Supplementary Table 2. Colorectal cancer patient characteristics

		Overall survival		
		N (%)	**P**	**HR (95%CI)**
Age	≤ 60	50 (33%)	<0.0009	1.045 (1.018, 1.073)
	> 60	101 (67%)		
Gender	Female	62 (41%)	NS	1.049 (0.596, 1.848)
	Male	89 (59%)		
T stage	T1 + T2	29 (19%)	0.036	0.287 (0.089, 0.924)
	T3 + T4	123 (81%)		
N stage	N0	100 (66%)	0.0008	0.383 (0.218, 0.673)
	N1 + N2 + N3	51 (34%)		
M stage	M0	125 (83%)	0.016	0.456 (0.241, 0.864)
	M1	26 (17%)		
Differentiation	Well/moderate	144 (95%)	NS	0.894 (0.379, 2.106)
	Poor	5 (3%)		
Localization of tumor	Distal colon	25 (17%)	NS	1.476 (0.763, 2.854)
	Proximal colon	100 (66%)		0.916 (0.440, 1.906)
	rectum	26 (17%)		
DOT1L	Low (< median)	65	0.0018	2.820 (1.469, 5.413)
	High (≥ median)	86		
H3K79me2	Low (< median)	70	0.045	1.835 (1.015, 3.316)
	High (≥ median)	74		

Table S2, related to Figure 6

Supplementary Table 3. Colorectal cancer patient characteristics

		DOT1L		H3K79me2	
		Low	**High**	**Low**	**High**
Age	< 60	25 (38%)	25 (29%)	24 (34%)	22 (30%)
	> 60	40 (62%)	61 (71%)	46 (66%)	52 (70%)
Gender	Female	22 (34%)	40 (47%)	28 (40%)	33 (45%)
	Male	43 (66%)	46 (43%)	42 (60%)	41 (55%)
T stage	T1 + T2	8 (12%)	20 (23%)	12 (17%)	16 (20%)
	T3 + T4	57 (88%)	66 (77%)	58 (83%)	59 (80%)
N stage	N0	39 (60%)	61 (71%)	41 (59%)	52 (70%)
	N1 + N2 + N3	26 (40%)	25 (29%)	29 (41%)	22 (30%)
M stage	M0	53 (82%)	72 (84%)	58 (83%)	61 (82%)
	M1	12 (18%)	14 (16%)	12 (17%)	13 (18%)
Differentiation	Well/moderate	64 (98%)	82 (95%)	68 (97%)	73 (97%)
	Poor	1 (2%)	4 (5%)	2 (3%)	2 (3%)
Localization of tumor	Distal colon	27 (42%)	20 (23%)	24 (34%)	22 (30%)
	Proximal colon	22 (34%)	25 (29%)	27 (39%)	19 (26%)
	Rectum	16 (25%)	41 (48%)	19 (27%)	33 (45%)

Table S2, related to Figure 6

Supplementary Table 4. Multivariate survival analysis

	P	HR (95%CI)
Age	<.0001	1.069 (1.037, 1.102)
Gender	NS	0.775 (0.420, 1.432)
T stage	NS	0.420 (0.122, 1.444)
N stage	0.001	0.333 (0.170, 0.651)
M stage	0.051	0.479 (0.228, 1.004)
Differentiation	NS	1.141 (0.455, 2.864)
DOT1L	0.0008	3.347 (1.646, 6.806)
H3K79me2	NS	1.003 (0.997, 1.010)

Table S4, related to Figure 6

Supplementary Table 5. Real-Time PCR Primers

Target Gene	Forward primer	Reverse primer
GAPDH	CTGCCCCCTCTGCTGATG	TCCACGATACCAAAGTTGTCATG
NANOG	AGGAAGACAAGGTCCCGGTCAA	GGTGCTGAGGCCTTCTGCGT
SOX2	ATGCACCGCTACGACG	CTTTTGCACCCCTCCCATTT
POU5F1	GAAGGTATTCAGCCAAACGA	AAATTCTCCAGGTTGCCTCT
huIL22	CTGGCCAGGCTCAGCAACAGG	CTTTGCTCTGGTCAAATGCAGGC
mIL22	CCTACATGCAGGAGGTGGTG	AAACAGCAGGTCCAGTTCCC
mbactin	CATTGCTGACAGGATGCAGAAGG	TGCTGGAAGGTGGACAGTGAGG
DOT1L	TCAAGATGACCGACGACGAC	CACCTCAGGACCAAAGGCAA
AFF4	CCTTCCCACCTAGCTCTCCT	GGTGGTACTGTAGGCTTGGG
MLLT10	CTCTCACCCACACAACCGTA	TGTGTTGGTCCAGGGATCTG

Table S5, related to Figures 1-4

Supplementary Table 6. ChIP Primers

Target Gene	Forward primer	Reverse primer
***SOX2* (-8kb)**	AGTGCCCAACTTTCTAGGGC	TGGGCCACCCACTTGTTTAG
***SOX2* (-7kb)**	TCAGACGGGCAGATAAGCAC	TGGGCTCAATGGTGTCAAGT
***SOX2* (-4kb)**	TAAGGCCTTTTGGCTAGGGC	TTCCACGGGCAACAAAAAGC
***SOX2* (4kb)**	ACCAGGAACCAACAATCGGG	CTCAAGGGTGGAAGACGCTG
***SOX2* (6kb)**	TTGTGGGAGCTCCATTGACG	GCGGTATTTCTTGTCCCCCT
***NANOG* (-9kb)**	ATAGGAGACCAAACGCGAGAA	TCTCGGCTATGACGGTTGCT
***NANOG* (-7kb)**	GACTCATCACTTTTGTGTAGCACC	ATTATGTGTTGACTACTTGGCCCT
***NANOG* (-1kb)**	TGTTAGTGCTGGAACCCCAC	AGACTACTCCGTGCCCATCT
***NANOG* (0.5kb)**	GGCACCTGCCCTTTGAACTA	TTTCCACCATGCCTAAGCCC
***NANOG* (4kb)**	TACCTCAGCCTCCAGCAGAT	GAGGCGATGTACGGACACAT
***POU5F1* (-9kb)**	GACGGCTCTGACTTTCACTCA	TTGTCCACAAGAGATGGCCC
***POU5F1* (-7kb)**	GTTGTGTGCCATGATGCTCC	GCAAGCCCCTTACGAAGTCT
***POU5F1* (-4kb)**	TAGCGGGACAGGGAAAAGTG	CTAAAGCCCTGGTGTGGAGG
***POU5F1* (-2kb)**	CCTGGCAGATTGAGGGATGT	ACTAGACCAGCAGCATGAGC
***POU5F1* (1kb)**	AATTGCTTACACTTGTCGCCT	CCTGGCACCCCTTGTAGAAA
***DOT1L* (-0.5kb)**	GGACCGTGACTCTTATGGGG	GGTTCGAATCCCGTCCCAG
***DOT1L* (4.5kb)**	GAGCAGTTCGGAAGGGGTTT	CATGCGGGGCTTGAGAAAAG
***DOT1L* (7.2kb)**	ATCCCTGCATTGAGCCCTTC	TTTGAACCTTCCAGGGCCAG
***DOT1L* (7.5kb)**	GGGGGTAGCCCTGCTTTTAT	GCAGCAAGCCTGGATCTTG
***AFF4* (1kb)**	CACGCACAAGCTACTGGGAT	CTTCGACTTGCGAAGGGGAT
***MLLT3* (10.7kb)**	AGTCGTCAACCACTGCAAAC	CCCCAATTCACCCAGTGGTAA
***MLLT10* (186kb)**	CAGTCAGCCTACTCATGGGAC	ACCGCCAGTAAGGCTTCAAT
FOS	GCAGCCCGCGAGCAGTT	GCCTTGGCGCGTGTCCTAATC

Table S6, related to Figures 3-5

Figure S1, related to Figure 1

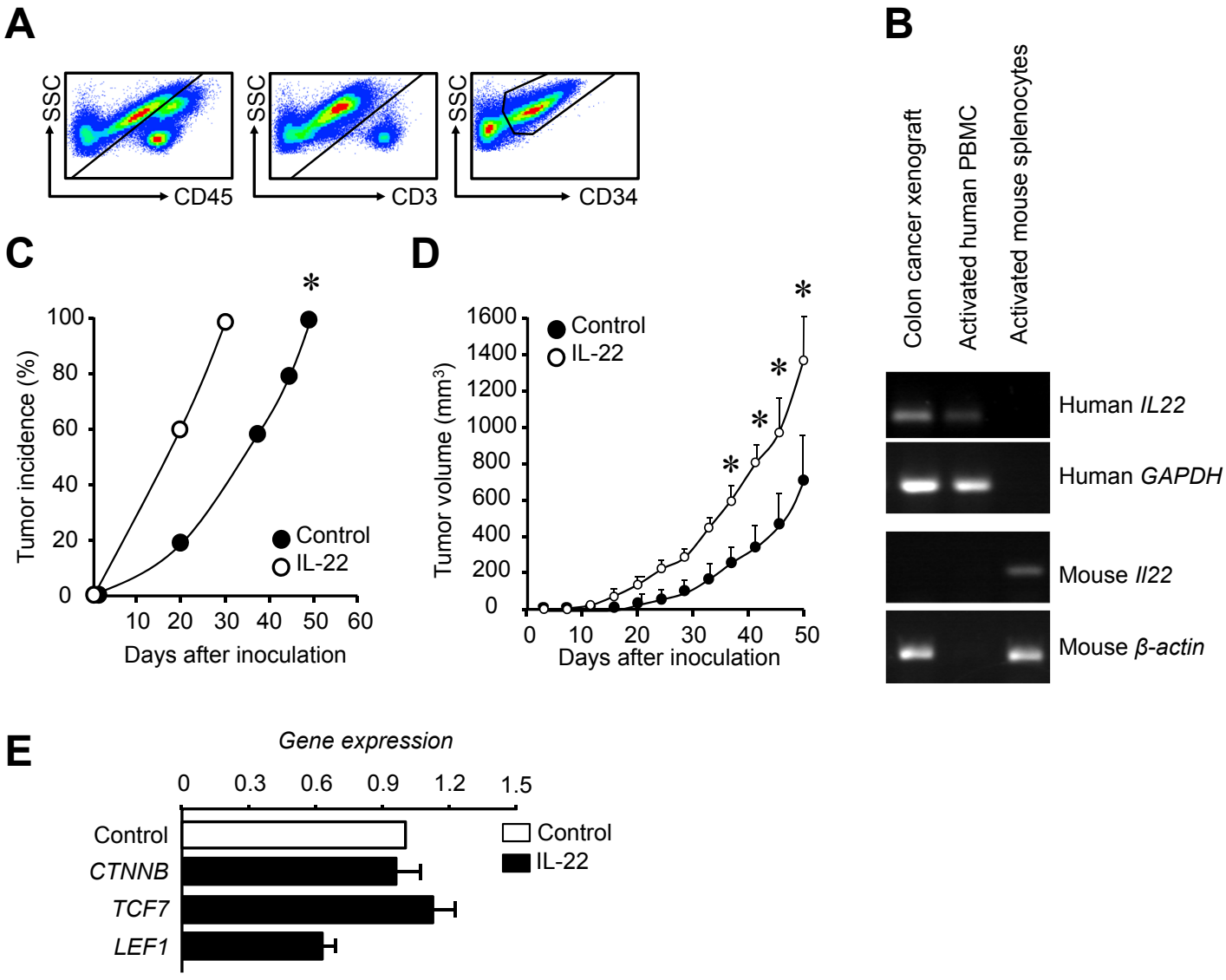


Figure S2, related to Figure 2

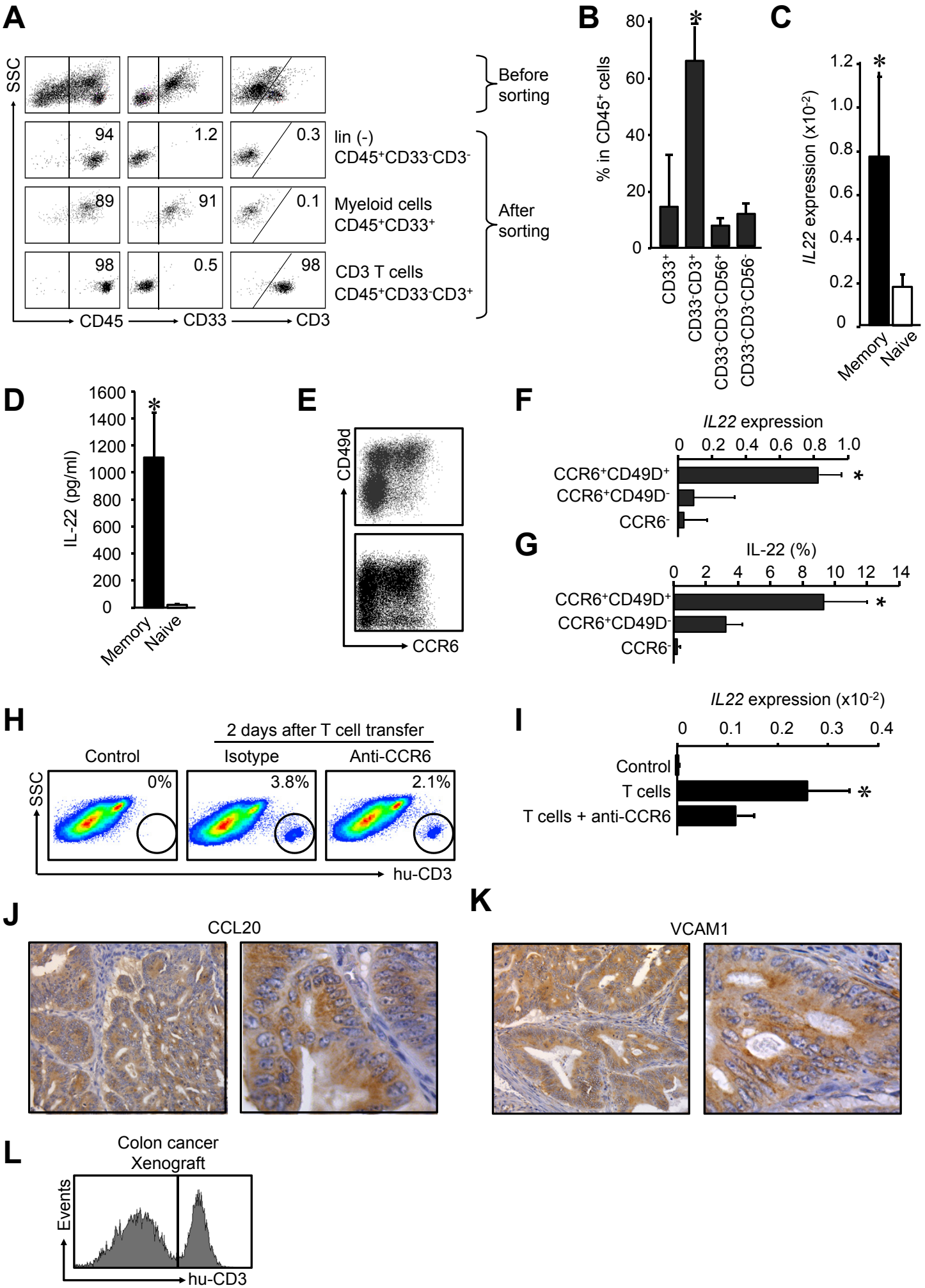


Figure S3, related to Figure 3

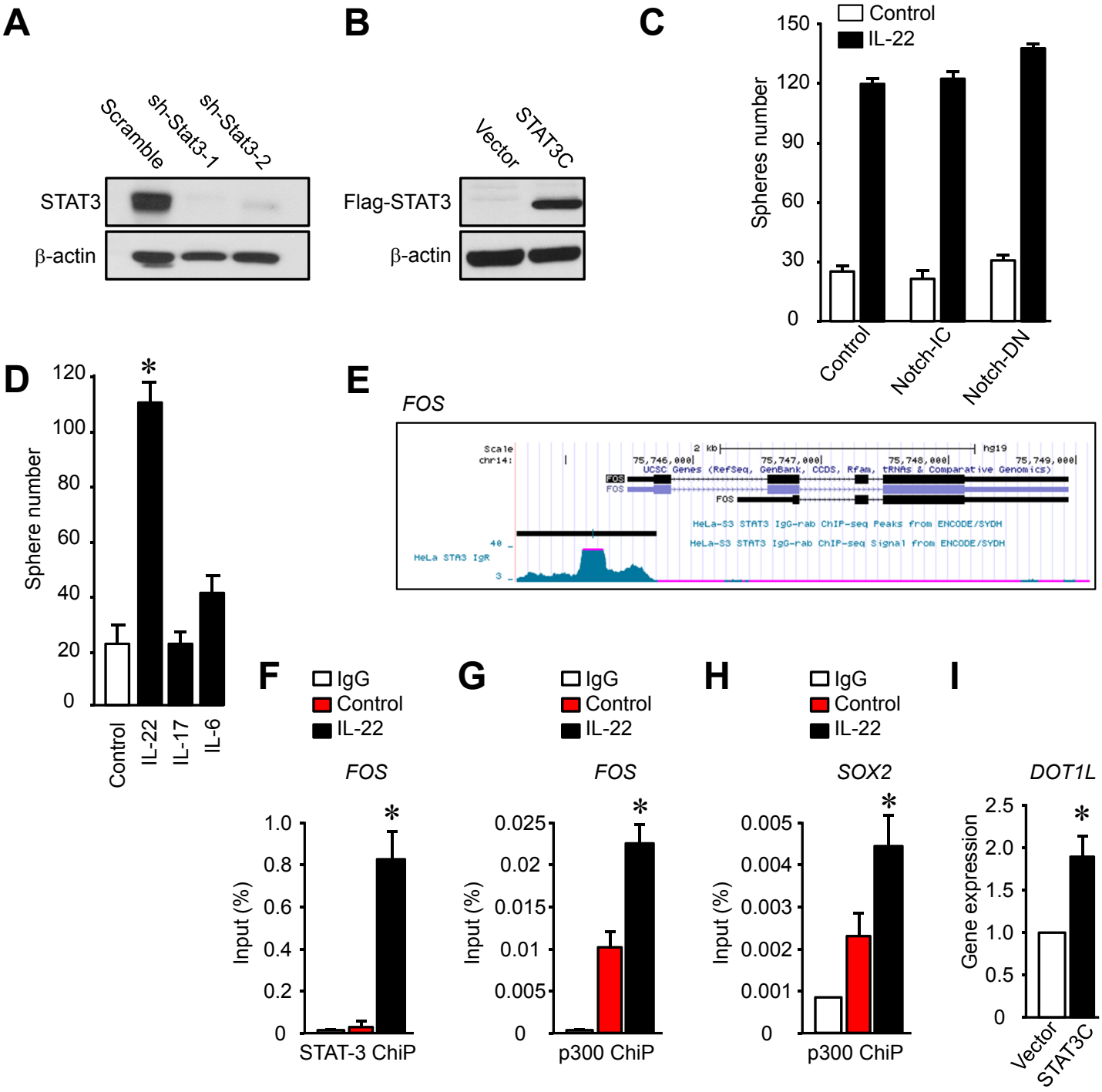


Figure S4, related to Figure 4

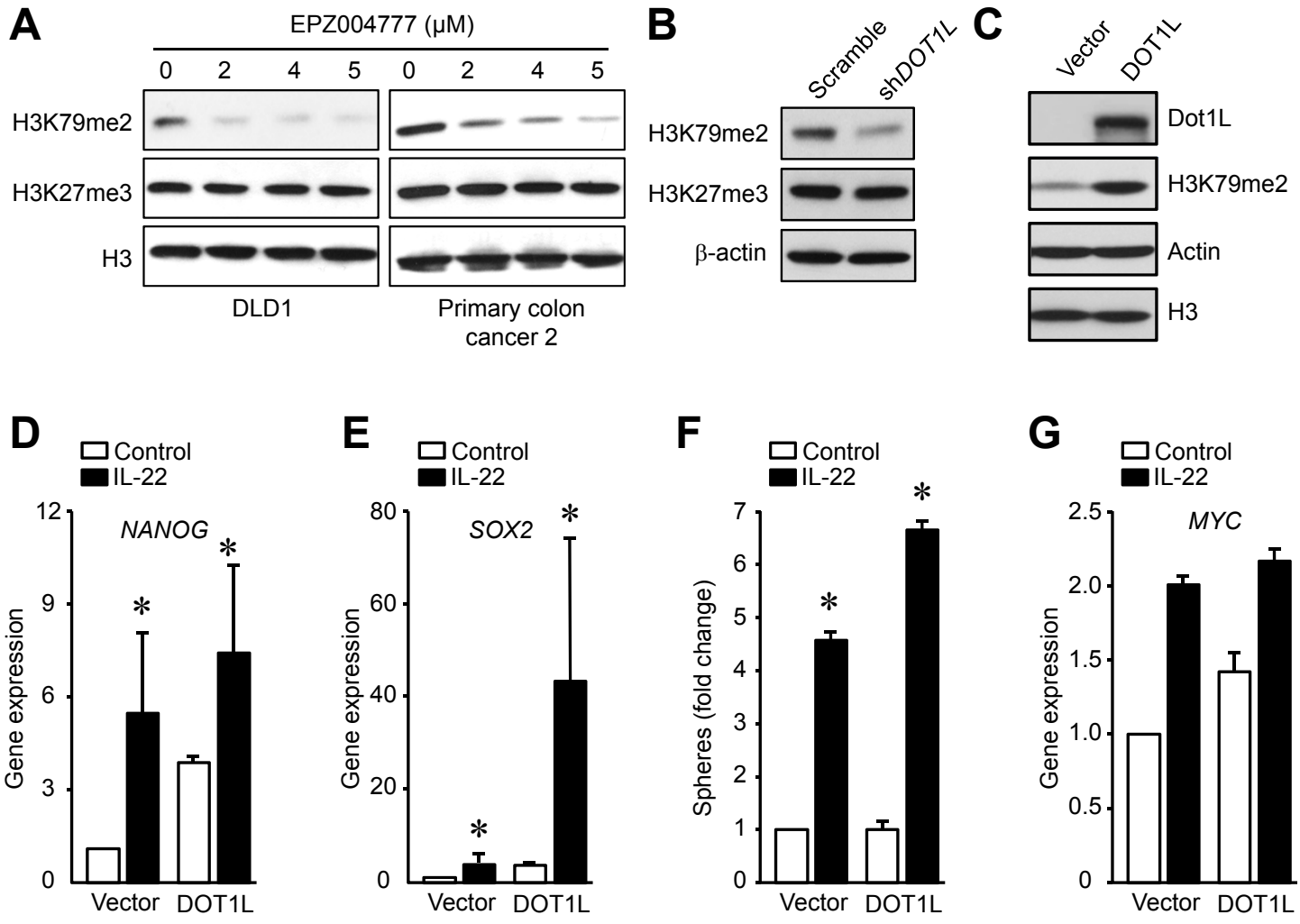
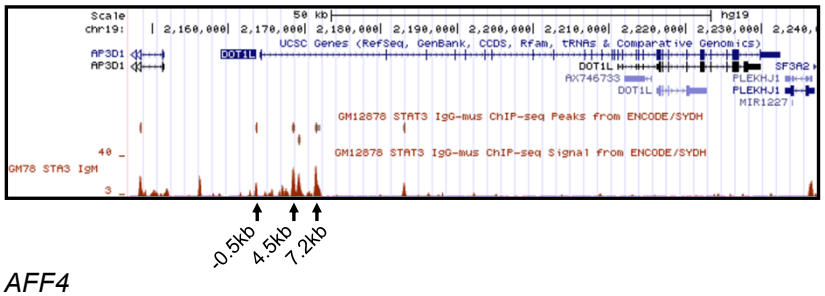
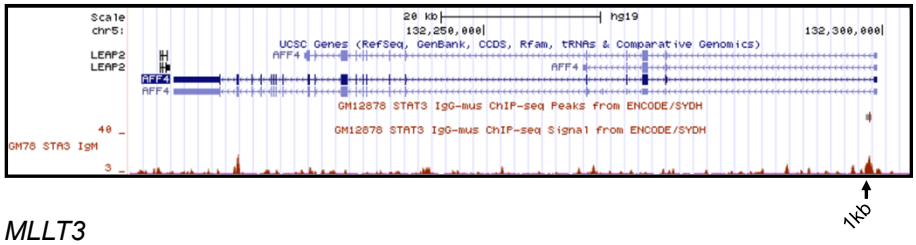


Figure S5, related to Figure 5

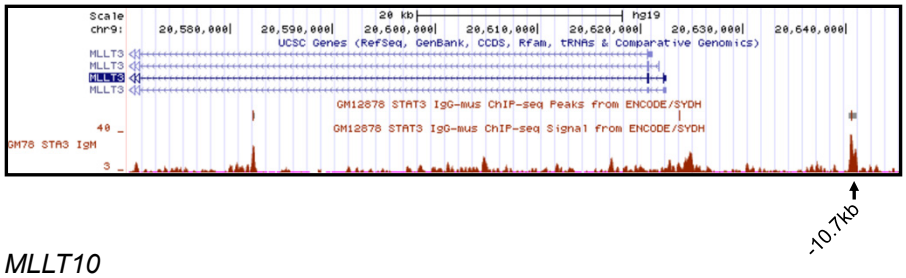
A *DOT1L*



B *AFF4*



C *MLLT3*



D *MLLT10*

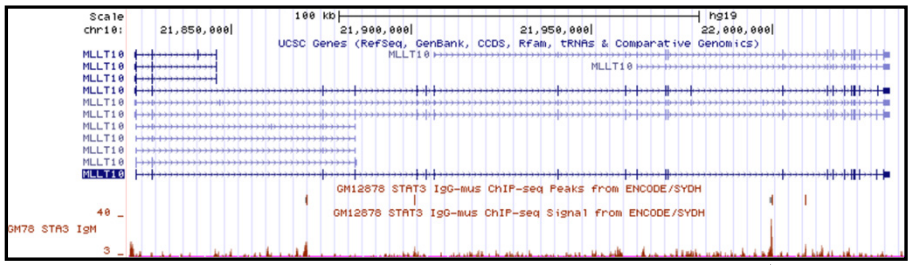
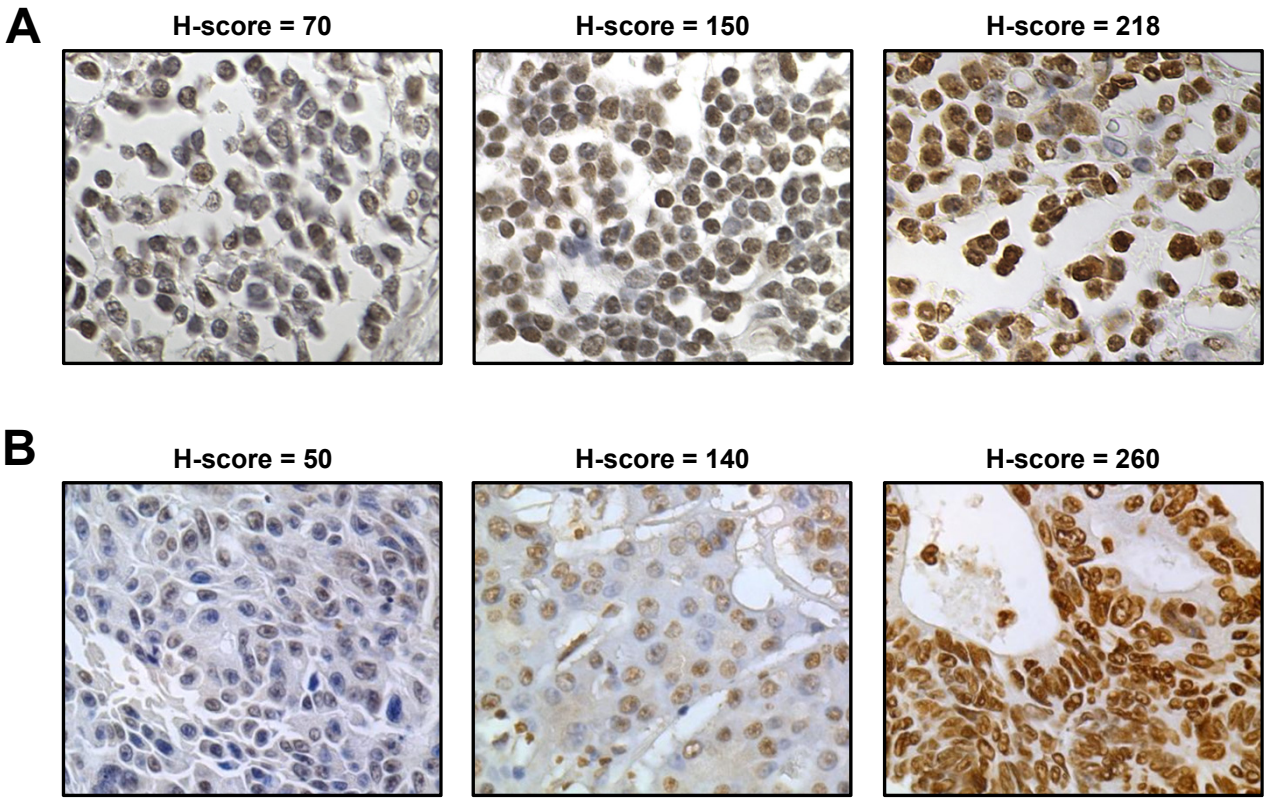


Figure S6, related to Figure 6



Supplementary Information

Supplementary figure legends

Figure S1, related to Figure 1 Expression and function of IL-22 expression in colon cancer tissues

(A) Major cellular components in fresh human colon cancer microenvironment. Single cells were prepared from fresh human colon cancer tissues and were stained for the following markers: immune cells (CD45), endothelial cells (CD33), and T cells (CD3). One of 15 is shown.

(B) Human and mouse IL-22 expression in the xenografted human colon cancer tissues. RT-PCR was conducted with human specific IL-22 primers or mouse specific IL-22 primers in the grafted human colon cancer tissues from NSG mice, activated human PBMCs, and wild type mouse splenocytes. One of 3 samples is shown.

(C, D) Effects of IL-22 on colon cancer development and growth. IL-22 treated 10^6 DLD-1 cells were inoculated into NSG mice. Tumor incidence (A) and tumor growth (B) were recorded. n = 5 per group, $P < 0.05$.

(E) Effects of IL-22 on β-catenin/Wnt signaling gene expression. The expression of β-catenin and Wnt signaling gene expression was measured by real time PCR in IL-22-treated DLD-1 cells. n = 5 per group, $P > 0.05$.

Figure S2, related to Figure 2 Phenotype and migration of IL-22$^+$CD4$^+$ T cells in the colon cancer microenvironment

(A, B) Single cells were prepared from fresh colon cancer tissues and were stained for the following markers: immune cells (CD45), myeloid cells (CD33), T cells (CD3), and NK/NKT cells (CD56). **(A)** Immune cell subsets were shown before and after sorting. n = 5. **(B)** Percentage of immune cell subsets is shown as the mean ± SEM. n = 10, *$P < 0.05$.

(C, D) CD45RA$^+$CD45RO$^-$ (naïve) and CD45RA$^-$CD45RO$^+$ (memory) CD4$^+$ T cells were sorted from peripheral blood. **(C)** The expression of *IL22* was measured by real time PCR in freshly sorted cells. **(D)** The sorted cells were activated for 3 days with anti-CD3 and anti-CD28. IL-22 was detected in the culture supernatants by ELISA. n = 5, * $P < 0.05$.

(E) CCR6 and CD49D are co-expressed in CD4$^+$ T cells. Expression of CCR6 and CD49D was analyzed in CD4$^+$ T cells by FACS. One of 4 donors is shown.

(F, G) CCR6$^+$CD49d$^+$CD4$^+$ T cells expressed IL-22. CCR6$^+$CD49d$^+$ T cells were sorted and activated with anti-CD3 and anti-CD28 and antigen presenting cells (APC). IL-22 mRNA was detected by real time PCR **(F).** IL-22 protein was detected by intracellular staining **(G).** 5 different donors ($P < 0.05$)

(H, I) $CCR6^+CD49d^+CD4^+$ T cells migrated into colon cancer in vivo. $CCR6^+CD49d^+$ T cells were injected into DLD-1 bearing NSG mice with anti-CCR6 or isotype. After 48 hours, tumor infiltrating human T cells were detected by FACS **(H)** and human IL-22 mRNA **(I)** was detected by real time PCR in the DLD-1 tissue. 3 independent experiments ($P < 0.05$).

(J, K) Colon cancer tissues express CCL20 and VCAM-1. Immunohistochemistry staining was performed with anti-CCL20 **(J)** and anti-VCAM-1 **(K)** mAbs. n = 5. Magnification: x100 (left panel) and x400 (right panel).

(L) Human T cells were detected in the grafted colon cancer in NSG mice. Fresh human colon cancer environmental cells were transferred into NSG mice. Human T cells were detected by FACS in the grafted human colon cancer tissues on day 54. One of 3 independent experiments is shown.

Figure S3, related to Figure 3 Manipulation of STAT3 expression with shRNA and overexpression.

(**A, B**) STAT3 manipulation. DLD1 cells were transfected with control vectors or sh-STAT3-1 and sh-STAT3-2 (A) or STAT3 active domain (STAT3C) (B). STAT3 expression was detected by western blotting.

(**C**) Effects of NOTCH signaling on IL-22-mediated colon cancer sphere formation. Sphere assays were performed with DLD-1 cells expressing Notch active domain (Notch-IC) or Notch negative domain (Notch-DN) in the presence of IL-22. Results are shown as the mean numbers of spheres in triplicates. ($n = 4$. $P > 0.05$).

(**D**) Effects of IL-6, IL-17 and IL-22 on colon cancer sphere formation. Sphere assays were performed with DLD-1 cells in the presence of cytokines. Results are shown as the mean numbers of spheres in triplicates. ($n = 4$. *, $P < 0.01$).

(**E**) ENCODE STAT3-ChIP-Seq data base analysis. Based on the ENCODE STAT3-ChIP-Seq data base (Access number: GSM935557), STAT3 occupancy on the proximal promoter area of *FOS* is shown.

(F, G) IL-22 causes higher occupancy of STAT3 and p300 on the *FOS* promoter. STAT3-ChIP and p300-ChIP assays were performed in DLD-1 colon cancer cells cultured with IL-22. One of 3 experiments is shown. (*, $P < 0.01$).

(H) IL-22 causes higher occupancy of p300 on the SOX2 promoter. p300-ChIP assay was performed in DLD-1 colon cancer cells cultured with IL-22. One of 3 experiments is shown. (*, $P < 0.01$).

(I) STAT3 activation causes increased DOT1L expression. STAT3C expressing colon cancer cells were cultured with IL-22. DOT1L expression was quantified by real-time PCR. One of 3 experiments is shown. (*, $P < 0.01$).

Figure S4, related to Figure 4 Manipulation of DOT1L and H3K79me2.

(A) EPZ004777 specifically inhibits H3K79 dimethylation. DLD-1 and primary colon cancer cells (#2) were cultured with different concentrations of EPZ004777. Histone marks were detected by western blotting. One of 4 experiments is shown.

(B) DOT1L expression is knocked down via shRNA. Colon cancer cells were transfected with control vectors or the shDOT1L vector. Histone marks were detected by western blotting.

(C) DOT1L overexpression. Colon cancer cells were transfected with control vectors or the DOT1L vector. Histone marks were detected by western blotting.

(D, E) Effects of DOT1L overexpression on core stemness genes. Colon cancer cells were transfected with control vectors or the DOT1L vector, and treated with IL-22. Core stemness genes were quantified by real-time PCR. n = 3. *, $P < 0.05$.

(F) Effects of DOT1L overexpression on sphere formation. Sphere assay was conducted with the vector or the DOT1L vector expressing colon cancer cells in the presence of IL-22. Sphere numbers were recorded. n = 3. *, $P < 0.05$.

(**G**) Effects of DOT1L on *MYC* expression. Colon cancer cells were transfected with control vectors or the DOT1L vector, and treated with IL-22. *MYC* expression was quantified by real-time PCR. One of 3 experiments is shown. ($P > 0.05$).

Figure S5, related to Figure 5 ENCODE STAT3-ChIP-Seq data base analysis.

Based on the ENCODE STAT3-ChIP-Seq data base (Access number: GSM935557), STAT3 occupancy on the promoter areas of *DOT1L* (A), *AFF4* (B), *MLLT3* (C) and *MLLT10* (D) is shown.

Figure S6, related to Figure 6 Immunohistochemical DOT1L and H3K79me3 scoring in colorectal cancers tissues

Colon cancer tissues were stained with anti-DOT1L antibody **(A**) and anti-H3K79me2 antibody **(B)**. Nuclear $DOT1L^+$ and $H3K79me2^+$ tumor cells were quantified with the H-score as described in Methods. < the median values were considered low DOT1L expression, and ≥ the median values were considered high DOT1L and H3K79me2 expression. Magnification: 400X