## Supplementary Materials

Chimera	RNA Fusion Junction Sequence
MUC1-TRIM46	See Table S5
SH2D4A-CSGALNACT1	TTTCCTCATCGATGCCTCTGCAGACGCCTACAGCTTCCTGGGCGTGGACC
	AGCTACAGCATGCCACCTTGGCGGATTTGGTGGAATATCACAAG*GTGAG
	CCTCTATCTCGTGCCTGAACACCAGCATGCCCAGCTGGCCGTGG
CCDC19-VSIG8	CAGCAGAAGGAAGTGCAGAACCGGATTGCCACCTTTGAGGAGGGCCGG
	CGCCTCAAAGAGGAGGCCCAGAAACGCCGTGAGCGCATCGATGAGATC
	AAGAGGAAAAAGCTTGAAGAGCTGAG*CACTGCTGTCTGCTGCGGAT
	CAACGGGGATGGACAGGAGGTCCTGTACCTGGCAGAAGGTGATAATGTG
	AGGCTGGGCTGCCCTACGTCCTGGACCCTGAGGACTATGGTCCCAATG
	GGCTGGACATCGAGTGGATGCAGGTCAACT
SFPQ-AL831889	TGGTGGTGGCATAGGTTATGAAGCTAATCCTGGCGTTCCACCAGCAACC
	ATGAGTGGTTCCATGATGGGAAGTGACATG*GTAAGAATGATTGATGTTG
	GCTGATATTGGAGTGCTCATTCACATGAAGTGGATAGATA
	CATCACACAGCGTGAGTCAATCAAGGAGGGAAGCCACAAGCAGACTGA
	CAACGTTTCTAGGATCAGGTGAGCTGTGTCCAGAAA
CS-CNPY2	CTGCTCACCTTGGTTCCCAAGAGCCGGGCGGCCGCA*CTCCGACTCAGAT
	TTTGAACGGATGTTTGCATGTATACCACAGTGCTACATTAAAAACCAGTC
	ATGTGGATGTGGAGAAGGGTGCTGATCAGGACTCCCTGTTTGCTAACAA
	ATTCTACAGCAGAAACAAAGTATTAAGGTTTGCAGCCAGGAAGCAAGTG
	CATGGGGAA
SLC16A3-CSNK1D	TTCAAGGAGCTCATACAGGAGTTTGGGATCGGCTACAGCGACACAGCCT
	GGATCTCCTCCATCCTGCTGGCCATGCTCTACGGGACAG*GGGTCCTGCT
	GTCACCCAGGCTGGAAGGCAGTGGCACCATCACAGCTCCCTGCAGCCTT
	AACCTCCCGAACTCTAGCCATCCTCCTGCCTCAGCCTCCAGAGTAGCTGA
	GACTACAGATGTAAGC
SPINT2-YIF1B	CCCACGCTGGTACTTTGACGTGGAGAGGAACTCCTGCAATAACTTCATCT
	ATGGAGGCTGCCGGGGCAATAAGAACAGCTACCGCTCTGAGGAGGCCT
	GCATGCTCCGCTGCTTCCGCCAGCAGGAGAATCCTCCCCTGCCCTTGGC
	TCAAAGG*CCCTGGAGCCCACCTCCCAGAAGCCCGGTGTGGG
CTSC-RAB38	TTGCCAACATCTTGGGACTGGAGAAATGTTCATGGTATCAATTTTGTCAG
	TCCTGTTCGAAACCAAG*GTCAAGAAAGATTTGGAAACATGACGAGGGT
	CTATTACCGAGAAGCTATGGGTGCATTTATTGTCTTCGATGTCACCAGGC
	CAGCCACATTTGAAGCAGTGGCAAAGTGGAAAA
JAK3-INSL3	CCGCCTCTTGGAACTGCTGGAGGAGGGCCAGAGGCTGCCGGCGCCTCCT
	GCCTGCCCTGCTGAG*GTGAGTTGCTACAGTGGCTGGAGAGACGACATCT
	GCTCCATGGGCTGGTGGCCGACAGTAATCTCACGCT

Table S1. RNA junction sequences of identified chimeric RNAs.

The nine validated chimeric RNAs along with their RNA fusion junctions are shown. The RNA fusion junction is represented by "\*".

**Table S2.** Recurrence profile of identified chimeric RNAs. The validated chimeric RNAs are shown along with their recurrence in the 130 TCGA sample cohort by bioinformatics analysis. Since TCGA transcriptome sequencing collection does not contain organ specific controls, we examined these nine chimeric RNAs using our in-house non-cancerous ovary samples by RT-PCR to answer the question of whether these candidates are also present in non-cancerous tissue.

Chimera	Type	Chromosome	Recurrence among 130 TCGA Samples	Recurrence among 11 Non-Cancer Donor Ovary Samples
MUC1-TRIM46	Intrachromosomal	1	9	0
SH2D4A-CSGALNACT1	Intrachromosomal	8	22	11
CCDC19-VSIG8	Intrachromosomal	1	10	1
SFPQ-AL831889	Intrachromosomal	1	98	11
CS-CNPY2	Intrachromosomal	12	23	n.d.
SLC16A3-CSNK1D	Intrachromosomal	17	18	5
SPINT2-YIF1B	Intrachromosomal	19	55	11
CTSC-RAB38	Intrachromosomal	11	6	11
JAK3-INSL3	Intrachromosomal	19	7	11

n.d.: not determined.

**Table S3.** Paired chimeric reads for *MUC1-TRIM46* identified from TCGA database. These 25 reads were obtained from nine patient samples. These nine patient sample IDs are: TCGA-04-1348, TCGA-10-0931, TCGA-13-0720, TCGA-13-1403, TCGA-13-1482, TCGA-23-1122, TCGA-24-0979, TCGA-24-1560 and TCGA-24-1604.

Read 1 (Aligning to MUC1 Gene)	Read 2 (Aligning to TRIM46 Gene)
CATTTCAAACCTCCAGTTTAATTCCTCTCTGGA	TGCGACAGGGAGGAAATAATGCGTCGAGGG
AGATCCCAGCACCGACTACTACCAAGAGCTGC	GGCATCCTCGGCCACCGCCCTGGTCCTCAGG
AGAGAGACAT	CTTTCCTCGTACCT
CGAGTACCCCACCTACCACACCCATGGGCGCT	GAGGAAATAATGCGTCGAGGGGGGCATCCTC
ATGTGCCCCCTAGCAGTACCGATCGTAGCCCC	GGCCACCGCCCTGGTCCTCAGGCTTTCCTCGT
TATGAGAAGAG	ACCTGGTGCGTCT
GATACCTACCATCCTATGAGCGAGTACCCCAC	GCAGCGCCCTCTTGCGACAGGGAGGAAATA
CTACCACACCCATGGGCGCTATGTGCCCCCTA	ATGCGTCGAGGGGGGCATCCTCGGCCACCGCC
GCAGTACCGAT	CTGGTCCTCAGGCT
ATCCTATGAGCGAGTACCCCACCTACCACACC	TCGAGGGGGCATCCTCGGCCACCGCCCTGGT
CATGGGCGCTATGTGCCCCCTAGCAGTACCGA	CCTCAGGCTTTCCTCGTACCTGGTGCGTCTGC
TCGTAGCCCCT	TGCCACTCTTCT
GAAGAATGCTTTTAATTCCTCTCTGGAAGATCC	AGGGGGCATCCTCGGCCACCGCCCTGGTCCT
CAGCACCGACTACTACCAAGAGCTGCAGAGA	CAGGCTTTCCTCGTACCTGGTGCGTCTGCTGC
GACATTTCTGA	CACTCATTTCAG
ATCCTATGAGCGAGTACCCCACCTACCACACC	CCTGGACGCAGCGCCCTCTTGCGACAGGGAG
CATGGGCGCTATGTGCCCCCTAGCAGTACCGA	GAAATAATGCGTCGAGGGGGGCATCCTCGGCC
TCGTAGCCCCT	ACCGCCCTGGTCC
CGGGATACCTACCATCCTATGAGCGAGTACCC	GTCGAGGGGGCATCCTCGGCCACCGCCCTGG
CACCTACCACACCCATGGGCGCTATGTGCCCC	TCCTCAGGCTTTCCTCGTACCTGGTGCGTCTG
CTAGCAGTACC	CTGCCACTCTTC
GAATTAAACTGGAGGTTTGAAATGTGAAAAGA	GGTACGAGGAAAGCCTGAGGACCAGGGCGG
CAGGAAAAAGAAAGAGACCCCAGTAGACAAC	TGGCCGAGGATGCCCCCTCGACGCATTATTT
TGGGGAGAAGTG	CCTCCCTGTCGCAA
GAGTACCCCACCTACCACACCCATGGGCGCTA	TCTCCTCTTTCTTCCTCTCAGTGCCTGGACGC
TGTGCCCCCTAGCAGTACCGATCGTAGCCCCT	AGCGCCCTCTTGCGACAGGGAGGAAATAAT
ATGAGAAGAGT	GCGTCGAGGGGGC

Table S3. Cont.

Read 1 (Aligning to MUC1 Gene)	Read 2 (Aligning to TRIM46 Gene)
GTCAGTGCCGCCGAAAGAACTACGGGCAGCT	GGACGCAGCGCCCTCTTGCGACAGGGAGGA
GGACATCTTTCCAGCCCGGGATACCTACCATC	AATAATGCGTCGAGGGGGGCATCCTCGGCCAC
CTATGAGCGAG	CGCCCTGGTCCTCA
AGCTGGACATCTTTCCAGCCCGGGATACCTAC	GGAAATAATGCGTCGAGGGGGGCATCCTCGG
CATCCTATGAGCGAGTACCCCACCTACCACAC	CCACCGCCCTGGTCCTCAGGCTTTCCTCGTAC
CCATGGGCGCT	CTGGTGCGTCTGC
	GCGCTCAGCGCTGCTTCTCCCTCTCTTTGTTTG
CCACCTGGGGACAGGATGTCACCTCGGTCCCA	TAGGCTGTGGCCACCGCGGACTCTGCTCCGG
GTCACCAGGCCAGCCCTGGGCTCCACCACCC	AGCACCCCAGG
GTACCCCACCTACCACACCCATGGGCGCTATG	CTCTTGCGACAGGGAGGAAATAATGCGTCGA
TGCCCCTAGCAGTACCGATCGTAGCCCCTAT	GGGGGCATCCTCGGCCACCGCCCTGGTCCTC
GAGAAGAGTGG	AGGCTTTCCTCGT
GGCTACGATCGGTACTGCTAGGGGGGCACATAG	TGCGTCGAGGGGGCATCCTCGGCCACCGCCC
CGCCCATGGGTGTGGTAGGTGGGGTACTCGCT	TGGTCCTCAGGCTTTCCTCGTACCTGGTGCGT
CATAGGATGG	CTGCTGCCACTC
TGTCTTTTCACATTTCAAACCTCCAGTTTAATTC	GAAATAATGCGTCGAGGGGGGCATCCTCGGCC
CTCTCTGGAAGATCCCAGCACCGACTACTACC	ACCGCCCTGGTCCTCAGGCTTTCCTCGTACCT
AAGAGCTGC	GGTGCGTCTGCT
ACCACACCCATGGGCGCTATGTGCCCCCTAGC	AGGGAGGAAATAATGCGTCGAGGGGGGCATC
AGTACCGATCGTAGCCCCTATGAGAAGAGTGG	CTCGGCCACCGCCCTGGTCCTCAGGCTTTCCT
CAG	CGTACCTGGTGCG
CCTACCACACCCATGGGCGCTATGTGCCCCCT	GCCCTCTTGCGACAGGGAGGAAATAATGCGT
AGCAGTACCGATCGTAGCCCCTATGAGAAGAG	CGAGGGGGCATCCTCGGCCACCGCCCTGGTC
TGGCAG	CTCAGGCTTTCCT
CCATCCTATGAGCGAGTACCCCACCTACCACA	CAGTGCCTGGACGCAGCGCCCTCTTGCGACA
CCCATGGGCGCTATGTGCCCCCTAGCAGTACC	GGGAGGAAATAATGCGTCGAGGGGGGCATCC
GATCGTAGCCC	TCGGCCACCGCCCT
ATGAGCGAGTACCCCACCTACCACACCCATGG	GTCGAGGGGGCATCCTCGGCCACCGCCCTGG
GCGCTATGTGCCCCCTAGCAGTACCGATCGTA	TCCTCAGGCTTTCCTCGTACCTGGTGCGTCTG
GCCCCTATGAG	CTGCCACTCTTC
	ATGCGTCGAGGGGGGCATCCTCGGCCACCGCC
ACAATGGCCAGCGCAACCAGAACACAGACCA	CTGGTCCTCAGGCTTTCCTCGTACCTGGTGCG
GCACCAGCAGCGCGATG	TCTGCTGCCACT
GTCTTTTCACATTTCAAACCTCCAGTTTAATTCC	AATAATGCGTCGAGGGGGGCATCCTCGGCCAC
TCTCTGGAAGATCCCAGCACCGACTACTACCA	CGCCCTGGTCCTCAGGCTTTCCTCGTACCTGG
AGAGCTGCA	TGCGTCTGCTGC
ACCATCCTATGAGCGAGTACCCCACCTACCAC	
ACCCATGGGCGCTATGTGCCCCCTAGCAGTAC	GGAGGAAATAATGCGTCGAGGGGGGCATCCT
CGATCGTAGCC	CGGCCACCGCCTGGTCCTCAGGCTTTCCTC
CTACCCAGAGAAGTTCAGTGCCCAGCTCTACT	ATAATGCGTCGAGGGGGGCATCCTCGGCCACC
GAGAAGAATGCTTTTAATTCCTCTCTGGAAGAT	GCCCTGGTCCTCAGGCTTTCCTCGTACCTGGT
CCCAGCACCG	GCGTCTGCTGCC
GACATCTTTCCAGCCCGGGATACCTACCATCCT	GGAAATAATGCGTCGAGGGGGGCATCCTCGG
ATGAGCGAGTACCCCACCTACCACACCCATGG	CCACCGCCCTGGTCCTCAGGCTTTCCTCGTAC
GCGCTATGTG	CTGGTGCGTCTGC
CTCCAGTTTAATTCCTCTCTGGAAGATCCCAGC	GCCTGGACGCAGCGCCCTCTTGCGACAGGGA
ACCGACTACTACCAAGAGCTGCAGAGAGACA	GGAAATAATGCGTCGAGGGGGGCATCCTCGG
TTTCTGAAATG	CCACCGCCCTGGTC

Primers	Sequence 5'-> 3'
TRIM46-MUC1 F1	AAGATCCCAGCACCGACTACT
TRIM46-MUC1 R1	AGGGAGGAAATAATGCGTCGA
MUC1 fwd	GCCCGCTCCACCTCTCAAGCAGCCAGCGCCTGCCTGAATCTGTTCT
KRTCAP2 R9	GTCAGTTTCTCTTGCTCT
E-S-N muc1 fwd new	GGCAAAGAATTCATAATTAACCGCGGGCGGCCGCCCGCTCCACCTC
	TCAAGCAGCCAGCGCCTGCCTGAATCTGTTCT
Bam-Stop-FLAG Muc1 rev	TTGGATCCCGTCATTACTTGTCGTCATCGTCTTT
-	GTAGTCTTGAAGGCAGTGAGCGAGAACA

**Table S5.** MUC1-TRIM46-KRTCAP2 isoforms and RNA fusion junctions. MUC1 sequence is in red, TRIM46 in blue and KRTCAP2 sequence is in green. Lowercase letters represent 5' and 3' UTR sequences. Start and stop codons are shown in Bold. *"\*"* Indicates fusion junction between MUC1 and TRIM46 while "^" indicates fusion junction between TRIM46 and KRTCAP2.

	acgctccacctctcaagcagccagcgcctgcctgaatctgttctgccccctccccaccca
	GCACCCAGTCTCCTTTCTTCCTGCTGCTGCTCCTCACAGTGCTTACAGCTACCACAG
	CCCCTAAACCCGCAACAGTTGTTACGGGTTCTGGTCATGCAAGCTCTACCCCAGGT
	GGAGAAAAGGAGACTTCGGCTACCCAGAGAAGTTCAGTGCCCAGCTCTACTGAGA
	AGAATGCTTTGTCTACTGGGGTCTCTTTTCTTTTTCCTGTCTTTTCACATTTCAAACCTC
	CAGTTTAATTCCTCTCTGGAAGATCCCAGCACCGACTACTACCAAGAGCTGCAGAG
MUC1 TRIMAC	AGACATTTCTGAAATG*AGTGGC
WUCI-INIVI40-	AGCAGACGCACCAGGTACGAGGAAAGCCTGAGGACCAGGGCGGTGGCCGAGGAT
KRICAP2 ISOIOFM I	GCCCCCTCGACGCATTATTTCCTCCCTGTCGCAAGAGGGCGCTGCGTCCAGGCACT
complete CDNA:	GAGAGGAAGAAAGAGGAGAACGCGAGGAGTAGCAAGTCCGCG^TGGTGGGTACG
	GGCACCTCGCTGGCGCTCTCCTCCTCCTGTCCCTGCTGCTCTTTGCTGGGATGCAG
	ATGTACAGCCGTCAGCTGGCCTCCACCGAGTGGCTCACCATCCAGGGCGGCCTGCT
	TGGTTCGGGTCTCTTCGTGTTCTCGCTCACTGCCTTCAATAAtctggagaatcttgtctttggcaaa
	${\it ggattccaagcaaagatcttccctgagattctcctgtgcctcctgttggctctctttgcatctggcctcatccaccgagtctgtgtcacca$
	cctgcttcatcttctccatggttggtctgtactacatcaacaagatctcctccaccctgtaccaggcagcagctccagtcctcacaccagc
	caaggtcacaggcaagagcaagaagaagaaactgaccctgaatgttcaataaagttgattctttgt
	cgctccacctctcaagcagccagcgcctgcctgaatctgttctgccccctccccaccca
	CACCCAGTCTCCTTTCTTCCTGCTGCTGCTCCTCACAGTGCTTACAGCTACCACAGC
	CCCTAAACCCGCAACAGTTGTTACGGGTTCTGGTCATGCAAGCTCTACCCCAGGTG
	GAGAAAAGGAGACTTCGGCTACCCAGAGAAGTTCAGTGCCCAGCTCTACTGAGAA
	GAATGCTTTTAATTCCTCTCTGGAAGATCCCAGCACCGACTACTACCAAGAGCTGC
	AGAGAGACATTTCTGAAATGGCTGTCTGTCAGTGCCGCCGAAAGAACTACGGGCA
	GCTGGACATCTTTCCAGCCCGGGATACCTACCATCCTATGAGCGAGTACCCCACCT
MUC1-TRIM46-	ACCACACCCATGGGCGCTATGTGCCCCCTAGCAGTACCGATCGTAGCCCCTATGAG
KRTCAP2 isoform 2	AAG*AGTGGCAGCAGACGCACCAGGTACGAGGAAAGCCTGAGGACCAGGGCGGT
complete cDNA:	GGCCGAGGATGCCCCCTCGACGCATTATTTCCTCCCTGTCGCAAGAGGGCGCTGCG
	TCCAGGCACTGAGAGGAAGAAAGAGGAGAACGCGAGGAGTAGCAAGTCCGCG^T
	GGTGGGTACGGGCACCTCGCTGGCGCTCTCCTCCTCCTGTCCCTGCTGCTCTTTGC
	TGGGATGCAGATGTACAGCCGTCAGCTGGCCTCCACCGAGTGGCTCACCATCCAG
	GGCGGCCTGCTTGGTTCGGGTCTCTTCGTGTTCTCGCTCACTGCCTTCAATAAtctggag
	aatcttgtctttggcaaaggattccaagcaaagatcttccctgagattctcctgtgcctcctgttggctctctttgcatctggcctcatccaggattccaggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcatccaggattctggcctcaggattctggcctcaggattctggcctcatccaggattctggcctcatccaggattctggcctcaggattctggcctcaggattctggcctcaggattctggcctcaggattctggcctcaggattctggcctcaggattctggcctcaggattctggcctcaggattctggcctcaggattctggcctcaggattctggcctcaggattctggcctcctgggcctcaggattctggcctcaggattctggcctcaggattctggcctcaggattctggcctcctgggcctcaggattctggcctcaggattctgggcctcaggattctggcctcctggggcccccaggattctgggcctccaggattctgggcctcaggattctgggcctcaggattctgggcctcaggattctgggcctccctgggggtggccccgggattctggggccccgggattctgggccggattctgggggattctgggggggg
	ccgagtctgtgtcaccacctgcttcatcttctccatggttggt
	ccagtcctcacaccagccaaggtcacaggcaagagcaagaagaagaaactgaccctgaatgttcaataaagttgattctttgt

<b>I ubic</b> 00. Com.	Tab	le S5	6. Cont.
------------------------	-----	-------	----------

	cgctccacctctcaagcagccagcgcctgcctgaatctgttctgccccctccccaccca
	CACCCAGTCTCCTTTCTTCCTGCTGCTGCTCCTCACAGTGCTTACAGCTACCACAGC
	CCCTAAACCCGCAACAGTTGTTACGGGTTCTGGTCATGCAAGCTCTACCCCAGGTG
	GAGAAAAGGAGACTTCGGCTACCCAGAGAAGTTCAGTGCCCAGCTCTACTGAGAA
	GAATGCTTTTAATTCCTCTCTGGAAGATCCCAGCACCGACTACTACCAAGAGCTGC
	AGAGAGAGACATTTCTGAAATGTCTGGGGGCTGGGGTGCCAGGCTGGGGCATCGCGCT
	CCTCCTCCTCCTCTCTCTCCCTCCCCCCCCCCCCCCCCC
	GTCTGTCAGTGCCGCCGAAAGAACTACGGGCAGCTGGACATCTTTCCAGCCCGGG
MUC1 TRIMAG	
VPTC AP2 isoform 2	
KRICAP2 ISOIOIIII 3	
complete CDINA:	
	GCGCICICCICCIGICCCIGCIGCICITIGCIGGGAIGCAGAIGIACAGCCGI
	CAGCTGGCCTCCACCGAGTGGCTCACCATCCAGGGCGGCCTGCTTGGTTCGGGTCT
	CTTCGTGTTCTCGCTCACTGCCTTCAA <b>TAA</b> tctggagaatcttgtctttggcaaaggattccaagcaaagatc
	tt ccct gag att ctcct gt gcct cct gt g gct ctctt t g catct g g cct catcca c c g g g t c g g g c c c a c c g g g c c c c c g g g c c c c
	ttggtctgtactacatcaacaagatctcctccaccctgtaccaggcagcagctccagtcctcacaccagccaaggtcacaggcaaga
	gcaagaagagaaactgaccctgaatgttcaataaagttgattctttgt
	Gcgcctgcctgaatctgttctgccccctccccaccatttcaccaccatgacaccgggcacccagtctcctttcttcctgctgctgct
	cct cac agt gct tac agg tg agg gg cac gag gt gg gg gg gg gg gg cct gcct g
	ccctggcagatggcaccatgaagttaagctaccacagcccctaaacccgcaacagttgttacgggttctggtcatgcaagctctacccccatgaagttaagctaccacagcccctaaacccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctaccccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctaccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctaccccgcaacagttgttacgggttctggtcatgcaagctctacccccgcaacagttgttacgggttctggtcatgcaagctctaccccgcaacagttgttacgggttctggtcatgcaagctctaccccgcaacagttgttacgggttctggtcatgcaagctctaccccgcaagctgtgttacgggttctggtcatgcaagctctacccgcaacagttgttacgggttctggtcatgcaagctctaccccgcaagctctaccccgcaacagttgttacgggttctgggttctggtcatgcaagctctaccccgcagcqccccgaacagttgttacgggttctggtcatgcaagctctacccgcagcqccccqaagctgtgtacgggttctggtcatgcaagctctacccqcagcqcccqaagctgtgtacgggttctggtcatgcaagctctacccgcagcqcqcagagcqcqcagagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
	caggtggagaaaaggagacttcggctacccagagaagttcagtgcccagctctactgagaagaatgcttttaattcctctctggaagaagaatgcttttaattcctctctggaagaagaagaagaagaagaagaagaagaagaagaaga
	$atcccagcaccgactactaccaagagctgcagagagacatttctgaa { {                                $
	ggggctggggtgccaggctggggcatcgcgctgctggtgctggtctgtgttctggttgcgctggccattgtctatctcattgccttgg
MUC1-TRIM46-	ctgtctgtcagtgccgccgaaagaactacgggcagctggacatctttccagcccgggatacctacc
KRTCAP2 isoform 4	cctaccacccatgggcgctatgtgccccctagcagtaccgatcgtagcccctatgagaag*agtggcagcagcagcagcagcaggta
complete cDNA:	cgaggaaagcctgaggaccagggcggtggccgaggatgccccctcgacgcattatttcctccctgtcgcaagagggcgctgcgtc
1	
	ctccctcctgtccctgctgctctttgctgggatgcagatgtacagccgtcagctggcctccaccgagtgggtcaccatccaggggggg
	ctocttoottootottcotottctoctcactoccttcaataatctooaaaatcttotctttoocaaagaattccaagcaaagattt
	ccdragattetectotocetectottooctectttocatetooceteatecaceoaotetototocacetoetteatettetecatoott
	ggtetgtactacateaacaagateteetectgaccagacagacagactecagtectcacagacaagateacaggacaagag
	aggaggaggaggaggtgactoraccologaatgttcaatgaggttgattctttgt
	ataattaacgacgaggaggaggaggaggaggaggaggaggaggagg
<i>MUC1-TRIM46-</i> <i>KRTCAP2</i> isoform 5	
complete cDNA:	ATTICCTCCCIGTCGCAAGAGGGCGCTGCGTCCAGGCACTGAGAGGAAGAAAGA
compicie eDivis.	GAGAACGCGAGGAGTAGCAAGTCCGCG^1GGTGGGTACGGGCACCTCGCTGGCGC
	TCTCCTCCTCCTGTCCCTGCTGCTCTTTGCTGGGATGCAGATGTACAGCCGTCAGC
	TGGCCTCCACCGAGTGGCTCACCATCCAGGGCGGCCTGCTTGGTTCGGGTCTCTTCG
	$TGTTCTCGCTCACTGCCTTCAA {\bf TAA} tctggagaatcttgtctttggcaaaggattccaagcaaagatcttccctga$
	gattetectgtgcetectgttggetetetttgcatetggeeteatccacegagtetgtgtcaceacetgetteatettetecatggttggtetggte
	tacta cat caa caa gat ctcct ccaccct gt ac cag g cag ctc cag t cct cac a c cag g c caa g g c a a gag caa g a ga a g
	gagaaactgaccctgaatgttcaataaagttgattctttgt

## Table S5. Cont.

	Cgetecaceteteaageagecagegeetgeatetgttetgeeeeecececececaceaecaceatgacacegggeaece
	agt ctcctttcttcctgctgctgctcctcacagtgcttacagctaccacagcccctaaacccgcaacagttgttacgggttctggtcatgcatg
	a a get ctacccc a g g g g g a a a g g g g a ctt c g g ctaccc a g g g a g g t c g g g a g g a g g a g g g g g g g g g
	$caccgactactaccaagagctgcagagagacatttctgaa { {                                $
	TTTCTGGGCCTCTCCAATATTAAGTTCAGGCCAGGATCTGTGGTGGTACAATTGACT
	CTGGCCTTCCGAGAAGGTACCATCAATGTCCACGACGTGGAGACACAGTTCAATC
	AGTATAAAACGGAAGCAGCCTCTCGATATAACCTGACGATCTCAGACGTCAGCGT
	GAGTGATGTGCCATTTCCTTTCTCTGCCCAGTCTGGGGCTGGGGTGCCAGGCTGGG
	GCATCGCGCTGCTGGTGCTGGTCTGTGTTCTGGTTGCGCTGGCCATTGTCTATCTCAT
MUC1-TRIM46-	TGCCTTGGCTGTCTGTCAGTGCCGCCGAAAGAACTACGGGCAGCTGGACATCTTTC
KRTCAP2 isoform 6	CAGCCCGGGATACCTACCATCCTATGAGCGAGTACCCCACCTACCACACCCATGG
complete cDNA:	GCGCTATGTGCCCCCTAGCAGTACCGATCGTAGCCCCTATGAGAAG*AGTGGCAGC
	AGACGCACCAGGTACGAGGAAAGCCTGAGGACCAGGGCGGTGGCCGAGGATGCC
	CCCTCGACGCATTATTTCCTCCCTGTCGCAAGAGGGCGCTGCGTCCAGGCACTGAG
	AGGAAGAAAGAGGAGAACGCGAGGAGTAGCAAGTCCGCG^TGGTGGGTACGGGC
	ACCTCGCTGGCGCTCTCCTCCTGTCCCTGCTGCTCTTTGCTGGGATGCAGATG
	TACAGCCGTCAGCTGGCCTCCACCGAGTGGCTCACCATCCAGGGCGGCCTGCTTGG
	TTCGGGTCTCTTCGTGTTCTCGCTCACTGCCTTCAATAAtctggagaatcttgtctttggcaaaggatt
	ccaagcaaagatcttccctgagattctcctgtgcctcctgttggctctctttgcatctggcctcatccaccgagtctgtgtcaccacctgct
	tcatcttctccatggttggtctgtactacatcaacaagatctcctccaccctgtaccaggcagcagcagctccagtcctcacaccaggcagg
	tcacaggcaagagcaagagagaaactgaccctgaatgttcaataaagttgattctttgt

**Table S6.** Frequency of occurrence of different MUC1-TRIM46-KRTCAP2 isoforms determined using in-house HGSC samples. RT-PCR data from Figure 3A was utilized for densitometric analysis and then the samples positive for each isoform were determined.

MUC1-KRTCAP2	Frequency of Occurrence
Isoform	among 59 HGSC Samples
Isoform 1	64%
Isoform 2	47%
Isoform 3	17%
Isoform 4	34%
Isoform 5	34%
Isoform 6	20%

For Isoform 1, the following HGSC samples are considered positive: S4,5,7,11,12,14,15,16,20,21,22,23, 46,49,50,51,52,53,54,57,58,60,61,62,63,64,65,66,67,70,71,72,73,74; For Isoform 2, the following samples are considered positive: S1,2,3,4,5,7,9,11,12,15,23,49,53,54,55,61,62,63,64,65,66,67,71,73,74,75,78,82; For Isoform 3, the following samples are considered positive: S46,49,53,55,61,62,63,71,73,75; For Isoform 4, the following samples are considered positive: S1,3,4,5,7,12,15,20,23,46,49,52,55,57,61,62,72,73,75,82; For Isoform 5, the following samples are considered positive: S1,3,4,5,7,12,15,20,23,46,49,52,55,57,61, 62,72,73,75,82; For Isoform 6, the following samples are considered positive: S1,3,4,5,7,12,15,20,23,46,49,52,55,57,61, 62,72,73,75,82.

**Table S7.** MUC1-TRIM46-KRTCAP2 isoforms protein sequence. MUC1 sequence is in red, TRIM46 in blue and KRTCAP2 sequence is in green. *"\*"* Indicates fusion junction between MUC1 and TRIM46 while "^" indicates fusion junction between TRIM46 and KRTCAP2.

	MTPGTQSPFFLLLLLTVLTATTAPKPATVVTGSGHASSTPGGEKETSATQRSSVPSSTEKNAL
Isoform 1: 210 aa	STGVSFFFLSFHISNLQFNSSLEDPSTDYYQELQRDISEM*SGSRRTRYEESLRTRAVAEDAPST
	HYFLPVARGRCVQALRGRKRRTRGVASPR^GGYGHLAGALLPPVPAALCWDADVQPSAG
	LHRVAHHPGRPAWFGSLRVLAHCLQ
	MTPGTQSPFFLLLLLTVLTATTAPKPATVVTGSGHASSTPGGEKETSATQRSSVPSSTEKNAF
Isoform 2: 242 aa	NSSLEDPSTDYYQELQRDISEMAVCQCRRKNYGQLDIFPARDTYHPMSEYPTYHTHGRYVP
150101111 2. 2 <del>1</del> 2 dd	PSSTDRSPYEK*SGSRRTRYEESLRTRAVAEDAPSTHYFLPVARGRCVQALRGRKRRTRGVAS
	PR^GGYGHLAGALLPPVPAALCWDADVQPSAGLHRVAHHPGRPAWFGSLRVLAHCLQ
	MTPGTQSPFFLLLLLTVLTATTAPKPATVVTGSGHASSTPGGEKETSATQRSSVPSSTEKNAF
	NSSLEDPSTDYYQELQRDISEMSGAGVPGWGIALLVLVCVLVALAIVYLIALAVCQCRRKNY
Isoform 3: 272 aa	GQLDIFPARDTYHPMSEYPTYHTHGRYVPPSSTDRSPYEK*SGSRRTRYEESLRTRAVAEDAP
	STHYFLPVARGRCVQALRGRKRRTRGVASPR^GGYGHLAGALLPPVPAALCWDADVQPSA
	GLHRVAHHPGRPAWFGSLRVLAHCLQ
Isoform 4: It is	
predicted to yield no	
protein since the	
annotated start codon is	
immediately followed	
by a stop codon.	
Isoform 5: 210 aa	YGQLDIFPAKDTYHPMSEYPTYHTHGKYVPPSSTDKSPYEK*SGSKKTKYEESLKTKAVAEDA
	P51HYFLPVARGKCVQALKGKKKKIKGVASPK*GGYGHLAGALLPPVPAALCWDADVQPS
Isaharma (+ 262 aa	
150101111 0: 203 aa	
	HITGREAWFG5LKVLAHCLQ
	AC
	28 . KV
	CPT 64
	$lsoform$ 1 2 3 5 6 $(N_N)^2$
	70 kD

**Figure S1.** MUC1-TRIM46-KRTCAP2 isoforms are not secreted into the cellular media. MUC1-TRIM46-KRTCAP2 cDNA constructs with FLAG tag were transfected into OVCAR8. 48 hours after transfection, the cellular media was collected and was analyzed by western blotting with a FLAG antibody. None of the isoforms are detectable in the cellular media. WDFY2-FLAG is used as a positive control for the western blot.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under



the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).

50 kD 37 kD

25 kD