# Supplementary Data

## The Kub5-Hera/RPRD1B interactome: A novel role in preserving genetic stability by regulating DNA mismatch repair

Praveen L. Patidar[1], Edward A. Motea[1], Farjana J. Fattah[1], Yunyun Zhou[2], Julio C. Morales[1] Yang Xie[2], Harold R. Garner[3] and David A. Boothman[1*]

[1]Departments of Pharmacology and Radiation Oncology, Program in Cell Stress and Cancer Nanomedicine, [2]Quantitative Biomedical Center, Department of Clinical Science, Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, ND2.210K, Dallas, TX 75390-8807, USA; [3]College of Osteopathic Medicine and the MITTE Office, Virginia Tech, Blacksburg, VA

*Correspondence: *David.Boothman@UTSouthwestern.edu*; Phone: +1 214 645 6371; Fax: +1 214 645 6347

### Material and Methods

***Cloning, expression and purification of His-tagged K-H-*** K-H protein was purified in our laboratory using a comparable strategy as described previously with indicated modifications (1). Human *k-h* gene cDNA was cloned into bacterial expression vector pET28a vector expressing 6X His-tagged K-H. Correct plasmid DNA sequences were confirmed by DNA sequencing. BL21 bacterial cells were transformed to express His-K-H. Early log phase cells were induced with 5 mM IPTG to induce expression of His-K-H and induction was continued for 3 h at 37$^{o}$C. Cells were harvested, flash frozen in liquid nitrogen and stored at -20$^{o}$C until further use. Then, cells were lysed in binding buffer (50 mM Tris-HCl [pH 7.5], 500 mM NaCl, 5 mM imidazole, 10% glycerol, 1X protease inhibitor cocktail, 1 mM DTT, 10 units/ml of DNase I and final concentration of 0.5 µg/mL lysozyme. Cell debris was removed by centrifugation and supernatant was incubated with Ni-NTA beads. Beads were washed three times with (50 mM Tris-HCl [pH 7.5], 500 mM NaCl, 30 mM imidazole, 5% glycerol, 1 mM protease inhibitor). Then, proteins were eluted with elution buffer (50 mM Tris-HCl [pH

7.5], 500 mM NaCl, 300 mM imidazole, 1% NP-40, 1X protease inhibitor cocktail) with gentle rotation at 4$^{o}$C for 30 min. Proteins were subjected to another round of purification using Superdex 200 gel-filtration chromatography. Purified protein samples were dialyzed to remove traces of imidazole using dialysis buffer (20 mM Tris-HCl [pH 7.5], 150 mM NaCl, 10% glycerol). Protein concentrations were measured using BCA assay and proteins were flash frozen in liquid nitrogen and stored at -80$^{o}$C.

**Legend**

TABLE S1. **List of proteins identified by mass spectrometric analysis of TAP-K-H/RPRD1B pull-down.** Proteins with distinguished peptide identity, peptide sequences ≥5, PSM ≥5, % coverage ≥5, either exclusively present in the TAP-K-H fraction or enriched in TAP-K-H pull-downs (TAP-K-H/TAP ratio >1.0) are shown. This list represents known, as well as novel, protein binding partners of K-H. Proteins are represented by their UniProt accession ID, along with protein description. Also included are PSMs (number of spectra assigned to peptides that contributed to inference of a protein), peptide sequences (number of different unique peptide sequences, or modified variants of sequences identified for the protein), % sequence coverage (percentage of protein sequence covered by peptides identified for a specific protein), spectral index (MIC Sin) and enrichment ratio (TAP-K-H/TAP ratio).

TABLE S2. **Bioinformatics analyses using DAVID for functional annotation of proteins associated with K-H/RPRD1B.** Functional categories determined by DAVID v6.7 for proteins identified in TAP-MS analyses of K-H. Functional categories with –log10(FDR) cutoff value of >5 considered significant and *p*-values for these categories are included.

TABLE S3. **List of select proteins associated with K-H/RPRD1B and analyzed by STRING.** Select proteins from **Table S1** involved in RNA metabolism, DNA repair and replication were used as input in STRING analysis to create interactome of K-H.

FIGURE S1. **Tandem affinity purification of K-H/RPRD1B.** ***A.*** Schematic representation of the N-terminal TAP tag fused to K-H protein. This method requires fusion of a specific tag to the N- or C-terminus of the protein of interest, and introduction of this construct into host cells. The TAP tag consists of two immunoglobin G (IgG) binding domains of Protein A (ProtA) from *Staphylococcus aureus* and a Calmodulin binding peptide (CBP) separated via a linker that is cleavable by TEV protease. ***B.*** General outline of the TAP purification strategy to identify the binding partners of K-H. A large-scale culture of 293T cells stably expressing either TAP-empty vector (as negative control) or TAP-K-H was prepared. Cells were resuspended in NP-40 buffer and lysed by repeated freeze-thaw cycles followed by sonication to release the soluble proteins. Cell

debris was removed by centrifugation and soluble fraction was subjected to two sequential steps of affinity purification. A complex mixture of purified TAP or TAP-K-H along with its potential binding partners was examined by mass spectrometric analyses. ***C.*** A 4-20% gradient SDS-PAGE for purified TAP and TAP-K-H proteins stained with coomassie. Lane 1 represents the protein marker with indicated molecular weight in kDa. Lanes 2 & 3 represent purified TAP and TAP-K-H along with other associated proteins, respectively.

FIGURE S2. **Validation of novel K-H/RPRD1B binding partners.** ***A.*** Whole cell lysates prepared from 293T cells knockdown for endogenous K-H using 3'-UTR sik-h and transiently expressing mycK-H were used for co-IP analyses. Similar co-IP procedure as described in "Materials and Method" was used except cell lysate were treated with 100 units of DNase I for 30 min prior to immunoprecipitation. Antibodies against myc-tag or normal mouse IgG were then used for pull-down proteins and Western blot analyses were performed to detect specific proteins using respective antibodies. Lane 1 represents input for IP and lanes 2 & 3 represent IP by IgG and anti-myc antibodies, respectively. ***B.*** Empty vector control co-IP of 293T cells transfected with siScr (to keep endogenous untagged K-H protein level intact) and pCMV-myc empty vector followed by pull-down with myc-antibodies. Western blot were done using anti-myc (for mycEV) or specific antibody against indicated proteins. Lane 1 represents input for IP and lanes 2 & 3 represent IP by IgG and anti-myc antibodies, respectively.

FIGURE S3. **K-H/RPRD1B depletion results in defective $G_2$ cell cycle checkpoint arrest responses, similar to MMR-deficient cells in response to 6-TG exposure.** Stable shScr or shk-h 231 cells were exposed to 0.5 μM 6-TG for 24, 48 or 72 h and analyzed by flow cytometry for changes in cell cycle distributions. Quantification of cell cycle phase distributions of shScr (***A***) or shk-h (***B***) 231 cells after mock (open bars) or 6-TG exposure (shaded bars) are shown for cells treated for 24, 48 or 72 h as indicated. Shown are means, ±SD for experiments performed at-least three times and *p*-values determined using two-tailed student's t-tests. ***, $p \leq 0.005$; *, *p*-value < *0.05* and ns, not significant.

FIGURE S4: **Hypothetical model showing how K-H/RPRD1B loss may lead to the 'leaky' regulation of *Apaf1*, and subsequence increase in basal level activation of caspases-3 that, in turn, mediate MLH1-PMS2 degradation.**
***A-B,*** **K-H physically interacts with Ku.** Recombinant Ku70-Ku86 (Ku) complex was a generous gift from David Chen laboratory at UTSW (2,3) and K-H proteins was purified in our laboratory. SDS-PAGE and Western blot analyses are presented. Lane 1 and 2 represent protein marker (in kDa) and purified proteins, respectively. Co-IP experiments were then conducted using these proteins and K-H polyclonal antibodies. All pull-down studies were performed as described in 'Materials and Methods' except that purified proteins (Ku and K-H, 5 μg each) were used instead of cell lysates. Although some background Ku and K-H was noted with IgG pull-downs but the co-IP of Ku with K-H antibody was considerably enriched **(~2-4 fold**) over the control. These data support a direct

interaction between Ku and K-H, and are consistent with prior evidence showing that K-H and Ku form higher order complexes (this study (**Fig. 1)** and **(4)**) that might regulate Ku function, such as regulating basal level Apaf1 activities. ***C-D,*** **K-H loss stimulates a 'leaky' *Apaf1* expression-caspase-3 activation-MLH1/PMS2 degradation.** Since evidence in **Fig. S4B** supports a direct interaction with Ku, and other evidence also supports possible interaction with Ku (heterodimer of Ku86/Ku70) (**Fig.1 and (4)**), we favor a hypothesis whereby K-H functions to increase the efficiency of Ku complex binding to a specific regulatory region (gray box) within the *Apaf1* promoter (white box). Prior data demonstrated that Ku binding within this region can inhibit *Apaf1* promoter activity and Apaf1 protein expression, resulting in suppression of caspase-3 activation (5). Data presented here are consistent with the hypothesis that the direct interaction of K-H and Ku (**B**) is involved in the formation of a K-H-Ku complex that actively suppresses *Apaf1* expression, and works to prevent the activation of caspase-3 (**C**). Upon K-H loss (due to one copy number loss or expression knockdown), leaky promoter activity and expression of *Apaf1* constitutively activates caspase-3 resulting in MLH1 and subsequently PMS2 protein degradation (**D**). MLH1 harbors a caspase-3 proteolytic cleavage site (6). Degradation of MLH1 leads to concomitant loss of PMS2 (**Fig. 5**) resulting in a MMR defect. The MMR repair defect presented as 6-TG damage tolerance (**Fig. 2),** defective $G_2$ cell cycle checkpoint responses (**Fig. 3**), and an overall MSI+ mutator genotype (**Fig. 4**). z-VAD, a pan-caspase inhibitor, prevents caspase-3 activity and restores MLH1-PMS2 expression in K-H knockdown cells (**Fig. 5**). Further work is necessary to provide proof of this proposed hypothesis.

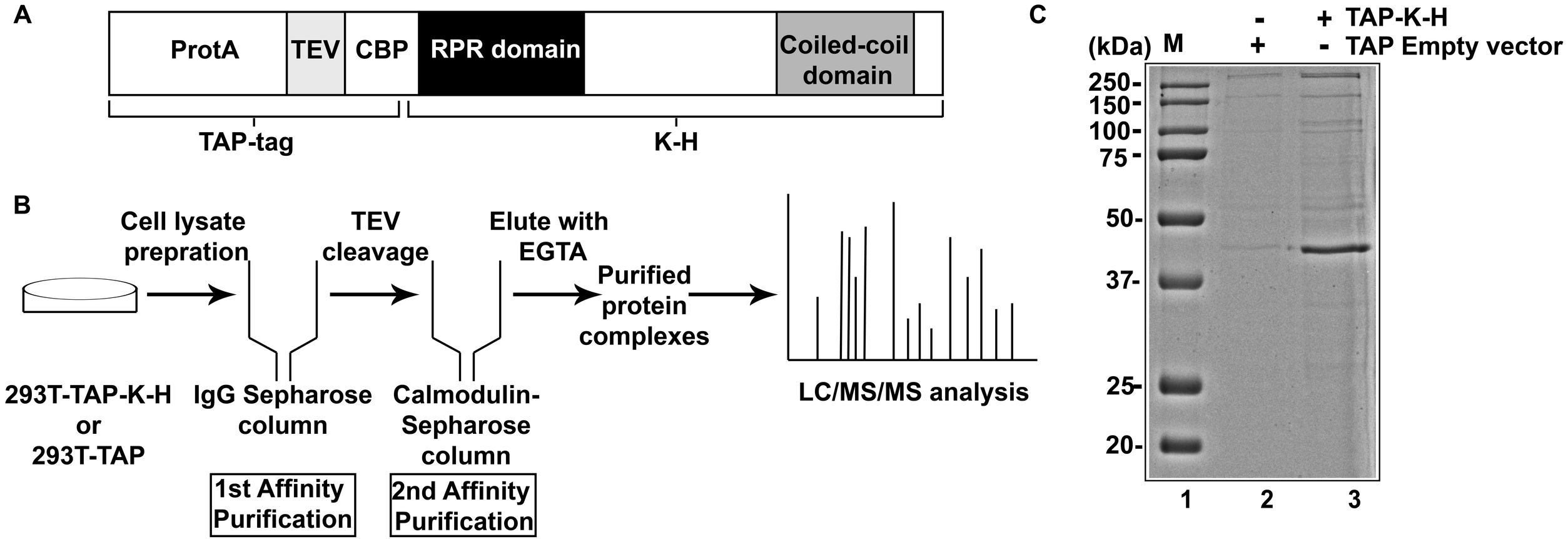
FIGURE S5. **Colon cancer tissue samples displaying changes in mRNA levels and copy number for *k-h* (*Rprd1b*) gene.** Data deposited in Oncomine depository by several individual studies was extracted to analyze changes in copy number (7-9) and mRNA ((10) and TCGA) of *k-h* gene. ***A-C.*** The changes in *k-h* gene copy number in colon/colorectal carcinoma. ***D-E.*** The mRNA level of *k-h* gene in normal colon (left) and colon/colorectal carcinoma (right). ***, *p*-value < *0.005;* *, *p*-value < *0.05*.

## References

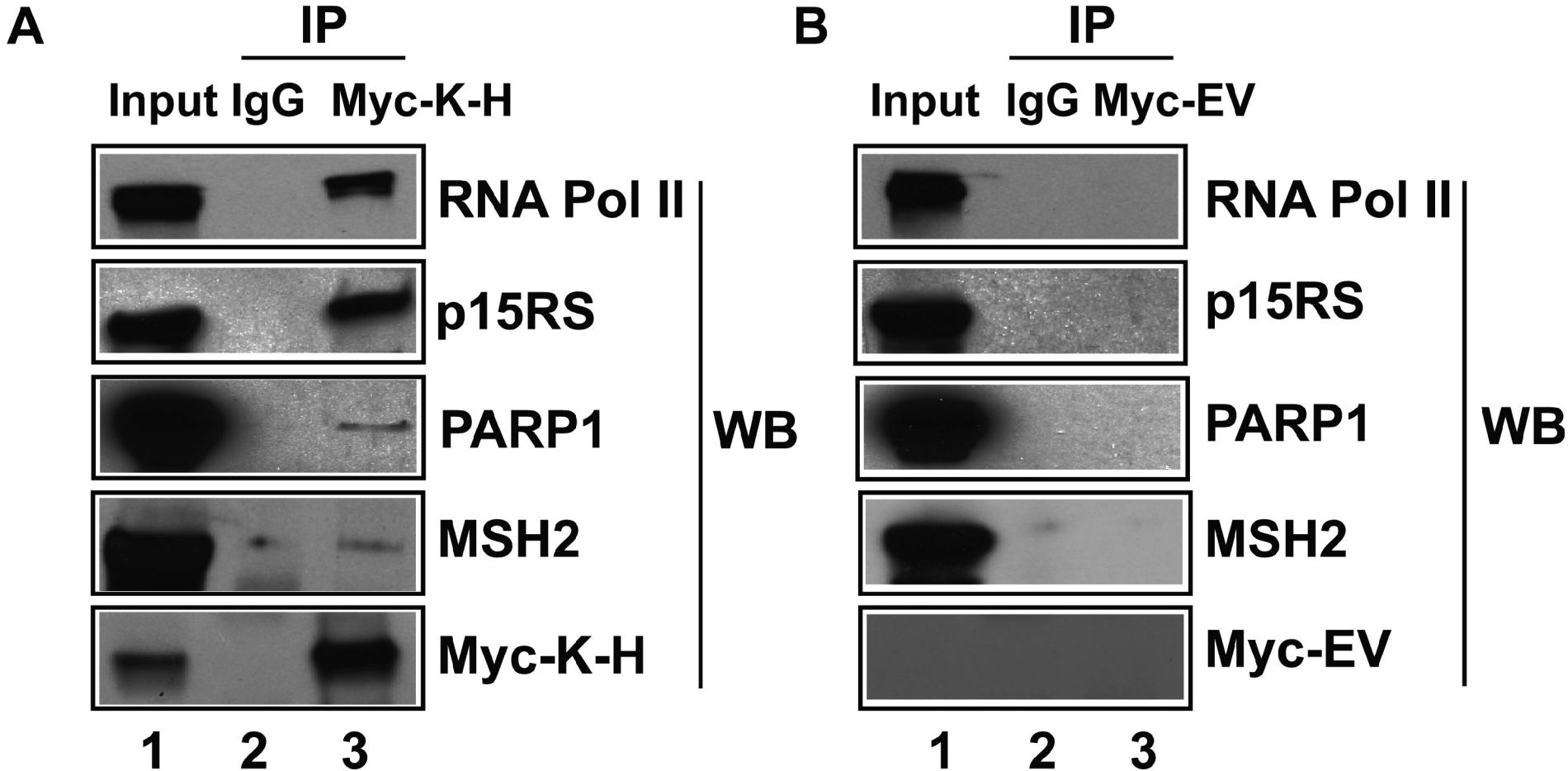

1. Ni, Z., Xu, C., Guo, X., Hunter, G.O., Kuznetsova, O.V., Tempel, W., Marcon, E., Zhong, G., Guo, H., Kuo, W.H. *et al.* (2014) RPRD1A and RPRD1B are human RNA polymerase II C-terminal domain scaffolds for Ser5 dephosphorylation. *Nature structural & molecular biology*, **21**, 686-695.
2. Cary, R.B., Peterson, S.R., Wang, J., Bear, D.G., Bradbury, E.M. and Chen, D.J. (1997) DNA looping by Ku and the DNA-dependent protein

kinase. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 4267-4272.
3. Hsu, H.L., Gilley, D., Galande, S.A., Hande, M.P., Allen, B., Kim, S.H., Li, G.C., Campisi, J., Kohwi-Shigematsu, T. and Chen, D.J. (2000) Ku acts in a unique way at the mammalian telomere to prevent end joining. *Genes & development*, **14**, 2807-2812.
4. Morales, J.C., Richard, P., Rommel, A., Fattah, F.J., Motea, E.A., Patidar, P.L., Xiao, L., Leskov, K., Wu, S.Y., Hittelman, W.N. *et al.* (2014) Kub5-Hera, the human Rtt103 homolog, plays dual functional roles in transcription termination and DNA repair. *Nucleic acids research*, **42**, 4996-5006.
5. De Zio, D., Bordi, M., Tino, E., Lanzuolo, C., Ferraro, E., Mora, E., Ciccosanti, F., Fimia, G.M., Orlando, V. and Cecconi, F. (2011) The DNA repair complex Ku70/86 modulates Apaf1 expression upon DNA damage. *Cell death and differentiation*, **18**, 516-527.
6. Chen, F., Arseven, O.K. and Cryns, V.L. (2004) Proteolysis of the mismatch repair protein MLH1 by caspase-3 promotes DNA damage-induced apoptosis. *The Journal of biological chemistry*, **279**, 27542-27548.
7. Reid, J.F., Gariboldi, M., Sokolova, V., Capobianco, P., Lampis, A., Perrone, F., Signoroni, S., Costa, A., Leo, E., Pilotti, S. *et al.* (2009) Integrative approach for prioritizing cancer genes in sporadic colon cancer. *Genes, chromosomes & cancer*, **48**, 953-962.
8. Firestein, R., Bass, A.J., Kim, S.Y., Dunn, I.F., Silver, S.J., Guney, I., Freed, E., Ligon, A.H., Vena, N., Ogino, S. *et al.* (2008) CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity. *Nature*, **455**, 547-551.
9. Lu, X., Zhang, K., Van Sant, C., Coon, J. and Semizarov, D. (2010) An algorithm for classifying tumors based on genomic aberrations and selecting representative tumor models. *BMC medical genomics*, **3**, 23.
10. Zou, T.T., Selaru, F.M., Xu, Y., Shustova, V., Yin, J., Mori, Y., Shibata, D., Sato, F., Wang, S., Olaru, A. *et al.* (2002) Application of cDNA microarrays to generate a molecular taxonomy capable of distinguishing between colon cancer and normal colon. *Oncogene*, **21**, 4855-4862.

# Figure S1


A
ProtA
TEV
CBP
RPR domain
Coiled-coil domain
TAP-tag
K-H
B
Cell lysate prepration
TEV cleavage
Elute with EGTA
Purified protein complexes
293T-TAP-K-H or 293T-TAP
IgG Sepharose column
Calmodulin-Sepharose column
1st Affinity Purification
2nd Affinity Purification
LC/MS/MS analysis
C
- + TAP-K-H
(kDa) M + - TAP Empty vector
250-
150-
100-
75-
50-
37-
25-
20-
1
2
3

# Figure S2


A
IP
Input IgG Myc-K-H
RNA Pol II
p15RS
PARP1
MSH2
Myc-K-H
WB
1 2 3
B
IP
Input IgG Myc-EV
RNA Pol II
p15RS
PARP1
MSH2
Myc-EV
WB
1 2 3

# Figure S3

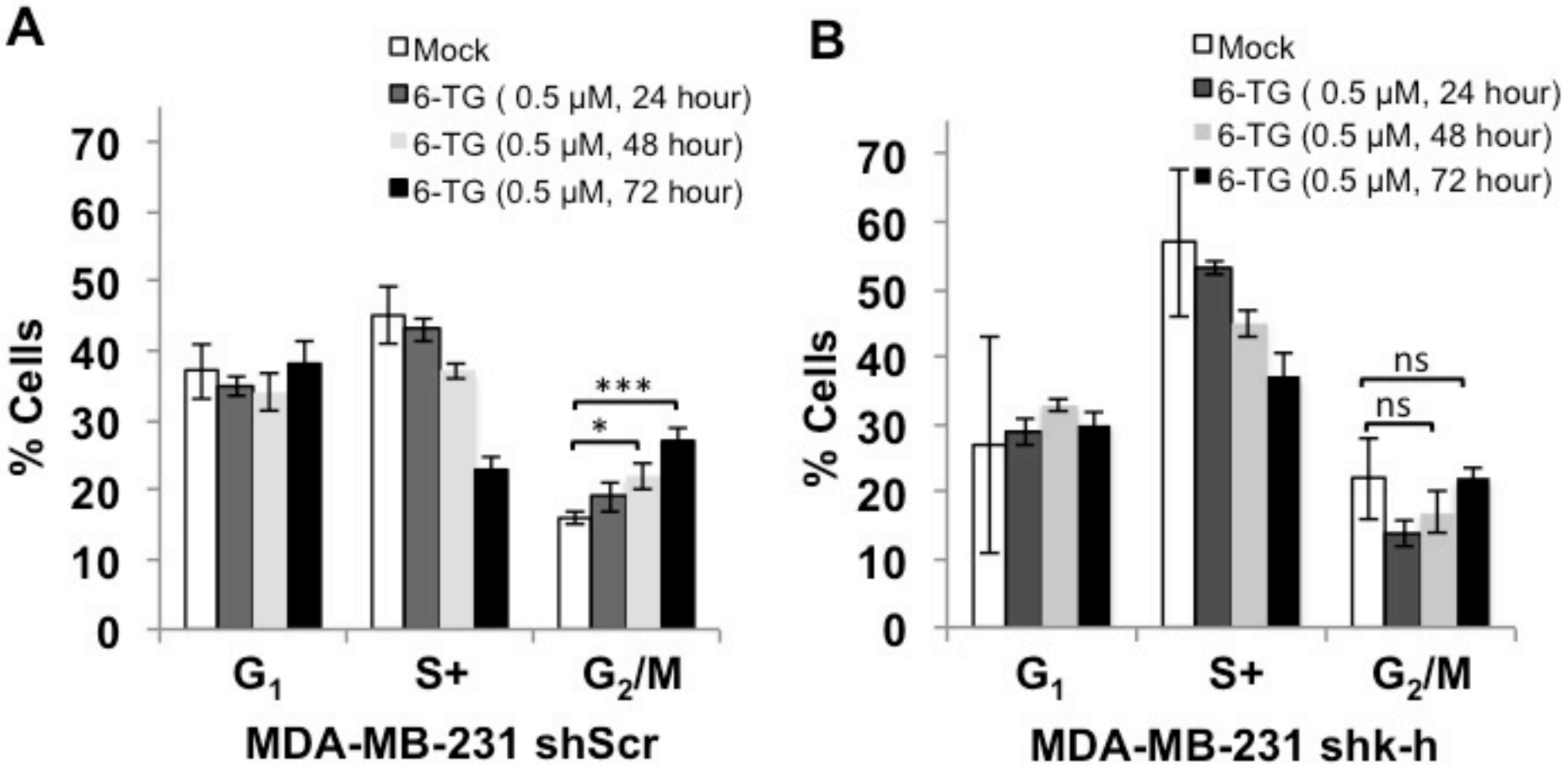

A
Mock
6-TG ( 0.5 μM, 24 hour)
6-TG (0.5 μM, 48 hour)
6-TG (0.5 μM, 72 hour)
% Cells
70
60
50
40
30
20
10
0
***
*
G1
S+
G2/M
MDA-MB-231 shScr
B
Mock
6-TG ( 0.5 μM, 24 hour)
6-TG (0.5 μM, 48 hour)
6-TG (0.5 μM, 72 hour)
% Cells
ns
ns
G1
S+
G2/M
MDA-MB-231 shk-h

Figure S4

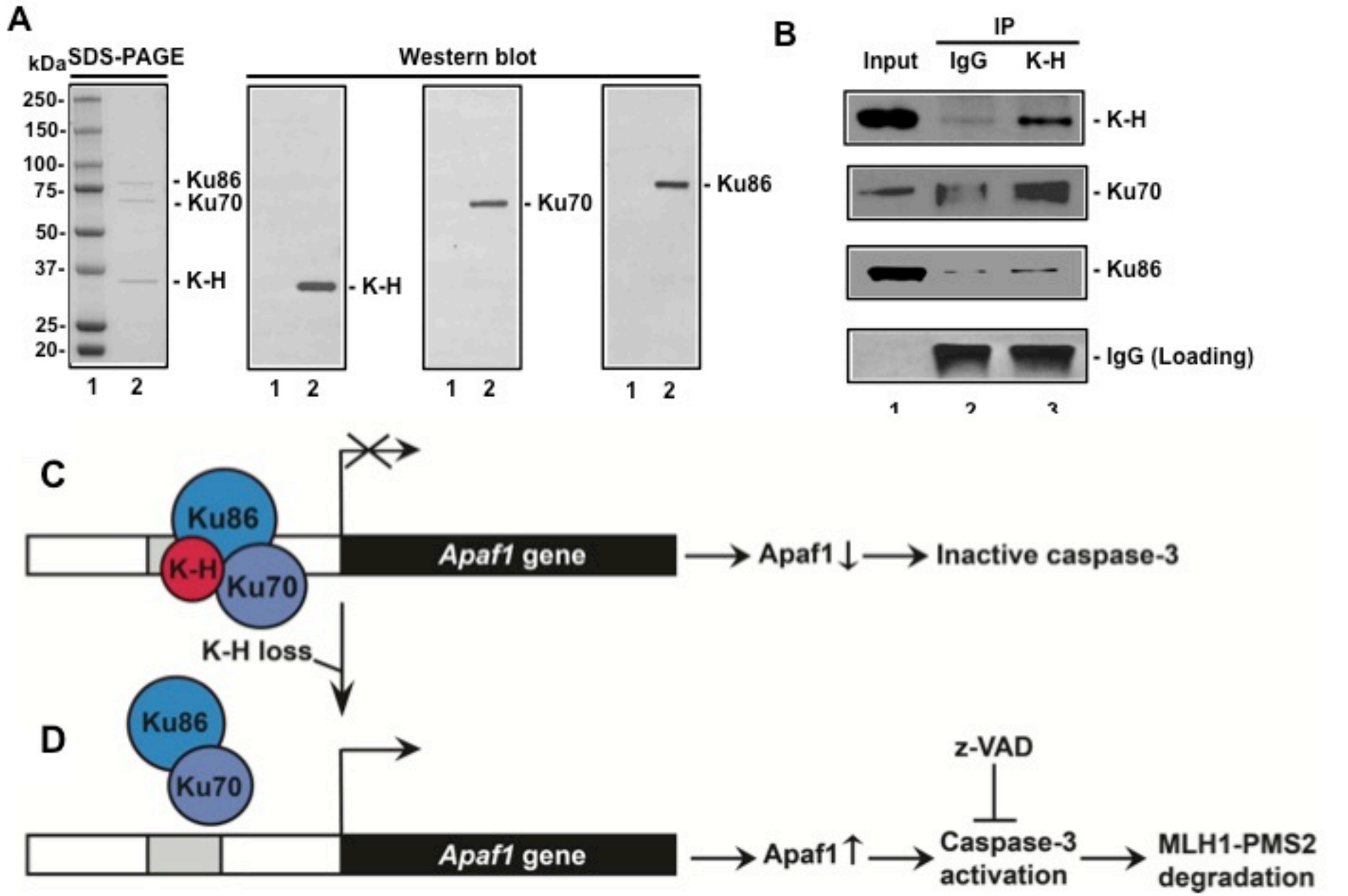
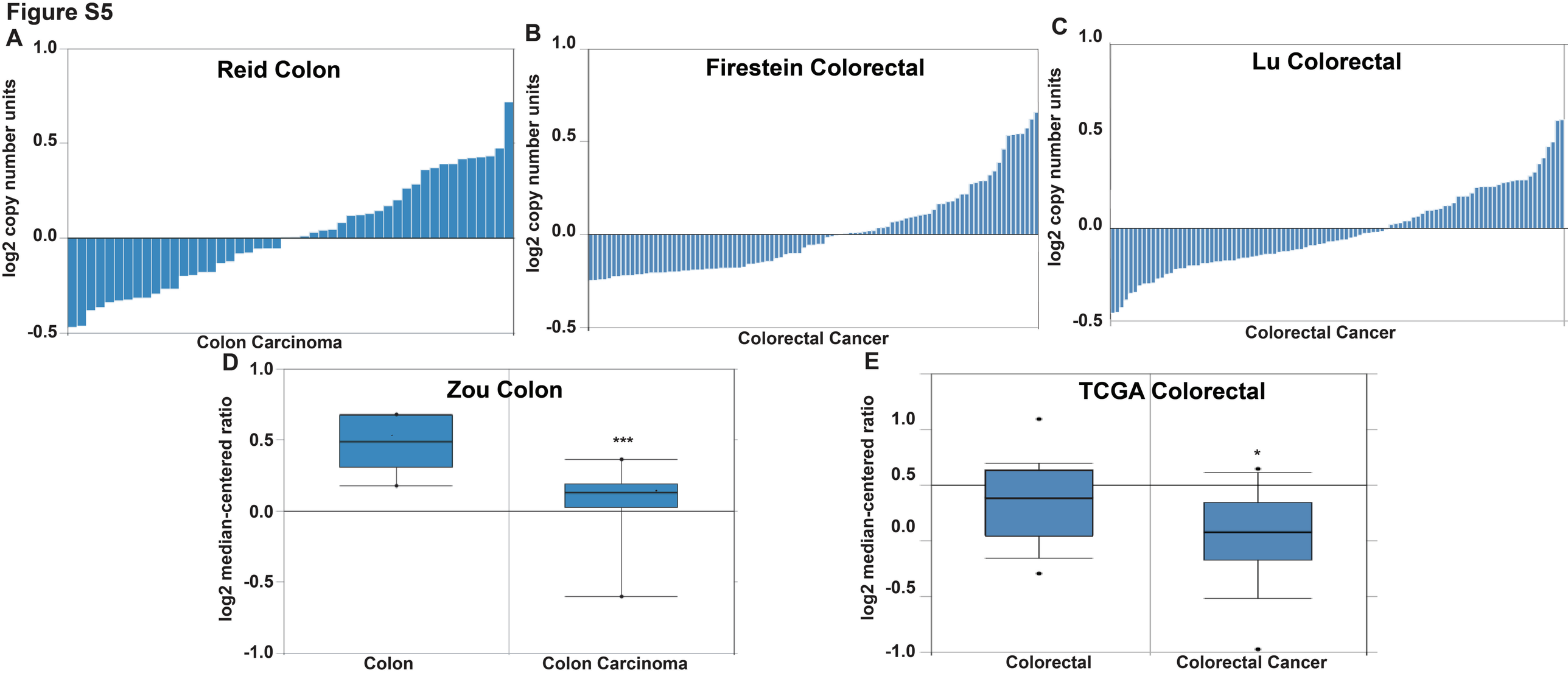

A
kDa
SDS-PAGE
Western blot
250-
150-
100-
75-
50-
37-
25-
20-
- Ku86
- Ku70
- K-H
1 2
B
IP
Input
IgG
K-H
- K-H
- Ku70
- Ku86
- IgG (Loading)
1 2 3
C
Ku86
K-H
Ku70
Apaf1 gene
Apaf1 ↓
Inactive caspase-3
K-H loss
D
Ku86
Ku70
Apaf1 gene
Apaf1 ↑
z-VAD
Caspase-3 activation
MLH1-PMS2 degradation

Figure S5


A
Reid Colon
log2 copy number units
1.0
0.5
0.0
-0.5
Colon Carcinoma
B
Firestein Colorectal
log2 copy number units
1.0
0.5
0.0
-0.5
Colorectal Cancer
C
Lu Colorectal
log2 copy number units
1.0
0.5
0.0
-0.5
Colorectal Cancer
D
Zou Colon
log2 median-centered ratio
1.0
0.5
0.0
-0.5
-1.0
***
Colon
Colon Carcinoma
E
TCGA Colorectal
log2 median-centered ratio
1.0
0.5
0.0
-0.5
-1.0
*
Colorectal
Colorectal Cancer