Table W1. Clinical Data from the Two Patient Cohorts.

	Series 1	Series 2	Combined
No. of patients	92	14	106
Age			
Mean	65.3	75.4	66.6
Gender			
Women	51	5	56
Men	41	9	50
Tumor localization			
Colon	54	14	
Rectum	38		
T stage			
T2	50	2	52
T3	42	11	53
T4	0	1	1
MSI status			
MSI	14	6	20
MSS	77	8	85
NA	1	0	1

A total of 106 patients, from two cohorts, were used in validating the fusions. The cohorts were enriched for patients with stage II or stage III CRC.

Table W2. Number of Paired RNA Sequence Reads that Passed the Illumina Chastity Filtering and the Number of Reads that Aligned to the Reference Genome hg19, Using TopHat v1.3.3 and Bowtie 0.12.7, for Each of the Seven Cell Lines on the Illumina Genome Analyzer IIx Instrument.

Cell Line	Filtered Clusters (M)	Reads Aligned (M)
HT29	33.9	61.3
HCT15	31.3	43.8
SW480	33.5	54.6
SW48	23.4	32.6
HCT116	32.1	46.8
LS1034	31.4	34.1
RKO	37.6	61.7

Table W3. Oligonucleotide Primers Used in RT-PCR to Validate the Fusion Transcripts.

Fusion Transcript	Cell Line with Indication from RNA Sequence Data	Forward Primer	Reverse Primer
MGRN1-C16orf96	HCT15	TGCAGGACTTGCTCACTGAT	GCTGCAGTGCAACCATCTT
NCOA3-SPINT1	HCT15	TAAAGCTGAGCTGCGAGGAA	GTGCACAGAAGCCACAAGG
PRMT1-FLT3LG	HCT15	TACCGTCAAGGTGGAAGACC	GAAGTTCTGGCGAGTGATCC
COMMD10-AP3S1	HCT116	CAGCTTTGCAGCAGCAATTA	TGTGCATCAATTTGTGTAACAATC
*COMMD10-AP3S1**	HCT116	CAAAGCTGAAGCATTTGTCA	CCCATCACCATTTCTGCAAG
COMMD10-AP3S1†	HCT116	GCCACCCAGTTTTATGCTTG	GCAAGACCAAGAGGTTTTGATG
SPAG9-MBTD1	HCT116	GTGCTGGAGAACCTGGACTC	CCATTCCCCACTCTGGTCTA
SLC39A14-TSPAN15	HT29	CTGCCTGGACCTCCTCTTTC	GGACTACCGAGATTGGAGCA
AKAP13-PDE8A	SW480	GCGAAGGTGAAGAGTTGTCC	AACCCTGCTTTTTCACATGC
*AKAP13-PDE8A**	SW480	TGTGCTCTGCCGAGACTG	AGCCTGAGCCTCCTTGGTA
CTB-35F21.1-PSD2	SW480	AGAGCTCCTCACCTGTTCCA	TCCTCAGGCACTGCAGATAA
*CTB-35F21.1-PSD2**	SW480	GTGTCTCCTGAGCCCACTCC	TCACTGGCCATCCCATTC
FAM96A-STIM1	SW480	CTGGAAGTGGTCTCGGAAAG	TCAGCCACTGTACCACCTCA
GRIN2B-CYP4F3	SW480	ATTCCCAACATGCTCACTCC	GCATCACTCCTGGGTGTCTT

*Primers used in nested RT-PCR.
†Primers used for testing the genomic breakpoint.

Table W4. By Using Five Filtering Steps, the Number of Candidate Fusions was Reduced from 3391 to 11.

Sample	Unfiltered Fusions	Normal	Ribosomal	Promiscuous	Intrachromosomal <100 kb	Not Exon Boundaries
(A) The number of potential fusions in the sequenced cell lines, before and after the five filtering steps						
RKO	999	880	397	45	40	1
HCT15	705	586	358	36	27	1
HCT116	693	545	207	41	31	2
SW480	506	406	220	43	37	6
LS1034	381	293	128	31	26	0
HT29	75	34	15	14	10	1
SW48	32	3	2	2	2	0
All	3391	2747	1328	212	173	11
(B) The number of potential fusions in the sequenced cell lines applying each filter separately						
RKO	999	880	418	72	897	96
HCT15	705	586	390	71	635	75
HCT116	693	545	228	59	586	79
SW480	506	406	235	136	433	64
LS1034	381	293	137	55	332	42
HT29	75	34	20	35	56	14
SW48	32	3	3	6	17	6
All	3391	2747	1431	434	2956	376

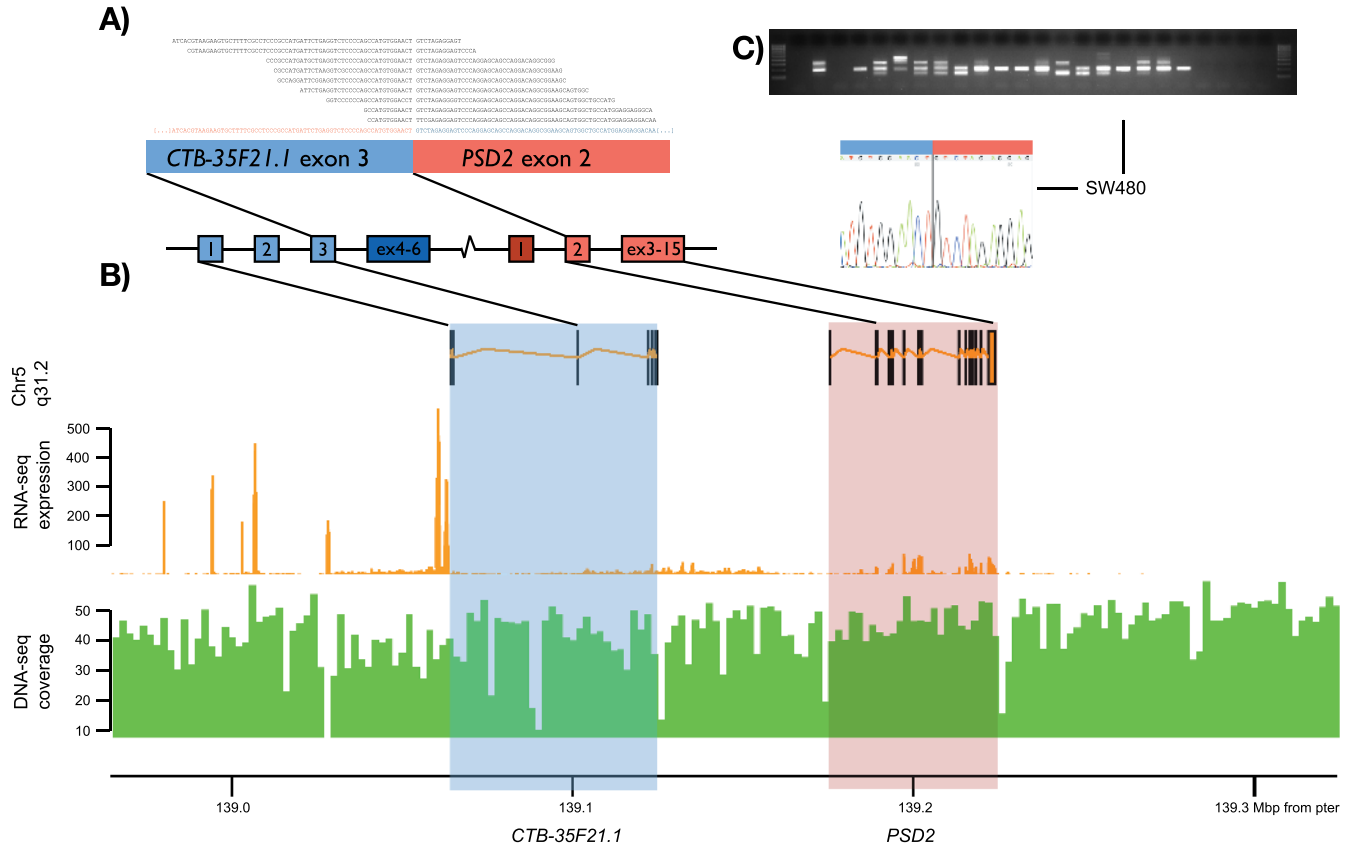

Figure W1. The *CTB-35F21.1-PSD2* fusion transcript is recurrent in CRC, and the genomic locus is rearranged in the SW480 cell line. (A) Nine sequence reads spanned the chimeric breakpoint, passing from exon 3 of *CTB-35F21.1* (ENST00000515296) to exon 2 of *PSD2* (ENST00000274710). Dark colors indicate exons that are not part of the fusion transcript. (B) Genomic view of the rearranged locus, from the top showing annotated exons of the fused genes (exons belonging to genes located between and within *CTB-35F21.1* and *PSD2* were removed for clarification) and relative RNA expression levels and DNA copy numbers, both based on high-throughput sequencing data from the SW480 cell line. (C) The *CTB-35F21.1-PSD2* fusion transcript was initially detected in the SW480 cell line, but nested RT-PCR demonstrated detectable levels in 17 additional colon cancer cell lines. The identity of the *CTB-35F21.1-PSD2* fusion transcript was verified by Sanger sequencing.

Table W5. Number of Positive Samples for Each of the 10 Experimentally Verified Fusion Transcripts.

Fusion Transcript	Colon Cancer Cell Lines	Clinical CRCs	Normal Colonic Mucosa
n	21	106	14
AKAP13-PDE8A	18	19	4
COMMD10-AP3S1	20	61	4
CTB-35F21.1-PSD2	18	49	2
FAM96A-STIM1	1	0	0
GRIN2B-CYP4F3	1	0	0
MGRN1-AC023830.2	2	0	0
SPAG9-MBTD1	1	4	1
NCOA3-SPINT1	2	0	1
PRMT1-FLT3LG	2	0	0
SLC39A14-TSPAN15	1	0	0

Table W6. The Presence of Fusion Transcripts in the 14 Matched Tumor-Normal Samples from Series 2.

Series 2 Pair	*AKAP13-PDE8A*		*CTB-35F21.1-PSD2*		*COMMD10-AP3S1*	
	Tumor	Normal	Tumor	Normal	Tumor	Normal
1	N	N	N	N	Y	N
2	N	N	N	N	Y	Y
3	N	N	Y	Y	Y	N
4	N	N	Y	N	Y	N
5	Y	N	N	N	N	N
6	N	N	Y	N	Y	N
7	N	N	N	N	Y	Y
8	N	N	Y	N	Y	N
9	Y	Y	N	Y	Y	N
10	N	N	Y	N	Y	Y
11	N	Y	Y	N	Y	N
12	N	N	N	N	N	N
13	N	Y	N	N	Y	Y
14	Y	Y	Y	N	Y	N

Table W7. Sequence Validation of Nested PCR Products was Performed on Altogether 36 Samples Positive for either of the Identified Fusion Transcripts.

Fusion Transcript	Positive Samples Confirming the Identity of Both Partner Genes
COMMD10-AP3S1	7/7 (100%)
CTB-35F21.1-PSD2	6/6 (100%)
AKAP13-PDE8A	5/6 (83%)
SPAG9-MBTD1	2/4 (50%)
NCOA3-SPINT1	2/3 (67%)
PRMT1-FLT3LG	7/8 (88%)
MGRN1-C16orf96	2/2 (100%)

Table W8. The Presence of Fusion Transcripts in 20 Normal RNA Samples.

Sample	*AKAP13-PDE8A*	*COMMD10-AP3S1*	*CTB-35F21.1-PSD2*
Adipose	N	N	Y
Bladder	N	N	Y
Brain	N	N	Y
Cervix	N	N	Y
Colon	Y	Y	Y
Esophagus	N	Y	Y
Heart	N	N	N
Kidney	N	N	Y
Liver	N	Y	Y
Lung	Y	N	Y
Ovary	N	Y	Y
Placenta	N	Y	Y
Prostate	N	N	Y
Skeletal muscle	N	N	Y
Spleen	N	Y	Y
Stomach	N	N	N
Testes	Y	Y	Y
Thymus	N	Y	Y
Thyroid	N	Y	N
Trachea	N	Y	Y