Electronic Supplementary Material (ESI) for Analyst.
This journal is © The Royal Society of Chemistry 2021

Supplementary Information

Evolution of DNA aptamers against esophageal squamous cell carcinoma using cell-SELEX

Xinhuan Chen [a,b,c], Yangyang Zhang [a], Yanli Shi [a], Tingting Niu [a], Bo Li [a,b,d], Linyan Guo [a], Yan Qiao [a,b,c], Jimin Zhao [a,b,c], Baoyin Yuan [a,b,c] *, Kangdong Liu [a,b,c,d,e]

[a] Department of Pathophysiology, School of Basic Medical Sciences, Zhengzhou University, Zhengzhou, Henan, China

[b] Henan Provincial Cooperative Innovation Center for Cancer Chemoprevention, Zhengzhou, Henan, China

[c] State Key Laboratory of Esophageal Cancer Prevention & Treatment, Zhengzhou University, Zhengzhou, Henan, China

[d] China-US Hormel (Henan) Cancer Institute, Zhengzhou, Henan, China

[e] Cancer Chemoprevention International Collaboration Laboratory, Zhengzhou, Henan, China

Corresponding Author

* E-mail: yuanbaoyin@zzu.edu.cn

Table S1. Top ten sequences by copy number.[a]

Name	Sequence of random region (5'to3')	Copies	Percentage (%)
S1	GTAGGTAGGTGTCCATATGGTTTGGCACAGATGCGAGATG	57389	28.99
S2	CCCTGAGGGGTGATGTGAATCCTAACAGTAAGGAGGAAGG	48271	24.38
S3	CTTCTTCGTTAGAAGCCTGTCAATGGCCAGGGGGGAGGGG	8995	4.54
S4	GGTGTGCTTCGGTTCTTGCGAGTCGCAACGAGCCTGTCAT	2779	1.40
S5	GTAGGTAGGTGTCCATATGGTTTGGCGCAGATGCGAGATG	2476	1.25
S6	GGACCTCTTTGCTCGCCTGAAGCGATATCGACCCTCTCGT	1641	0.83
S7	GTAGGTAGGTGTCCATATGGTTTGGCACAGATGCGAGGTG	1612	0.81
S8	AGGCCCTACACGCCTTTTAGCGAGGGCGTTCCGTGGAGGA	1413	0.71
S9	CCTGGCCGGGACCTGTAGACTGTTGAGACTACGGTTCTTG	1277	0.65
S10	GACTGTTGAGACTGCCCAGTCGCGTTCGTCCAAGCTGCGC	1202	0.61
Total	S1-10	127055	64.2

[a] A total of 19,7960 ssDNA sequences were sequenced. The copy number and percentage of the top ten sequences were listed.

Table S2. Sequences of truncated aptamers. [a]

Name	Sequence (5'to3')
S2-1	CCCTGAGGGGTGATGTGAATCCTAACAGTAAGGAGGAAGG
S2-2	TCCTAACAGTAAGGAGGAAGGTGTAGGA
S3-1	CTTCTTCGTTAGAAGCCTGTCAATGGCCAGGGGGGAGGGG
S3-2	ATGGCCAGGGGGGAGGGGTGT
S3-3	TACTTCTTCGTTAGAAGCCTGTCAATGGCCAGGGGGGAGGGGTGTAGGAGG G TGCGGAAGTA
S8-1	AGGCCCTACACGCCTTTTAGCGAGGGCGTTCCGTGGAGGA
S8-2	GCGTCGGTGTGGTAAGGCCCTACACGCCTTTTAGCGAGGGC
S8-3	TTCCGTGGAGGATGTAGGAGGGTGCGGAAGTA
S3-2-1	GGCCAGGGGGGAGGGG
S3-2-2	CAGGGGGGAGG
S3-2-3	ATGGCCAGGGGGGAGGGG

[a] The bases marked in red are primer sequences.

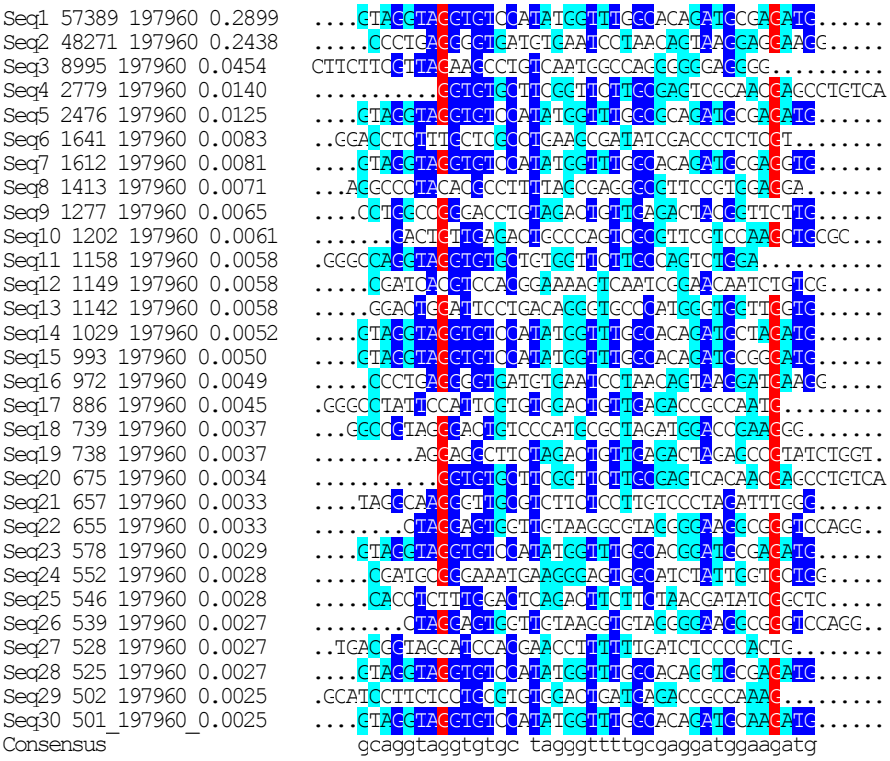

Fig. S1 Homology analysis of the top thirty sequences. The 40 nt of random region of the selected sequences were analyzed. The sequences were named by ranking number, copy number, total sequence number and percentage (for example, Seq1_57389_197960_0.2899). The different colors represent the extent of sequence homology, the red＞75%, the blue＞50%, the cyan＞33%. The results of homology analysis demonstrated that some bases of the sequences were highly conservative, revealed that the sequences were successfully evolved in the selection process.

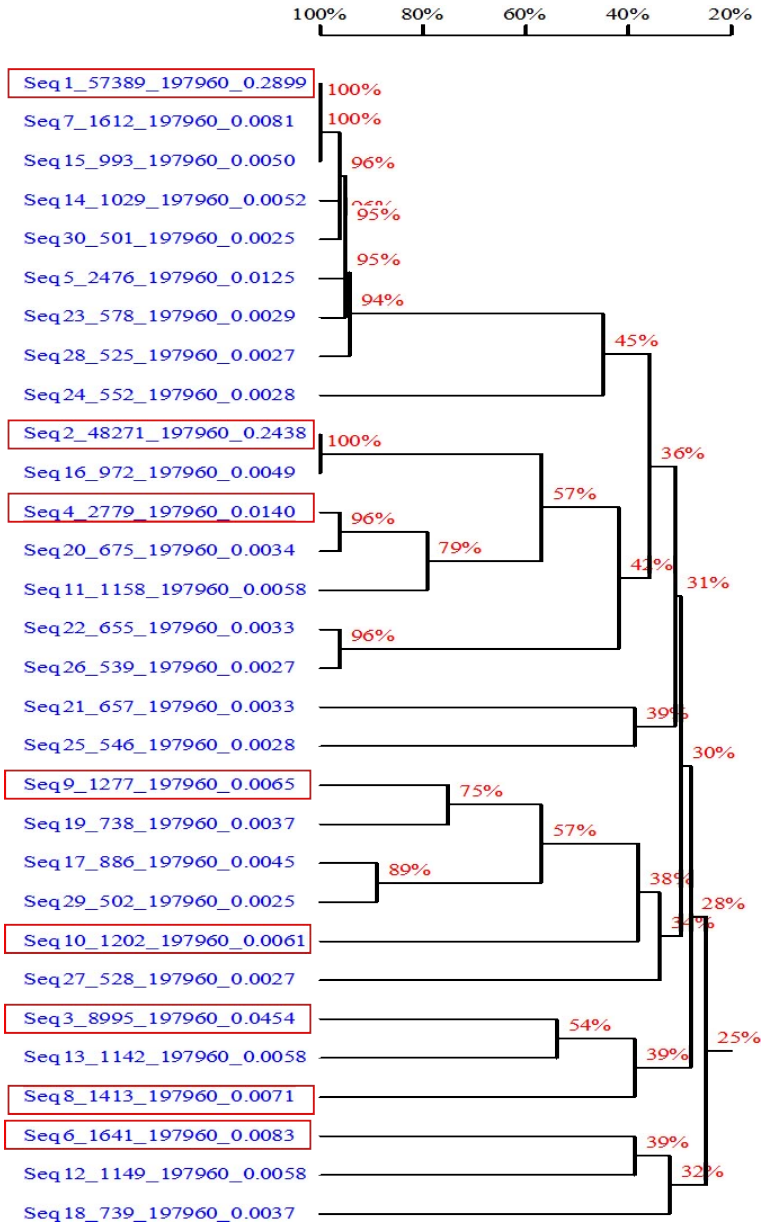

Fig. S2 Phylogenetic tree analysis of the top thirty sequences. From the results of phylogenetic tree analysis, several families were presented, and some sequences in the same family were highly homologous. For developing distinct aptamers, aptamers candidates were selected from different families with high copy number as a reference factor. Therefore, S1-4, S6, S8, S9 and S10 (displayed in red frames) were selected for binding ability analysis to ESCC cells.

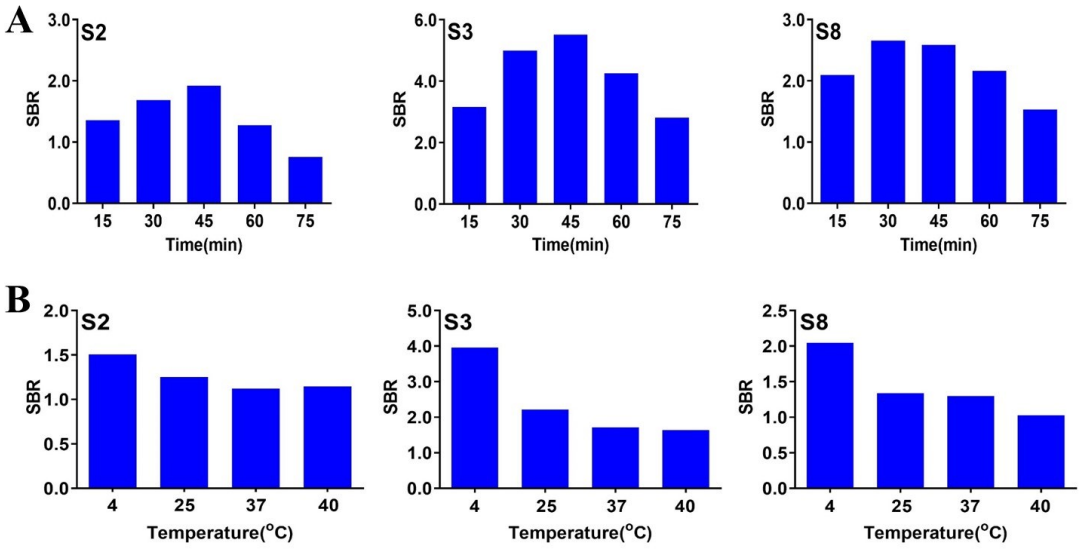

Fig. S3 Optimization of incubation conditions. (A) SBR of FAM-labeled aptamers binding to KYSE150 cells with different times (15, 30, 45, 60 and 75 min). (B) SBR of FAM-labeled aptamers binding to KYSE150 cells at different temperatures (4, 25, 37 and 40 °C). The FAM fluorescence intensity was detected by flow cytometry. The concentration of the aptamers or random DNA was 250 nM. FAM-labeled random DNA was used as controls.

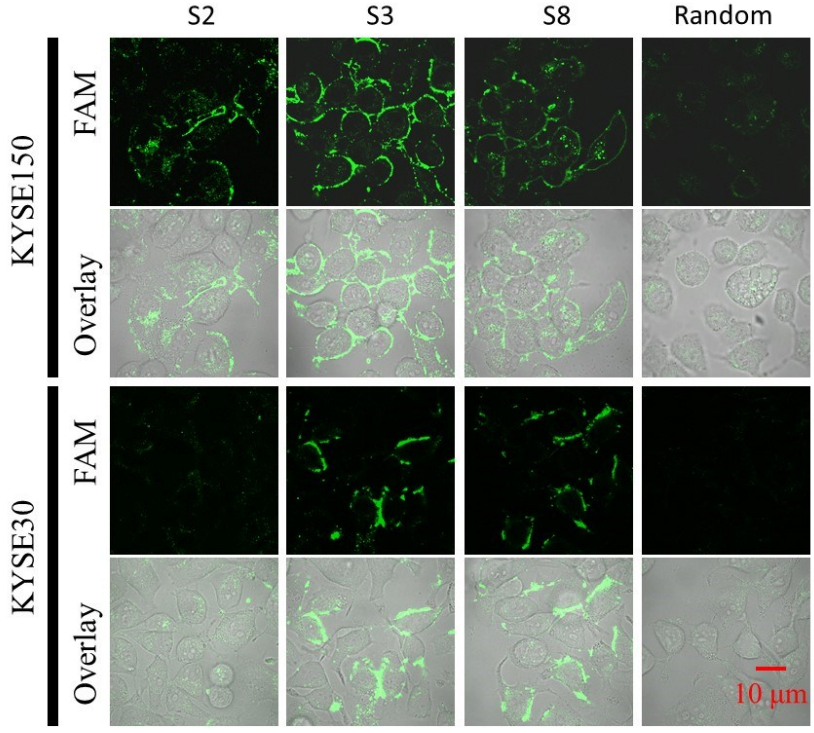

Fig. S4 Laser scanning confocal microscope images of KYSE150 and KYSE30 cells incubated with FAM-labeled aptamers (S2, S3 and S8) or random DNA. FAM-labeled random DNA was used as controls. The first and third rows are fluorescence images, the second and fourth rows are the overlays of the fluorescence channel and the bright-field channel. The concentration of the aptamers or random DNA was 250 nM. The fluorescence signal was collected by a 100× objective (fluorescence channel: EX 488 nm, EM 525 nm long-pass). Scale bar is 10 μm.

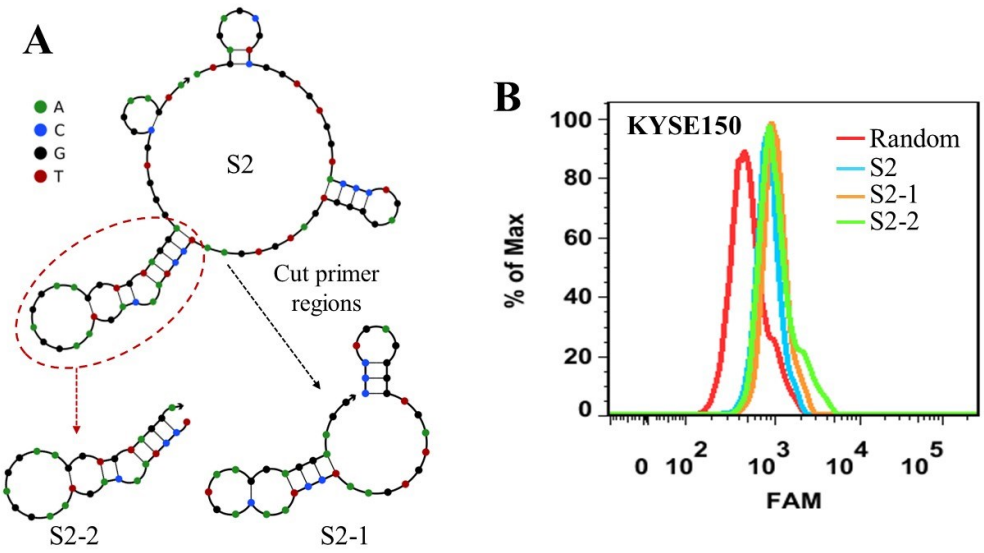

Fig. S5 (A) Truncating strategies for sequence optimization of S2. The secondary structures were predicted by NUPACK. (B) Flow cytometry assays of FAM-labeled random sequences, S2, S2-1and S2-2 to KYSE150 cells. FAM-labeled random DNA was used as controls. The fluorescence signal was collected through FAM channel (EX 488 nm, EM 525 nm band pass).

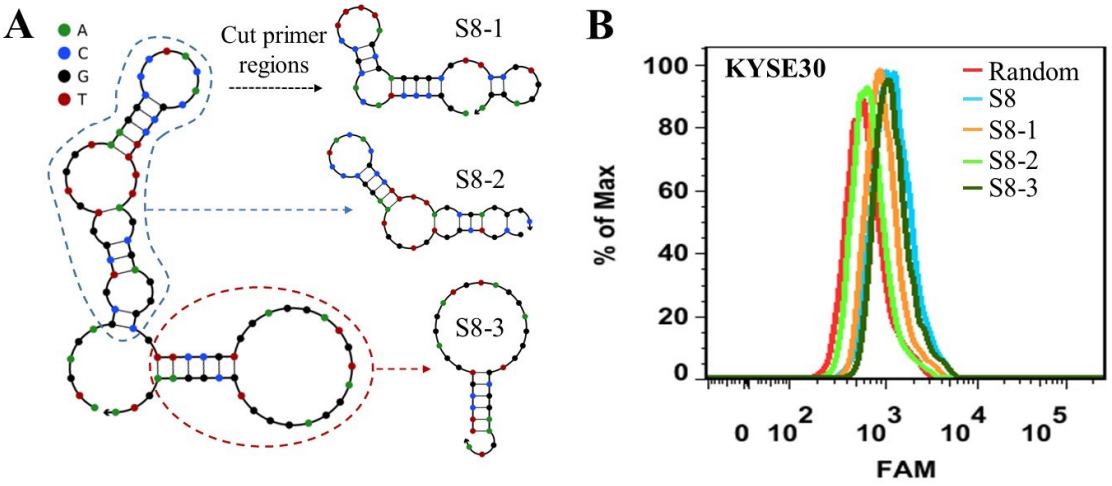

Fig. S6 (A) Truncating strategies for sequence optimization of S8. The secondary structures were predicted by NUPACK. (B) Flow cytometry assays of FAM-labeled random sequences, S8, S8-1, S8-2 and S8-3 to KYSE30 cells. FAM-labeled random DNA was used as controls. The fluorescence signal was collected through FAM channel (EX 488 nm, EM 525 nm band pass).

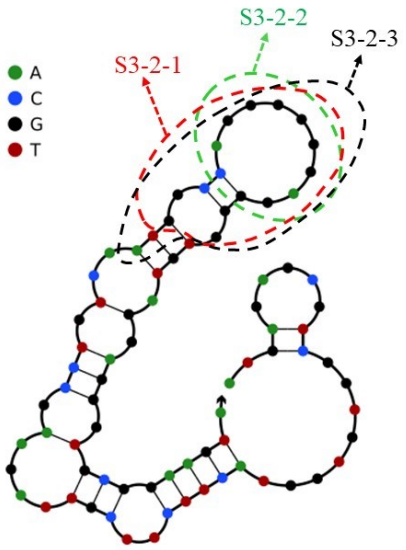

Fig. S7 Truncating strategies for sequence optimization of S3-2. S3-2-1 was generated by removing redundant non-G bases at ends, S3-2-2 was generated by retaining the small G-rich hairpin structure, and S3-2-3 was generated by removing 3 bases belong to the primer region. The secondary structures were predicted by NUPACK.

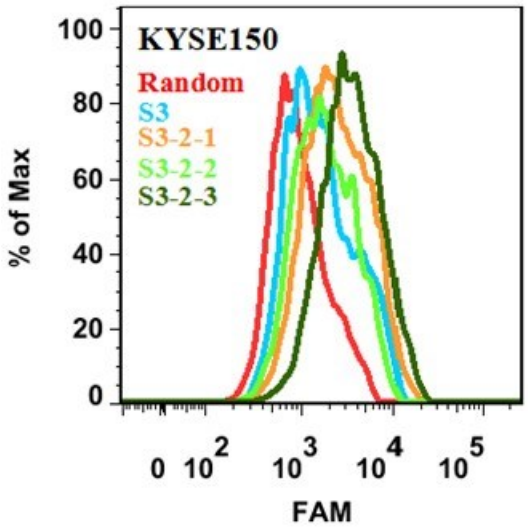

Fig. S8 Flow cytometry assays of FAM-labeled random sequences, S3, S3-2-1, S3-2-2 and S3-2-3 to KYSE150 cells. FAM-labeled random DNA was used as controls. The fluorescence signal was collected through FAM channel (EX 488 nm, EM 525 nm band pass).

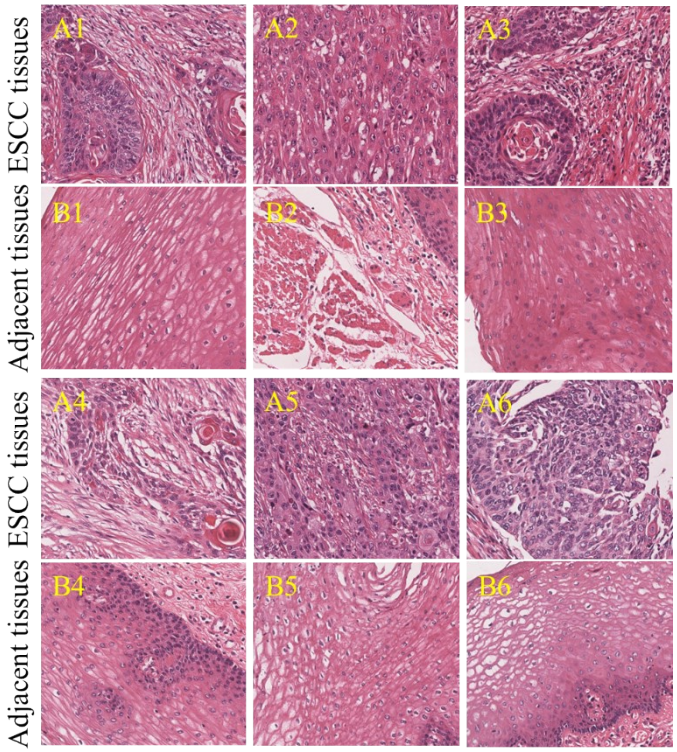

Fig. S9 Images of pathologic sections of ESCC tissues (A1-A6) and their paired adjacent tissues (B1-B6) stained with haematoxylin and eosin (H&E). The results displayed that sections from A1 to A6 presented the typical features and distribution of cancer cells, while sections from B1 to B6 presented obvious features and distribution of normal cells, suggesting the tissues are reliable and can be used for tissue imaging. The TNM stages of six patients (A1-A6) are T3N1M0, T3N1M0, T3N0M0, T3N1M0, T3N0M0 and T3N1M0, respectively. All the patients have no distant metastasis and are mainly characterized by primary tumor.

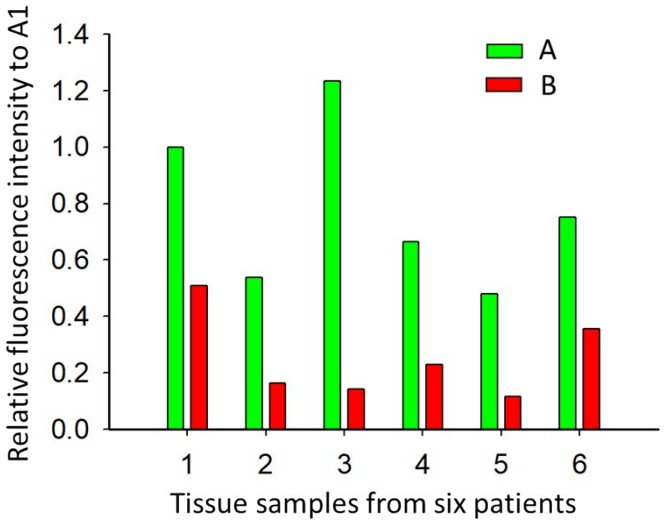

Fig. S10 Corresponding fluorescence intensity statistics of ESCC tissues and their paired adjacent tissues from Fig. 5. The green (A) represents signals of ESCC tissues, and the red (B) represents signals of the paired adjacent tissues. The signal intensity was calculated by ImageJ software.

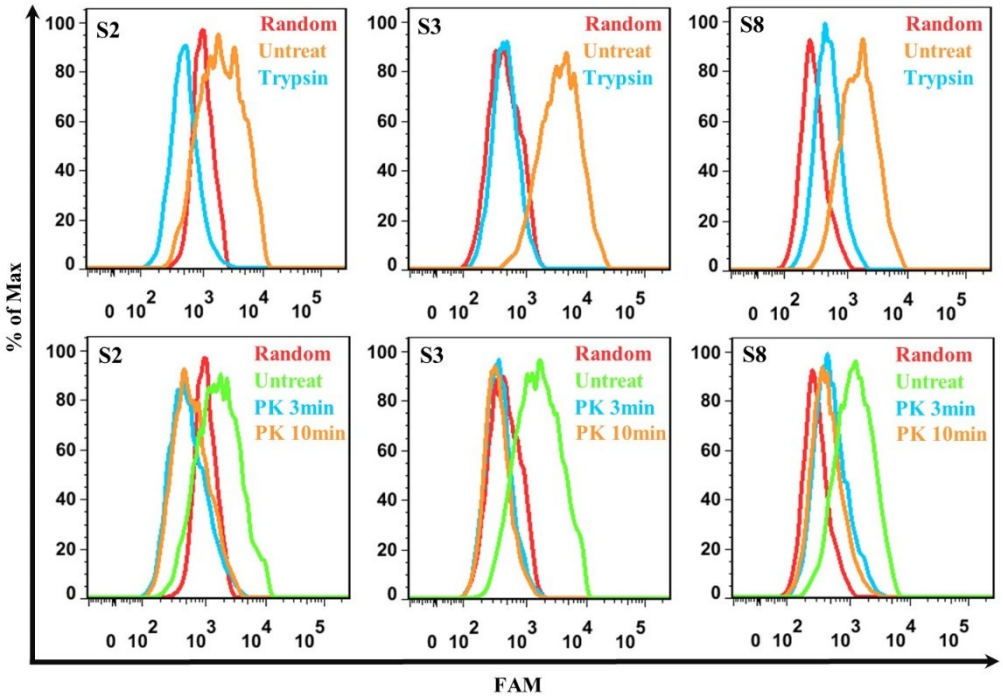

Fig. S11 Target type determination of the selected aptamers (S2, S3 and S8) using flow cytometry. KYSE150 cells were treated with trypsin for 3 min and proteinase K (PK) for 3 or 10 min for membrane protein degradation. KYSE150 cells inculcated with FAM-labeled random sequences were used as negative controls, and untreated cells inculcated with FAM-labeled aptamers were used as positive controls. The fluorescence signal was collected through FAM channel (EX 488 nm, EM 525 nm band pass).

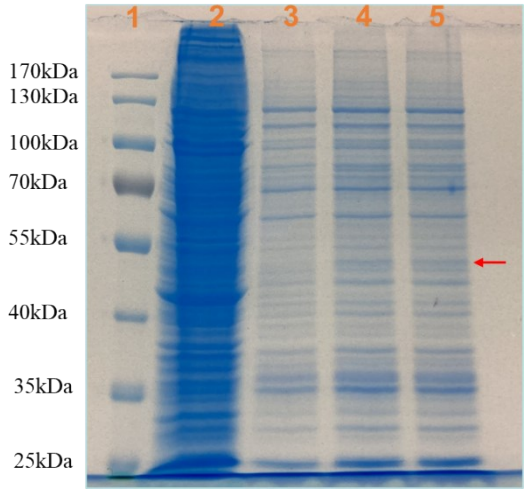

Fig. S12 S3-2-3 aptamer was used to capture proteins for SDS-PAGE analysis. SDS-PAGE gel was stained by coomassie blue. 1, protein markers; 2, total proteins; 3, proteins captured with random sequences; 4 and 5, proteins captured with S3-2-3 aptamer. Red arrow indicates the location of specific bands.