Supplemental Information

Supplemental Materials and Methods

Gene expression analysis

The National Institutes of Gene Expression Omnibus (GEO) database was used to obtain gene expression data in all kinds of cancers (https://www.ncbi.nlm.nih.gov/geo/). The probe set ID of APOBEC3B in kidney cancer (GSE66270), cervical cancer (GSE63514), non-small cell lung cancer (GSE27262) and hepatocellular carcinoma (GSE14520) was 206632_s_at. The probe set ID of APOBEC3B in head and neck squamous cell cancer (GSE58911) was 8073062. The probe set ID of APOBEC3B in ovarian cancer (GSE12470) was 22344. The probe set ID of APOBEC3B in bladder cancer (GSE37815) was ILMN_1691457. The probe set ID of APOBEC3B in breast cancer (GSE109169) was 3945545. The probe set ID of APOBEC3B in anaplastic thyroid cancer (GSE65144) was 206632_s_at. The probe set ID of APOBEC3B in papillary thyroid cancer (GSE50901) was 24480. The probe set ID of APOBEC3B and PD-L1 (CD274) in ESCC (GSE17351) and colon cancer (GSE606970 were 206632_s_at and 227458_at, respectively. The probe set ID of APOBEC3B, PD-1 and PD-L1 (CD274) in Renal cell carcinoma (GSE67501) and Melanoma (GSE79691) were ILMN_2219466, ILMN_1806725 and ILMN_1701914, respectively. The probe set ID of mouse APOBEC3 and mouse PD-L1 (CD274) in AOM/DSS induced colon cancer were A_65_P17700 and A_51_P248666, respectively.

Cell lines

Human esophageal squamous cell carcinoma cell lines (KYSE70, KYSE150), human colon cancer cell line (RKO), human umbilical vein endothelial cell line (HUVEC) and immortalized normal esophageal cell line (HET-1A) were cultured in Roswell Park Memorial Institute (RPMI) 1640 complete medium (Gibco, Grand Island, USA) supplemented with 10% fetal bovine serum (FBS, BI, USA), 100 U/mL penicillin (Solarbio, China) and 100 μg/mL streptomycin (Solarbio, China). Human embryonic kidney cell line (HEK-293T) was cultured in Dulbecco Modified Eagle Medium (DMEM) (Gibco, Grand Island, USA) supplemented with 10% FBS, 100 μg/mL streptomycin and 100 U/mL penicillin. All cells were incubated in a sterile incubator with 5% CO_2 at 37°C.

Plasmid construction and transfection

The full-length nucleic acid sequence of APOBEC3B was optimized by Shanghai Ziben Biotechnology Co., Ltd (Shanghai, China) and cloned into pLVX-Tetone-GP3 vector through *Age* I and *Not* I restriction sites. Subsequently, APOBEC3B overexpression vector and empty vector were transfected into KYSE150 cells with PowerTrans 293 (Sixiang Biological, SX-TR293-001, China) according to the protocol. Finally, KYSE150 cells overexpressing APOBEC3B were induced by doxycycline (dox).

Western blot

The total cell lysates were prepared using protein lysis buffer. Then protein was

fractionated by 8% SDS PAGE, transferred to polyvinylidene fluoride (PVDF) membrane (Merck Millipore, IPVH00010, USA), and then blocked with 5% defatted milk dissolved in PBS (pH7.2) containing 0.1% Tween 20 for 2 hours at room temperature. The PVDF membranes were incubated with antibody of human APOBEC3B (Abcam, ab184990, UK) overnight at 4 °C. The reference antibody was GAPDH (Servicebio, GB11002, China). After incubation of PVDF membranes with secondary antibody (Sangon Biotech, D110011-0100, China), the blots were visualized using ECL system (Azure C600, USA).

Cell viability assay

Effect of compounds on the proliferation of cells were determined by MTT assay. Briefly, KYSE70, HEK-293T, HUVEC and HET-1A cells were seeded into 96 plates (3000 cells/well) and grew overnight. Cells were then synchronized by starvation with serum-free RPMI 1640 medium or DMEM for 8 hours, followed by treating with different concentrations of compounds in complete medium for 24 hours, 48 hours and 72 hours, respectively. MTT (Sigma, USA) dissolved in PBS were then added and cultured for 4 hours at 37°C. Formazan crystals formed by viable cells were dissolved by 150 μL of DMSO, and the absorbance was measured by using a multi-functional microplate detector at 490 nm. Cell viability was calculated as the formula:

Cell viability (%) = $(OD_{Experimental\ group} - OD_{Blank\ group}) / (OD_{Ctrl\ group} - OD_{Blank\ group}) \times 100\%$

Microscale Thermophoresis (MST) assay

The detailed method was described in materials and methods in the main text. The

experimental concentration of APOBEC3A protein (MedChemExpress, HY-P72080, USA) was 5 μM. The experimental concentration of APOBEC3G protein (Abnova, H00060489-P01, China) was 3 μM.

Fluorescence-based single-stranded DNA cytosine deamination assay

The detailed method was described in materials and methods in the main text. The final reaction concentration of APOBEC3A and APOBEC3G was 0.4 μM.

H&E staining and Immunofluorescence analysis

The APOBEC3B antibody in immunofluorescence analysis was purchased from the Absin Bioscience Inc. (Shanghai, China), and the dilution ratio was 1:100 for immunofluorescence analysis. The H&E staining and immunofluorescence analysis were accomplished by the Wuhan Service Biotechnology company in China.

***In vivo* toxicity analysis**

The fresh whole blood was taken from mice, quickly added to the buffer containing heparin sodium, and then sent to the laboratory department of hospital for blood routine analysis containing WBC, RBC, HGB, PLT and so on. The serum was collected for blood biochemistry (AST and ALT) analysis according to the protocol.

T cell migration assay

The chemotaxis of $CD8^+$ T cells was measured by transwell assay. The aperture of

chamber (Corning, 3421, USA) was 0.5 μm. The peripheral blood mononuclear cells (PBMCs) from human donors were isolated and activated with 100 units IL-2, 1 μg/mL anti-CD3, and 1 μg/mL anti-CD28 stimulatory antibodies for 2 days. Then, activated PBMCs were harvested, centrifuged and resuspended in serum-free RPMI 1640 medium. 200 μL PBMCs suspensions (1×10^5 cells) were seeded into the upper chamber. Then, 600 μL RPMI 1640 complete medium, supernatant of KYSE70 (vehicle group) and supernatant of KYSE70 pretreated with 10 μM 3, 5-diiodotyrosine for 48 hours were added into the lower chamber, respectively. Three days later, the migration number of $CD8^+$ T cells was analyzed by flow cytometry.

Assessment of disease activity index in colon cancer mice model

The disease activity index (DAI) of mice was assessed based on changes in body weight (such as whether body weight increased or decreased), appearance of stool, and hematochezia of stool.

115

116

Supplementary Figures and Tables

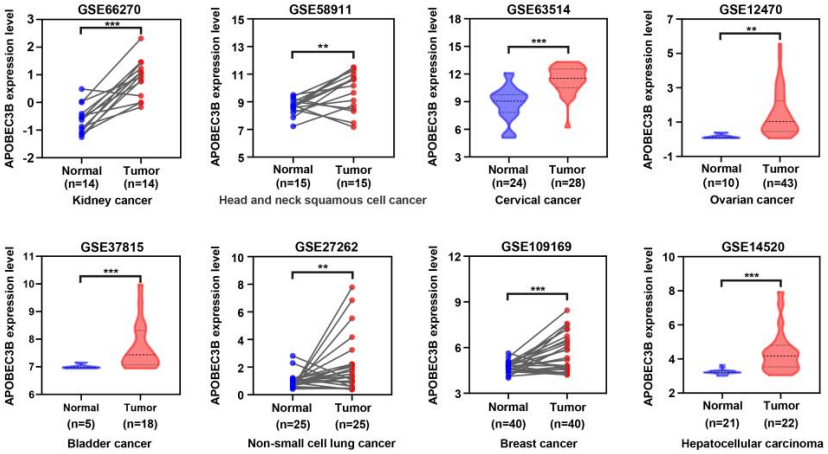


Figure S1. The expression level of APOBEC3B in normal and tumor tissues of different cancer types from GEO database. Data from GEO database have performed background correction and normalization. The statistical analysis was performed by one-tailed Student's *t* test (**$P < 0.01$, ***$P < 0.001$).

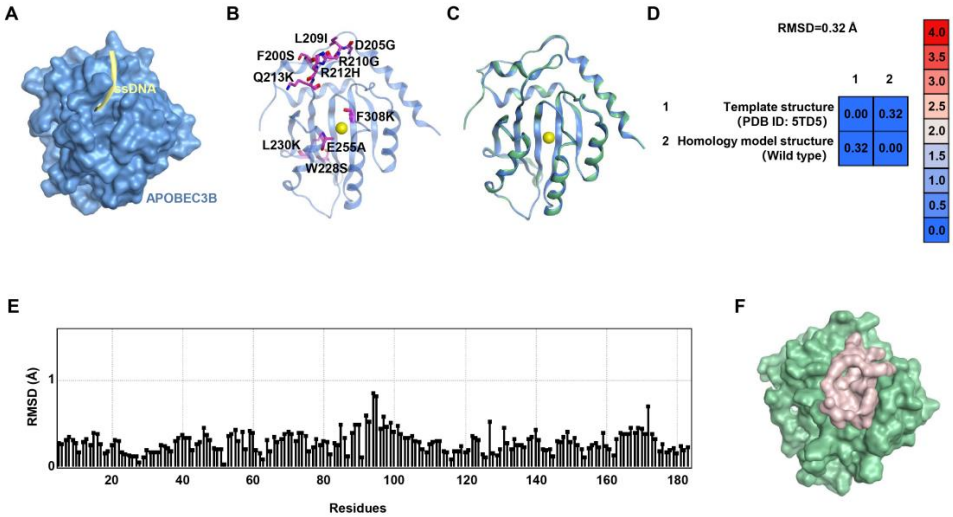


Figure S2. The structure of APOBEC3B. (A) The reported X-ray cocrystal structure of APOBEC3B (blue) with its substrate ssDNA (yellow) from PDB (PDB ID: 5TD5). (B) The mutated sites (purple) of APOBEC3B in cocrystal structure. The yellow dot showed the zinc ion in APOBEC3B structure. (C, D) The homology model structure (wild type, green) was superimposed with template structure (blue) with an RMSD of 0.32 Å by MOE. (E) The all-residue RMSD of homology model structure compared to template structure. (F) The pocket surface (pink) of wild type APOBEC3B suitable for screening compounds.

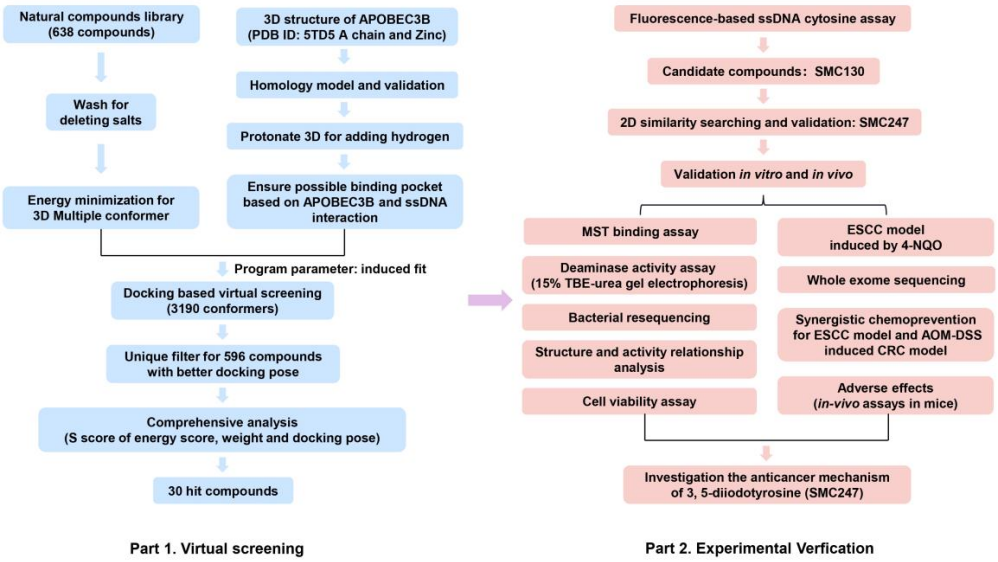


Figure S3. The comprehensive workflow chart in this research. The blue part on the left was virtual screening process using computer, and the pink part on the right was experimental verification process.

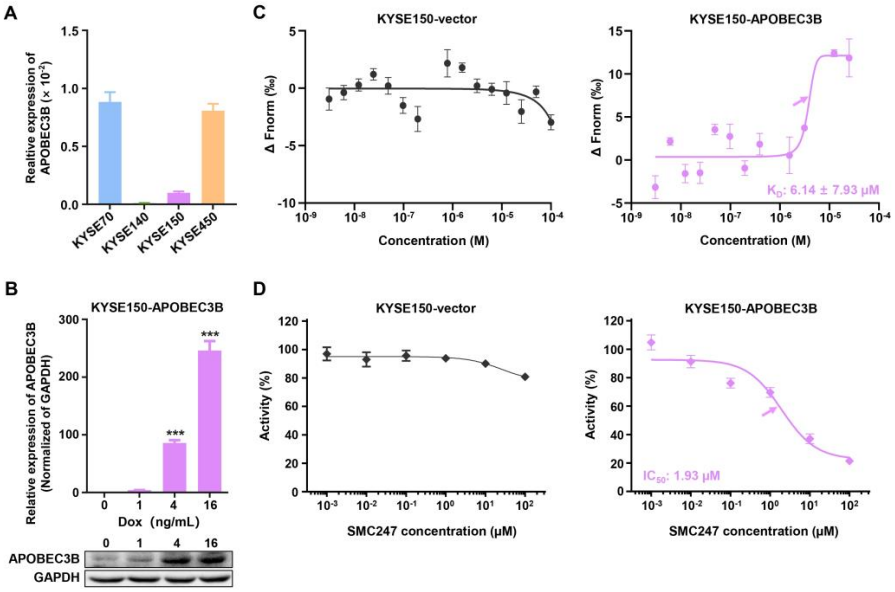


Figure S4. The interaction between 3, 5-diiodotyrosine and APOBEC3B in cancer cells. (A) The mRNA expression level of APOBEC3B in different ESCC cell lines. GAPDH was considered as the reference gene. (B) The mRNA and protein expression levels of APOBEC3B after KYSE150 cells were transfected with expression vector and induced by different concentrations of doxycycline (dox). (C, D) The interaction between 3, 5-diiodotyrosine and APOBEC3B were measured within KYSE150-vector (Empty vector was stably expressed in KYSE150 cells) and KYSE150-APOBEC3B (APOBEC3B vector was stably expressed in KYSE150 cells) cells by MST and fluorescence-based ssDNA deaminase analysis. Arrows showed the K_D or IC_{50} concentration values. Data are representative of at least three independent experiments.

179

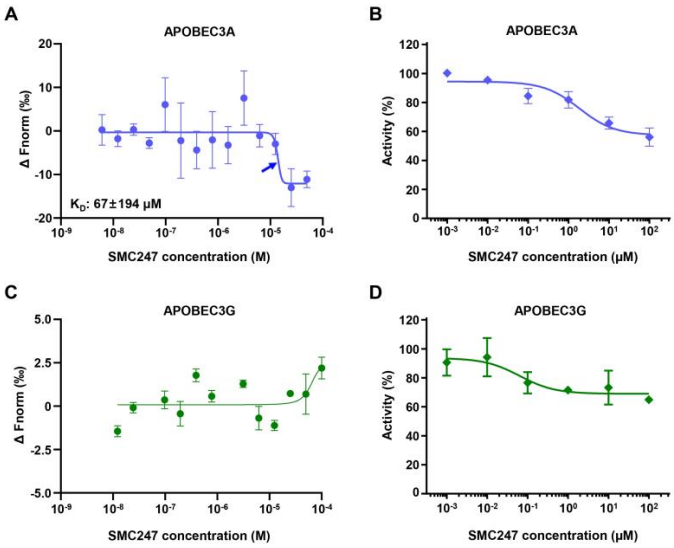


Figure S5. The interaction of 3, 5-diiodotyrosine with APOBEC3A and APOBEC3G. (A) The curves of 3, 5-diiodotyrosine binding to human APOBEC3A protein and APOBEC3G protein. (B) The curves of 3, 5-diiodotyrosine affecting on APOBEC3A and APOBEC3G deaminase activity. Arrows showed the K_D or IC_{50} concentration values. The absence of label indicated that K_D or IC_{50} could not be fitted. The data are representative of at least three independent experiments.

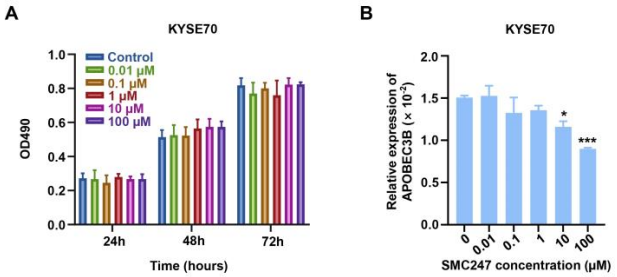


Figure S6. The effects of 3, 5-diiodotyrosine on the proliferation of KYSE70 cell and expression level of APOBEC3B. (A) Effects of different concentrations (0, 0.01, 0.1, 1, 10, 100 μM) of 3, 5-diiodotyrosine on the proliferation of KYSE70 tumor cell were measured by MTT. (B) The mRNA expression levels of APOBEC3B after 3, 5-diiodotyrosine treatment. KYSE70 cells with high expression of APOBEC3B were treated with different concentrations (0, 0.01, 0.1, 1, 10 and 100 μM) of 3, 5-diiodotyrosine for 48 hours, and then the mRNA expression levels of APOBEC3B were measured by qRT-PCR. (Mean ± SD were shown for triplicate reactions normalized to GAPDH, $*P < 0.05$ and $***P < 0.001$ by two-tailed Student's t test). Data are representative of at least three independent experiments.

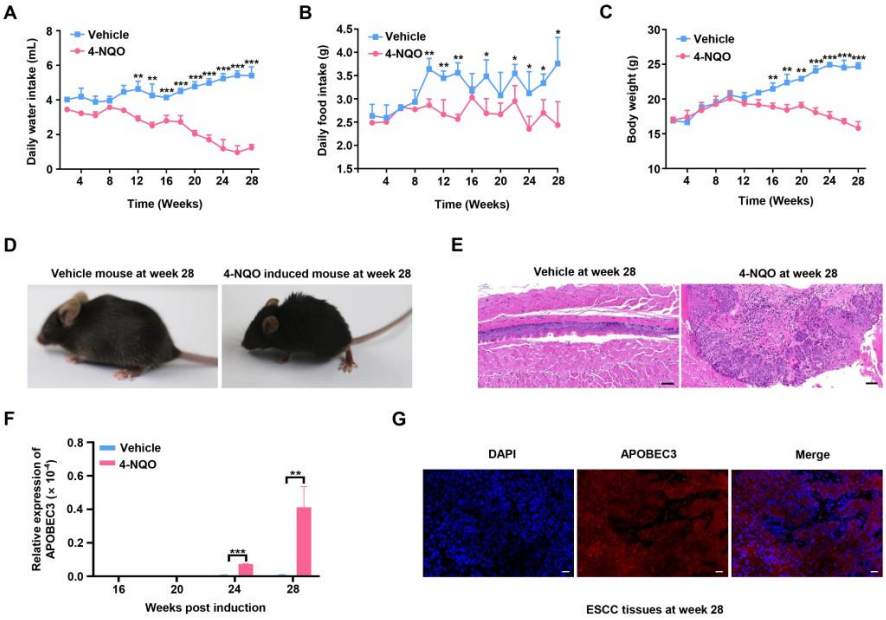


Figure S7. Model establishment and physiological indexes of 4-NQO induced ESCC mice. (A) Water intake, (B) Food intake and (C) body weight of vehicle mice and 4-NQO induced ESCC mice (**P* < 0.05, ***P* < 0.01 and ****P* < 0.001 by two-tailed Student's *t* test). (D) Representative physiological state images of vehicle mice and 4-NQO induced C57BL/6J mice at 28th weeks. (E) Histopathological condition of vehicle and 4-NQO induced mice at 28 weeks was determined by H&E staining assay (Scale bars: 50 μm). (F) The mRNA expression level of APOBEC3 was detected in esophagus tissues of vehicle and 4-NQO induced ESCC mice at different periods (***P* < 0.01, ****P* < 0.001). (G) Protein expression of APOBEC3 in esophagus tissues of 4-NQO induced ESCC mice at week 28 was detected by immunofluorescence analysis (Scale bars: 20 μm).

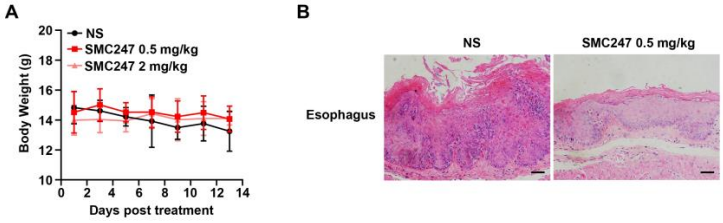


Figure S8. The physiological assessment of mice. (A) Body weight changes of mice (n = 5). (B) Histopathological assessment of esophagus tissues in NS and 3, 5-diiodotyrosine groups were determined by H&E staining assay (Scale bars: 100 μm).

248

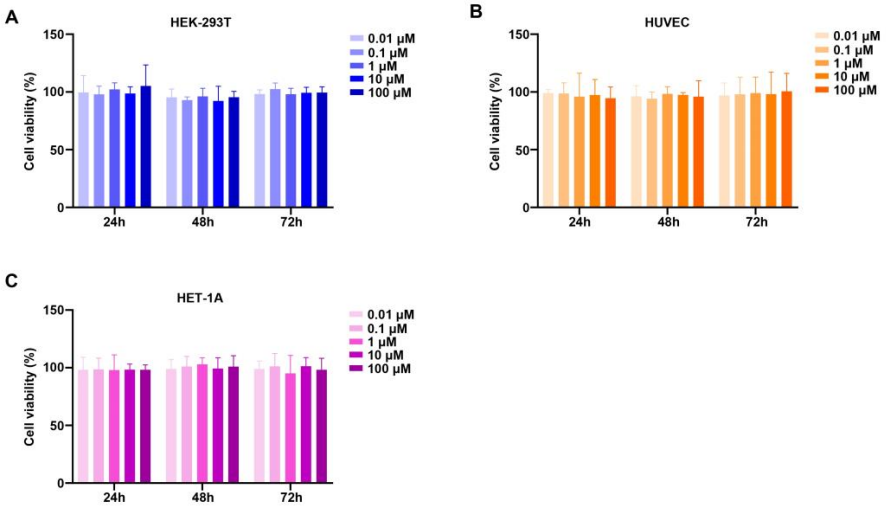


Figure S9. The effect of 3, 5-diiodotyrosine on cell viability. (A-C) HEK-293T, HUVEC and HET-1A cells were treated with different concentrations of 3, 5-diiodotyrosine for 24, 48 and 72 hours, and cell viability was measured by MTT analysis. Data are representative of at least three independent experiments.

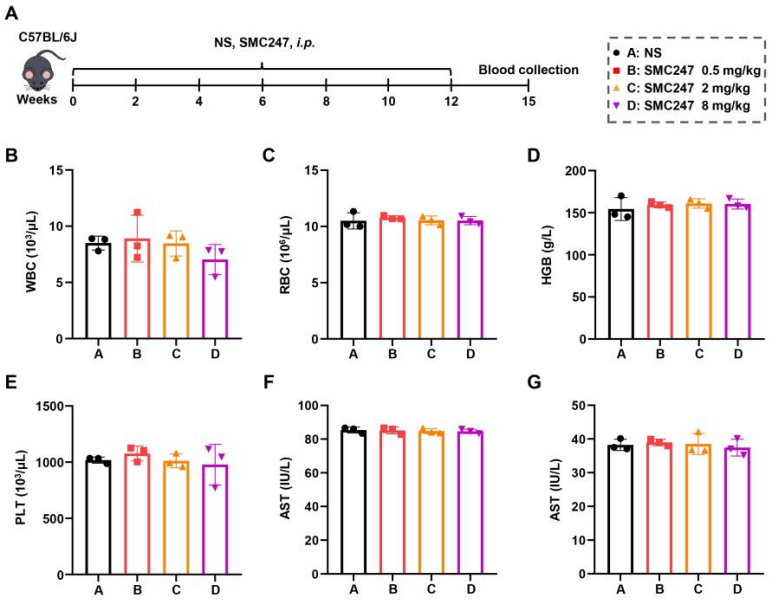


Figure S10. Toxicity evaluation of 3, 5-diiodotyrosine *in vivo.* (A) Seven-week-old naive female C57BL/6J mice were *i.p.* injected with 200 μL normal saline and different dosages of 3, 5-diiodotyrosine (0.5 mg/kg, 2 mg/kg, and 8 mg/kg), respectively, every two days for 7 times. (B-E) Blood physiology indexes (WBC, RBC, HGB and PLT) of mice were analyzed by blood routine assay. (F, G) Biochemistry indexes of mice (AST and ALT) were analyzed by AST and ALT kit, respectively (n = 3).

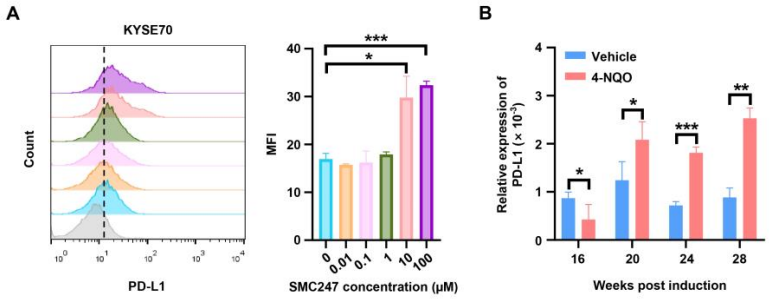


Figure S11. The expression level of PD-L1 in ESCC cell line and tumor tissue. (A) The expression of PD-L1 in KYSE70 cell after treatment with different concentrations (0, 0.01, 0.1, 1, 10, 100 μM) of 3, 5-diiodotyrosine for 48 hours was measured by flow cytometry. (B) The mRNA expression of PD-L1 in esophageal tissues of vehicle mice and 4-NQO induced ESCC mice at different periods was detected by qRT-PCR (*P < 0.05, **P < 0.01, ***P < 0.001).

299

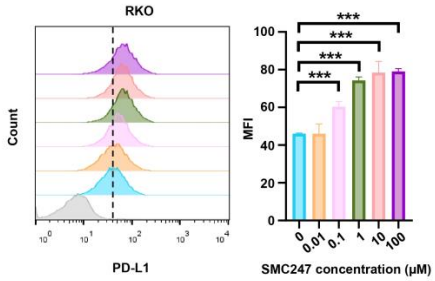


Figure S12. The expression of PD-L1 in colon cancer cell line. Flow cytometry analysis of PD-L1 expression in RKO cell after treatment with different concentrations (0, 0.01, 0.1, 1, 10, 100 μM) of 3, 5-diiodotyrosine for 48 hours. The statistical analysis was conducted by one-tailed Student's *t* test (***$P < 0.001$).

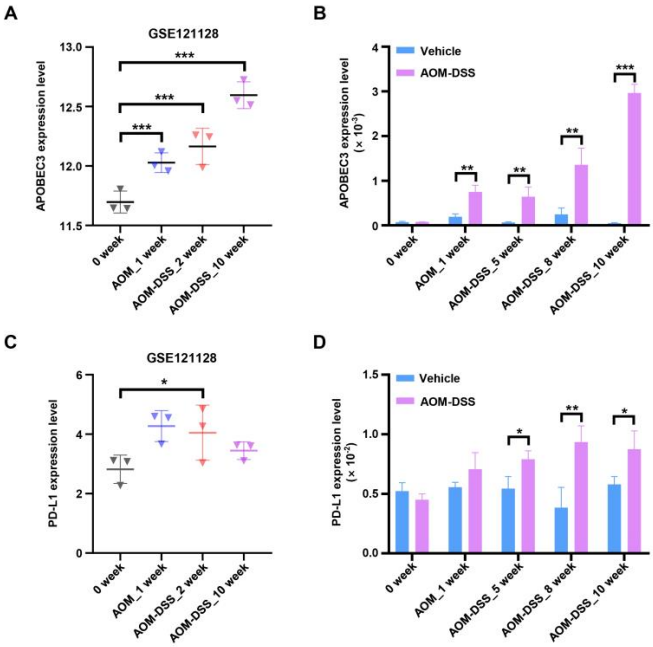


Figure S13. The expression levels of APOBEC3 and PD-L1 in colon tissues of vehicle mice and AOM/DSS induced colon cancer mice at different periods. (A, C) The expression levels of APOBEC3 and PD-L1 in colon tissues at 0, 1, 2 and 10 weeks after AOM/DSS induction were analyzed from GEO database (GSE121128) (n = 3). (B, D) The expression levels of APOBEC3 and PD-L1 in colon tissues of AOM / DSS induced colon cancer mice and vehicle at different time points (0, 1, 5, 8, 10 weeks) were analyzed (n = 3, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

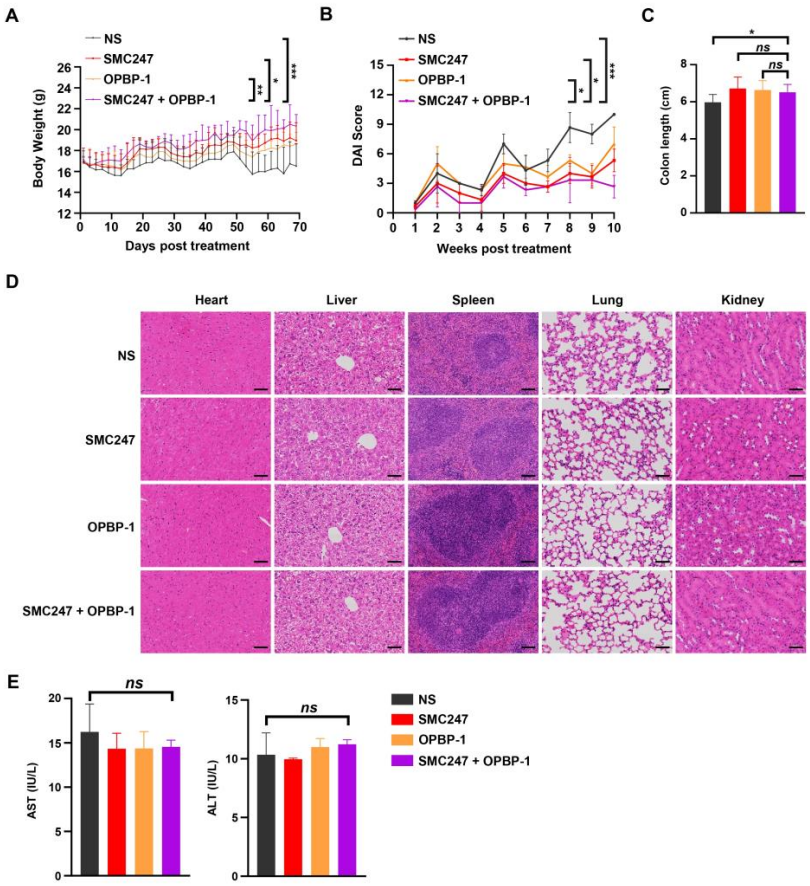


Figure S14. Model establishment of colon cancer mice induced by AOM/DSS and changes of physiological indexes after 3, 5-diiodotyrosine combined with OPBP-1 treatment. (A-B) The changes of body weight and DAI score in each group (n = 6 or 7, **P* < 0.05, ***P* < 0.01, ****P* < 0.001). (C) Colon length in each group. (D) The H&E staining of mice organs in each group after treatment (Scale bars: 100 μm for heart, liver, lung and kidney, 50 μm for spleen). (E) Analysis of hepatic damage by the level of AST and ALT in serum (n = 3).

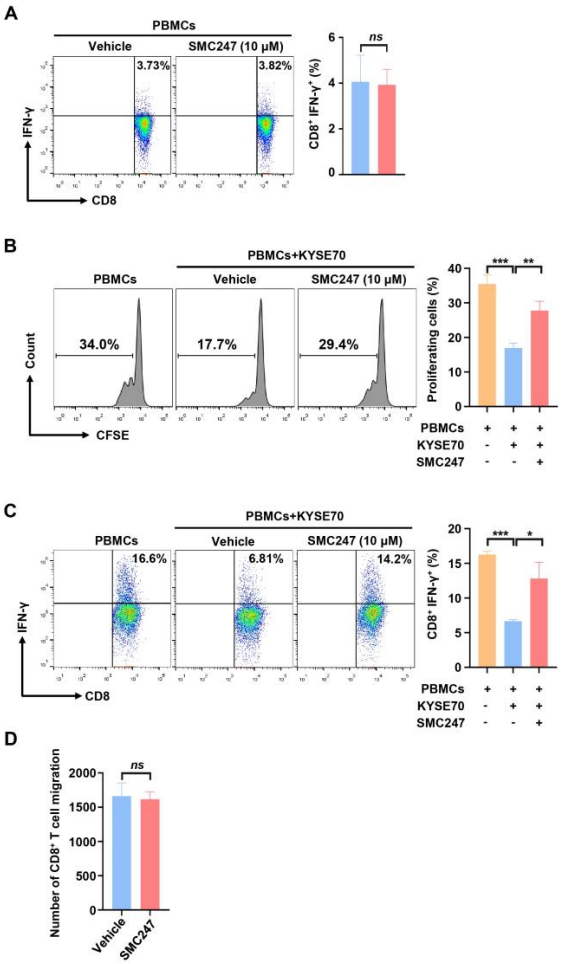


Figure S15. The effect of 3, 5-diiodotyrosine on $CD8^+$ T cell function. (A) PBMCs were isolated and activated with 100 units IL-2, 1 μg/mL anti-CD3, and 1 μg/mL anti-CD28 stimulatory antibodies and directly treated with 10 μM 3, 5-diiodotyrosine for 3 days. Then proportion of IFN-γ^+CD8^+ T cells were analyzed by flow cytometry. (B, C) Activated PBMCs were cocultured with KYSE70 cells within the addition of 10 μM 3, 5-diiodotyrosine. The proliferation of $CD8^+$ T cells and proportion of IFN-γ^+CD8^+ T cells were analyzed by flow cytometry. (D) The migration number of $CD8^+$ T cells was analyzed by flow cytometry. Data are representative of at least three independent experiments (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

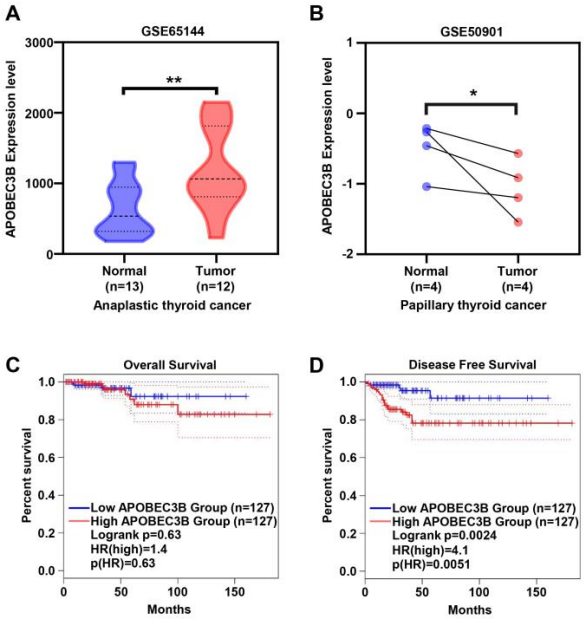


Figure S16. The expression and prognosis of APOBEC3B in thyroid cancer. (A, B) Expression levels of APOBEC3B in normal and tumor tissues of human anaplastic cancer and papillary thyroid cancer from GEO database were analyzed, $*P < 0.05$, $**P < 0.01$. (C, D) The overall survival curve and disease free survival curve in low expression of APOBEC3B and high expression of APOBEC3B in THCA were analyzed from GEPIA database.

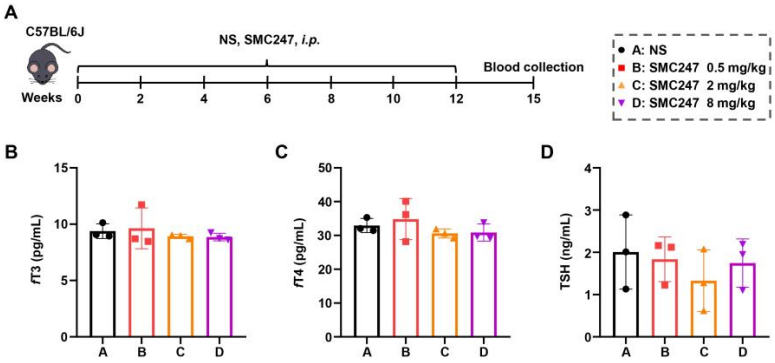


Figure S17. Effects evaluation of 3, 5-diiodotyrosine on thyroid function *in vivo*. (A) Seven-week-old naive female C57BL/6J mice were *i.p.* injected with 200 μL normal saline and different dosages of 3, 5-diiodotyrosine (0.5 mg/kg, 2 mg/kg, and 8 mg/kg), respectively, every 2 days for 2 weeks. (B-D) The main indexes of thyroid function (*f*T3, *f*T4 and TSH) in mice were analyzed (n = 3).

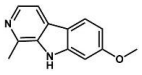
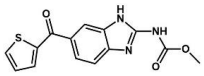
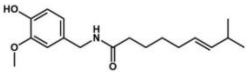
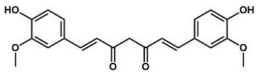
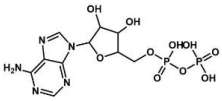
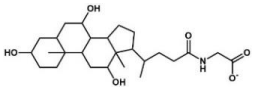
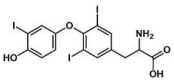
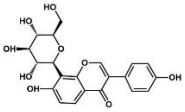
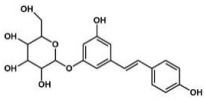
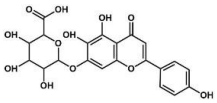
Table S1. The primer sequences for qRT-PCR.

Gene	Species	Forward	Reverse
APOBEC3B	Human	CCATCCTCTATGGTCGGAGC	GAGGCTTGAAATACACCTGGC
IL-7	Human	CCTCCCCTGATCCTTGTTCTG	ACCAATTTCTTTCATGCTGTCCA
IL-15	Human	ACAGAAGCCAACTGGGTGAAT	TGCTGTTACTTTGCAACTGGG
GAPDH	Human	GGAGTCCCTGCCACACTCA	GCCCCTCCCCTCTTCAAG
APOBEC3	Mouse	TGCTACATCTCGGTCCCTTC	TCCTCTTCACTTAGCGGGTC
PD-L1	Mouse	TCACTTGCTACGGGCGTTTAC	AGTTGCTGTGCTGAGGCTTA
GAPDH	Mouse	GCATCCACTGGTGCTGCC	TCATCATACTTGGCAGGTTTC

Table S2. Specific information for the crystal structures of human APOBEC3B.

NO.	PDBID	Notation	Resolution (Å)	Length	Mutation and missing	Reference
1	2NBQ	monomer	-	187-378	-	Biochemistry, 2016
2	5CQD	monomer	2.08	187-378	F200S, W228S, L230K, Y250S, F308K; Missing: A242-F249	J Biol Chem, 2015
3	5CQH	monomer	1.73	187-378	F200S, W228S, L230K, Y250S, F308K; Missing: A242-F249	J Biol Chem, 2015
4	5CQI	monomer	1.68	187-378	F200S, W228S, L230K, Y250S, F308K; Missing: A242-F249	J Biol Chem, 2015
5	5CQK	monomer	1.88	187-378	F200S, W228S, L230K, Y250S, F308K	J Biol Chem, 2015
6	5SXG	monomer	1.93	191-378	F200S, W228S, L230K, Y250S, F308K; Missing: D224-M231, A242-F249	Sci Rep, 2017
7	5SXH	monomer	1.78	191-378	F200S, W228S, L230K, Y250S, F308K; Missing: N225-M231, A242-F249	Sci Rep, 2017
8	**5TD5**	**APOBEC3B-ssDNA complex**	**1.72**	**191-378**	**F200S, D205G, R212H, Q213K, W228S, L230K, R210G, L209I, Y250S, E255A, F308K; missing: P206-V208, A242-F249**	**Nat Struct Mol Biol, 2017**
9	5TKM	monomer	1.90	1-191	Y83D, W127S, Y162D	Nucleic Acids Res, 2017
10	6NFK	monomer	1.86	187-378	F200S, W288S, L230K, Y250S, E255Q, F308K, Y315D, D316Q, P317G, L318R, Y319C, K320Q; Missing：A242-F249	FASEB Bioadv, 2019
11	6NFL	monomer	1.73	187-378	F200S, W288S, L230K, Y250S, F308K, Y315D, D316Q, P317G, L318R, Y319C, K320Q; Missing：A242-F249	FASEB Bioadv, 2019
12	6NFM	monomer	2.53	187-378	F200S, W288S, L230K, Y250S, F308K, Y315D, D316Q, P317G, L318R, Y319C, K320Q; Missing：A242-F249	FASEB Bioadv, 2019

Table S3. Structure and inhibition rate of 10 hit compounds in enzyme assay.

Compounds	Mol ID	2D structure	S score	Molecule weight	Log (O/W)	RMSD (Å)	Inhibition rate (%)
SMC111	T1711		-5.47	212.25	2.92	3.35	53.64
SMC115	T2802		-6.45	301.33	2.31	1.27	25.94
SMC116	T1062		-6.76	305.42	3.56	1.99	32.27
SMC118	T1516		-7.07	368.40	3.72	1.53	75.23
SMC121	T1723		-6.87	427.20	-4.09	1.34	41.76
SMC123	T2831		-7.11	464.62	3.15	2.08	23.67
SMC130	T1653		-5.54	650.98	3.45	2.09	70.19
SMC132	T2815		-7.45	416.38	-0.3	1.95	37.51
SMC139	T3427		-7.16	390.39	1.43	1.97	31.47
SMC140	T3242		-7.48	462.36	0.47	1.56	41.25

Inhibition rate indicates that the inhibition effect of compounds on APOBEC3B activity at the concentration of 100 μM.

S score indicates the comprehensive score for docking minimum energy.

Table S4. Representative 2D similarity searching results and docking results of SMC130.

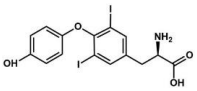
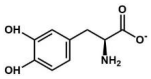
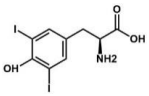
Compounds	Mol ID	2D structure	Molecular name	S score	Molecule weight	RMSD (Å)	Similarity score
SMC245	T4461		3,5-Diiodo-L-thyronine	-7.70	525.08	2.76	0.95
SMC246	T0848		Levodopa	-8.97	197.19	2.00	0.70
SMC247	T2760		3,5-Diiodotyrosine	-9.08	432.98	0.91	0.79

Table S5. The ssDNA sequences in fluorescence-based single-stranded DNA cytosine deamination assay.

Gene	Sequence
APOBEC3A	5′-6-FAM-ATTATTATTATTCTAATGGATTTATTTATTTATTTATTTATTT-TAMRA-3′
APOBEC3B	5′-6-FAM-ATTATTATTATTCAAATGGATTTATTTATTTATTTATTTATTT-TAMRA-3′
APOBEC3G	5′-6-FAM-ATTATTATTATTCCAATGGATTTATTTATTTATTTATTTATTT-TAMRA-3′

Table S6. The detailed DAI scoring criteria.

Score	Range of body weight loss	Stool appearance	Hematochezia
0	None	Normal	Normal
1	1-5%	—	—
2	5-10%	Light diarrhoea	—
3	10-20%	—	—
4	>20%	Severe diarrhoea	Severe hematochezia