

A

	Poorly differentiated				Well-differentiated		
	HCT116	SW480	RKO	SW40	SW1222	LS180	C80
CEACAM5	0.0	0.0	0.0	0.0	106.7	81.1	220.6
CDX1	0.0	0.0	0.0	0.0	110.9	97.5	200.1
CDX2	4.8	9.4	5.4	4.8	46.8	16.2	41.7
KRT20	2.0	0.3	0.0	1.5	154.5	2.8	155.4
VIL1	0.0	2.6	0.0	0.0	69.0	55.3	94.5

B

FPKM

>gi|305410872|ref|NR_036581.1| Homo sapiens long intergenic non-protein coding RNA 675 (LINC00675), long non-coding RNA

GTGGCTCCAAGAAGCGCCAGCTCGCTCATTTGCTCACACCCAGCAGGCAGAGAAGGCAGCAGCAGGCAGGACCGCCACCCTCCCATGCAAATCA
CCCCGGGAGTGCGAGCTGGGCTCCTCCCGCTCCTCCTGGGCAATGCTCCTGGGGAGTCTGTGGGGAAGATGCCATCCAGGGCGCTGTGCGCTC
TTCCTCATCCTCGCCCTCCTGCTGGACGCGGTGCGCCTGGTCCTTTTGCTGCTGGGGATCTTGGCCCCCTGAGTTCCTGGGACTTCTTCATCTA
CACAGGTGCCCTGATCCTGGCTCTCAGCCTACTGCTCTGGATCATCTGGTATTCCTCAACATTGAGGTGTCTCCTGAAAACTGGACCTGTAATT
TGGCCATGGGAAGAGGAGAAGAGACGCAGGTGCTGTATGCAGACATGTCTGTGAACCTGGGGCTCTTGGGCAGCAACACGTTGCAGCTTCCACC
TAGCAAGCCACGCCGGGACCAGGTCCCATCTGATGGAGGAGAATCAATTCAAGGAGATGCCCTTCCCTTACAGAACACCCTTTAACAGCATCCAG
GAGGAACGAGAGGCTGCAATACTGAGGCTTTCAAAGTACTCACGAGGATGTCCGAGAATGGCTGTGATGCCAGGCTTCTGGCAGGTTCCAGACT
CCATCACAAAGCCCAGCATCCCTGCACCAGATCTGACATCGCTGCTGTTGTGCCAGCTGTTTATGAAGGGCCTGAGTAGCTAGCAGGTTTTATCA
GGAGCCCTGCTGGGGGCTTAGACACCAAAAGAGAAGACTCATCCTCTGTAGTTCTTCTTGTGAATGTCCTTTAGAAAAACAATTAGAACCAACCA
CAAGCACCAAAGTCCTAATGGGATCTCCTGCGAGCACATATCAAGCAGGATTGTTGCTATTTCTCCTCACTGGCTCTTTGGACAGACTGTGTGAGC
TCCTGGAGGGTTCCACTGTATCTACCCTTTGACACTGACCAGTTGGCACATGGTGACATATTCCAATGTGTTGATTGCAAATGTGAACGTACAGCC
AGTGCTGTGTGCGGGAGGACTCTCTCCTCCTCAGTGGGGCCACACCGTGCATTAATGGAGCCCCACTCCTTTGCACAGCCTGGCCATGCAG
TGGCTCATATTGAGGTTTTAGCCAACTGAAATCTCCCGTGCATTTTTCTGACAAGCCAGCTAGGCCTCTGCTATGCTGTCTTGTCTTTCAATTTG
ATGACCTTAAGGGTGGGACTGTTTTATCTTAAGTTACAGGTGGTCAAGTCCAGCCCAAGGACAGCAACTCTGAGGGTCAAGCCTCATAGGCTAAC
TGGATAGATGTTCTCTGCTTTGCCACCCACTGGAGCCCCGACCTGCCCCACTAATTTATATTTCCCTGGTCTCATTTTGTACTTTTTATTATAATTC
ACCCTTAAAGTGATGTGTCTCTTATAAGCTGCCTCCGATCTTTCATGGTATGAGGTGGTTACCTAAATAAAGAAGGAGATTGGCCTTTGTTTTTA
TGTA

Figure 1-figure supplement 2.

(A) RNA-seq was used to determine the expression of specific pro-differentiation genes in poorly differentiated CRC lines (green) vs well-differentiated CRC lines (blue). See Supplementary file 2 for more details.
(B) Sequence of *FORCP* transcript annotated as *LINC00675* by Refseq is shown. Polyadenylation signal (AATAAA) is shown in red near the 3'end of the transcript.