA and RT-qPCR

Cell type	Splice isoform	SpliceRCA (copy number/ cell)	Ct	RT-qPCR (copy number/ cell)	Detection efficiency using SpliceRCA (%)[a]
jurkat	GAPDH	71.03	18.25	658.24	10.79
	CD45RO	3.70	23.18	21.44	17.26
	CD45RA	5.98	22.29	39.74	15.05
	CD45RB	11.07	20.88	105.60	10.48
MCF-7	GAPDH	115.86	15.98	918.08	12.62
	$BRCA1_{wt}$	8.02	22.39	45.65	17.57
	Δ(9,10,11q)	4.20	21.35	22.20	18.92
	Δ(11q)	5.29	21.97	34.14	15.50

[a] The detection efficiency of SeqEA and in situ RCA was calculated on the basis of a comparison to RT-qPCR data.

Table S6. The coefficient of variation (CV)[15-16] of CD45 splice isoforms in the jurkat T cells measured by SpliceRCA in resting and PMA stimulated T cells.

Splice isoform	CV(Resting)	CV(PMA stimulated)
CD45RO	0.416	0.328
CD45RA	0.404	0.361
CD45RB	0.456	0.283
Average	0.425	0.324

References

1. Budhram-Mahadeo, V.; Ndisang, D.; Ward, T.; Weber, B. L.; Latchman, D. S., The Brn-3b POU family transcription factor represses expression of the BRCA-1 anti-oncogene in breast cancer cells. *Oncogene* **1999,** *18* (48), 6684-6691.

2. Orban, T. I.; Olah, E., Expression profiles of BRCA1 splice variants in asynchronous and in G1/S synchronized tumor cell lines. *Biochem. Biophys. Res. Commun.* **2001,** *280* (1), 32-38.

3. Lee, K.; Cui, Y.; Lee, L. P.; Irudayaraj, J., Quantitative imaging of single mRNA splice variants in living cells. *Nat. Nanotechnol.* **2014,** *9* (6), 474-480.

4. Wang, H.; Wang, H.; Duan, X.; Sun, Y.; Wang, X.; Li, Z., Highly Sensitive and Multiplexed Quantification of mRNA Splice Variants by Direct Ligation of DNA Probes at Exon Junction and Universal PCR Amplification. *Chem. Sci.* **2017,** *8*, 3635-3640.

5. Lohman, G. J.; Zhang, Y.; Zhelkovsky, A. M.; Cantor, E. J.; Evans, T. C., Efficient DNA ligation in DNA–RNA hybrid helices by Chlorella virus DNA ligase. *Nucleic Acids Res.* **2014,** *42* (3), 1831-1844.

6. Jin, J.; Vaud, S.; Zhelkovsky, A. M.; Posfai, J.; McReynolds, L. A., Sensitive and specific miRNA detection method using SplintR Ligase. *Nucleic. Acids. Res.* **2016,** *44* (13), e116.

7. Wee, E. J.; Trau, M., Simple Isothermal Strategy for Multiplexed, Rapid, Sensitive, and Accurate miRNA Detection. *ACS Sensors* **2016,** *1* (6), 670-675.

8. Ke, R.; Mignardi, M.; Pacureanu, A.; Svedlund, J.; Botling, J.; Wählby, C.; Nilsson, M., In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **2013,** *10* (9), 857-860.

9. Ke, R.; Mignardi, M.; Pacureanu, A.; Svedlund, J.; Botling, J.; Wahlby, C.; Nilsson, M., In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **2013,** *10* (9), 857-860.

10. Chen, K. H.; Boettiger, A. N.; Moffitt, J. R.; Wang, S.; Zhuang, X., Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **2015,** *348* (6233), aaa6090.

11. Lubeck, E.; Cai, L., Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat. Methods* **2012,** *9* (7), 743-748.

12. Coskun, A. F.; Cai, L., Dense transcript profiling in single cells by image correlation decoding. *Nat. Methods* **2016,** *13* (8), 657-660.

13. Clausson, C. M.; Allalou, A.; Weibrecht, I.; Mahmoudi, S.; Farnebo, M.; Landegren, U.; Wahlby, C.; Soderberg, O., Increasing the dynamic range of in situ PLA. *Nat. Methods* **2011,** *8* (11), 892-893.

14. Oberdoerffer, S.; Moita, L. F.; Neems, D.; Freitas, R. P.; Hacohen, N.; Rao, A., Regulation of CD45 alternative splicing by heterogeneous ribonucleoprotein, hnRNPLL. *Science* **2008,** *321* (5889), 686-691.

15. Ergun, A.; Doran, G.; Costello, J. C.; Paik, H. H.; Collins, J. J.; Mathis, D.; Benoist, C.; ImmGen, C., Differential splicing across immune system lineages. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110* (35), 14324-14329.
16. Waks, Z.; Klein, A. M.; Silver, P. A., Cell-to-cell variability of alternative RNA splicing. *Mol. Syst. Biol.* **2011,** *7*, 506.