SUPPLEMENTARY METHODS SM2

Inputs for apoptosis modelling-based signatures

The input concentrations for the APOPTO-CELL signatures are summarized in **Supplementary Methods 2 Table SM2.1**. A detail description of the measurements and data processing for each of the cohorts examined in this study is provided below. Additionally, Procaspase-3 expression was used as input for the APOPTO-CELL-PC3 signature.

Supplementary Methods 2 Table SM2.1. Measurement types for APOPTO-CELL inputs.

APOPTO-CELL signature inputs	Cohorts		
	Discovery	Expansion	Validation
Procaspase-3	protein expression (RPPA)	mRNA expression (Affimetrix HG-U133 Plus 2.0)	mRNA expression (Illumina GA and/or HiSeq RNASeqV2)
Procaspase-9	protein expression (RPPA)	mRNA expression (Affimetrix HG-U133 Plus 2.0)	mRNA expression (Illumina GA and/or HiSeq RNASeqV2)
SMAC	protein expression (RPPA)	mRNA expression (Affimetrix HG-U133 Plus 2.0)	protein expression (RPPA)
XIAP	protein expression (RPPA)	mRNA expression (Affimetrix HG-U133 Plus 2.0)	protein expression (RPPA)

1. Discovery cohort

The protein expression of Procaspase-9, XIAP, SMAC and Procaspase-3 was determined by Reverse Phase Protein Arrays (RPPAs).

1.1 Protein extraction

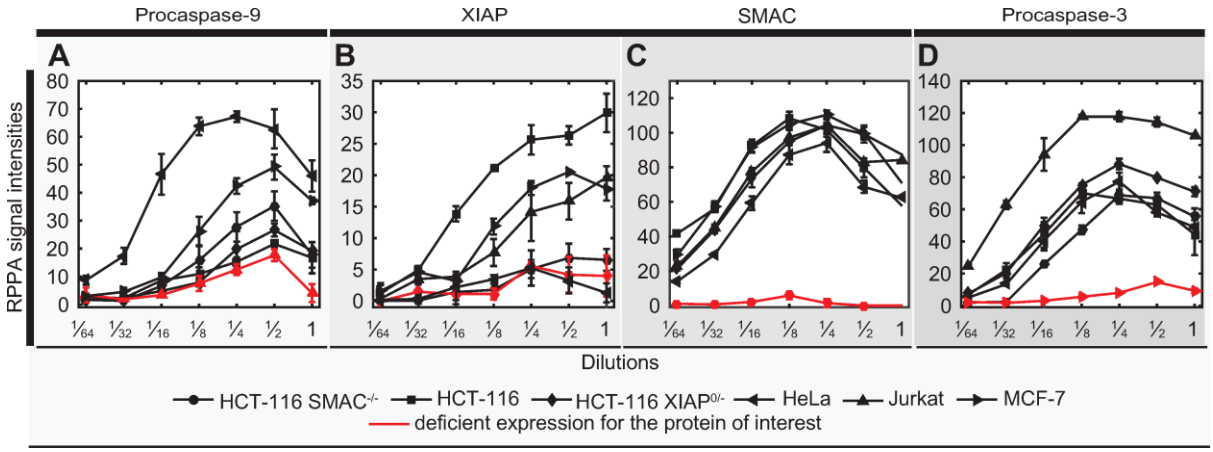
Protein lysates were prepared from FFPE samples as previously described (1). In brief, samples were histopathologically confirmed to contain at least 70% tumor and 3–20 serial 20 μm sections of tumor were then adhered to uncharged slides using nuclease-free water. One additional adjacent 5 μm section was hematoxylin and eosin (H&E) stained and the tumor area was outlined by an expert histopathologist. Slide overlays with the H&E stained section then allowed to isolate tumor-dense target areas with sterile razor blades. Isolated tumor material was deparaffinized using xylene/ethanol and incubated in protein extraction buffer (20 mM Tris buffer (pH 9) containing 2% SDS and protease inhibitors) for 2 hours at 80°C.

1.2 Reverse phase protein arrays, immunostaining and data acquisition

RPPAs were prepared and analyzed as described previously (2). In brief, protein lysates were concentration-normalized to 1 μg/μl, as assessed by bicinchoninic acid assay (DC Protein Assay, Bio-Rad, CA, USA). Three parts of protein lysates were mixed with 1 part of an SDS buffer (40% Glycerol, 8% SDS, 0.25 M Tris-HCl, pH6.8 plus Bond-Breaker TCEP Solution (Pierce Biotechnology, IL, USA) at 1/10 of the volume) and boiled. Lysates were then serially diluted four-fold with SDS buffer (10% Glycerol, 2% SDS, 0.0625 M Tris-HCl, pH6.8 plus Bond-Breaker TCEP Solution (Pierce Biotechnology, IL, USA) at 1/40 of the volume). A QArray 2 arrayer (Molecular Device, UK) was used to create a 378 sample array on Oncyte Avid nitrocellulose-coated slides (Grace Bio-Labs, OR, USA). The slides were stored together with desiccant (Drierite, OH, USA) at −20°C prior to immunostaining. In controls for antibody specificity, cell line lysates were spotted in 6-fold serial dilutions. Lysates included extracts from the following cell lines: HeLa, MCF-7 (Procaspase-3 deficient), Jurkat (Procaspase-9 deficient), HCT-116, HCT-116 $XIAP^{0/-}$, HCT-116 $SMAC^{-/-}$ (3,4).

Immunostaining was performed on an automated slide stainer (Dako Link 48, Dako, CA, USA) according to the manufacturer's instructions (CSA kit, Dako, CA, USA). Each slide was incubated with a single primary antibody at room temperature for 30 min. Primary antibodies were purchased from Cell Signalling for Caspase-3 (cat. no. 9662), Caspase-9 (cat. no. 9502), SMAC (cat. no. 2954), XIAP (cat. no. 2042) and from Sigma-Aldrich for β-actin (cat. no. A5441). The secondary antibodies used were goat anti-rabbit IgG (1:7500, Vector Laboratories, CA, USA) or rabbit anti-mouse IgG (1:10, Dako, CA, USA). Dako secondary antibodies were used as a starting point for tyramide signal amplification according to the manufacturer's instructions (Dako, CA, USA).

TIFF images of scanned slides were analyzed using Microvigene software version 5.1 (VigeneTech Inc., MA, USA) to obtain spot signal intensities. For quantification, the QRPPA module of Microvigene, using a 4 parameter logistic-log model ("SuperCurve" algorithm), was used (5). Signal intensities were normalized to β-actin to correct for differences in loading. The specificity of the signals obtain for the different antibodies was validated in cell line extracts deficient in expression of the target proteins (**Suppl. Methods 2 Fig. SM2.1**).



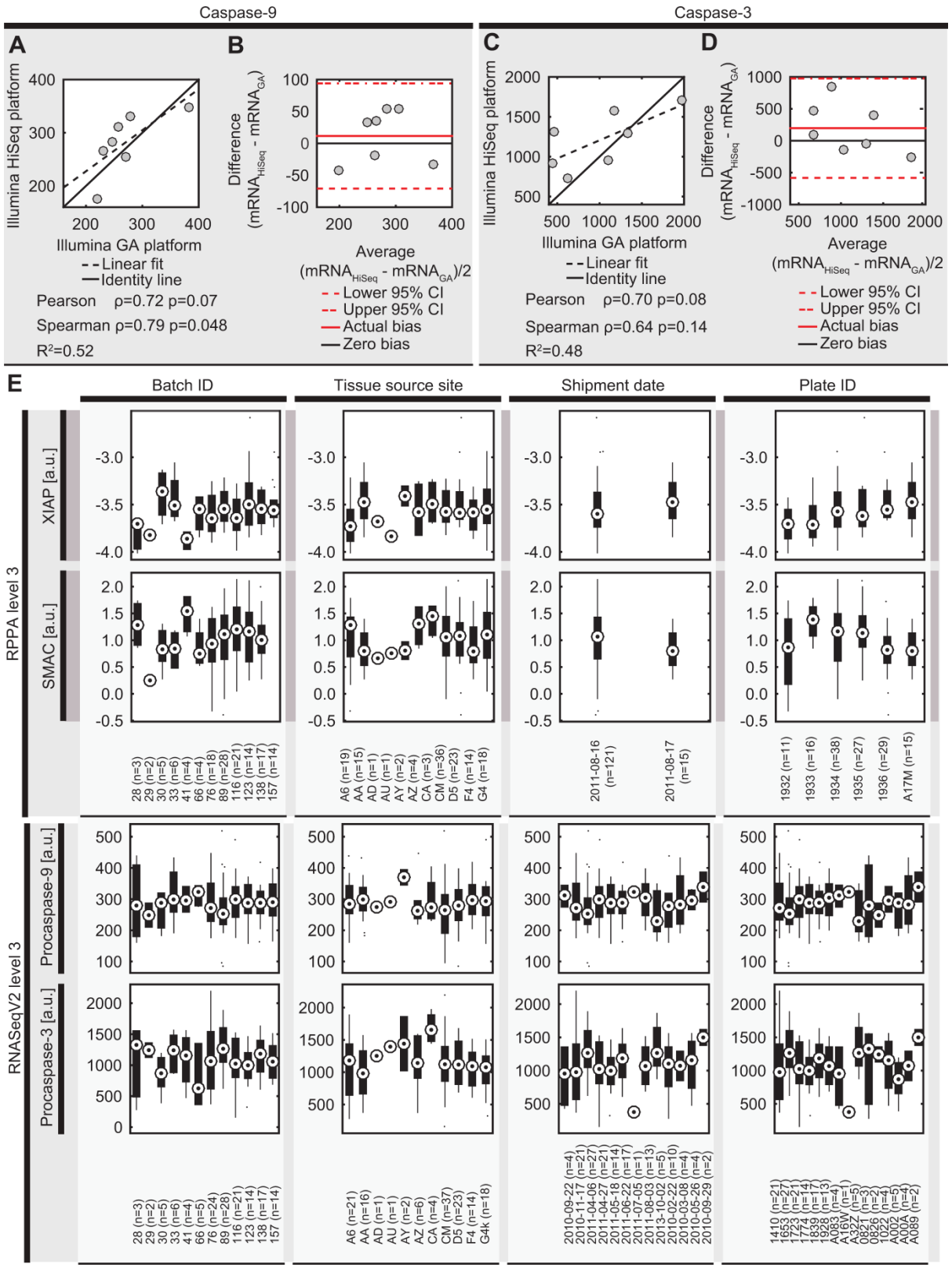
Supplementary Methods 2 Figure SM2.1. **RPPA antibody validation.** Antibody specificity was tested in a panel of human cancer cell lines that were proficient in all (HeLa, HCT-116 cells) or deficient in expression of the respective proteins of interest (SMAC- and XIAP-deficient HCT-116 cells, MCF-7 cells endogenously deficient in Procaspase-3, Jurkat cells deficient in Procaspase-9*;* red solid lines). Data represent mean ± S.D. of three technical replicates assayed on the same RPPA array as the patient samples of the discovery cohort.

2. Expansion cohort

Protein measurements were not available and thus gene expression was used as surrogate. Gene expression was measured in FF samples by Affymetrix U133 Plus 2.0 chips (6) and resulting data processed by the original authors (file "GSE39582_series_matrix.txt.gz") were downloaded from the GEO repository (http://www.ncbi.nlm.nih.gov/geo/). The list of probes mapping to each of the genes of interest was obtained *via* the "GeneAnnot" tool (http://www.genecards.org). We mapped multiple probes measurements onto a single gene by applying Principle Component Analysis (PCA) and by retaining the first component. Next, protein concentrations were estimated from the first component *via* the pipeline established in the discovery cohort. Again, since the reference molar distributions were obtained from stage II/III patients (7), we used data for stage II/III participants to determine the cumulative distribution functions.

3. Validation cohort

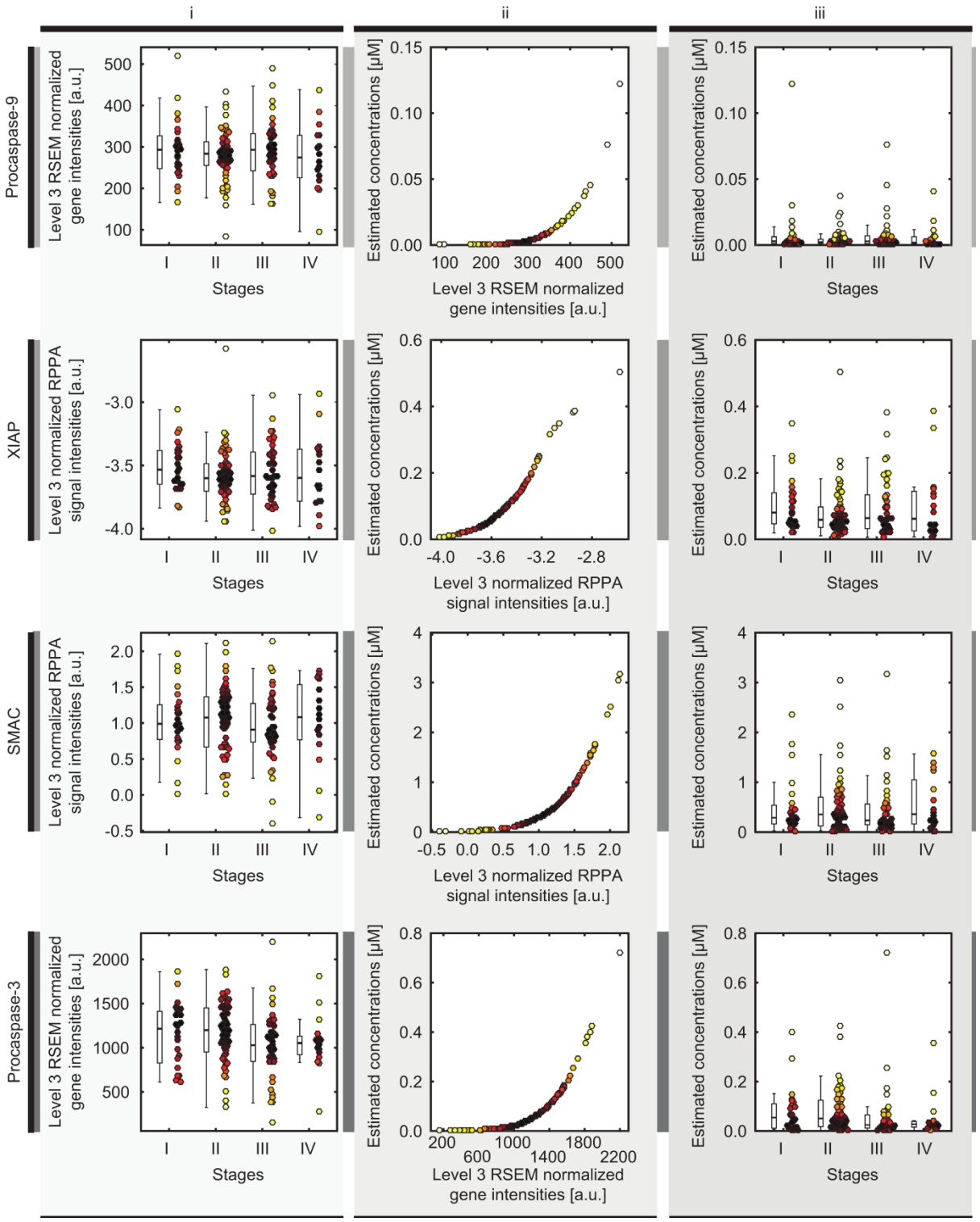
For the validation cohort, the protein measurements (performed by RPPA) were available only for SMAC and XIAP, but not for Procaspase-3 and Procaspase-9. Thus, input data for APOPTO-CELL systems modeling consisted of level 3 normalized RPPA protein expression data for SMAC and XIAP and RSEM normalized gene expression data for caspase-3 and caspase-9. Transcript abundance was assessed by two different Illumina platforms: HiSeq and GA. For the patients that also met the inclusion criteria, mRNA expression was measured with HiSeq and GA platforms in n=104 and n=39 samples, respectively. For 7 patients, both measurements were available. For these, both caspase-9 and -3 transcripts were strongly correlated (Pearson $\rho \geq 0.70$) and quantitatively similar between the datasets (**Suppl. Methods Fig. SM2.2A-D**). The two datasets were therefore merged without any platform-related batch adjustment. Duplicate measurements from both platforms were averaged. Notable batch effects attributable to batch ID, tissue source site, shipment date or plate ID were not detected in the RPPA or RNA datasets, therefore no batch effect correction was applied (**Suppl. Methods 2 Fig. SM2.2E**). The workflow described for the discovery cohort was used to obtain estimates for protein concentrations in the validation cohort. Analogously to the expansion cohort, only stage II/III COAD measurements were used to define the cumulative distribution functions. These functions were then used to obtain protein estimates for the entire stage I-IV COAD dataset (**Suppl. Methods 2 Fig. SM2.3**).



Supplementary Methods 2 Figure SM2.2 **Analysis of the RPPA and RNASeqV2 data in the validation cohort.**

A-D. Comparison of the RNASeqV2 transcripts in n=7 patient samples measured by both HiSeq and GA platforms. High correlation (Pearson and Spearman, **A, C**) and agreement (Bland-Altman plots, **B, D**) between mRNA measurements was identified.

E. Batch effects analysis. The presence of potential batch effects was assessed both in RPPA and RNASeqV2 expression data with respect to Batch ID, Tissue Source Site, Shipment date and Plate ID. Batch effects were not detected.



Supplementary Methods 2 Figure SM2.3 **Estimation of the absolute concentrations of the APOPTO-CELL proteins in tumor samples of the validation cohort.**

Level 3 normalized expression (RPPA and RSEM for XIAP/SMAC and Procaspase-3/Procaspase-9, respectively) were acquired (**i**) and the corresponding absolute concentrations (**iii**) were estimated *via*

a conversion function (**ii**) determined with the same pipeline developed for the discovery cohort. Further details on these procedures are provided in the Materials and Methods section.

References

1. Guo H, Liu W, Ju Z, Tamboli P, Jonasch E, Mills GB, et al. An efficient procedure for protein extraction from formalin-fixed, paraffin-embedded tissues for reverse phase protein arrays. Proteome Sci. Proteome Science; 2012;10:56.
2. Hennessy BT, Lu Y, Gonzalez-Angulo AM, Carey MS, Myhre S, Ju Z, et al. A Technical Assessment of the Utility of Reverse Phase Protein Arrays for the Study of the Functional Proteome in Non-microdissected Human Breast Cancers. Clin Proteomics. 2010;6:129–51.
3. Cummins JM, Kohli M, Rago C, Kinzler KW, Vogelstein B, Bunz F. X-linked inhibitor of apoptosis protein (XIAP) is a nonredundant modulator of tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-mediated apoptosis in human cancer cells. Cancer Res. 2004;64:3006–8.
4. Kohli M, Yu J, Seaman C, Bardelli A, Kinzler KW, Vogelstein B, et al. SMAC/Diablo-dependent apoptosis induced by nonsteroidal antiinflammatory drugs (NSAIDs) in colon cancer cells. Proc Natl Acad Sci U S A. 2004;101:16897–902.
5. Hu J, He X, Baggerly KA, Coombes KR, Hennessy BTJ, Mills GB. Non-parametric quantification of protein lysate arrays. Bioinformatics. 2007;23:1986–94.
6. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. PLoS Med. 2013;10:e1001453.
7. Hector S, Rehm M, Schmid J, Kehoe J, McCawley N, Dicker P, et al. Clinical application of a systems model of apoptosis execution for the prediction of colorectal cancer therapy responses and personalisation of therapy. Gut. 2012;61:725–33.