1   **KH-type splicing regulatory protein controls colorectal cancer cell growth and**

2   **modulates the tumor microenvironment**

3   Francesco Caiazza[*1,2,8], Katarzyna Oficjalska[1,2], Miriam Tosetto[1], James J. Phelan[3],

4   Sinéad Noonan[1], Petra Martin[1], Kate Killick[4], Laura Breen[5], Fiona O'Neill[5], Blathnaid

5   Nolan[1], Simon Furney[6], Robert Power[2,#], David Fennelly[1,2], Charles S. Craik[7], Jacintha

6   O'Sullivan[3], Kieran Sheahan[1,2], Glen A. Doherty[1,2] and Elizabeth J. Ryan[1,2§]

7

8   [1]Centre for Colorectal Disease, St. Vincent's University Hospital, Elm Park, Dublin,

9   Ireland. [2]School of Medicine, University College Dublin, Ireland. [3]Trinity Translational

10   Medicine Institute, Department of Surgery, Trinity College Dublin, Ireland. [4]Systems

11   Biology Ireland, University College Dublin, Ireland. [5]National Institute for Cellular

12   Biotechnology, Dublin City University, Ireland. [6]Department of Physiology and Medical

13   Physics, Royal College of Surgeons in Ireland. [7]Department of Pharmaceutical

14   Chemistry, University of California San Francisco, San Francisco, U.S.A.

15

16   *Corresponding author: Francesco Caiazza, Department of Pharmaceutical Chemistry,

17   University of California San Francisco, 600 16th street, Room S-514, San Francisco,

18   California, U.S.A. francesco.caiazza@ucsf.edu

19   [§]Present address: Dept. of Biological Sciences, School of Engineering and Science,

20   University of Limerick, Limerick, Ireland.

21   [#]Present address: Personalis Inc., Menlo Park, California, U.S.A.

22

23   **Running title:** Role of KHSRP in colorectal cancer

29   **Disclosures:** None declared.

30

31   **Page count:** 43

32   **Total number of figures:** 6, **Total number of tables:** 1

33

34   **Abbreviations:** KHSRP, KH-type splicing regulatory protein; CRC, colorectal cancer;

35   IHC, immunohistochemistry; FFPE, formalin fixed paraffin embedded; RBP, RNA-

36   binding protein; UTR, untranslated region; ARE, AU-rich element; TCGA, The cancer

37   genome atlas; DAMP, damage-associated molecular pattern; PAMP, pathogen-associated

38   molecular pattern.

39   **ABSTRACT**

40

41   KH-type splicing regulatory protein (KHSRP) is a multifunctional nucleic acid binding

42   protein implicated in key aspects of cancer cell biology: inflammation and cell-fate

43   determination. However, the role KHSRP plays in colorectal cancer (CRC) tumorigenesis

44   remains largely unknown. Using a combination of *in silico* analysis of large datasets, *ex*

45   *vivo* analysis of protein expression in patients, and mechanistic studies using *in vitro*

46   models of CRC, we investigated the oncogenic role of KHSRP. We demonstrated

47   KHSRP expression in the epithelial and stromal compartments of both primary and

48   metastatic tumors. Elevated expression was found in tumor versus matched normal

49   tissue, and we validated these findings in larger independent cohorts *in silico.* KHSRP

50   expression was a prognostic indicator of worse overall survival (HR=3.74, 95% CI =

51   1.43-22.97, p=0.0138). Mechanistic data in CRC cell line models supported a role of

52   KHSRP in driving epithelial cell proliferation in both a primary and metastatic setting,

53   through control of the G1/S transition. Additionally, KHSRP promoted a pro-angiogenic

54   extracellular environment by regulating the secretion of oncogenic proteins involved in

55   diverse cellular processes such as migration and response to cellular stress. Our study

56   provides novel mechanistic insight into the tumor-promoting effects of KHSRP in CRC.

57 **INTRODUCTION**

58 Almost all the major oncogenic signaling pathways result in the reprogramming of

59 translation. This impacts on both cancer initiation and progression by directing selective

60 translation of specific tumor-promoting mRNAs [1]. RNA-binding proteins (RBP) such as

61 KH-type splicing regulatory protein (KHSRP, also known as KSRP or FUBP2) are

62 capable of recognizing sequence-specific regulatory elements in the 3'UTR, such as AU-

63 rich elements (ARE) that regulate transcript specific translation promoting mRNA

64 splicing and stability [2]. KHSRP regulates translation by mediating mRNA decay through

65 direct binding and recruitment of proteins involved in RNA degradation, including the

66 Poly(A)-Specific Ribonuclease (PARN), exosome components and decapping enzymes

67 [3,4]. Alternatively, KHSRP can regulate translation independently of mRNA decay, either

68 by promoting translational inhibition [5] or indirectly by promoting maturation of selected

69 miRNA precursors [6]. Adding a further layer of complexity, KHSRP also acts as a

70 transcription factor, inducing the expression of the *MYC* oncogene [7].

71 The KHSRP-mediated control of mRNA translation and stability has been extensively

72 studied in the context of innate immunity [5,8]. Pro-inflammatory cytokines (e.g. IL-6, IL-

73 8, TNF-α, IL-1β) and inflammatory mediators such as iNOS are directly targeted and

74 negatively regulated by KHSRP [9,10]. However, the role of KHSRP is most likely context

75 dependent, with the balance between KHSRP and other RBP with similar or divergent

76 regulatory effects an important consideration [11].

77 Direct evidence for the involvement of KHSRP in cancer is accumulating. KHSRP has

78 been implicated in the pathogenesis of small cell lung cancer [12], osteosarcoma [13] and

79 breast cancer [14]. A number of distinct mechanisms have been proposed e.g. regulation of

80 cell differentiation in P19 mouse teratocarcinoma cells [15], deregulation of

81 oncosuppressive miRNAs such as let-7a and miR-30c [16] or control of TGF-β-dependent

82 regulation of epithelial-to-mesenchymal transition [17].

83 The role of KHSRP in colorectal cancer (CRC), however, remains underappreciated.

84 While other RBPs have been implicated in cellular transformation in the intestinal

85 epithelium [namely, ESRP1 [18], Apobec-1 [19], Musashi (MSI)-1 [20], and MSI-2 [21]], KHSRP

86 remains largely under-investigated in this indication. Interestingly, adenomatous

87 polyposis coli (APC), a tumor suppressor frequently mutated in CRC, is itself a RBP

88 capable of regulating the translation of mRNAs associated with cell adhesion, motility

89 and other cellular processes crucial for carcinogenesis [22], suggesting the importance of

90 RBP-mediated translation control in the gut. Here we use a combination of bioinformatic,

91 *in vitro* and *ex vivo* approaches to dissect the role of KHSRP in both regulation of cell

92 proliferation and inflammatory environment in CRC and provide evidence for a novel

93 prognostic role of this RBP in intestinal tumorigenesis.

94

95 **MATERIALS AND METHODS**

96

97 *Bioinformatic analysis of publicly available datasets*

98 CRC datasets in the Oncomine database (www.oncomine.org, accessed 11/2014, Thermo

99 Fisher Scientific, Ann Arbor, MI) were searched using the differential expression module

100 for cancer vs. normal tissue; output of the resulting database searches was downloaded

101 for offline plotting, along with detailed gene expression results for 2 representative

102 cohorts (GSE9348 and TCGA). A similar analysis was conducted without restricting for

103   cancer type. Differential expression analysis of RNA-Seq data for 12 tumor-normal pairs

104   in the TCGA database was previously described and made available by the authors [23].

105   Gene expression (RNA-Seq) and the frequency of genetic alterations for KHSRP in the

106   TCGA cohort were analyzed using the cBio portal (www.cbioportal.org, accessed

107   11/2016) [24]. All other datasets were downloaded from the Gene Expression Omnibus

108   (https://www.ncbi.nlm.nih.gov/geo/), using the indicated accession number.

109

110   *Patient Samples and Tissue Microarrays*

111   Fresh-frozen tumor and adjacent normal tissue samples were obtained from 16 patients

112   and were used for protein isolation and Western blotting (WB cohort) as described

113   thereafter. Tissue microarrays (TMA) were accessed from a previously described cohort

114   of formalin-fixed paraffin-embedded (FFPE) archival resected tumor tissue samples [25].

115   Patient details for both cohorts are reported in Table 1. For immunohistochemistry, TMA

116   sections   were   incubated   with   a   1:250   dilution   of   a   validated

117   (https://www.proteinatlas.org/ENSG00000088247-KHSRP/antibody, accessed 04/2019)

118   rabbit polyclonal anti-KHSRP antibody (Cat# HPA034739, RRID:AB_10601582,

119   Sigma-Aldrich, Saint Louis, MO, USA), and secondary antibody horseradish peroxidase

120   (DAKO Agilent, Santa Clara, CA, USA). Endogenous peroxidase activity was blocked

121   using 3% hydrogen peroxide, and non-specific binding was blocked with casein buffer.

122   Diaminobenzidine was used to visualize staining and sections were counter-stained with

123   haematoxylin, dehydrated and mounted. Stained TMA sections were acquired with a

124   ScanScope XT high throughput scanning system (Aperio Technologies, Leica

125   Biosystems, Buffalo Grove, IL, USA) and images were scored by two independent

126   investigators, blinded to the clinical data. Both epithelial and stromal cells were assessed

127   for staining intensity and percent positivity; positive cell count was scored as a

128   continuous variable (0-100%), while intensity was scored using a scale of 0 to 3

129   (correlating with negative, weak, medium and strong staining, respectively). The mean

130   value of each parameter from the two investigators was calculated, and parameters were

131   averaged from all tissue cores related to the same patient; the two parameters were finally

132   combined using the semi-quantitative *quickscore* method [26], by multiplying the averaged

133   intensity and positivity scores for epithelial and stromal cells separately. The quickscore

134   values for tumor tissue were further divided by the quickscore values for normal tissue, to

135   obtain a tumor-to-normal (T/N) ratio score that was used to quantify the amplitude of the

136   tumor-specific changes in protein expression. A similar strategy was used to obtain

137   metastasis-to-normal (M/N) scores. There was overall a good correlation between the

138   scores obtained by the two independent assessors (weighted Cohen Kappa coefficient =

139   0.63).

140

141   *Cell Culture, transient and stable knockdown of KHSRP*

142   SW480 and SW620 cells were obtained from the American Tissue Culture Collection

143   (ATCC, Manassas, VA, USA) and maintained in DMEM supplemented with 10% fetal

144   bovine serum, 1% penicillin/streptomycin and 1% Fungizone (all from Invitrogen Life

145   Technologies, Carlsbad, CA, USA). Cell line identity was confirmed by analysis of Short

146   Term Repeat loci, and cells were routinely tested for mycoplasma infection. Cells were

147   seeded in 24-well plates to be transfected with 30 pmol KHSRP siRNA (Silencer Select,

148   Thermo Fisher) or non-targeting scramble siRNA using 2 μl Lipofectamine 2000

149   (Invitrogen) in serum-free and antibiotic-free growth media for 6h, followed by

150   incubation in standard growth media for a total of 48h post-transfection. Alternatively,

151   cells were transfected with 50 nM KHSRP siRNA (SMARTpool: ON-TARGET plus)

152   and corresponding Non-Targeting Pool (Dharmacon, Lafayette, CO, USA). To generate

153   stable KHSRP knockdown cell lines, we took advantage of an ultracomplex pooled

154   shRNA library targeting each annotated human protein-coding gene with 25 independent

155   shRNAs on average (as well as 500 negative control shRNAs), which was previously

156   described and used to perform genome-wide genetic interaction screens in mammalian

157   cells [27]. From this original library, 3 independently validated shRNA sequences targeting

158   KHSRP were selected, along with one negative control scramble shRNA sequence; for

159   each sequence, top and bottom oligonucleotides were synthesized (Integrated DNA

160   Technologies, Coralville, Iowa, USA). Two shRNA-expressing derivative cell lines were

161   generated from SW480 and SW620 following a previously published protocol [28]. The

162   lentiviral plasmid used for shRNA expression was pMK1201, which is a modified

163   version of pMK1200 [28] based on the design of pINDUCER10 for tetracycline-induced

164   expression of the shRNA coupled with a fluorescent reporter (tRFP) and a Puromycin

165   resistance sequence for positive selection [29]. To clone each individual shRNA into the

166   lentiviral backbone, 2 μl of top and bottom oligonucleotides were annealed at 95°C for 5

167   min in a buffer containing 100 mM potassium acetate, 30 mM HEPES-KOH (pH=7.4), 2

168   mM magnesium acetate. Annealed oligonucleotides (0.01 μM) were ligated in 25 ng of

169   the pMK1201 vector pre-digested with BstXI (New England Biolabs, Ipswich, MA,

170   USA) and gel purified; ligation was carried out at RT for 2h using 2000 U of T4 DNA

171   ligase (New England Biolabs). DH5α cells (Thermo Fisher) were transformed and plated

172   onto ampicillin-containing LB plates overnight at 37°C. Single colonies were picked and

173   expanded for plasmid DNA mini prep (Qiagen, Germantown, MA, USA). Correct

174   insertion of shRNA sequences was confirmed by sequencing using the 5' pSico-Eco-

175   insert-seq [28]. For lentivirus preparation, the second-generation virion packaging vector

176   psPAX2 (plasmid #12260, Addgene, Cambridge, MA, USA) and the VSV-G envelope

177   plasmid pMD2.G (plasmid #12259, Addgene) from Didier Trono (Ecole Polytechnique

178   Fédérale de Lausanne) were used. The producer cell line was 293T, a highly-

179   transfectable derivative of human embryonic kidney cell 293 (ATCC), which was

180   maintained in DMEM supplemented with 10% fetal bovine serum, 1%

181   penicillin/streptomycin and 1% Fungizone (Invitrogen). 293T cells were transfected in 10

182   cm plates using the calcium phosphate method: 2.5 μg of scramble shRNA vector, or a

183   pooled mix of 0.83 μg each of the three KHSRP shRNA vectors, were combined with

184   0.58 μg of pMD2.G and 1.92 μg of psPAX2 and transfected with 12.5 μM chloroquine

185   diphosphate and 12.5 mM of calcium chloride. Transfection medium was replaced after

186   16h with full DMEM, and virus-containing conditioned medium was harvested 2 times at

187   48h and 72h post transfection; pooled harvests from each condition were pre-cleared by

188   centrifugation, filtered through a 0.45 μm membrane, and used immediately to transduce

189   SW480 and SW620. Cells were selected with 1 μg/ml puromycin for 2 weeks, expanded

190   and stock-frozen. A fresh stock of cells was used each time for experiments.

191

192   *Protein isolation, immunoblotting, immunofluorescence and ELISA*

193   Cells or tissue homogenate were lysed in RIPA buffer (150 mM NaCl, 50 mM Tris-HCl,

194   1% Triton, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with a protease and

8

195   phosphatase inhibitor cocktail (Roche Applied Science, Burgess Hill, UK) and 1 mM

196   PMSF (Sigma Aldrich). Total protein concentration was determined using the BCA

197   Protein Assay Kit (Thermo Fisher). Total proteins (10-20μg) were separated on 10% or

198   4-20% gradient SDS–PAGE gels and wet-transferred to PVDF using the mini-

199   PROTEAN® TGX™ system (Bio-Rad, Hercules, CA, USA). Membranes were pre-

200   blocked with 5% low-fat dry milk in TBS-T and incubated with a rabbit polyclonal anti-

201   KHSRP antibody (Sigma-Aldrich Cat# HPA034739, RRID:AB_10601582), or a

202   monoclonal anti-c-Myc antibody (D84C12, Cell Signaling, Danvers, MA, USA), or a

203   monoclonal anti-EphB2 (clone D2X2I, Cell Signaling), or a monoclonal anti-S100A11

204   (clone EPR11171B, Abcam, Cambridge, MA, USA), and either anti-rabbit or anti-mouse

205   horseradish peroxidase-conjugated secondary antibody (Bio-Rad). Protein loading

206   normalization was carried out with anti-β-actin or anti-GAPDH antibodies (Cell

207   Signaling), or Coomassie stain (Thermo Fisher). Proteins were visualized by

208   chemiluminescence with ECL substrate (Thermo Fisher) and semi-quantified using

209   ImageJ software (US National Institute of Health, Bethesda, MD, USA;

210   http://imagej.nih.gov/ij/, version 2.0.0-rc-69/1.52n) with normalization against β-actin

211   (Sigma-Aldrich). For immunofluorescence, cells were fixed on chamber slides with 4%

212   Formaldehyde, permeabilized with 0.5% Tween in PBS, and stained using the same anti-

213   KHSRP antibody and a Cy3-anti-rabbit, with DAPI and Phalloidin counter-stain

214   (DAKO); slides were imaged on a Leitz DM40 microscope (Leica Microsystems)

215   equipped with the AxioCam system and AxioVision 3.0.6 for image acquisition. Human

216   interleukin (IL)-8 and vascular endothelial growth factor (VEGF) were quantified in cell

217   culture supernatants using DuoSet® ELISA (R&D Systems, Minneapolis, MN, USA),

218   following manufacturer's instructions.

219

220   *RNA isolation and qRT-PCR*

221   Total RNA was isolated from cells using the EZNA Total RNA Kit (Omega, VWR,

222   Radnore, PA, USA). 1 μg of RNA was reverse-transcribed into cDNA using the RT

223   Omniscript cDNA kit (Qiagen), and 20 ng of cDNA template was amplified on a

224   LightCycler-480 (Roche, Indianapolis, IL, USA) using a FAST SYBR Green Master Mix

225   (Invitrogen) under the following conditions: pre-incubation step at 95 °C for 5 min was

226   followed by 45 cycles of denaturation at 95 °C for 10 sec, annealing at 60 °C for 10 sec,

227   and elongation at 72 °C for 10 sec. GAPDH was used as the housekeeping gene. The

228   mRNA expression levels of all samples were normalised to the housekeeping gene, and

229   the delta/delta Ct method was used to calculate relative expression (treatment vs. control).

230   The   following   primers   were   used   in   qPCR: KHSRP   Forward: 5'-

231   CCGCTTACTACGGACAGACC-3';   KHSRP   Reverse: 5'-

232   ATTCATTGAGCCTGCTGCTGT-3';   GAPDH   Forward: 5'-

233   CGGATTTGGTCGTATTGGGCGCCTG-3';   GAPDH   Reverse: 5'-

234   CAGCATCGCCCCACTTGATTTTGGA-3'. For validation of microarray candidate

235   genes, single strand cDNA synthesis and PCR amplification were carried out in a 1-step

236   reaction using the Brilliant II QRT-PCR Master Mix (Agilent) and TaqMan gene

237   expression assays (Life Technologies). 25 ng of RNA was loaded in 96-well plate format

238   and amplified with 0.9 mM of primers and 0.25 mM of MGB probe. The reaction was

239   carried out in a CFX Connect instrument (Bio-Rad) with the following parameters: an RT

10

240   step at 50 □C for 30 min was followed by a preincubation step at 95 □C for 10 min, 50

241   cycles of denaturation at 95 □C for 15 s, and annealing/extension at 60 □C for 1 min.

242   Relative quantification of gene expression was calculated with the comparative cycle

243   threshold method, normalizing for GAPDH expression levels. TaqMan gene expression

244   assays for FABP3 (GeneID: 2170, Assay ID: Hs00997360_m1), SMG1 (GeneID: 23049,

245   Assay ID: Hs00979691_m1), MT1F (GeneID: 4494, Assay ID: Hs00744661_sH), and

246   GAPDH (GeneID: 2597, Assay ID: Hs02786624_g1) contained MGB probes with a

247   FAM reporter dye.

248

249   *Gene Expression Microarray*

250   Total RNA was extracted from cells using the RNeasy Plus kit (Qiagen); the integrity and

251   concentration of RNA was confirmed using the RNA 6000 Nano kit on a Bioanalyzer

252   2100 (Agilent), with reported RNA Integrity Numbers (RIN) > 9. Amplified cDNA for

253   gene expression analysis was prepared with the Ovation® PicoSL WTA System V2

254   (Nugen, San Carlos, CA, USA); labeled cDNA targets were generated with the Encore®

255   Biotin Module (Nugen) and hybridized to a GeneChip® Human Transcriptome Array 2.0

256   (Affymetrix, Thermo Fisher) following manufacturer's instructions. CEL files were

257   analysed with the Affymetrix® Expression Console™ Software, using the Affymetrix

258   Human Transcripome Array 2.0 library files and the HTA-2_0.na35.2.hg19 annotation

259   files. An RMA workflow was performed which used a quantile normalization procedure

260   and a general background correction. The resulting normalized CHP files were then

261   imported in the Affymetrix® Transcriptome Analysis Console software to test for

262   differential expression using a paired One-Way Repeated Measure (ANOVA) approach,

263 and the default filter criteria (fold change > |2|, and p-value < 0.05). The full data have

264 been deposited in NCBI's Gene Expression Omnibus, and can be accessed with the GEO

265 Series accession number GSE112329. Over-representation analysis was used to identify

266 canonical pathways and functional processes of biological importance within the list of

267 differentially expressed genes; the analysis was performed using GeneTrail2 (version 1.5,

268 https://genetrail2.bioinf.uni-sb.de, accessed 04/2016) [30]. Network analysis was performed

269 using String (version 10.0, https://string-db.org/, accessed 04/2016) [31]. STRING is a

270 database of known and predicted protein-protein interactions, including both direct

271 (physical) and indirect (functional) associations, inferred from experimental data and

272 computational predictions. STRING computes a global score by combining the

273 probabilities from the different types of evidence and correcting for the probability of

274 randomly observing an interaction. Proteins in the network are then clustered based on

275 the distance matrix obtained from the global scores, using the KMEANS clustering

276 algorithm [32].

277

278 *Cell-based assays*

279 Cell proliferation was measured by Neutral Red (NR) staining following an established

280 protocol [33]. Briefly, cells were seeded in 24-well plates (20,000 cells/well) 24h after

281 transfection (for the transient knockdown) or 24h after doxycycline treatment (for the

282 stable knockdown), and allowed to grow for 7 days. Cells were stained with 33 mg/L

283 neutral red (Sigma) for 2h, then washed in PBS and imaged (for representative

284 photographs of the plates); cells were de-stained using a solution of 50% ethanol and 1%

285 glacial acetic acid in water, and absorbance measured at 540nm to quantify staining.

286   Real-time cell proliferation was measured using an E-plate 16 and a RTCA DP Analyser

287   (xCelligence System, Roche), which uses gold microelectrodes fused to the bottom

288   surface of a microtiter plate well. The impedance of electron flow at the interface

289   between electrodes and culture medium caused by adherent cells is reported using a

290   unitless parameter called the Cell Index (CI). Cells were seeded at 10,000 cells/well and

291   the recording of the CI occurred every 15 min during the first six hours, and every hour

292   for the rest of the period. Cells seeded at low density (5,000 cells/well in 6-well plates)

293   for colony-formation assay were grown for 7 days, then fixed in methanol/citric acid (4:1

294   v/v), stained in 2% crystal violet and counted with ImageJ. The plating efficiency was

295   calculated by dividing the number of colonies by the original seeding density, and the

296   surviving fraction was determined by comparing the plating efficiency of transfected vs

297   control wells. The total colony area was calculated for each biological replicate by

298   averaging the area of all colonies in replicate wells. Spheroids were generated by seeding

299   20,000 cells/well in 24-well plates on growth factor-reduced Matrigel (BD Biosciences,

300   San Jose, CA, USA) following the 3D on-top protocol [34]. After 7 days, spheroids were

301   stained with neutral red (as outlined above), imaged (averaging 3-4 fields per well at 10X

302   magnification) and quantified via ImageJ. Scratch-wound assays were performed in

303   culture inserts 2-well μ-dishes (Ibidi, Martinsried, Germany), quantified with ImageJ

304   (using MiToBo plugin) and representative images prepared with WimScratch software

305   (Ibidi). Invasion assays were performed using Matrigel-coated Boyden chambers (8-μm

306   pore size; BioCoat, BD Biosciences) for 72h against 10% FBS as chemo-attractant.

307   Invaded cells were fixed using 1% glutaraldehyde, stained with 0.1% crystal violet, and

308   imaged with a light microscope at a 10X magnification, averaging five random fields per

13

309   insert. Crystal violet staining was quantified with ImageJ. For tube-formation assays,

310   Human dermal microendothelial cells (HDEC, a gift from Prof. Ursula Fearon, TCD,

311   Dublin, Ireland) were seeded on Matrigel-coated 96-well plates (10,000 cells/well) in

312   MCDB131 medium (Invitrogen) with 40% conditioned medium from SW620 cells. The

313   tube analysis was performed manually after 6h by counting five sequential fields (10X

314   magnification) with a focus on the surface of the Matrigel, and a connecting branch

315   between two discrete endothelial cells was counted as 1 tube. Representative images were

316   prepared with WimTube (Ibidi).

317

318   *Flow Cytometry*

319   Cell cycle analysis was carried out by propidium iodide (PI) staining. Cells were

320   resuspended at $1\text{-}2\text{x}10^6$ cells/ml in PBS and fixed in 70% ethanol for 2h at -20°C, washed

321   twice in PBS, and re-suspended in staining solution [0.1% (v/v) Triton X-100, 10 μg/mL

322   PI, 100 μg/mL DNase-free RNase A in PBS] for 30' at R.T. Data was acquired on a

323   CyAn ADP (Beckman Coulter, Brea, CA, USA) using 488nm excitation laser, and data

324   de-convolution was performed in FlowJo[®] (Becton Dickinson, Franklin Lakes, NJ,

325   USA). To quantify the mitotic index, fixed cells were permeabilized with 0.25% Triton

326   X-100 in PBS for 20' on ice, blocked with 10% BSA in PBS and incubated with 1μg of

327   an Alexa-Fluor[®] 488-conjugated anti-phospho-histone H3 (Ser10) antibody (Cat#

328   650804, BioLegend, San Diego, CA, USA). After washing, cells were re-suspended in

329   PBS containing 50 μg/ml PI for 15' on ice prior to analysis on a CyAn ADP using 488nm

330   excitation laser. Mitotic cells with 4N DNA and elevated levels of phosphorylated

331   histone H3 can be detected by plotting FL1-H versus FL3-A.

332

333   *Shotgun Proteomics*

334   Stably transfected SW620 cells where treated with vehicle or 500ng/ml doxycycline for 4

335   days to induce shRNA-mediated knockdown of KHSRP, as outlined above. Cells were

336   then incubated in serum-free medium for 16h, and triplicate conditioned media for each

337   condition were harvested, buffer exchanged and concentrated in a final volume of 300 μl

338   PBS using Amicon centrifugal filter units with a 3 kDa cut off membrane

339   (MilliporeSigma, Burlington, MA, USA). From each sample, 10 μg total proteins were

340   denatured with 6 M urea, disulfide bonds were reduced with 10 mM dithiothreitol (DTT),

341   and free thiols were then alkylated with 12.5 mM iodoacetamide. The final volume was

342   diluted 3-fold into 25 mM ammonium bicarbonate, and trypsin digestion was performed

343   with a 1:50 mass ratio of sequencing-grade trypsin (Promega) to total protein for 16h at

344   37°C. Samples were acidified to approximately pH=2 with formic acid, and peptides

345   desalted with C18 Desalting Tips (Rainin, Oakland, CA, USA), lyophilized, and

346   rehydrated in 0.1% formic acid. Peptide sequencing by LC-MS/MS was performed on an

347   LTQ Orbitrap Velos mass spectrometer (Thermo Fisher) equipped with a nanoACQUITY

348   (Waters, Milford, MA, USA) ultraperformance LC (UPLC) system and an EASY-Spray

349   PepMap C18 column (Thermo Fisher, ES800; 3- m bead size, 75 m by 150 mm) for

350   reversed-phase chromatography. Peptides were eluted over a linear gradient over the

351   course of 60 min from 2% to 50% (vol/vol) acetonitrile in 0.1% formic acid. MS and

352   MS/MS spectra were acquired in a data-dependent mode with up to six higher-energy

353   collisional dissociation (HCD) MS/MS spectra acquired for the most intense parent ions

354   per MS. For data analysis, MS peak lists were generated with ProteoWizard MSConvert

15

355   [58], and database searches were performed with Protein Prospector software, v.5.16.1

356   (http://prospector.ucsf.edu/prospector/mshome.htm, UCSF) [35] against the SwissProt

357   human protein database (downloaded November 1, 2017). The database was

358   concatenated with an equal number of fully randomized entries for estimation of the

359   false-discovery rate (FDR). Database searching was carried out with tolerances of 20

360   ppm for parent ions and 0.8 Da for fragment ions. For database searches, peptides

361   sequences were matched to tryptic peptides with up to two missed cleavages.

362   Carbamidomethylation of cysteine residues was used as a constant modification and

363   variable modifications included oxidation of methionine, N-terminal pyroglutamate from

364   glutamine, N-terminal acetylation, and loss of N-terminal methionine. The following

365   Protein Prospector score thresholds were selected to yield a maximum protein FDR of

366   less than 1%: a minimum "protein score" of 15 and a minimum "peptide score" of 10

367   were used; maximum expectation values of 0.01 for protein and 0.05 for peptide matches

368   were used. The list of proteins was further manually curated to include only proteins

369   identified in at least 2 out of 3 biological replicates, and with a minimum of two unique

370   peptides identified in at least one biological replicate. The t-test analysis was performed

371   on the NSAF values for proteins, calculated as previously described [36]: NSAF is the

372   number of spectral counts (SpC, the total number of MS/MS spectra) identifying a

373   protein, divided by the protein's length (L), divided by the sum of SpC/L for all proteins

374   in the experiment. GO pathway enrichment analysis was performed using the DAVID 6.8

375   database (https://david.ncifcrf.gov, accessed 01/2018) [37]. Label-free quantitation was

376   used to compare relative abundance of selected proteins. The MaxQuant software

377   package [59] was used to obtain normalized peak areas for precursor ions from extracted

16

378   ion chromatograms, using the MaxLFQ algorithm [60]. Raw mass spectrometry data files

379   and peak list files have been deposited with accession number MSV000082206 at

380   ProteoSAFE (http://massive.ucsd.edu).

381

382   *Statistics and Data Availability*

383   Statistical significance (p-values) was determined with paired Student's *t*-test or one-way

384   ANOVA for differences between experimental group mean values. Where data required

385   non-parametric statistics, the Wilcoxon signed-rank test or the Mann-Whitney U test

386   were used for paired or unpaired experimental groups (respectively). The statistical test

387   used in each experiment is specified in the relevant figure legends. Survival curves were

388   calculated according to the Kaplan–Meier method and compared using a univariate Cox–

389   Mantel log-rank test. Cox regression multivariate analysis was used to determine if

390   KHSRP expression was an independent predictor of survival (variables included: age,

391   gender, size, stage), using SPSS version 20.0 (IBM, Armonk, New York, USA). Two-

392   tailed p-values < 0.05 were considered statistically significant. Data are represented as

393   mean $\pm$ S.E.M., unless stated otherwise. Analyses were performed using Prism 6

394   (Graphpad Software, La Jolla, CA, USA) or R Studio version 1.1.463 (RStudio, Inc.,

395   Boston, MA. http://www.rstudio.com/).

396   Gene expression microarray data from this study have been deposited in the Gene

397   Expression Omnibus under accession number GSE112329. Raw data from the proteomic

398   analysis is publicly available for download on the MassIVE database

399   (ftp://massive.ucsd.edu/MSV000082206).

400

17

401 **RESULTS**

402 *KHSRP is overexpressed in Colorectal cancer*

403 To profile *KHSRP* expression in CRC, we performed a meta-analysis of CRC microarray

404 datasets from the Oncomine database. *KHSRP* expression was increased in tumor

405 compared to normal tissue across 18 CRC patient cohorts from 8 different datasets, with

406 a median adjusted p-value of 4.06×10^{-6} (Figure 1-A, and supplementary table 1). *KHSRP*

407 ranked in the top 5% of overexpressed genes in two of these analyses (from GSE9348

408 and GSE6988), and in the top 25% for 8 other analyses. In the remaining 8 analyses

409 *KHSRP* was not one of the most overexpressed genes (<25%). Only 4 of the 18 analyses

410 (from GSE5261 and GSE9689) fell below the threshold of significance. Figure 1-B

411 shows individual data from the GSE9348 cohort[38], in this dataset *KHSRP* expression was

412 increased by 2-fold on average in 70 patients with early stage (stage II) colorectal

413 carcinoma compared to healthy controls ($p=4.6 \times 10^{-11}$). In the TCGA cohort [39] (also

414 shown in Figure 1-B) *KHSRP* expression increased in 211 samples of colorectal

415 adenocarcinoma (any stage) compared to normal tissue (pooled analysis, +1.36 fold

416 change, p<0.0001). In an additional analysis of the GSE6988 cohort [40], *KHSRP*

417 expression increased in samples from either primary colon adenocarcinoma (n=25) or

418 liver metastases (n=13) compared to matched normal colon mucosa and normal liver

419 tissue, respectively (Figure 1-C).

420 Increased *KHSRP* expression in tumor tissue was not limited to CRC but remained

421 significant across a meta-analysis of 59 cancer types in the Oncomine database (median

422 adjusted p-value=0.004, supplementary table 2). In addition, to look at expression of

423 *KHSRP* across different cancer types we utilized RNA-Seq transcriptomic data available

18

424   in the TCGA database, from which differentially expressed genes were derived for 12

425   different cancer types against their respective normal tissue [23]: *KHSRP* was significantly

426   overexpressed to varying degrees in 10 cancer types (Figure 1-D), with CRC being

427   among the most highly significant (Fold change overexpression=1.53, FDR-adjusted p-

428   value=5.02×10^{-11}). The colon has also generally a high tissue-specific cancer expression

429   of *KHSRP* across the entire spectrum of cancers represented in the TCGA

430   (Supplementary Figure 1). However, the frequency of genetic alterations (mutations,

431   deletions or amplifications) in *KHSRP* is less than 2% in the CRC TCGA dataset

432   (Supplementary Figure 2), which is low considering that the three most commonly

433   mutated genes in CRC (APC, TP53 and KRAS) have mutation frequencies of 81%, 60%

434   and 43%, respectively [39]. Collectively, these *in silico* data provided a framework to

435   support additional investigations into the putative role of KHSRP in the pathogenesis of

436   CRC.

437

438   *KHSRP is a marker of poor prognosis in colorectal cancer*

439   Next, we investigated KHSRP protein expression in primary tissue samples. Protein

440   extracts were obtained from fresh frozen colorectal tumor and normal adjacent tissue of

441   16 patients following surgical resection. WB analysis revealed a two-fold increased

442   expression (p=0.0009) of KHSRP in stage IV CRC, compared to matched uninvolved

443   (*adjacent normal*) tissue control (Figure 2-A). Analysis of stage II-III patients showed a

444   similar trend (Supplementary Figure 3-A). The KHSRP antibody used for WB detected a

445   single band at ~82 kDa (Supplementary Figure 3-B), the predicted size for the protein,

446   confirming the specificity of this antibody which has been previously orthogonally

447   validated (Human Protein Atlas).

448   To expand on these initial results, and to compensate for the limited number of samples

449   analyzed by WB, we then performed immunohistochemistry (IHC) to investigate KHSRP

450   expression in tissue obtained from a cohort of 62 patients with late stage (III-IV) CRC [25].

451   KHSRP was expressed in both the epithelial and stromal compartments (Figure 2-B).

452   KHSRP expression was significantly higher in tumor versus matched normal tissue in

453   both epithelium and stroma (Figure 2-B and 2-C). This was mainly due to differences in

454   percentage of positive cells, while the intensity values were similar overall

455   (Supplementary Figure 3-C). There were more KHSRP positive cells in the epithelium

456   than in the stroma (Figure 2-C), however the increased expression in tumor compared to

457   normal was more prominent in the stroma, as indicated by a higher tumor-to-normal

458   (T/N) ratio of KHSRP staining in the stroma (Supplementary Figure 3-D). When T/N

459   ratios are plotted separately for each patient, the majority had increased KHSRP

460   expression in the tumor compared to matched normal tissue (Supplementary Figure 3-C

461   and D). High KHSRP expression (more than 20% increase in tumor vs. normal tissue)

462   was associated with decreased overall survival (Figure 2-D, HR=3.74, 95% CI = 1.43-

463   22.97, p=0.0138) in a univariate analysis. Multivariate Cox regression analysis

464   demonstrated that higher KHSRP expression in tumor epithelium was an independent

465   predictor of poor survival, associated with a 2.5-fold increase in disease-associated death

466   ($p = 0.009$). As most of the patients in this cohort were diagnosed with metastatic

467   disease, relapse-free survival was not included in the analysis. There was also a weak but

20

468 significant correlation of the KHSRP T/N ratio with tumor size in both epithelium

469 (Pearson r=0.43, p=0.0016) and stroma (Pearson r=0.38, p=0.0051).

470 For a limited number of cases (N=10 patients), tissue was also available for the

471 corresponding liver metastasis. In both the epithelium and the stroma, the KHSRP score

472 was significantly higher in the metastatic tumor tissue compared to matched normal liver

473 (Supplementary Figure 4-A). Overall the metastasis-to-normal (M/N) ratio was positive

474 for all except one patient, but the difference between metastasis and normal was vastly

475 larger in the stroma than in the epithelium, with higher M/N ratios (Supplementary Figure

476 4-B, C, and D); this can be explained by the low expression of KHSRP in the normal

477 liver stroma. Comparison of matched primary tumor and liver metastasis (n=10 patients)

478 showed that the expression of KHSRP is unchanged in the epithelial cells upon

479 migration, however the metastatic stroma has a higher expression compared to the stroma

480 in the primary tumor (Supplementary Figure 4-E and F).

481 To validate the prognostic role of *KHSRP* in an independent cohort, we used a published

482 dataset of 566 patients with stage I-IV CRC in the French *Cartes d'Identité des Tumeurs*

483 program [41]. *KHSRP* expression was increased in tumor compared to uninvolved (*normal*)

484 tissue, in accordance with our previous data (Figure 3-A). Survival analysis was restricted

485 to stage II-III patients following recommendations from the original authors: high

486 *KHSRP* expression was associated with a decreased 5-year relapse-free survival in this

487 cohort (Figure 3-B, HR=1.95, 95% CI = 1.21-2.60, p=0.0038). *KHSRP* expression was

488 not associated with other tumor features such as loss of mismatch repair (MMR) proteins,

489 or mutations in *KRAS* or *BRAF* (Figure 3-C).

490   To independently validate the expression of KHSRP in epithelial and stromal

491   compartments from the IHC data, we took advantage of a transcriptomic dataset of 13

492   CRC and normal colonic mucosa samples, in which epithelial cells and stromal cells had

493   been micro-dissected by laser capture and profiled separately [42]. *KHSRP* was

494   significantly overexpressed in the micro-dissected epithelial tumor compared with tumor

495   stromal areas, and to a lesser extent in the epithelium compared with the stroma of the

496   normal colonic mucosa, in accordance with our IHC data (Figure 3-D). Although there

497   was a trend in increased *KHSRP* expression in tumor versus normal tissue for both the

498   epithelium and the stroma, these comparisons were not statistically significant, probably

499   due to the low number of samples. Finally, we used transcriptomic data from purified cell

500   populations that were isolated from 14 dissociated human primary CRC samples [43].

501   Fluorescence-activated cell sorting (FACS) was used to enrich for epithelial cancer cells

502   (EpCAM$^+$), leukocytes (CD45$^+$), endothelial cells (CD31$^+$) and cancer-associated

503   fibroblasts (CAFs) (FAP$^+$). The expression of *KHSRP* was higher in the Epcam$^+$

504   epithelial cell population compared to the stromal cells, with the CD31$^+$ endothelial cells

505   showing the lowest *KHSRP* expression (Figure 3-E).

506

507   *KHSRP is involved in promoting the growth and survival of CRC cells*

508   Next, we turned to an *in vitro* model of CRC to investigate the mechanistic role of

509   KHSRP in epithelial cancer cells. We used the SW480 cell line, which was derived from

510   a primary tumor of a moderately differentiated Duke B colon adenocarcinoma [44]. These

511   cells constitutively express KHSRP with a diffuse intracellular pattern (Supplementary

512   Figure 5). An initial screen with a pool of 4 different siRNA sequences targeting *KHSRP*

22

513   effectively silenced mRNA expression by 50% (Figure 4-A), and stably reduced

514   proliferation when cells were monitored in real-time for 6 days (Figure 4-B). We then

515   confirmed and extended these initial results using an independent siRNA that effectively

516   knocked down KHSRP protein expression (Figure 4-C): silencing of KHSRP reduced

517   cell growth as measured by neutral red uptake (Figure 4-D), reduced clonogenic potential

518   measured by both number of colonies (surviving fraction) and total colony area (Figure

519   4-E), and reduced spheroid formation in 3D Matrigel cultures (Figure 4-F). Similar

520   results were obtained by silencing KHSRP in an independent cell line model, SW620,

521   which was isolated from a lymph node after multiple metastases occurred in the same

522   patient (Supplementary Figure 6). Furthermore, we also developed a conditional

523   (doxycycline-dependent) shRNA-based stable knockdown model from both SW480 and

524   SW620 cell lines, which further confirmed results obtained with the transient silencing

525   model (Supplementary Figure 7).

526

527   *KHSRP regulates cell cycle and affects the tumor-promoting microenvironment*

528   To investigate the mechanisms involved in the tumor-promoting effects of KHSRP, we

529   started by analyzing the transcriptomic profiles of siRNA-transfected KHSRP

530   knockdown SW480 cells, compared to scramble control transfection (Supplementary

531   Figure 8). This preliminary analysis identified 135 genes differentially regulated upon

532   KHSRP knockdown (Supplementary Table 3), which are enriched for functional

533   categories related to cell growth and survival (specifically, cell cycle and apoptosis), as

534   well as cell signaling and known KHSRP-related functions (translation, protein

535   metabolism), as shown by GO terms analysis (Supplementary Figure 8-B). Three

23

536   representative genes (*FABP3, MTF1, SMG1*) were further validated by qRT-PCR

537   (Supplementary Figure 8-C). The protein interactions inferred from the differentially

538   expressed transcripts (Supplementary Figure 8-D) [31] included gene networks for cell

539   cycle, transcription and protein synthesis, further suggesting a role for KHSRP in cell

540   cycle control. This prompted us to investigate the effects of KHSRP knockdown on the

541   cell cycle distribution of SW480 cells. KHSRP siRNA caused an increase in the G0/G1

542   population (Figure 5-A), concomitant with a decrease in the mitotic index, measured as

543   the number of cells positively stained by phospho-Histone H3 (Figure 5-B), indicative of

544   a reduced proliferation rate due to a delayed G1/S transition. These data, together with

545   that presented in Figure 4, suggest an involvement of KHSRP in driving epithelial cells

546   into mitosis, thereby increasing proliferation. Furthermore, stable knockdown of KHSRP

547   resulted in reduced expression of c-Myc (Supplementary Figure 7-C).

548   KHSRP has been involved in regulating cell migration and invasion in different non-

549   colorectal tumor models [72]. However, silencing of *KHSRP* did not influence cell

550   migration (as measured by a wound-healing assay) in SW480 cells (Figure 5-C), and did

551   not impact on the ability of the metastatic SW620 cells to invade through a Matrigel layer

552   (Figure 5-D). Nevertheless, the cell culture supernatant from SW620 cells with silenced

553   KHSRP was able to reduce endothelial cell growth in a tube-forming assay (Figure 5-E),

554   suggesting that KHSRP may regulate the secretion of specific extracellular signaling

555   mediators promoting angiogenesis. In support of this, CRC cells with silenced KHSRP

556   expression secreted reduced levels of IL-8 and VEGF (Figure 5-F).

557

558   *KHSRP regulates the secretion of multiple proteins by tumor cells*

24

559   To further investigate the involvement of KHSRP in regulating the extracellular

560   microenvironment, we employed the conditional KHSRP knockdown cell line model

561   (described in Supplementary Figure 7) to probe the secretome of the metastatic SW620

562   cell line upon doxycycline-induced, shRNA-mediated knockdown of KHSRP. Shotgun

563   proteomics analysis identified 191 proteins, of which 27 where unique for the control

564   samples (i.e. undetected in the doxycycline-treated samples), and 16 where unique for the

565   doxycycline-treated samples (undetected in the control samples). Statistical analysis of

566   spectral counts data showed that 40 proteins were significantly under-expressed or over-

567   expressed by more than 50% upon doxycycline treatment; the majority of these

568   differentially regulated proteins (33 out of 40) were down-regulated (Figure 6-A, and

569   Supplementary Table 4). GO term analysis of the 40 highly differentially regulated

570   proteins revealed an enrichment for known terms related to KHSRP function (e.g. RNA

571   binding, spliceosome) and expected terms related to the cellular location of the probed

572   sample (e.g. membrane and extracellular space); interestingly, there was also a significant

573   enrichment for terms related to the pro-tumorigenic extracellular microenvironment (e.g.

574   exosomes, focal adhesion, cell migration), as well as terms broadly related to the immune

575   response and specifically to myeloid cell activation (Figure 6-B). Using label-free

576   quantitation, we validated specific examples of candidate proteins whose secretion is

577   deregulated upon KHSRP knockdown (Figure 6-C). Downregulated proteins included:

578   the Rho-specific guanine nucleotide dissociation inhibitor-α (RhoGDIα), a master

579   regulator of Rho GTPases that promotes cell migration [45]; the S100 calcium binding

580   protein A11 (S100-A11), an endoplasmic reticulum-associated calcium-binding protein

581   that regulates cell death [61] ; Ephrin-B2, a membrane-anchored proteins which activates

582 paracrine tyrosine kinase signaling impacting proliferation, migration and response to

583 therapy [62]. Conversely, upregulated proteins included: the actin-binding protein Villin-1,

584 a protein which upon cellular stress induces intestinal epithelial cell death by necroptosis,

585 as well as increased inflammation [46]; ADAMTS-like protein 2, an extracellular matrix-

586 residing protein with roles in matrix remodeling and modulation of metalloprotease

587 activity [63]. The KHSRP-dependent downregulation of protein S100-AA11 and Ephrin-B2

588 was further confirmed by Western blot (Figure 6-D).

589

590 **DISCUSSION**

591

592 The role of KHSRP in the pathogenesis and progression of CRC has not been extensively

593 studied so far. Here, we provide orthogonal evidence supporting a pro-tumorigenic role

594 of KHSRP in CRC, based on extensive mining of multiple publicly available data sets,

595 analysis of patient samples and *in vitro* cell line models. Our results are in agreement

596 with recent reports in osteosarcoma [13], lung cancer [47,12], papillary thyroid cancer [67], and

597 esophageal squamous cell carcinoma [48].

598 Our IHC data provide evidence for KHSRP overexpression in the context of both primary

599 tumor and metastasis. However, we surprisingly found that KHSRP remained unchanged

600 in the metastatic epithelial cells infiltrating the liver compared to that of the primary site,

601 while stromal expression was strongly upregulated from a very low basal expression in

602 the normal hepatic stroma. This suggests that invading epithelial cells are capable of

603 modifying the local microenvironment, or alternatively that a local increase in KHSRP

604 expression pre-exists and might contribute to the metastatic niche [49]. In fact, KHSRP is

26

605   involved in the transient expression of the chemokine CX3CL1 in liver epithelial cells in

606   response to IFN-γ stimulation [50], contributing to the control of cellular homeostasis

607   during inflammation and macrophage infiltration. Our proteomic data provide additional

608   evidence in support of a putative role for KHSRP in regulating the tumor

609   microenvironment, with implications for tumor progression and metastasis. To the best of

610   our knowledge, this represents the first report of such an effect for KHSRP in shaping the

611   extracellular environment of epithelial cells. Although preliminarily, our data also

612   suggest a possible role for KHSRP in regulating the immune response, both at the gene-

613   expression level and at the secretome level (as reported here by microarray and proteomic

614   data). This suggests the intriguing idea that the reported negative prognostic role of

615   KHSRP could be partly ascribed to a regulation of immune evasion mechanisms, and

616   provides a rationale for reconciling the oncogenic roles of KHSRP with its other known

617   functions in shaping innate immunity during viral infection and inflammation. In fact,

618   recent work from Bezzi and colleagues using genetically engineered mouse models of

619   prostate cancer driven by the loss of specific tumor suppressor genes has shown that the

620   genetic landscape of epithelial tumor cells can directly influence the composition of

621   immune cell infiltrates in primary tumors, through mechanisms based on distinct

622   chemokine pools resulting from aberrant transcriptional and signaling programs [51].

623   Furthermore, knockout mice studies have shown a requirement of KHSRP for the post-

624   transcriptional control of type-1 and 3 IFN genes [52,68]. Recent mechanistic evidence

625   proved that KHSRP can directly associate and negatively regulate retinoic acid-inducible

626   gene I (RIG-I) receptor signaling during viral infection; this exert a negative control on

627   the activation of type-I IFN innate immune responses during the recognition of pathogen-

628   associated molecular patterns (PAMPs) encoded by viral RNA [53]. This could have

629   potential consequences for the control of anti-tumor innate immune responses, as similar

630   mechanisms are involved in the recognition of damage-associated molecular pattern

631   (DAMPs) and the immunogenicity of tumor cells [54]. Consistently, KHSRP has been

632   implicated in the innate immune response to oncogenic pathogens such as *H. pylori* in the

633   stomach [69] and HPV in the uterus [70].

634   One question that remains open is how overexpression of KHSRP is controlled, given

635   that our data suggest it is not the product of a driver mutation event. A possible

636   explanation could come from studying another RBP which is commonly overexpressed in

637   CRC, Musashi (MSI)-1 [64]. Spears and colleagues have demonstrated a reciprocal

638   interaction between MSI-1 and APC through a double-negative feedback loop, whereby

639   MSI-1 contributes to APC inactivation, an early initiating driver event in the majority of

640   sporadic CRC [65]. This suggests the intriguing hypothesis that increased expression of

641   RBPs (perhaps not limited to MSI-1) might co-evolve as a passenger non-mutational

642   event early in the adenoma-to-carcinoma sequence. Another hypothesis, related to the

643   effect of KHSRP on the tumor microenvironment that our data begins to suggest, is that

644   deregulation of RBPs could be selected for indirectly through negative regulation of

645   immune surveillance, a relevant force shaping cancer evolution. In support of an early

646   event hypothesis, our *in silico* data suggest that KHSRP expression is unchanged in

647   hypermutated, microsatellite-instable CRC, a type of tumor characterized by a

648   proportionally increased burden of passenger mutations [66]. Nevertheless, specific

649   mutations in KHSRP could still impact on protein functions and show increased

650   frequencies in some cancer types [72]. Furthermore, although limited to a very small

651 fraction of the dataset, the analysis from the Oncomine database provides information on

652 colorectal mucinous carcinoma, a distinct form of CRC found in 10-15% of patients and

653 characterized by distinct molecular features and mutation events [71] in this subset of

654 patients, *KHSRP* was not one of the most overexpressed genes, suggesting that it might

655 be an event specifically related to adenocarcinoma.

656 There is also considerable interest in the suitability of different members of the

657 translational control machinery as potential therapeutic targets in cancer, although the

658 evidence so far is at the preclinical stage [55]. Small molecule inhibition of the RBP HuR in

659 APC^{Min} mice led to a reduction in small intestinal tumor formation, and a concomitant

660 reduction in c-myc expression [56]. Another recent report showed that inhibition of

661 mTORC1 using rapamycin caused regression of established APC-deficient intestinal

662 tumors, suggesting that inhibition of translation elongation using existing clinically

663 approved drugs might benefit patients at high risk of developing CRC [57]. Our loss-of-

664 function data in *in vitro* CRC models suggest that KHSRP could also have putative

665 therapeutic implications and warrants further investigation into the suitability of such an

666 approach.

667 In conclusion, our report sheds light onto the molecular role of KHSRP in CRC,

668 providing for the first time comprehensive data in support of a pro-tumorigenic role of

669 this RBP through direct modulation of epithelial cell phenotype and indirect modulation

670 of the tumor microenvironment.

671

672 **ADDITIONAL INFORMATION**

673

674   **Ethics Approval and Consent to Participate:** All human tissue used in this study was

675   obtained with the informed written consent of the patient; ethical approval was granted

676   by the Research and Ethics Committee of St. Vincent's University Hospital (Dublin,

677   Ireland). The study was performed in accordance with the Declaration of Helsinki.

678   **Authors contribution:** Study concept and design: FC, EJR, JOS, GAD; acquisition of

679   data: FC, KO, MT, LB, FON, RP, SN, PM; analysis and interpretation of data: FC, KO,

680   JP, KK, SF, KS, JOS, EJR; drafting and critical revision of the manuscript: FC, KO, JOS,

681   EJR, CSC; funding and support: EJR, KS, GAD, DF, CSC; administrative, technical or

682   material support: MT, BN; study supervision: EJR, GAD.

683   **Acknowledgments:** We gratefully acknowledge Dr. Giselle Knudsen (UCSF), Dr.

684   Sudipto Das (RCSI), Dr. Janet McCormack (UCD), Dr. Barry Moran (TCD), and Dr.

685   Dimitri Scholz (UCD) for technical assistance at various stages of this project. We also

686   wish to thank Dr. Martin Kampmann (UCSF) for providing advice and material for the

687   generation of the conditional knockdown cell lines, and Dr. Quing Xiong (Southwest

688   University, Chongqing, China) for providing access to their data. The bioinformatic

689   expertise of Dr. Tim Hacker and Sam Ivry (UCSF) is also acknowledged. Mass

690   spectrometry was performed in collaboration with the UCSF Mass Spectrometry Facility

691   (directed by Prof. Alma Burlingame).

692   **Availability of Data and Material:** Gene expression microarray data from this study

693   have been deposited in the Gene Expression Omnibus under accession number

694   GSE112329. Raw data from the proteomic analysis is publicly available for download on

695   the MassIVE database (ftp://massive.ucsd.edu/MSV000082206).

696

697
698 References

699

700 1. Truitt ML, Ruggero D: New frontiers in translational control of the cancer
701 genome. Nature reviews Cancer 2016, 16:288–304

702 2. Gerstberger S, Hafner M, Tuschl T: A census of human RNA-binding proteins.
703 Nature reviews. Genetics 2014, 15:829–845

704 3. Gherzi R, Lee K-Y, Briata P, Wegmüller D, Moroni C, Karin M, Chen C-Y: A
705 KH Domain RNA Binding Protein, KSRP, Promotes ARE-Directed mRNA
706 Turnover by Recruiting the Degradation Machinery. Molecular cell 2004, 14:571–
707 583

708 4. Chou CF, Mulky A, Maitra S, Lin WJ, Gherzi R, Kappes J, Chen CY: Tethering
709 KSRP, a Decay-Promoting AU-Rich Element-Binding Protein, to mRNAs Elicits
710 mRNA Decay. Molecular and cellular biology 2006, 26:3695–3706

711 5. Dhamija S, Kuehne N, Winzen R, Doerrie A, Dittrich-Breiholz O, Thakur BK,
712 Kracht M, Holtmann H: Interleukin-1 activates synthesis of interleukin-6 by
713 interfering with a KH-type splicing regulatory protein (KSRP)-dependent
714 translational silencing mechanism. The Journal of biological chemistry 2011,
715 286:33279–33288

716 6. Trabucchi M, Briata P, Garcia-Mayoral M, Haase AD, Filipowicz W, Ramos A,
717 Gherzi R, Rosenfeld MG: The RNA-binding protein KSRP promotes the
718 biogenesis of a subset of microRNAs. Nature 2009, 459:1010–1014

719 7. Guo L, Zaysteva O, Nie Z, Mitchell NC, Amanda Lee JE, Ware T, Parsons L,
720 Luwor R, Poortinga G, Hannan RD, Levens DL, Quinn LM: Defining the essential
721 function of FBP/KSRP proteins: Drosophila Psi interacts with the mediator
722 complex to modulate MYC transcription and tissue growth. Nucleic acids
723 research 2016,

724 8. Winzen R, Thakur BK, Dittrich-Breiholz O, Shah M, Redich N, Dhamija S,
725 Kracht M, Holtmann H: Functional analysis of KSRP interaction with the AU-rich
726 element of interleukin-8 and identification of inflammatory mRNA targets.
727 Molecular and cellular biology 2007, 27:8388–8400

728 9. Gherzi R, Chen C-Y, Ramos A, Briata P: KSRP controls pleiotropic cellular
729 functions. Seminars in cell & developmental biology 2014, 34:2–8

730 10. Briata P, Chen C-Y, Ramos A, Gherzi R: Functional and molecular insights
731 into KSRP function in mRNA decay. Biochimica et biophysica acta 2013,
732 1829:689–694

733   11. Suswam EA, Nabors LB, Huang Y, Yang X, King PH: IL-1beta induces
734   stabilization of IL-8 mRNA in malignant breast cancer cells via the 3' untranslated
735   region: Involvement of divergent RNA-binding factors HuR, KSRP and TIAR.
736   International journal of cancer Journal international du cancer 2005, 113:911–
737   919

738   12. Tong L, Luo Y, Wei T, Guo L, Wang H, Zhu W, Zhang J: KH-type splicing
739   regulatory protein (KHSRP) contributes to tumorigenesis by promoting miR-26a
740   maturation in small cell lung cancer. Molecular and cellular biochemistry 2016,
741   422:61–74

742   13. Pruksakorn D, Teeyakasem P, Klangjorhor J, Chaiyawat P, Settakorn J,
743   Diskul-Na-Ayudthaya P, Chokchaichamnankit D, Pothacharoen P, Srisomsap C:
744   Overexpression of KH-type splicing regulatory protein regulates proliferation,
745   migration, and implantation ability of osteosarcoma. International journal of
746   oncology 2016, 49:903–912

747   14. Santarosa M, Del Col L, Viel A, Bivi N, D'Ambrosio C, Scaloni A, Tell G,
748   Maestro R: BRCA1 modulates the expression of hnRNPA2B1 and KHSRP. Cell
749   cycle (Georgetown, Tex) 2010, 9:4666–4673

750   15. Giovarelli M, Bucci G, Pasero M, Gherzi R, Briata P: KSRP silencing favors
751   neural differentiation of P19 teratocarcinoma cells. Biochimica et biophysica acta
752   2013, 1829:469–479

753   16. Walerych D, Lisek K, Sommaggio R, Piazza S, Ciani Y, Dalla E, Rajkowska
754   K, Gaweda-Walerych K, Ingallina E, Tonelli C, Morelli MJ, Amato A, Eterno V,
755   Zambelli A, Rosato A, Amati B, Wiśniewski JR, Del Sal G: Proteasome
756   machinery is instrumental in a common gain-of-function program of the p53
757   missense mutants in cancer. Nature cell biology 2016, 18:897–909

758   17. Puppo M, Bucci G, Rossi M, Giovarelli M, Bordo D, Moshiri A, Gorlero F,
759   Gherzi R, Briata P: miRNA-Mediated KHSRP Silencing Rewires Distinct Post-
760   transcriptional Programs during TGF-β-Induced Epithelial-to-Mesenchymal
761   Transition. Cell reports 2016, 16:967–978

762   18. Fagoonee S, Picco G, Orso F, Arrigoni A, Longo DL, Forni M, Scarfò I,
763   Cassenti A, Piva R, Cassoni P, Silengo L, Tolosano E, Aime S, Taverna D,
764   Pandolfi PP, Brancaccio M, Medico E, Altruda F: The RNA-binding protein
765   ESRP1 promotes human colorectal cancer progression. Oncotarget 2017,
766   8:10007–10024

767   19. Blanc V, Henderson JO, Newberry RD, Xie Y, Cho S-J, Newberry EP,
768   Kennedy S, Rubin DC, Wang HL, Luo J, Davidson NO: Deletion of the AU-rich
769   RNA binding protein Apobec-1 reduces intestinal tumor burden in Apc(min) mice.
770   Cancer Research 2007, 67:8565–8573

771   20. Sureban SM, May R, George RJ, Dieckgraefe BK, McLeod HL, Ramalingam
772   S, Bishnupuri KS, Natarajan G, Anant S, Houchen CW: Knockdown of RNA
773   Binding Protein Musashi-1 Leads to Tumor Regression In Vivo. Gastroenterology
774   2008, 134:1448–1458.e2

775   21. Wang S, Li N, Yousefi M, Nakauka-Ddamba A, Li F, Parada K, Rao S,
776   Minuesa G, Katz Y, Gregory BD, Kharas MG, Yu Z, Lengner CJ: Transformation
777   of the intestinal epithelium by the MSI2 RNA-binding protein. Nature
778   communications 2015, 6:6517

779   22. Preitner N, Quan J, Nowakowski DW, Hancock ML, Shi J, Tcherkezian J,
780   Young-Pearse TL, Flanagan JG: APC is an RNA-binding protein, and its
781   interactome provides a link to neural development and microtubule assembly.
782   Cell 2014, 158:368–382

783   23. Peng L, Bian X-W, Li DK, Xu C, Wang GM, Xia QY, Xiong Q: Large-scale
784   RNA-Seq Transcriptome Analysis of 4043 Cancers and 548 Normal Tissue
785   Controls across 12 TCGA Cancer Types. Scientific reports 2015, 5:13413

786   24. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y,
787   Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N: Integrative
788   Analysis of Complex Cancer Genomics and Clinical Profiles Using the
789   cBioPortal. Science Signaling 2013, 6:pl1–pl1

790   25. Noonan SA, Morrissey ME, Martin P, Biniecka M, Ó'Meachair S, Maguire A,
791   Tosetto M, Nolan B, Hyland J, Sheahan K, O'Donoghue D, Mulcahy H, David F,
792   O'Sullivan J: Tumour vasculature immaturity, oxidative damage and systemic
793   inflammation stratify survival of colorectal cancer patients on bevacizumab
794   treatment. Oncotarget 2018, 5:

795   26. Detre S, Saclani Jotti G, Dowsett M: A "quickscore" method for
796   immunohistochemical semiquantitation: validation for oestrogen receptor in
797   breast carcinomas. Journal of clinical pathology 1995, 48:876–878

798   27. Bassik MC, Kampmann M, Lebbink RJ, Wang S, Hein MY, Poser I,
799   Weibezahn J, Horlbeck MA, Chen S, Mann M, Hyman AA, Leproust EM,
800   McManus MT, Weissman JS: A systematic mammalian genetic interaction map
801   reveals pathways underlying ricin susceptibility. Cell 2013, 152:909–922

802   28. Kampmann M, Bassik MC, Weissman JS: Functional genomics platform for
803   pooled screening and generation of mammalian genetic interaction maps. Nature
804   protocols 2014, 9:1825–1847

805   29. Meerbrey KL, Hu G, Kessler JD, Roarty K, Li MZ, Fang JE, Herschkowitz JI,
806   Burrows AE, Ciccia A, Sun T, Schmitt EM, Bernardi RJ, Fu X, Bland CS, Cooper
807   TA, Schiff R, Rosen JM, Westbrook TF, Elledge SJ: The pINDUCER lentiviral
808   toolkit for inducible RNA interference in vitro and in vivo. Proceedings of the
809   National Academy of Sciences 2011, 108:3665–3670

810   30. Stöckel D, Kehl T, Trampert P, Schneider L, Backes C, Ludwig N, Gerasch A,
811   Kaufmann M, Gessler M, Graf N, Meese E, Keller A, Lenhof H-P: Multi-omics
812   enrichment analysis using the GeneTrail2 web service. Bioinformatics (Oxford,
813   England) 2016, 32:1502–1508

814   31. Damian Szklarczyk JHMHCMKSWMSASNTDARPBLJJCVM: The STRING
815   database in 2017: quality-controlled protein–protein association networks, made
816   broadly accessible. Nucleic acids research 2017, 45:D362

817   32. Sylvain Brohée JVH: Evaluation of clustering algorithms for protein-protein
818   interaction networks. BMC Bioinformatics 2006, 7:488

819   33. Repetto G, del Peso A, Zurita JL: Neutral red uptake assay for the estimation
820   of cell viability/cytotoxicity. Nature protocols 2008, 3:1125–1131

821   34. Lee GY, Kenny PA, Lee EH, Bissell MJ: Three-dimensional culture models of
822   normal and malignant breast epithelial cells. Nature methods 2007, 4:359–365

823   35. Chalkley RJ, Baker PR, Medzihradszky KF, Lynn AJ, Burlingame AL: In-
824   depth analysis of tandem mass spectrometry data from disparate instrument
825   types. Molecular & cellular proteomics : MCP 2008, 7:2386–2398

826   36. Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP:
827   Statistical analysis of membrane proteome expression changes in
828   Saccharomyces cerevisiae. Journal of proteome research 2006, 5:2339–2347

829   37. Huang DW, Sherman BT, Lempicki RA: Systematic and integrative analysis
830   of large gene lists using DAVID bioinformatics resources. Nature protocols 2009,
831   4:44–57

832   38. Hong Y, Downey T, Eu KW, Koh PK, Cheah PY: A "metastasis-prone"
833   signature for early-stage mismatch-repair proficient sporadic colorectal cancer
834   patients and its implications for possible therapeutics. Clinical & experimental
835   metastasis 2010, 27:83–90

836   39. Cancer Genome Atlas Network: Comprehensive molecular characterization
837   of human colon and rectal cancer. Nature 2012, 487:330–337

838   40. Ki DH, Jeung H-C, Park CH, Kang SH, Lee GY, Lee WS, Kim NK, Chung
839   HC, Rha SY: Whole genome analysis for liver metastasis gene signatures in
840   colorectal cancer. International journal of cancer Journal international du cancer
841   2007, 121:2005–2012

842   41. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, Etienne-
843   Grimaldi M-C, Schiappa R, Guenot D, Ayadi M, Kirzin S, Chazal M, Fléjou J-F,
844   Benchimol D, Berger A, Lagarde A, Pencreach E, Piard F, Elias D, Parc Y,
845   Olschwang S, Milano G, Laurent-Puig P, Boige V: Gene expression classification

846 of colon cancer into molecular subtypes: characterization, validation, and
847 prognostic value. PLoS medicine 2013, 10:e1001453

848 42. Nishida N, Nagahara M, Sato T, Mimori K, Sudo T, Tanaka F, Shibata K, Ishii
849 H, Sugihara K, Doki Y, Mori M: Microarray analysis of colorectal cancer stromal
850 tissue reveals upregulation of two oncogenic miRNA clusters. Clinical Cancer
851 Research 2012, 18:3054–3070

852 43. Calon A, Espinet E, Palomo-Ponce S, Tauriello DVF, Iglesias M, Céspedes
853 MV, Sevillano M, Nadal C, Jung P, Zhang XHF, Byrom D, Riera A, Rossell D,
854 Mangues R, Massagué J, Sancho E, Batlle E: Dependency of colorectal cancer
855 on a TGF-β-driven program in stromal cells for metastasis initiation. Cancer cell
856 2012, 22:571–584

857 44. Leibovitz A, Stinson JC, McCombs WB, McCoy CE, Mazur KC, Mabry ND:
858 Classification of human colorectal adenocarcinoma cell lines. Cancer Research
859 1976, 36:4562–4569

860 45. Zhao L, Wang H, Li J, Liu Y, Ding Y: Overexpression of Rho GDP-
861 Dissociation Inhibitor Alpha Is Associated with Tumor Progression and Poor
862 Prognosis of Colorectal Cancer. Journal of proteome research 2008, 7:3994–
863 4003

864 46. Roy S, Esmaeilniakooshkghazi A, Patnaik S, Wang Y, George SP, Ahrorov
865 A, Hou JK, Herron AJ, Sesaki H, Khurana S: Villin-1 and Gelsolin Regulate
866 Changes in Actin Dynamics That Affect Cell Survival Signaling Pathways and
867 Intestinal Inflammation. Gastroenterology 2017,

868 47. Bikkavilli RK, Zerayesus SA, Van Scoyk M, Wilson L, Wu P-Y, Baskaran A,
869 Tang K, Raheem S, Samuelson BA, Reddy NM, Reddy SP, Cool CD, Kosmider
870 B, Avasarala S, Winn RA: KSRP promotes post-transcriptional destabilization of
871 Spry4 transcripts in non-small cell lung cancer. The Journal of biological
872 chemistry 2017,

873 48. Fujita Y, Masuda K, Hamada J, Shoda K, Naruto T, Hamada S, Miyakami Y,
874 Kohmoto T, Watanabe M, Takahashi R, Tange S, Saito M, Kudo Y, Fujiwara H,
875 Ichikawa D, Tangoku A, Otsuji E, Imoto I: KH-type splicing regulatory protein is
876 involved in esophageal squamous cell carcinoma progression. Oncotarget 2017,
877 8:101130–101145

878 49. Descot A, Oskarsson T: The molecular composition of the metastatic niche.
879 Experimental cell research 2013, 319:1679–1686

880 50. Xia Z, Lu Y, Li X, Mao T, Chen X-M, Zhou R: Upregulation of KSRP by miR-
881 27b provides IFN-γ-induced post-transcriptional regulation of CX3CL1 in liver
882 epithelial cells. Scientific reports 2015, 5:17590

883   51. Bezzi M, Seitzer N, Ishikawa T, Reschke M, Chen M, Wang G, Mitchell C, Ng
884   C, Katon J, Lunardi A, Signoretti S, Clohessy JG, Zhang J, Pandolfi PP: Diverse
885   genetic-driven immune landscapes dictate tumor progression through distinct
886   mechanisms. Nature medicine 2018, 24:165–175

887   52. Lin W-J, Zheng X, Lin C-C, Tsao J, Zhu X, Cody JJ, Coleman JM, Gherzi R,
888   Luo M, Townes TM, Parker JN, Chen C-Y: Posttranscriptional control of type I
889   interferon genes by KSRP in the innate immune response against viral infection.
890   Molecular and cellular biology 2011, 31:3196–3207

891   53. Soonthornvacharin S, Rodriguez-Frandsen A, Zhou Y, Galvez F, Huffmaster
892   NJ, Tripathi S, Balasubramaniam VRMT, Inoue A, de Castro E, Moulton H, Stein
893   DA, Sánchez-Aparicio MT, De Jesus PD, Nguyen Q, König R, Krogan NJ,
894   García-Sastre A, Yoh SM, Chanda SK: Systems-based analysis of RIG-I-
895   dependent signalling identifies KHSRP as an inhibitor of RIG-I receptor
896   activation. Nature microbiology 2017, 2:17022

897   54. Green DR, Ferguson T, Zitvogel L, Kroemer G: Immunogenic and tolerogenic
898   cell death. Nature reviews. Immunology 2009, 9:353–363

899   55. Silvera D, Formenti SC, Schneider RJ: Translational control in cancer. Nature
900   reviews Cancer 2010, 10:254–266

901   56. Lang M, Berry D, Passecker K, Mesteri I, Bhuju S, Ebner F, Sedlyarov V,
902   Evstatiev R, Dammann K, Loy A, Kuzyk O, Kovarik P, Khare V, Beibel M, Roma
903   G, Meisner-Kober N, Gasche C: HuR small molecule inhibitor elicits differential
904   effects in adenomatosis polyposis and colorectal carcinogenesis. Cancer
905   Research 2017, canres.1726.2015

906   57. Faller WJ, Jackson TJ, Knight JRP, Ridgway RA, Jamieson T, Karim SA,
907   Jones C, Radulescu S, Huels DJ, Myant KB, Dudek KM, Casey HA, Scopelliti A,
908   Cordero JB, Vidal M, Pende M, Ryazanov AG, Sonenberg N, Meyuhas O, Hall
909   MN, Bushell M, Willis AE, Sansom OJ: mTORC1-mediated translational
910   elongation limits intestinal tumour initiation and growth. Nature 2015, 517:497–
911   500

912   58. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S,
913   Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman
914   N, Frewen B, Baker TA, Brusniak MY, Paulse C, Creasy D, Flashner L, Kani K,
915   Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J,
916   Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick
917   T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M,
918   Tabb DL, Mallick P: A cross-platform toolkit for mass spectrometry and
919   proteomics. Nat Biotechnology 2012, 30(10):918-20

920   59. Tyanova S, Temu T, Cox J: The MaxQuant computational platform for mass
921   spectrometry-based shotgun proteomics. Nature Protocols 2016, 11(12):2301-
922   2319

923   60. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M: Accurate proteome-
924   wide label-free quantification by delayed normalization and maximal peptide ratio
925   extraction, termed MaxLFQ. Molecular and Cellular Proteomics. 2014,
926   13(9):2513-26

927   61. Wang Y, Kuramitsu Y, Baron B, Kitagawa T, Akada J, Tokuda K, Cui D,
928   Nakamura K: PERK/CHOP contributes to the CGK733-induced vesicular calcium
929   sequestration which is accompanied by non-apoptotic cell death. Oncotarget
930   2015, 6(28):25252-65

931   62. Pasquale EB: Eph receptors and ephrins in cancer: bidirectional signalling
932   and beyond. Nature Reviews Cancer 2010, 10(3):165-80

933   63. Mead TJ, Apte SS: ADAMTS proteins in human disorders. Matrix Biology
934   2018, 71-72:225-239

935   64. Kudinov AE, Karanicolas J, Golemis EA, Boumber Y: Musashi RNA-Binding
936   Proteins as Cancer Drivers and Novel Therapeutic Targets. Clinical Cancer
937   Research 2017, 23(9):2143-2153

938   65. Spears E, Neufeld KL: Novel double-negative feedback loop between
939   adenomatous polyposis coli and Musashi1 in colon epithelia. Journal of
940   Biological Chemistry 2011, 286(7):4946-50

941   66. Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P,
942   Davies H, Stratton MR, Campbell PJ: Universal Patterns of Selection in Cancer
943   and Somatic Tissues. Cell 2017, 171(5):1029-1041.e21.

944   67. Gou Q, Gao L, Nie X, Pu W, Zhu J, Wang Y, Liu X, Tan S, Zhou JK, Gong Y,
945   He J, Wu K, Xie Y, Zhao W, Dai L, Liu L, Xiang R, Wei YQ, Zhang L, Peng Y:
946   Long Noncoding RNA AB074169 Inhibits Cell Proliferation via Modulation of
947   KHSRP-Mediated CDKN1a Expression in Papillary Thyroid Carcinoma. Cancer
948   Research 2018, 78(15):4163-4174

949   68. Schmidtke L, Schrick K, Saurin S, Käfer R, Gather F, Weinmann-Menke J,
950   Kleinert H, Pautz A: The KH-type splicing regulatory protein (KSRP) regulates
951   type III interferon expression post-transcriptionally. Biochemical Journal 2019,
952   31;476(2):333-352

953   69. Li N, Cao M, Yi S, Cheng J, Wang L, Tao Y, Wu D, Peng J, Zhang M, Qi P,
954   Zhao J: Effects of the RNA-binding protein, KSRP, on innate immune response
955   against Helicobacter pylori infection in mice. Biochemical and Biophysical
956   Research Communications 2018, 495(2):1573-1579

957   70. Xu J, Liu H, Yang Y, Wang X, Liu P, Li Y, Meyers C, Banerjee NS, Wang HK,
958   Cam M, Lu W, Chow LT, Xie X, Zhu J, Zheng ZM: Genome-Wide Profiling of
959   Cervical RNA-Binding Proteins Identifies Human Papillomavirus Regulation of
960   RNASEH2A Expression by Viral E7 and E2F1. MBio 2019, 10(1): e02687-18

961   71. Hugen N, Brown G, Glynne-Jones R, de Wilt JH, Nagtegaal ID: Advances in
962   the care of patients with mucinous colorectal cancer. Nature Reviews Clinical
963   Oncology 2016, 13(6):361-9

964   72. Briata P, Bordo D, Puppo M, Gorlero F, Rossi M, Perrone-Bizzozero N,
965   Gherzi R: Diverse roles of the nucleic acid-binding protein KHSRP in cell
966   differentiation and disease. Wiley Interdisciplinary Reviews RNA 2016, 7(2):227

967

968
969
970

971 **Table 1:** Summary of clinical characteristics for the patients included in the two cohorts
972 analyzed by either Western blotting (WB) or Tissue Microarrays (TMA).
973
974

TABLE 1: clinical characteristics of CRC patients		
Variable	**Number of Patients**	
	WB cohort	**TMA cohort**
Total Patients	16	62
Tissue Type	fresh frozen	FFPE
Age mean (range)	65.8 (50-82)	63.5 (25-84)
Sex		
M	9	34
F	7	28
Clinical stage		
I	0	1
II	3	2
III	4	22
IV	9	37

975

976 **Figure 1:** *In silico* analysis of publicly available CRC datasets reveals a consistent

977 tumor-specific over-expression of KHSRP. **(A)** Differential expression analysis of

978 KHSRP in the Oncomine database returns 18 analyses (each comprising a different

979 patient group compared to normal tissue) across 8 different CRC datasets. The analyses

980 are plotted by percentile ranking of KHSRP over-expression (as a percentage of all

981 overexpressed genes in the dataset); dots are sized in proportion to the fold change of

982 KHSRP expression in tumor vs. normal, and colored by p-value (further details of the

983 analyses are provided in supplementary table 1). **(B)** Differential expression of KHSRP

984 between tumor and normal tissue, showing individual data points along with the mean ±

985 SD, is reported for the two representative analyses indicated (with the color code

986 referring to panel A). *** $p<0.0001$ (Mann-Whitney test). **(C)** Differential expression of

987 KHSRP between paired matched tumor and normal tissue in primary adenocarcinoma

988 and liver metastasis. The mean ± SD is reported with individual points. *** $p<0.0001$

989 (Wilcoxon matched-pairs signed rank test). (**D**) Differential expression analysis for

990 KHSRP in RNA-Seq data for the indicated 12 cancer types in the TCGA dataset. For

991 each type, the fold change expression of KHSRP in tumor versus normal tissue and the

992 associated FDR-adjusted p-value are reported

993 **Figure 2:** Analysis of KHSRP expression in primary tissue samples reveals a tumor-

994 specific overexpression with prognostic value. **(A)** Western blot analysis of KHSRP in

995 lysates from fresh-frozen tissue samples of patients with stage IV CRC (N=9), comparing

996 tumor (T) and matched normal tissue (N) from each patient. A representative blot for 2

997 patients is shown along with semi-quantitative densitometric analysis of all patients (the

998 Box and whisker plot depicts interquartile range, median, minimum and maximum limits)

40

999   ***p<0.0001 (Student's *t* test) **(B)** Representative images from TMA sections of normal

1000   and CRC tissue stained for KHSRP. **(C)** Quantitation of stromal and epithelial KHSRP

1001   staining in normal colon vs. primary CRC tissue samples from the TMA. The mean ± SD

1002   is reported with individual points. ** p<0.01 and ***p<0.0001 (Wilcoxon). **(D)** T/N

1003   ratios in the epithelium were used to segregate patients, using 20% as a cutoff (T/N<1.2

1004   versus T/N>1.2). Survival curves were calculated according to the Kaplan–Meier method

1005   and compared using the Cox–Mantel log-rank test. The hazard ratio (HR) and associated

1006   confidence interval are shown.

1007   **Figure 3:** in silico analysis of three independent cohorts validates the overexpression and

1008   prognostic significance of KHSRP. **(A)** Log_2 normalized signal intensity of *KHSRP* in a

1009   cohort of 566 CRC patients and 19 normal controls profiled by gene expression

1010   microarrays. The mean ± SD is reported with individual points and statistical significance

1011   (Mann-Whitney test). **(B)** Survival curves for patients segregated into high/low groups

1012   using the 25th percentile of the *KHSRP* probe signal intensity. The hazard ratio (HR) and

1013   associated confidence interval are shown. **(C)** Distribution of patients with high/low

1014   *KHSRP* expression according to clinical characteristics: mismatch repair deficient

1015   (dMMR, i.e. MSI) or proficient (pMMR, i.e. MSS), KRAS (codon 12 and 13) mutational

1016   status, BRAF (p.V600E) mutational status. **(D)** Expression of *KHSRP* in laser-capture

1017   microdissected epithelial and stromal cells from 13 CRC and normal colonic mucosa

1018   samples. ** p<0.01 (Kruskal-Wallis) **(E)** Expression of *KHSRP* in a cohort of FACS-

1019   sorted cells from 14 CRC patients. EpCAM: epithelial cancer cells; CD45: leukocytes;

1020   CD31: endothelial cells; FAP: cancer-associated fibroblasts. *p<0.05 (Kruskal-Wallis).

1021   The Box and whisker plots in D and E depict interquartile range, median, minimum and

41

1022   maximum limits

1023   **Figure 4:** KHSRP is involved in growth and survival of an *in vitro* CRC cell line model.

1024   **(A)** qRT-PCR analysis of SW480 cells transfected with a pool of siRNAs targeting

1025   *KHSRP* (siKHSRP), a scramble control pool (Scramble), or transfection reagent control

1026   (Mock) for 48h. **(B)** Cells were transfected as in A, and continuously monitored for

1027   proliferation for the indicated time; full time-course growth traces are shown, along with

1028   quantification of cell index (arbitrary measure of cell proliferation based on impedance

1029   measurements) at the indicated time point. **(C)** WB analysis of KHSRP protein

1030   expression in SW480 cells transfected with an siRNA targeting *KHSRP*, or a scramble

1031   control siRNA, or a mock (transfection reagent) control for 48h. **(D)** Growth of cells

1032   transfected as in C, monitored after 7 days by neutral red assay; representative images

1033   from one experiment are shown, with staining quantification from triplicate assays.

1034   Growth is expressed as % of untreated control (cells with no transfection) **(E)** Clonogenic

1035   potential for cells transfected as in C, measured by both colony area and surviving

1036   fraction. representative images from one experiment are shown, with quantification from

1037   triplicate assays.   **(F)** Spheroids from cells transfected as in C, grown in Matrigel;

1038   representative images are shown with staining quantification from triplicate assays.

1039   Growth is expressed as % of untreated control (cells with no transfection) **(A-**

1040   **F)****p<0.001, and ***p<0.0001 (ANOVA) for all tests in the figure.

1041   **Figure 5:** KHSRP is involved in regulating cell cycle and mitosis, as well as influencing

1042   the pro-tumorigenic extracellular environment *in vitro*. **(A)** Cell cycle analysis of SW480

1043   cells transfected with siKHSRP or Scramble negative control siRNA; a representative

1044   histogram of propidium iodide staining from one experiment is shown, along with a bar

42

1045   graph of cell distribution into cell cycle phases across three replicate experiments.

1046   **p<0.001 (ANOVA) **(B)** Analysis of mitotic cells by quantitation of phosphor-histone

1047   H3 in SW480 cells transfected as in A; a representative histogram is shown with bar

1048   graph of three replicate experiments. **p<0.001 (Mann-Whitney). **(C)** Wound-healing

1049   assay of SW480 cells transfected as in A; representative images from one experiment are

1050   shown, with quantitation of wound area from two replicate experiments. **(D)** Invasion of

1051   SW620 cells transfected with siKHSRP or Scramble negative control siRNA, measured

1052   by the capacity to migrate through Matrigel; quantification of invaded cells from three

1053   replicate experiments is shown, along with representative images. **(E)** Endothelial tube

1054   formation of HDEC cells treated with conditioned media (CM) from SW620 cells

1055   transfected as in D. Tube formation is expressed as % of control (HDEC cells treated

1056   with the same media without previous conditioning in SW620 cells) *p<0.05 (Student's *t*

1057   test). **(F)** Expression of secreted IL-8 and VEGF in the CM of SW620 cells transfected as

1058   in D, measured by ELISA. *p<0.05, **p<0.001 (Student's *t* test).

1059   **Figure 6:** KHSRP regulates expression of key proteins that modulate the tumorigenic

1060   microenvironment. Differential expression of secreted proteins was analyzed by shotgun

1061   proteomics in conditioned media from SW620 cells treated with Doxycycline for 4 days

1062   to induce shRNA-mediated KHSRP knockdown. **(A)** Volcano plot showing differentially

1063   regulated proteins in the doxycycline-treated samples compared to control, colored in red

1064   (log$_2$ Fold Change > |1|, p-value < 0.05). Proteins that are differentially regulated above

1065   the fold change threshold with a non-significant p-value are colored in yellow **(B)** Dot

1066   plot of results from a GO analysis of the 40 highly differentially regulated proteins. GO

1067   terms are reported with their category (CC=cellular component, MF=molecular function,

43

1068   BP=biological process, UP=uniprot annotation, KP=Kegg pathway) and plotted by gene

1069   ratio (the number of genes in one GO term compared to the total). Dots are sized in

1070   proportion to the number of hits within a GO term, and colored by p-value. **(C)** Label-

1071   free quantitation of relative abundance of the indicated five proteins in the doxycycline-

1072   treated vs. control samples. ***p<0.0001 (Student's *t* test). **(D)** Western blot analysis of

1073   Ephrin B2 and protein S100-A11 in conditioned media from control or doxycycline-

1074   treated SW620 cells. A Coomassie stain is shown for protein loading control.

A Differential Expression Analysis
KHSRP in tumor vs. normal

B GSE9348   TCGA

C GSE6988

D

A STAGE IV

102 kDa →
76 kDa →

KHSRP

β-actin

| N | T | N | T |
| #679 | | #620 | |

Stage IV (N=9)

normalized KHSRP protein (fold difference)

***

Normal Colon   CRC

B

x10   x20

Normal Colon

200μm   100μm

CRC

C

● epithelium   ▲ stroma

KHSRP staining quickscore

**   ***

Normal Colon   CRC   Normal Colon   CRC

D

Percent survival

low KHSRP
high KHSRP

log-rank p = 0.0138
HR = 3.741 (95% CI: 1.434-22.97)

Months

A GSE39582

p=0.001

log$_2$ normalized KHSRP signal intensity

normal colon | CRC

B Stage II-III

Relapse-Free Survival

low KHSRP
high KHSRP

log-rank p = 0.0038
HR = 1.952 (95% CI: 1.206-2.595)

Delay (months)

C

high KHSRP
NA 10% dMMR 10%
pMMR 80%

low KHSRP
NA 7% dMMR 16%
pMMR 78%

high KHSRP
NA 4% mutKRAS 39%
wtKRAS 57%

low KHSRP
NA 3% mutKRAS 38%
wtKRAS 59%

high KHSRP
NA 10% mBRAF 6%
wtBRAF 84%

low KHSRP
NA 9% mBRAF 11%
wtBRAF 80%

D GSE35602

tumor normal

log$_2$ normalized KHSRP signal intensity

**

epithelium | stroma | epithelium | stroma

E GSE39396

log$_2$ normalized KHSRP signal intensity

*

Epcam | FAP | CD45 | CD31