Supplementary Materials and Methods

Deep sequencing primer design.

The following primer pairs were used (in different combinations) to amplify the different regions in C-terminal *AKD*. Forward A-Tag and reverse B-Tag sequences are shown in black and *AKD* sequences are in red. Nucleotide co-ordinates are noted in the primer name and lengths of *AKD* sequence captured in each amplicon are shown in parentheses.

Akd_2674_A: 5'-GCCTCCCTCGCGCCATCAGTTGGAAGAAAGATTAGGC-3'

Akd_2714_B: 5'-GCCTTGCCAGCCCGCTCAGCTTTTTGTATTTATGTGGATATGTA-3'

(41bp)

Akd_2786_A: 5'-GCCTCCCTCGCGCCATCAGGGAAACTGCACTGCAAACAAG-3'

Akd_2847_B: 5'-GCCTTGCCAGCCCGCTCAGTCTCACATCTTCACCTTCCTC-3'

(62bp)

Akd_2712_A: 5'-GCCTCCCTCGCGCCATCAGTCCCAGGGGACAATTGAAAGA-3'

Akd_2786_B: 5'-GCCTTGCCAGCCCGCTCAGCCTCAGTCCCTTCTTCCCCA-3'

(75bp)

Akd_2712_A: 5'-GCCTCCCTCGCGCCATCAGTCCCAGGGGACAATTGAAAGA-3'

Akd_2764_B: 5'-GCCTTGCCAGCCCGCTCAGCTTGTTTGCAGTGCAGTTTCC-3'

(53bp)

Akd_2888_A: 5'-GCCTCCCTCGCGCCATCAGCGTCTTCCATGTATGCACC-3'

Akd_2949_B: 5'-GCCTTGCCAGCCCGCTCAGCAGCTGGGCCTCAATGTCCAC-3'

(62bp)

Akd_2823_A: 5'-GCCTCCCTCGCGCCATCAGACTGAGGAAGACACAGAGGAAAA-3'

Akd_2891_B: 5'-GCCTTGCCAGCCCGCTCAGGGTCAACACACACTTGGTGG-3'

(69bp)

Akd_2674'_A: 5'-GCCTCCCTCGCGCCATCAGCATTTGGAAGAAAGATTAGGC-3'

Akd_2763_B: 5'-GCCTTGCCAGCCCGCTCAGTTGTTTGCAGTGCAGTTTCCT-3'

(90bp)

Akd_2820_A: 5'-GCCTCCCTCGCGCCATCAGGACTGAGGAAGACACAGAGGAAAA-3'

Akd_2869_B: 5'-GCCTTGCCAGCCCGCTCAGAAGGTGCATACATGGAAGACG-3'

(50bp)

Akd_2740_A: 5'-GCCTCCCTCGCGCCATCAGGTACATATCCACATAAATACAAAAAG-3'

Akd_2786_B: 5'-GCCTTGCCAGCCCGCTCAGCCTCAGTCCCTTCTTCCCCA-3'

(47bp)

Akd_2833_A: 5'-GCCTCCCTCGCGCCATCAGCACAGAGGAAAAATGTACTATC-3'

Akd_2891_B: 5'-GCCTTGCCAGCCCGCTCAGGGTCAACACACACTTGGTGG-3'

(59bp)

Akd_2868_A: 5'-GCCTCCCTCGCGCCATCAGTTTAGAGGAAGGTGAAGATGTGAG-3'

Akd_2916_B: 5'-GCCTTGCCAGCCCGCTCAGGCACTTCTTATTGGTAACCAA-3'

(49bp)

Akd_2724_A: 5'-GCCTCCCTCGCGCCATCAGGACAATTGAAAGATGTACATATCC-3'

Akd_2793_B: 5'-GCCTTGCCAGCCCGCTCAGTGTGTCTTCCTCAGTCCCTTC-3'

(71bp)

Construction of point mutations and deletions of AKD. The following primers were used for site-directed mutagenesis: Gln_STOP-For: 5'-CAATCGTGGAGCATCCTAGGGGACAATTGAAAG-3'; Gln_STOP-Rev: 5'-CTTTCAATTGTCCCCTAGGATGCTCCACGATTG-3'; Pro_Ser-For: 5'-GAAAGATGTACATATTCACATAAATACAAAA-3'; Pro_Ser-Rev: 5'-TTTTGTATTTATGTGAATATGTACATCTTTC-3'; Asp_Asn-For: 5'-ATTTTTCCTCTGTGTTTTCCTCAGTCCCTTC-3'; Asp_Asn-Rev: 5'-ATTTTTCCTCTGTGTTTTCCTCAGTCCCTTC-3'; Thr_Ile-For: 5'-

CAGAGGAAAAATGTATTATCTG-3' and Thr_Ile-Rev: 5'-ATAGACAAACAGATAATACATTTTTCCTCTG-3'. Fragments were subsequently PCR amplified and cloned in-frame into the expression plasmid pTriex2-GFP as *Xho*I-digested products using the following primer pair: 5'-GTCTCGAGCTTCTCAATGGACTCCTGAATAT-3' and 5'GTCTCGAGTCAACTTTCACTTGGCAGCTG-3'. The ΔRING and ΔNRG deletions of Arkadia were generated as previously described (1). The ΔTIER mutation was constructed by PCR and consists an internal deletion of the seven amino acids directly after the NRG domain.

RNA/Protein extraction from colon tissue and tumors. Tissue was simultaneously lysed and homogenized in buffer RLT (Qiagen) using a TissueLyser II machine (Qiagen). RNA and protein was subsequently extracted using the AllPrep RNA/Protein mini kit (Qiagen). Protein was resuspended in 5% SDS and quantified using the BCA assay kit (Pierce). Determination of RNA concentration/purity and cDNA synthesis was performed as previously described (2).

Quantitative PCR (qPCR). qPCR was performed as previously described (3). Primer sequences for human *SNON* and Alu repeats, and mouse *Akd*, *SnoN*, *p21*, *p15*, *Pai-1*, *Tmepai*, *Smad7*, c-*Myc*, *Histone H3* are as follows: *SNON*-For*:* 5'-GGCTGAATATGCAGGACAG-3'; *SNON*-Rev: 5'-TGAGTTCATCTTGGAGTTCTTG-3'; Alu-For*:* 5'-CATGGTGAAACCCCGTCTCTA-3'; Alu-Rev: 5'- GCCTCAGCCTCCCGAGTAG-3'; *Akd*-For*:* 5'-CCCATCCACATAGGATGCAC-3'; *Akd*-Rev: 5'-CAGTTCCCAGGCAGTTCTCT-3'; *SnoN*-For: 5'-TCCAATCAAAGACAGATACACCA-

3’; *SnoN*-Rev**:** 5’-CTGGGGTGTAAAAATGAATGTG-3’**;** *p21*^{WAF}-For: 5’-CAGACATTCAGAGCCACAGG-3’; *p21*^{WAF}-Rev: 5’-GCGCAACTGCTCACTGTC-3’; *p15*^{INK4b}-For: 5’-AGATCCCAACGCCCTGAAC-3’; *p15*^{INK4b}-Rev: 5’-TCGTGCACAGGTCTGGTAAG-3’; *Pai-1*-For: 5’-TCCTGCCTAAGTTCTCTCTGG-3’; *Pai-1*-Rev: 5’-CTGCTCTTGGTCGGAAAGAC-3’; *Tmepai*-For: 5’-GGAGATCACGGAGCTGGAGT-3’; *Tmepai*-Rev: 5’-CAGCTTGTAGTGGCTCAGCA-3’; *Smad7*-For: 5’-GAAACCGGGGGAACGAATTAT-3’; *Smad7*-Rev: 5’-CGCGAGTCTTCTCCTCCCA-3’; *c-Myc*-For: 5’-GCGACTCTGAAGAAGAGCAAG-3’; *c-Myc*-Rev: 5’-GATGGAGATGAGCCCGACT-3’; *H3*-For: 5’-CGTTACAGGCCTGGTACTGTG-3’ and *H3*-Rev: 5’-TCTGTTTTGAAGTCCTGAGCAA-3’.

References

1. Mavrakis KJ, Andrew RL, Lee KL, Petropoulou C, Dixon JE, Navaratnam N, et al. Arkadia enhances Nodal/TGF-beta signaling by coupling phospho-Smad2/3 activity and turnover PLoS Biol 2007;5:e67.

2. Guzman-Ayala M, Lee KL, Mavrakis KJ, Goggolidou P, Norris DP, Episkopou V. Graded Smad2/3 activation is converted directly into levels of target gene expression in embryonic stem cells PLoS One 2009;4:e4268.

3. Nagano Y, Mavrakis KJ, Lee KL, Fujii T, Koinuma D, Sase H, et al. Arkadia induces degradation of SnoN and c-Ski to enhance transforming growth factor-beta signaling J Biol Chem 2007;282:20492-501.

Supplementary Table S1.

Patient	Nucleotide change	Amino acid change	Frequency (%)	Reads
6	2695 (CAG>TAG)	Q899STOP	2.76	4167
6	2717 (ACA>ATA)	T906I	4.58	4167
6	2722 (CCA>TCA)	P908S	9.94	4167
6	2725 (CAT>TAT)	H909Y	3.31	4167
6	2773 (CAC>TAC)	H925Y	3.26	6857
32	2809 (GAC>AAC)	D937N	4.3	3833
32	2828 (ACT>ATT)	T943I	1.59	3833
6	2828 (ACT>ATT)	T943I	1.76	5282
32	2924 (ACC>ATC)	T975I	1.67	4856

Supplementary Table Legend

Table S1. Summary of significant mutations identified. Mutations that were absent in normal tissue but observed in tumor tissue are shown above. Position of nucleotide change in the open-reading frame (ORF) is given.

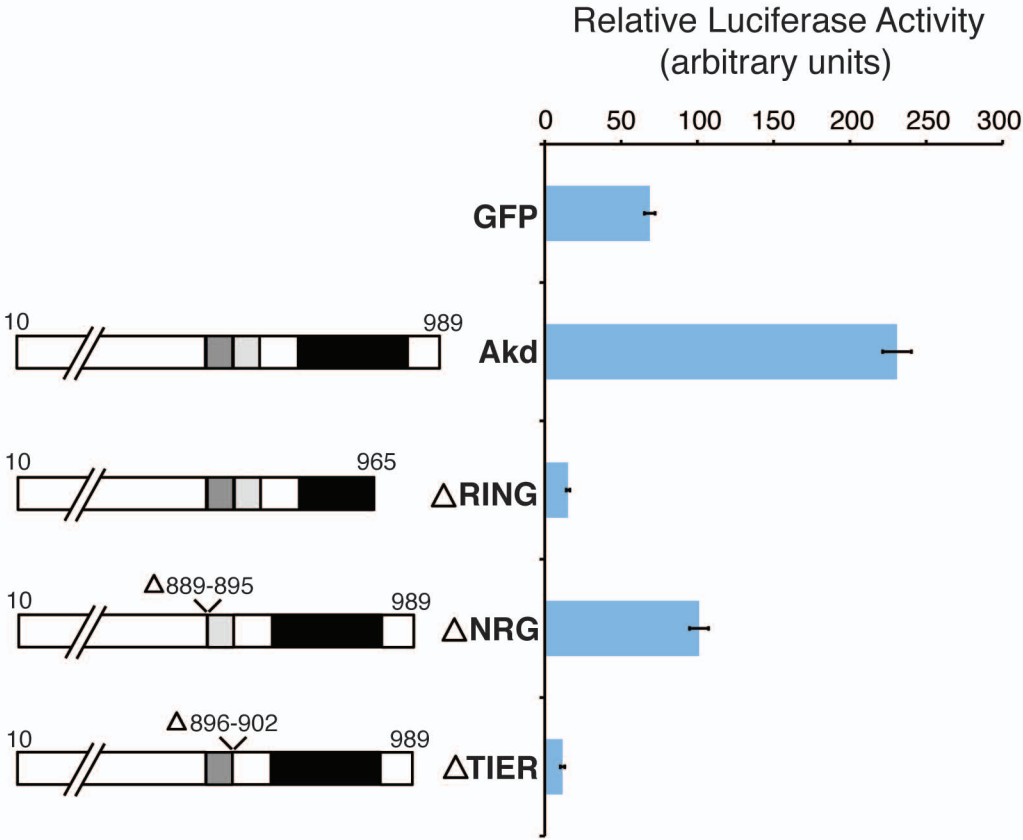
Supplementary Figure Legends

Figure S1. Deletion of RING, NRG and TIER domains of Arkadia in a Smad-dependent reporter assay. Luciferase reporter assay in HEK293T cells using the Smad-dependent $CAGA_{12}$-Luc reporter. Deletion of the RING (ΔRING) and TIER (ΔTIER) domains resulted in a potent dominant-negative effect. Deletion of the NRG (ΔNRG) domain reduced reporter

readout but did not exert a dominant-negative effect. Three biological repeat experiments were performed, each in quadruplicate.

Figure S2. **Example of a sequencing flowgram representing a single amplicon read from carcinoma tissue of patient 32**. Two putative mutations were deduced from this sample: D937N (4.3 %) and T943I (1.59%) from approximately 4000 reads of this amplicon (upper panel). The bottom three panels show the flowgram data set for a representative read containing the D937N causing mutation. The top flowgram shows the theoretical readout for the reference amplicon. The middle flowgram below shows the G>A substitution that results in the D937N mutation. The bottom flowgram shows reference minus actual read. Homopolymeric false-positives inherent with pyrosequencing are highlighted in pink boxes (upper panel).

Supplementary Fig. 1



Supplementary Fig. 2