The long noncoding RNA lnc-FANCI-2 intrinsically restricts RAS signaling and phosphorylation of Akt and Erk in HPV16-infected cervical cancer

Haibin Liu[1a,b], Lulu Yu[1a], Vladimir Majerciak[1], Thomas Meyer[2], Ming Yi[3], Peter F. Johnson[4], Maggie Cam[2], Douglas R. Lowy[5], and Zhi-Ming Zheng[1]*

[1]Tumor Virus RNA Biology Section, HIV Dynamics and Replication Program, CCR/NCI/NIH

[2]Collaborative Bioinformatics Resource, CCR/NCI/NIH

[3]NCI RAS Initiative, Cancer Research Technology Program, Frederick National Laboratory for Cancer Research, Frederick, MD, USA.

[4]Mouse Cancer Genetics Program, CCR/NCI/NIH

[5]Laboratory of Cellular Oncology, CCR/NCI/NIH

[a]*Co-first authors*

Running title: lnc-FANCI-2 regulates RAS signaling

[b]Present address: Center for Emerging Infectious Diseases, Wuhan Institute of Virology, and Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan, Hubei, 430207, China

*Corresponding: zhengt@exchange.nih.gov

ABSTRACT

We recently discovered increased expression of a long noncoding RNA, lnc-FANCI-2, coinciding with cervical lesion progression from CIN1, CIN2-3 to cervical cancer. Viral E7 of high-risk HPVs and host transcription factor YY1 are two major factors promoting lnc-FANCI-2 expression. To explore possible roles of lnc-FANCI-2 in HPV-induced cervical carcinogenesis, we ablated the expression of *lnc-FANCI-2* in the HPV16-positive cervical cancer cell line, CaSki. Knock-out (KO) single cell clones expressed HPV16 oncogenes normally but displayed altered cell morphology when compared with the parental cells. Proteomic profiling of cytosolic and secreted proteins from the parental and KO cells showed that lnc-FANCI-2 regulates expression of a subset of cell surface and adhesion-related proteins, including inhibition of MCAM, PODXL2 and ECM1 and increased levels of ADAM8 and TIMP2. RNA-seq analyses revealed that, relative to the parental cells, KO cells exhibited significantly increased RAS signaling but decreased IFN pathways. In KO cells, phosphorylated Akt and Erk1/2, two important RAS pathway effectors, were increased more than 3-fold, accompanied by increase of IGFBP3, MCAM, VIM, and CCND2 (cyclin D2) and decrease of RAC3. Accordingly, high levels of lnc-FANCI-2 and lower levels of MCAM in cervical cancer patients are associated with improved survival. We found that lnc-FANCI-2 in CaSki cells interacts specifically with 32 host proteins, including H13, HNRH1, K1H1, MAP4K4, and RNPS1, and knockdown of MAP4K4 led to increase phosphorylation of Akt and Erk1/2. In summary, a key function of lnc-FANCI-2 is to intrinsically regulate RAS signaling, thereby affecting cervical cancer outcome.

43 **Significance**

44 Expression of lnc-FANCI-2 is related to cervical lesion progression. Knock-out (KO) or knock-
45 down (KD) of lnc-FANCI-2 expression in HPV16-positive cervical cancer cells CaSki
46 significantly increases RAS signaling, phosphorylation of Akt and Erk1/2, and increase of
47 epithelial mesenchymal transition factors. lnc-FANCI-2 KO also regulates the expression of a
48 subset of cell surface and adhesion-related proteins IGFBP3 and MCAM. A high level of lnc-
49 FANCI-2 and lower level of MCAM in cervical cancer patients are associated with improved
50 survival. lnc-FANCI-2 in CaSki cells interacts specifically with 32 host proteins, including
51 MAP4K4. KD of MAP4K4 expression in CaSki cells led to increase phosphorylation of Akt and
52 Erk1/2. Thus, one lnc-FANCI-2 function is to intrinsically regulate RAS signaling to impact
53 cervical lesion progression.

54 **INTRODUCTION**

55 Cervical cancer is a leading cause of mortality in women worldwide. It is the second most
56 diagnosed cancer among women, with an estimated 530,000 new cases worldwide each year, and
57 the third most frequent cause of cancer-related death, accounting for 270,000 deaths annually[1].
58 Persistent infections in the cervical epithelium with the cancer-associated alpha human
59 papillomaviruses (high-risk HPVs), in particular HPV16 and HPV18 with the sustained
60 expression of viral oncoproteins, E6 and E7, contribute to squamous epithelial cell tumorigenesis
61 [2-4]. Approximate 1.4% of high-grade neoplastic cervical lesions progress to invasive cervical
62 cancer [5], while most lesions resolve spontaneously. The progression to cervical cancer
63 typically takes 15 to 20 years after initial infection (https://www.who.int/news-room/fact-
64 sheets/detail/cervical-cancer). These observations suggest that active host defense systems play
65 important roles in preventing malignant transformation of neoplastic cervical lesions, and
66 additional factors besides E6 and E7 are necessary for cervical cancer development. Indeed, one
67 tumor suppressor network, the Fanconi anemia (FA) pathway [6], can be activated by HPV-
68 induced DNA damage. FA proteins including FANCA, FANCD2 and FANCI are elevated upon
69 HPV infection [7-9] and the activated FA pathway restricts HPV replication and guards host
70 genome stability [10]. FANCD2 mutation stimulates HPV16 and HPV31 genome amplification
71 and also promotes cervical and vaginal cancer development in HPV16 E7 transgenic mouse [11,
72 12].

73 HPV-positive cervical cancer tissues in general exhibits wild-type p53, wild-type K-RAS,
74 and no overexpression of RAS genes[13-16]. Viral oncoprotein E7 plays a major role in the
75 immortalization of primary epithelial cells mainly by inactivation of pRb family of proteins [17],
76 whereas E6 mediates degradation of tumor suppressor p53 and activation of human telomerase

77 reverse transcriptase (hTERT) transcription [18-20]. However, high-risk E6 and E7 are necessary
78 but not sufficient to transform fully the immortalized cells into malignant cells [21-25] and
79 additional oncogenic stress is needed. For example, the $HRAS^{G12V}$, a constitutively activated
80 RAS GTPase mutation, triggers malignant transformation of E6/E7-expressing keratinocytes
81 [26-29]. Normal RAS regulates cell proliferation, cell differentiation and cell adhesion through
82 cellular signal transduction from growth-factor receptors such as insulin-like growth factors
83 (IGFs) and insulin-like growth factor binding proteins (IGFBP1-6) [30-32]. Thus, overactive
84 RAS signaling might facilitate transformation of E6/E7-expressing keratinocytes. In transgenic
85 mice, HPV16 E6 and E7 can promote the development of spontaneous epithelial skin tumors, but
86 not spontaneous tumors of the reproductive tract [33-36]. Prolonged estrogen treatment is
87 required [37]. Estrogen receptor (ER) activation stimulates the mitogen-activated protein kinase
88 (MAPK/Erk) and phosphoinositide 3-kinase (Pl3K/Akt) pathways, both of which are mediated
89 by RAS [38]. However, the molecular mechanism underlying estrogen-induced cervical cancer
90 in E6 and E7 transgene mice is not fully understood.

91 Long non-coding RNAs (lncRNAs) are RNAs over 200 nucleotides (nt) in length lacking
92 a coding capacity for translation into functional proteins. Although several lncRNAs have been
93 proposed to function in diverse biological processes, evidence is still lacking to support the
94 functionality of the majority of lncRNAs [8, 39]. Nevertheless, the identified lncRNA functions
95 include regulating: (1) transcription of neighboring or distant genes by recruiting histone
96 modifiers, as shown for Xist [40] and ANRIL [41], or by causing chromosome looping, as
97 proposed for PVT1 [42] and CCAT1-L [43]; (2) RNA splicing by acting as decoy of splicing
98 factors, as demonstrated for MALAT1 [44] and PNCTR [45]; (3) mRNA stability by directly
99 base pairing as antisense transcripts [46], by acting as RNA binding protein (RBP) decoy to

100 prevent mRNA-RBP interaction, such is the case for NORAD [47], or by functioning as miRNA
101 sponges to prevent miRNA-induced degradation of mRNAs, as shown for PTEN [48] and
102 PNUTS [49]; and (4) protein stability by directly binding, such as MaIL1 [50], or by acting as
103 protein decoy to prevent ubiquitin-mediated proteasomal degradation, such as FAST [51]. Thus,
104 lncRNA genomic position, subcellular localization and their interactions with DNA, RNA and
105 proteins are crucial determinants to the functionality of individual lncRNAs.

106 We recently reported an increased level of lnc-FANCI-2 expression in high-risk HPV-
107 positive cervical intraepithelial neoplasia (CIN) and invasive cervical cancer (ICC) tissues [8].
108 At least 14 RNA isoforms of lnc-FANCI-2, which is wrongly annotated by NCBI as a *MIR9-*
109 *3HG* in the recently updated RefSeq database (https://www.ncbi.nlm.nih.gov/refseq/), are
110 expressed by usage of two alternative promoters , which are ~10 kb downstream of a non-
111 expressible miR9-3 gene [8], alternative RNA splicing, and alternative selection of RNA
112 polyadenylation sites from the genomic locus adjacent to the FANCI on Chr5. Both lnc-FANCI-
113 2 and FANCI are up-regulated simultaneously in neoplastic cervical lesions and cervical cancer
114 by high-risk HPV infections. E6 and, in particular E7 are responsible for the enhanced
115 expression of lnc-FANCI-2, the transcription of which is also regulated by YY1[8]. HPV
116 infection increases YY1 levels but decreases the expression of p53-dependent miR-29a which
117 targets the YY1 3′ UTR. Viral E7 interacts with YY1 and facilitates YY1 transactivation of lnc-
118 FANCI-2 promoter [8]. In situ hybridization has revealed that lnc-FANCI-2 is preferentially
119 cytoplasmic, but its function in HPV infected cells has not been characterized. In this report, we
120 demonstrate that knockout (KO) or knockdown (KD) of lnc-FANCI-2 promotes RAS signaling
121 and phosphorylation of Akt and Erk.

122 **RESULTS**

123 **Differential expression of lnc-FANCI-2 in cervical cancer tissues and its derived cell lines.**
124 As lnc-FANCI-2 expression is up-regulated along with cervical lesion progression by high-risk
125 HPV infections [8], we further verified the increased expression of lnc-FANCI-2 in cervical
126 cancer tissues by RNAscope single molecule RNA in situ hybridization (RNA-ISH) using a lnc-
127 FANCI-2 antisense probe spanning over the major isoform of lnc-FANCI-2 nt 359-1713 region
128 (GenBank MT669800.1). We found both cytoplasmic and nuclear locations of the increased lnc-
129 FANCI-2 ISH signals within the tumor nest in HPV16-infected cervical cancer tissues, but not
130 much so in its adjacent normal tissue areas (Fig. 1A). The increased expression of lnc-FANCI-2
131 became obvious in human foreskin keratinocytes with HPV16 or HPV18 infection when
132 compared to the HFK cells without HPV infection (Fig. 1B).

133 However, we subsequently demonstrated the high expression of lnc-FANCI-2 in four
134 HPV16-infected cervical cancer cell lines (SiHa, CaSki, W12 20861 and W12 20863 cells), but
135 not in two HPV18-infected cervical cancer cell lines (HeLa and C4II cells), nor other HPV-
136 negative cell lines of HCT116 (colorectal cancer cells), BCBL-1 (body cavity B lymphoma
137 cells), HEK293 (Ad5 E1/E2-immortalized human kidney cells), and HaCaT (spontaneously
138 immortalized human epidermal cells) (Fig. S1A). Interestingly, the high expression of lnc-
139 FANCI-2 was observed in a HPV-negative cervical cancer cell line C33A cells with mutations of
140 both p53 and RB genes [52] (Fig. S1A). Northern blot confirmed the increased expression of lnc-
141 FANCI-2 in C33A cells and no expression of lnc-FANCI-2 in HeLa cells (Fig. S1B).

142 Subcellular distribution of lnc-FANCI-2 was characterized by cell fractionation and
143 Western blot assays using nuclear SRSF3 and cytoplasmic GAPDH protein as an indication of
144 fractionation efficiency (Fig. S1C). By RT-qPCR analysis of the fractionated total RNA, we
145 demonstrated that lnc-FANCI-2 RNA is mainly cytoplasmic in HPV16-positive CaSki cells, but

146 nuclear in HPV16-positive SiHa cells (Fig. S1C). The differential subcellular distributions of
147 lnc-FANCI-2 RNA from CaSki to SiHa cells were further confirmed by RNAscope RNA-ISH
148 using the lnc-FANCI-2 antisense probe described above (Fig. S1D).

149 **lnc-FANCI-2 regulates proliferation of HPV-transformed cervical cancer cells.**

150 lnc-FANCI-2 RNA is transcribed mainly from a proximal promoter TSS2, but also from
151 an alternative minor, distal promoter TSS1. Two highly conserved YY1 binding motifs upstream
152 of the TSS2 are essential for the TSS2 transcriptional activity [8] (Fig. 2A). To elucidate the
153 function of lnc-FANCI-2 in high-risk HPV-infected cervical cancer cells, we knocked out (KO)
154 lnc-FANCI-2 expression in HPV16-positive CaSki cells using CRISPR/Cas9 deletion of either a
155 3-kb promoter region encompassing both TSS1 and TSS2 or an 86-bp region containing two
156 YY1-binding motifs in the TSS2 promoter (Fig. 2A). Two tested gRNAs with high KO
157 efficiency were selected and cloned into a modified CRISPR/Cas9 expression vector to express
158 both 5′-specific gRNA and 3’-specific gRNA simultaneously for efficient genome editing [53].
159 CaSki cells stably transfected with the dual gRNA expression vector were generated. Through a
160 serial dilution, several single cell clones were isolated and verified for homozygous deletion by
161 PCR screening. PCR genotyping indicated successful homozygous deletion of the promoter or
162 YY1-binding motifs, respectively (Fig. 2B).

163 Loss of lnc-FANCI-2 expression in the single-cell clones was examined by RT-qPCR
164 (Fig. 2C). When compared to all control clones transfected and selected from an empty vector
165 containing no gRNA, all single cell clones with deletion of YY1-binding motifs (ΔYY1) showed
166 70-80% reduction of lnc-FANCI-2 expression, whereas the single cell clones with deletion of
167 both TSS1 and TSS2 promoters (ΔPr) showed at least 90% reduction (Fig. 2C). Northern blot
168 analysis confirmed the decrease in levels of the two major lnc-FANCI-2 isoforms (Fig. 2D), with

169 an abundant 4-kb lnc-FANCI-2 RNA derived from a distal polyadenylation site pA2 and a less
170 abundant 2-kb of lnc-FANCI-2 RNA derived from a proximal polyadenylation site pA1[8] (Fig.
171 2A and 2D). Reduced lnc-FANCI-2 RNA expression was also confirmed by RNAscope RNA-
172 ISH (Fig. 2E and 2F). The parental (WT) CaSki cells expressed ~10 copies of lnc-FANCI-2
173 RNA per cell [8], whereas the copy number of lnc-FANCI-2 RNA in the ΔYY1-D5 cells
174 dropped to ~5 copies and to ~2 copies per cell in the ΔPr-A9 cells (each averaged from 200
175 cells).

176 Although the WT CaSki cells grow as cell islands, the lnc-FANCI-2 KO cells displayed a
177 dispersed cell growth pattern and an irregular or spindle-like cell morphology which was more
178 obvious in the ΔPr-A9 cells (Fig. S2A). The ΔPr-A9 cells were further examined for effects on
179 HPV16 E6 and E7 expression. When compared with WT CaSki cells, the ΔPr-A9 cells showed
180 little change in E6, E7, p53 (E6 downstream target) and E2F1 (E7 downstream target) protein
181 levels (Fig. S2B).

182 **lnc-FANCI-2 regulates the expression and secretion of cell soluble receptors.**

183 We considered that altered expression of cell membrane proteins and secreted factors in
184 the lnc-FANCI-2 KO cells might contribute to the observed different cell morphology and
185 growth properties. Therefore, we performed a Proteome Profiler Human sReceptor Array
186 analysis to examine possible changes in expression of 105 well-characterized soluble protein
187 receptors using total cell lysates and cell culture supernatants from ΔPr-A9 and WT CaSki cells.

188 Using 30% cut-off (FC -/+0.3), we identified from the ΔPr-A9 cell lysate nine proteins
189 with increased expression compared to the WT CaSki cells, including PODXL2, ECM1,
190 NECTIN2, MCAM, ADAM9, ADAM10, CDH5, ITGA5 and NOTCH1, and six proteins with

191 decreased expression including ITGB6, CDH13, LGALS3BP, TIMP2, ADMA8 and SCARF2
192 (Fig. 3A and 3B). We also found from the ΔPr-A9 cell culture supernatant five proteins with
193 increased expression, including ADAM9, NECTIN2, ADAM10, CDH5 and ECM1, and the
194 decreased expression of four proteins, including CRELD2, SDC1, SDC4 and TIMP2 (Fig. 3A
195 and 3B). By immunoblot assays, we verified selectively the increase in PODXL2, MCAM, and
196 ECM1 and the decrease in ADAM8 and TIMP2 from the ΔPr-A9 cell lysates and the increase in
197 ECM1 and the decrease in TIMP2 in the ΔPr-A9 cell culture supernatant (Fig. 3C). As ADAM8
198 is proteolytically processed into two protein isoforms [54], we confirmed by immunoblot the
199 decreased expression of all three sizes of ADAM8 protein in the ΔPr-A9 cell lysate over the WT
200 CaSki cells (Fig. 3C).

201 **lnc-FANCI-2 regulates the expression of genes involved in RAS signaling.**

202 We next conducted genome-wide RNA-seq analyses of WT and lnc-FANCI-2 KO cells
203 (four samples/group) to determine the transcriptomic consequences of lnc-FANCI-2 deficiency.
204 We obtained ~120 million mappable RNA reads to the human reference genome hg38 from each
205 sample. Analysis of RNA-seq reads-coverage map by Integrative Genomics Viewer (IGV)
206 showed ~90% reduction of lnc-FANCI-2 expression in ΔPr-A9 cells and ~72% decrease in
207 ΔYY1-D5 cells compared to the WT CaSki cells (Fig. 4A), consistent with the results shown in
208 Fig. 2C-F.

209 Hierarchical clustering analysis showed more global transcriptional similarity between
210 ΔYY1-D5 and ΔPr-A9 cells (Fig. 4B). By applying a threshold of fold change (FC) ≥ 1.8 or FC
211 ≤ -1.8 with FDR ≤ 0.01, we found 1230 genes in ΔPr-A9 and 797 genes in ΔYY1-D5 cells with
212 significantly differential expression relative to the WT CaSki cells expressing ~15890 genes. The
213 most significantly affected genes are shown in Volcano plots (Fig. 4C and Table S1). Among

214 these, 211 were upregulated and 189 were downregulated in both ΔPr-A9 and Δ YY1-D5 cells
215 (Fig. 4D and Table S2). By applying more stringent criteria with FPKM ≥7 in at least one of four
216 RNA-seq samples as a cutoff, we profiled the genes with large expression differences in both
217 ΔPr-A9 and ΔYY1-D5 cells. The results displayed in the heatmap of Fig. 4E included 52
218 upregulated, and 47 downregulated genes excluding lnc-FANCI-2. Subsequently, we selectively
219 verified by RT-PCR the decreased expression of ITGB6, NLRP2, PLAC8, PSG4, PSMB9, and
220 SERPINB1 and increased expression of CFH, CNTN5, EMP3, IGFBP3, PTGS2, and
221 SERPINB2 in ΔPr-A9 cells (Fig. 4F), as well as the increased expression of IGFBP3 protein, a
222 RAS signaling driver [30-32], in ΔPr-A9 cells by immunoblot (Fig. 4G).

223 By performing the Gene Set Enrichment Analysis (GSEA) on the Hallmark gene sets,
224 which provide more refined and concise inputs for GSEA [55], we found the most significantly
225 upregulated pathways in both ΔPr-A9 and ΔYY1-D5 cells, when compared with the WT CaSki
226 cells, were KRAS signaling and epithelial mesenchymal transition (EMT). The most
227 significantly downregulated pathways in both ΔPr-A9 and ΔYY1-D5 cells over the WT CaSki
228 cells were interferon gamma (IFN-γ) and interferon alpha (IFN-α) responses (Fig. 5A). GSEA
229 plots for ΔPr-A9 shows 29 of 111 genes, including IGFBP3 [30-32], in RAS signaling and 39 of
230 130 genes in EMT (Fig. 5B and 5C, Table S3) were significantly upregulated, whereas 37 of 139
231 genes in IFN-γ response and 24 of 71 genes in IFN-α response were significantly downregulated
232 (Fig. 5D and 5E, Table S3). Similar enriched gene sets with increased RAS signaling and EMT
233 and decreased IFN-γ and IFN-α responses were observed in ΔYY1-D5 cells (Fig. S3, Table S3).

234 Since ΔPr-A9 cells exhibit higher KO efficiency of lnc-FANCI-2 and a higher number of
235 differentially expressed genes (DEGs) and show more severe phenotypic changes in cell

236 proliferation, migration, and colony formation than ΔYY1-D5 cells, the ΔPr-A9 cells were
237 primarily used for further focused studies on RAS signaling in this report.

238 **lnc-FANCI-2 regulates RAS GTPase activities to phosphorylate RAS signaling effectors**
239 **Akt and Erk.**

240 RAS activation triggers two major downstream signal transduction pathways,
241 Raf/Mek/Erk and PI3K/Akt (Fig. S4), to transduce signals from extracellular stimuli to the cell
242 nucleus where specific effector genes are activated for their corresponding functions [30-32, 56,
243 57]. To further confirm the GSEA data and verify that two major signaling pathways of RAS are
244 activated, we first examined and compared the RAS GTPase activities in ΔPr-A9 cells and the
245 WT CaSki cells. We found a significant increase in RAS GTPase activity in ΔPr-A9 cells (Fig.
246 6A, top bar graphs), indicating that the endogenous lnc-FANCI-2 RNA in CaSki cells suppresses
247 not only the expression of IGFBP3 (Fig. 4F and 4G), the most abundant IGFBP, but also RAS
248 activation (Fig. 6A, top bar graphs). Since ΔPr-A9 cells had undergone long term selection and
249 adaption during single cell screening, the increased RAS GTPase activity might result from
250 selection pressure for cell survival. However, by transient siRNA knockdown (KD) of lnc-
251 FANCI-2 expression in the WT CaSki cells, we obtained a similar, albeit weaker, increase of
252 RAS GTPase activity (Fig. 6A, lower bar graphs). This result suggests that the increased RAS
253 GTPase activity in ΔPr-A9 cells was unrelated to the persistent selection pressure, but rather the
254 deficiency of lnc-FANCI-2.

255 Next, we examined activation of the PI3K/Akt and Raf/Mek/Erk pathways [58-60]. We
256 observed a 3-fold increase in phosphorylated Akt (p-Akt) and 2.5-fold increase of p-Erk1/2
257 (p44/p42 MAPK), after being normalized to total Akt and Erk levels, in ΔPr-A9 cells over the
258 WT CaSki cells (Fig. 6B). These results were confirmed in ΔYY1-D5 cells (Fig. S5A). The

259 increased p-Akt and p-Erk (mostly p-Erk2/p42) was accompanied by elevated expression of
260 MCAM, VIM, and CCND2, and decreased expression of RAC3 (Fig. 6B). These randomly
261 chosen potential RAS effector proteins facilitate membrane receptor function, cell proliferation,
262 and EMT. Moreover, transient siRNA knockdown of lnc-FANCI-2 in the WT CaSki cells also
263 led to the increased levels of p-Akt and p-Erk (Fig. 6C) and increased expression of MCAM and
264 VIM at 48 h and 96 h post-transfection (Fig. 6C). These observations indicate that a transient
265 loss of lnc-FANCI-2 in CaSki cells is sufficient to trigger RAS signaling. The same siRNA
266 transfection in HeLa cells, a lnc-FANCI-2 negative cell line (Fig. S1A and S1B), exhibited no
267 effect on p-Akt or p-Erk1/2 levels (Fig. S5B), but in SiHa cells containing mainly nuclear lnc-
268 FANCI-2 (Fig. S1C and S1D) enhanced the expression of p-Akt and p-Erk1/2 (Fig. S5B),
269 verifying the specific effect of lnc-FANCI-2 depletion on RAS signaling pathways
270 independently of cellular distribution of predominant lnc-FANCI-2. The lnc-FANCI-2 KO-
271 mediated increase in p-Akt and p-Erk was sensitive to PI3K inhibitor LY294002 [61] (Fig. 6D,
272 lanes 4-6) and MEK1/2 inhibitor U0126 [62] (Fig. 6D, lanes 7-9), respectively. The inhibitors
273 also decreased the expression of VIM and MCAM and cell proliferation, in particular by
274 MEK1/2 inhibitor U0126 (Fig. 6D and 6E). The effects of LY294002 on cell proliferation were
275 similar from ΔPr-A9 to the WT CaSki cells.

276 **lnc-FANCI-2 inhibits the expression of IGFBP3 and MCAM (CD146 or MUC18).**

277 Given that IGFBP-3 is the most abundant IGFBP among all six IGFBP members in
278 potentiation of IGF action and PI3K/AKT activities [30] and a reduced expression of IGFBP-3
279 mRNA level is associated with progression to cervical cancer [63], we further investigated the
280 effect of lnc-FANCI-2 on the expression of IGFBP3 as lnc-FANCI-2 effector gene. As shown in
281 Fig. 4F-G and Fig. S6A, the increased RNA and protein expression of IGFBP3 appears in the

282 lnc-FANCI-2 KO ΔPr-A9 cells over the parental WT CaSki cells, indicating a suppressive effect
283 of lnc-FANCI-2 on IGFBP3 expression. This suppressive function of lnc-FANCI-2 was further
284 confirmed in lnc-FANCI-2 rescue experiments in the ΔPr-A9 cells (Fig. 7A and 7B). By
285 transient expression of one major isoform of lnc-FANCI-2 RNA (a-PA2, GenBank:
286 MT669800.1) [8] in the ΔPr-A9 cells, lnc-FANCI-2 in red and IGFBP3 RNA in green at 24 h
287 post transfection were detected by RNAscope RNA-ISH using each specific antisense RNA
288 probe (Fig. 7A). We demonstrated that the cells with rescued expression of cytoplasmic lnc-
289 FANCI-2 RNA displayed much reduced expression of cytoplasmic IGFBP3 RNA (Fig. 7B).

290 As shown in Fig. 4E and Fig. S6A, the increased RNA expression of MCAM (CD146 or
291 MUC18) [64] appears to be transcriptional or posttranscriptional in both ΔPr-A9 and ΔYY1-D5
292 cells. We confirmed by RT-PCR the lnc-FANCI-2 KO-mediated increase of MCAM RNA
293 expression in ΔPr-A9 cells (Fig. 7C). A rescue experiment was further performed by transient
294 expression of one major isoform of lnc-FANCI-2 RNA (a-PA2, GenBank: MT669800.1) [8] in
295 ΔPr-A9 cells to determine both nuclear and cytoplasmic MCAM expression [65, 66] in
296 individual lnc-FANCI-2 expressing cells. ΔPr-A9 cells ectopically expressing diffused
297 cytoplasmic lnc-FANCI-2 RNA, as detected by RNAscope RNA-ISH, exhibited a marked
298 reduction in nuclear MCAM protein (Fig. 7D-7E). Quantitative analyses of the ΔPr-A9 cells with
299 or without transient lnc-FANCI-2 RNA expression showed significant reduction of nuclear
300 MCAM protein in the 35 cells with rescued lnc-FANCI-2 expression when compared to the 54
301 cells expressing no lnc-FANCI-2 (Fig. 7F). Due to RNAscope procedures, the membrane-bound
302 and cytoplasmic MCAM protein was mostly removed by protease III treatment when RNA-ISH
303 was performed and thus only the nuclear signal of MCAM protein remained was detectable by IF
304 staining. Using cell fractionation, we detected a cleaved MCAM of ~46 kDa mainly in the

305 nuclear fraction by Western blot (Fig. 7G). The increased nuclear MCAM expression in ΔPr-A9
306 cells was proportional to the cytoplasmic and total MCAM levels when compared with the WT
307 CaSki cells (Fig. 7G). The nuclear MCAM signal appeared as a proteolytic cleavage product
308 presumably by the increased metalloprotease ADAM9 and ADAM10 proteins and decreased
309 metalloprotease inhibitor TIMP2 (Fig. 3), as observed in other studies [66].

310 The significance of the inverse correlation of lnc-FANCI-2 and MCAM in CaSki cells
311 was further investigated by analyzing their expression in 304 cervical cancer samples from the
312 TCGA dataset. There was a significant negative correlation in lnc-FANCI-2 (LINC00925 being
313 wrongly assigned by NCBI)[8] and MCAM RNA levels (Fig. 7H). The opposite effects of lnc-
314 FANCI-2 and MCAM on cervical cancer survival (Fig. 7I and Fig. S7) show that the cervical
315 cancer patients with a higher level of lnc-FANCI-2 [8] but a lower level of MCAM in the
316 cervical tissue (Fig. S7 and Fig. 7I) exhibited a better survival prognosis.

317 **Regulatory roles of lnc-FANCI-2-mediated increase of MCAM and IGFBP3 on RAS**
318 **signaling**

319 To further dissect the lnc-FANCI-2-associated expression of MCAM and IGFBP3 on
320 RAS signaling, we examined MCAM and IGFBP3 on phosphorylation of PI3K/Akt and
321 Erk1/Erk2 in CaSki cells in the presence or absence of lnc-FANCI-2. As shown in Figure 7J, we
322 found, although KD or KO of lnc-FANCI-2 promotes phosphorylation of Akt and Erk (Fig. 7J,
323 lane 2), that KD of MCAM expression in WT parental CaSki cells in the absence of lnc-FANCI-
324 2 (Fig. 7J, lane 3) or ΔPr-A9 cells (Fig. 7, lane 5) led to reduction of Akt and Erk
325 phosphorylation. Data suggest that MCAM is a lnc-FANCI-2 effector, but also could be a trigger
326 of signal transduction as reported [64, 67].

327 IGFBP3 protein has been viewed as a RAS signaling regulator [30-32], but recent studies
328 show that IGFBP3 has a variety of intracellular ligands involved in many unexpected
329 functions[68, 69]. Thus, KD or KO of lnc-FANCI-2-mediated IGFBP3 expression on RAS
330 signaling in WT parental CaSki cells or ΔPr-A9 cells was examined by Western blot after siRNA
331 KD of IGFBP3 expression. As shown in Figure 7K, KD of IGFBP3 expression was found to
332 increase phosphorylation of Erk1/2 by ~70% (lane 4) to ~60% (lane 6), but not much so for Akt
333 phosphorylation (lanes 4 and 6). Instead, KD of IGFBP3 expression could prevent KD of lnc-
334 FANCI-2 to enhance Akt phosphorylation in WT parental CaSki cells (Fig. 7K, compare lane 2
335 and lane 3). Data suggesting a separate role of IGFBP3 on phosphorylation of Erk1/2 from Akt
336 in the presence or absence of lnc-FANCI-2.

337 **MAP4K4 association with lnc-FANCI-2 RNA in CaSki cells regulates RAS signaling and**
338 **phosphorylation of Akt and Erk.**

339 We speculated lnc-FANCI-2 restriction on RAS signaling through interactions with
340 cellular proteins. Subsequently, comprehensive identification of lnc-FANCI-2-binding proteins
341 in the WT CaSki cells was performed by an IRPCRP protocol (modified from published ChIRP
342 [70]) in combination with mass spectrometry (Fig. 8A). In this protocol, the lnc-FANCI-2-
343 binding proteins in the WT CaSki cells were covalently crosslinked to lnc-FANCI-2 RNA via
344 UV irradiation and pulled down by pooled 30 antisense oligos crossing over the entire lnc-
345 FANCI-2. The RNA extracted from the pulldowns were verified to be lnc-FANCI-2 specific by
346 RT-PCR in the absence (-) or presence (+) of reverse transcriptase (RT) (Fig. 8B). The lnc-
347 FANCI-2-binding proteins in the pulldowns were then subjected to LC-MS/MS analyses. We
348 identified 32 specific lnc-FANCI-2-binding proteins, including H13, HNRH1, K1H1, MAP4K4,
349 and RNPS1 (Fig. 8C, Table S4).

350 MAP4K4, a serine/threonine protein kinase, stimulates cancer cell proliferation, invasion
351 and migration, and is recently characterized as a novel MAPK/ERK pathway regulator [71-73]
352 and negatively regulates RAS signaling by binding to Ras p21 protein activator 1 or RASA1
353 [74, 75]. Thus, whether MAP4K4 in association with lnc-FANCI-2 could regulate PI3K/Akt
354 signal transduction was investigated in the WT CaSki cells. We demonstrated that siRNA
355 knockdown of MAP4K4 expression in the WT CaSki cells, as seen for KD of lnc-FANCI-2, led
356 to ~40% or more increase of phosphorylation of both Akt and Erk1/2 (Fig. 8D, compare lane 2 to
357 lane 3). The data suggest that, through RNA-protein interactions, the increased lnc-FANCI-2
358 RNA in cells and their association with MAP4K4 and other cellular proteins could be one arm to
359 control RAS signaling and gene expression of its effectors (Fig. 8E).

360

361 **DISCUSSION**

362 HPV oncoproteins E6 and E7 are necessary but not sufficient for development of cervical
363 cancer. Indeed, early studies indicated that other oncogenic stress, such as activated RAS mutant,
364 is needed to trigger malignant transformation of E6 and E7 expressing cells [21, 26, 76, 77] and
365 tumorigenesis [27, 28, 78-82]. Moreover, E7 has been also reported to repress phosphorylation
366 of Akt and Akt mediated signaling [83]. In this study, we show that lnc-FANCI-2, whose
367 expression is highly dependent on E7 and YY1 [8], intrinsically restricts RAS GTPase activities
368 and Akt and Erk phosphorylation presumably by interacting with cellular factors, in HPV16-
369 positive CaSki cells. Interestingly, early reports indicated that p-Akt attenuation by E7 could be
370 abolished by introduction of H73E mutation in the E7 CR3 domain [83]. This E7 CR3 domain is
371 also essential to interact with transcription factor YY1 to activate lnc-FANCI-2 transcription [8].
372 Thus, it would be presumable that the reported p-Akt attenuation by E7 might be mediated by the

373 increased expression of lnc-FANCI-2. More importantly, these findings also suggest that
374 additional oncogenic stress, such as activated RAS mutant, is required to overcome the lnc-
375 FANCI-2 restriction on RAS signaling for the early observed malignant transformation [21, 26,
376 76, 77] and tumorigenesis [27, 28, 78-82] of E6 and E7 expressing cells. In CaSki cells, loss of
377 lnc-FANCI-2 in this report was found to promote RAS signaling but reduction of IFN responses.

378 Some lncRNAs function as RAS regulators [84, 85] by their ability to sequester RAS-
379 targeting miRNAs, including MALAT1/ miR-217, RMRP/miR-206, and KRAS1P/miR-143/let-
380 7 [85]. Conceivably, lncRNAs may also regulate RAS signaling through other mechanisms
381 beyond miRNAs. In this report, loss of lnc-FANCI-2 in HPV16-positive CaSki and SiHa cells
382 promotes RAS signaling and expression of RAS signaling effectors (Fig. 8E). By profiling the
383 proteins associated with lnc-FANCI-2 in CaSki cells, we identified MAP4K4, a serine/threonine
384 protein kinase and a negative RAS signaling regulator in the context of the early embryo [72, 74]
385 and lymphatic vascular development by interacting with RASA1[75] as one of major lnc-
386 FANCI-2-binding proteins that may partially mediate the lnc-FANCI-2 restriction on RAS
387 signaling. Indeed, silencing of MAP4K4 in CaSki cells led to increase of both p-Akt and p-
388 Erk1/2 as reported [74, 75]. Thus, our data suggest that MAP4K4 protein functions coordinately
389 with lnc-FANCI-2 RNA in the CaSki cells to restrict RAS signaling.

390 Although all HPV16-positive cervical cancer and pre-cancer cell lines examined, such as
391 CaSki, SiHa, W12 subclone cells 20861(HPV16^{+}, integrated) and 20863 (HPV16^{+}, episomal),
392 produce predominantly cytoplasmic or nuclear lnc-FANCI-2, unexpectedly, we found not all
393 cervical cancer cell lines, including HPV18-positive HeLa and C4II cells, express lnc-FANCI-2,
394 nor colorectal cancer cell line HCT116, adenovirus 5 DNA-transformed human embryonic
395 kidney cell line HEK293, spontaneously transformed aneuploid immortal keratinocyte cell line

396 HaCaT, and KSHV-infected B cell lymphoma cell line BCBL-1 cells which are all HPV-
397 negative cell lines. It remains to know what negatively regulates lnc-FANCI-2 expression in
398 those cell lines and more HPV18-positive cell lines should be examined for lnc-FANCI-2
399 expression. However, we did find that a HPV-negative cervical cancer cell line, C33A cells
400 expressing mutant p53 and mutant pRB [52], produce a high level of lnc-FANCI-2 (Fig. S1A
401 and S1B). In the lnc-FANCI-2 producing cells, we noticed that cellular location of lnc-FANCI-2
402 also varies from cell types, with lnc-FANCI-2 predominantly in the nucleus in SiHa cells, but in
403 the cytoplasm in CaSki cells, although both contribute to regulate RAS signaling. This suggests
404 that a variety of nuclear and cytoplasmic functions of lnc-FANCI-2 remains to be explored.

405 IGFBP3 is a major IGF carrier and enhances both IGF- and EGF-RAS signaling [31, 86]
406 to activate phosphatidylinositol 3-kinase (PI3K)/AKT and the mitogen-activated protein kinase
407 (MAPK)/ERK pathways [87]. Surprisingly, we find a suppressive role of IGFBP3 on Erk
408 phosphorylation, but not on Akt, in the presence or absence of lnc-FANCI-2 in HPV16-positive
409 CaSki cells. Although ~99% of circulating IGF-1 is bound to IGFBPs, predominantly to IGFBP-
410 3, which is the most abundant IGFBP in human serum, case-control studies showed significantly
411 lower IGF-1 and IGFBP-3 serum levels in patients with invasive cervical cancer over the control
412 group [88, 89]. A reduced expression of IGFBP-3 mRNA level appears to be associated with
413 progression to cervical cancer [63]. Consistently with lnc-FANCI-2's suppressive effect on
414 IGFBP3 expression in our study, an increased level of lnc-FANCI-2 expression is associated
415 step-wisely with high-risk HPV-positive cervical intraepithelial neoplasia and invasive cervical
416 cancer tissues [8].

417 It is obvious from our RNA-seq and human soluble receptor array analysis that KO of
418 lnc-FANCI-2 in CaSki cells affects the expression of a large set of genes responsible for RAS

419 signaling, EMT pathway, and IFN responses. In addition to the notified activation of PI3K/Akt
420 and Raf/Mek/Erk pathways, one of the major RAS signaling effectors in lnc-FANCI-2 KO or
421 KD cells was MCAM. MCAM binds various cellular surface receptors or co-receptors to trigger
422 signal transduction, cell proliferation and motility, tumor angiogenesis, and tumor cell metastasis
423 [64, 67]. It is an integral, highly glycosylated membrane protein with three common protein
424 isoforms [64, 90, 91]. Two membrane-anchored forms with a long (p113) or short (p76)
425 cytoplasmic tail produced by alternative RNA splicing and a proteolytically cleaved soluble form
426 without transmembrane and cytoplasmic regions [64, 91, 92]. In addition to its roles in cell signal
427 transduction [64, 67, 90, 91], the short isoform produced by an intramembrane cleavage may be
428 directed towards the nucleus to regulate gene transcription [66]. We found that CaSki cells
429 express mainly a cytoplasmic form of full-length MCAM (~113 kDa) and a small nuclear form
430 (~46 kDa), but not the short ~78 kDa isoform nor the soluble product. Moreover, lnc-FANCI-2
431 KO cells exhibit high MCAM levels, which was most likely a result of the increased RAS
432 signaling. Consistently, lnc-FANCI-2 and MCAM levels are negatively correlated in cervical
433 cancer patients, and low levels of MCAM and high levels of lnc-FANCI-2 indicate better
434 prognosis. This suggests that increased MCAM might promote cervical tumorigenesis in vivo.
435 Given that the cells with lnc-FANCI-2 KO exhibit increased levels of IGFBP3 and MCAM and
436 the ΔPr-A9 cells with rescued lnc-FANCI-2 RNA expression displayed remarkable reduction of
437 IGFBP3 and MCAM expression, our current model (Fig. 8E) highlights how the HPV16-
438 enhanced expression of lnc-FANCI-2 in cervical cancer cells might function as a negative
439 regulator by binding to various RNA-binding proteins to block RAS signaling and reduce the
440 expression of RAS signaling effectors such as IGFBP3 and MCAM, etc.

441 In conclusion, this study provides the first evidence that one of lnc-FANCI-2 functions is
442 to maintain epithelial functional integrity by repressing RAS signaling in preventing malignant
443 transformation of neoplastic cervical lesions. In addition, more extensive studies are undergoing
444 to elucidate the role of lnc-FANCI-2 in regulation of IFN responses, which are not covered in
445 this report. Since transactivation of lnc-FANCI-2 expression is largely dependent on host
446 transcription factor YY1 of which expression is regulated by HPV oncoproteins and host miR-
447 29a, dissecting their transcription network on lnc-FANCI-2 and its effectors may further provide
448 novel insights involved in host homeostasis against HPV infection-induced carcinogenesis. Our
449 observations also highlight additional, unexpected players beyond viral E6 and E7 that drive
450 cells toward malignant transformation.

451 **MATERIALS AND METHODS**

452 **Cell cultures and treatment**

453 CaSki, a HPV16-positive (HPV16^{+}) human cervical cancer cell line with wild-type p53, pRb,
454 HRAS, KRAS, and NRAS (https://depmap.org/portal/cell_line/ACH-001336?tab=mutation) was
455 obtained from the American Type Culture Collection (ATCC, Manassas, VA), maintained in
456 Dulbecco's modified Eagle medium (DMEM) (Thermo Fisher Scientific) supplemented with
457 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin, and grown in a cell culture
458 incubator at 37°C with 5% CO_2.

459 SiHa (HPV16^{+}), HeLa (HPV18^{+}), C4II (HPV18^{+}), C33A (no HPV), and other HPV-
460 negative cell lines, HEK293T (Ad5 E1/E2-immortalized human kidney cells), and HaCaT
461 (spontaneously immortalized human epidermal cells), were all obtained from ATCC and grown
462 in Dulbecco's modified Eagle's medium (DMEM) with 10% FBS at 37°C and 5% CO_2. HCT116

463 cell line, a colorectal cancer cell line, was a gift from Dr. Bert Vogelstein of Johns Hopkins
464 University and was grown in McCoy's 5A medium with 10% FBS at 37°C and in 5% CO_2. W12
465 subclones 20861 ($HPV16^+$, integrated) and 20863 ($HPV16^+$, episomal) were a gift from Dr. Paul
466 Lambert of University of Wisconsin, Madison, and grown on mitomycin C-pretreated NIH3T3
467 feeder cells in F12 medium at 37°C and in 5% CO_2. BCBL-1 (body cavity B lymphoma cells
468 with KSHV infection) was obtained from the AIDS Research and Reference Reagent Program,
469 Division of AIDS, NIAID, NIH and cultured in RPMI 1640 containing 10% FBS at 37°C and in
470 5% CO_2.

471 CaSki cells were treated with 20 μM phosphoinositide 3-kinase (PI3K) inhibitor,
472 LY294002 (Cell Signaling Technology, Danvers, MA, #9901) or 10 μM mitogen-activated
473 protein kinase kinase (MEK) inhibitor, U0126 (Cell Signaling Technology, Danvers, MA,
474 #9903) and examined for their proliferation and viability by CCK-8 assay and RAS effector
475 expression by Western blot.

476 **siRNAs, plasmid constructions and transfections**

477 Synthetic double-stranded siRNA targeting the lnc-FANCI-2 transcript is listed in a
478 supplementary Table S5. Human MAP4K4 siRNA (#M-003971), human IGFBP3 siRNA (#M-
479 004777) and non-targeting control siRNA(#D-001210-01) were purchased from Dharmacon, Inc
480 (Lafayette, CO). Each siRNA transfection into CaSki or other type of cells was carried out by
481 LipoJet in vitro transfection reagent (SignaGen, Frederick, MD, USA). Plasmid pHBL03 has a
482 short isoform of lnc-FANCI-2 cDNA (lnc-FANCI-2a-PA1, GenBank Acc. No. MT669801).
483 Overlapping PCR products generated from 5' and 3' RACE PCR products with a primer pair of
484 oHBL 17 and oHBL 18 were digested and inserted into pcDNA3 at Hind III and Xba I sites. The
485 insertion was verified by sequencing. Plasmid pHBL05 has a long isoform of lnc-FANCI-2 (lnc-

486 FANCI-2a-PA2, GenBank Acc. No. MT669800). PCR products generated from genomic DNA
487 with oHBL 19 and oHBL 23 were digested and swapped into pHBL03 to replace lnc-FANCI-2a-
488 PA1 at Bsmb I and Xba I sites.

489 CRISPR/Cas9 system was used to knockout lnc-FANCI-2 gene from CaSki cells. Two
490 lnc-FANCI-2 specific guide RNAs (gRNAs) were designed by an on-line CRISPR design tool
491 developed by the Zhang lab (http://crispr.mit.edu). A modified cloning strategy was used to
492 clone the two gRNA into pSpCas9(BB)-2A-Puro vector (Addgene, Watertown, MA, #62988)
493 and described in our previous publication [53]. This strategy allows all three components (two
494 gRNAs and Cas9) express from the same plasmid to ensure the transfected cells expressing two
495 gRNAs simultaneously knocking out the targeted gene [53]. The sequences of primers to make
496 guide RNAs were listed in Supplementary Table S5.

497 **Knockout of lnc-FANCI-2 gene in CaSki cells, PCR screening and genotyping**

498 Two gRNAs each under control by a U6 promoter in a pSpCas9(BB)-2A-Puro vector were used
499 to create deletion in CaSki cells. Plasmid pHBL25 which contains gRNA 1 and 6 was used to
500 delete the entire promoter region of lnc-FANCI-2 and plasmid pHBL27 which contains gRNA 3
501 and 4 was used to delete only two YY1-binding motifs spanning over an 87-bp region in the lnc-
502 FANCI-2 promoter. CaSki cells transfected with the gRNA-expressing vectors were then under 2
503 weeks of puromycin selection (0.1 μg/ml). The puromycin-resistant CaSki cells were diluted to a
504 final concentration of 0.5 cells per 100 μl, plated 100 μl of diluted cells into each well of a 96-
505 well plate, and grew under continuous purymycin selection. The colonies were inspected to
506 ensure the clones were grown from one cell, and refed or re-plated as needed in 3 weeks of
507 expansion.

508 A direct PCR was performed for number of single cell clones as described [53]. Briefly,
509 several hundreds of cells expanded from a single cell clone were resuspended in 50 µl of
510 phosphate-buffered saline (PBS) and then frozen-thawed on dry ice for three times. Cell lysates
511 were treated with 4 µl of Qiagen protease (QIAGEN, Hilden, Germany, #1017782,) at 56°C for
512 10 min, and the protease was inactivated at 95°C for 5 mins. The resulting product was directly
513 used for PCR screening. P1 + P5 were used for verification of promoter deletion and P3 + P4 for
514 YY1 motifs deletion. To ensure the selected single cell clones displaying homozygous knockout
515 of lnc-FANCI-2, total cell DNA was isolated using a QIAamp DNA Blood Minikit (QIAGEN,
516 Hilden, Germany, #51106) and then the homozygous deletion was screened and the cells with
517 homozygous KO were selected and confirmed by PCR. The primers used for genotyping are
518 listed in supplementary Table S5. At least two single cell clones either with lnc-FANCI-2
519 promoter deletion or with YY1 motif deletion were selected. Three single cell clones from the
520 cells transfected with an empty vector, pSpCas9(BB)-2A-Puro, were also selected and served as
521 controls for the genotyping screening.

522 **Cell proliferation and viability assay**

523 Cell proliferation and viability were determined by a CCK-8 assay (Dojindo Molecular
524 Technologies, Rockville, MD). For proliferation assay, 500 µl of cell suspension (50,000
525 cells/well) were dispensed into individual wells in a 24-well plate. Cells were then treated with
526 inhibitors (20 µM LY294002, or 10 µM U0126 respectively) after 24 h culture. At each time
527 point of 24 h, 48 h and 72 h, 50 µl of CCK-8 solution was added in three wells of each group and
528 incubated for 1 h in a cell culture incubator at 37°C. Cell proliferation was determined by
529 absorbance at 450 nm, and normal medium was used to subtract background. Percentage of cell

530 viability was calculated by the following formula: cell viability % = OD450 of inhibitors treated
531 sample / OD450 of untreated sample × 100%.

532 **Nuclear and cytoplasmic fractionation**

533 CaSki and SiHa cells were fractionated by using Nuclei EZ Prep Kit (Sigma-Aldrich, #NUC-
534 101) following the manufacturer's protocol. The details of the method could be found in our
535 previous publication [93]. Western blot analysis for nuclear protein SRSF3 and cytoplasmic
536 protein GAPDH was used to determine the fractionation efficiency. The fractionated cytoplasmic
537 and nuclear RNAs were used for detection of lnc-FANCI-2 RNA with human GAPDH RNA
538 serving as a control for RNA fractionation efficiency of the cytoplasmic RNAs.

539 **RT-qPCR**

540 Detection of lnc-FANCI-2 by RT-qPCR was performed as described [8]. Pre-designed primers
541 for lnc-FANCI-2 are listed in Table S5. Briefly, 2 µg total RNA was converted to cDNA using
542 Superscript First-stand Synthesis kit (Thermo Fisher Scientific, Waltham, MA, #11904018).
543 qPCR was performed using TaqMan gene expression Master Mix (Applied Biosystems,
544 Waltham, MA, #4369016) on a StepOne Plus Real-Time PCR system (Applied Biosystems).
545 TaqMan Gene Expression Assay (Applied Biosystems, #4331182) for GADPH
546 (Hs02758991_g1) was served as an internal control. Data were plotted as fold change over the
547 control group using the $2^{-\Delta\Delta Ct}$ method by which the data was normalized first to the values for
548 GAPDH and then to the median value for control samples. Data are presented as a bar graph with
549 mean ± SD for each group.

550 **Northern blot**

551 To validate lnc-FANCI-2 KO in CaSki cells, Northern blot was performed using polyA$^+$ mRNA
552 isolated from 100 µg of total CaSki RNA with PolyATtract mRNA Isolation Systems (Promega,
553 #Z5310) or 10 µg of total CaSki total RNA as described previously [8]. Briefly, RNA samples
554 were denatured in NorthernMax Formaldehyde loading dye (Thermo Fisher Scientific,
555 #AM8552) at 75°C for 15 min, then separated in an 1% formaldehyde-containing agarose gel
556 and transferred onto a GeneScreen Plus hybridization transfer membrane (Perkin Elmer,
557 Waltham, MA, #NEF987001PK). RNAs on the membrane were crosslinked by exposing to UV
558 light, and then prehybridized with PerfectHyb Plus hybridization buffer (Sigma-Aldrich, St.
559 Louis, MO, #H7003) for 2 h at 42°C. Specific oligos against lnc-FANCI-2 or GAPDH listed in
560 Table S5 were labeled with [γ-^{32}P]-ATP using T4 PNK (Thermo Fisher Scientific, Waltham,
561 MA, #18004–010) and added into hybridization buffer for overnight incubation at 42°C. The
562 membrane was then washed with 2× SSPE/0.5% SDS solution for 5 mins, followed by twice
563 washes with 0.5× SSPE/0.5% SDS each for 15 min at 42°C and then exposed to a
564 PhosphorImager screen.

565 **RNA-seq and data analysis**

566 Total RNA was from WT CaSki cells, ΔYY1-D5 and ΔPr-A9 cells were extracted using TRIzol
567 reagent (Thermo Fisher Scientific, Waltham, MA, #15596018). Total ribo-minus RNA-seq
568 libraries, four samples in each group, were prepared with TruSeq Stranded Total RNA Library
569 Kit and then subjected to pair-end sequencing using Illumina-HiSeq3000/4000 platform.

570 The obtained reads were processed using the CCBR Pipeliner utility
571 (https://github.com/CCBR/Pipeliner). Briefly, reads were trimmed from adapters and low-quality
572 bases using Cutadapt (version 1.18) (https://bioweb.pasteur.fr/packages/pack@cutadapt@1.18)
573 before alignment to the custom reference genome described below. The transcripts were aligned

574 using STAR v2.5.2b in two-pass mode [94]. Expression levels were quantified using RSEM
575 (version 1.3.0) [95] with a custom gene annotation described below.

576 The custom reference genome allowing quantification of both HPV16 and host
577 expression used in this alignment consisted of the human reference genome (hg38/Dec.
578 2013/GRCh38) with a HPV16 FASTA (https://pave.niaid.nih.gov/) sequence added as an
579 additional pseudochromosome. The custom gene annotation used for gene expression
580 quantification consisted of a concatenation of the hg38 GENCODE annotation version 30 [96]
581 and the HPV16 genome, with one other notable alteration. The hg38 v30 gene annotation is
582 incorrect at the location of the lnc-FANCI-2 locus in the hg38 genome at chr15:89378104-
583 89398487, which affects quantification of this gene. To correct this annotation and provide
584 accurate quantification, therefore, we first performed BLAST against hg38 using the sequences
585 of the 14 known isoform transcripts of lnc-FANCI-2 (MT669800 for lnc-FANCI-2a-PA2,
586 MT669801 for lnc-FANCI-2a-PA1, MT669802 for lnc-FANCI-2b-PA2, MT669803 for lnc-
587 FANCI-2b-PA1, MT669804 for lnc-FANCI-2c-PA2, MT669805 for lnc-FANCI-2c-PA1,
588 MT669806 for lnc-FANCI-2d-PA2, MT669807 for lnc-FANCI-2d-PA1, MT669808 for lnc-
589 FANCI-2e-PA2, MT669809 for lnc-FANCI-2e-PA1, MT669810 for lnc-FANCI-2f-PA2,
590 MT669811 for lnc-FANCI-2f-PA1, MT669812 for lnc-FANCI-2g-PA2, and MT669813 for lnc-
591 FANCI-2g-PA1). The results were used to determine the precise exon start and stop sites for
592 each of the isoforms. These were then used to manually adjust the default hg38 v30 gene
593 annotations at the locus indicated above to account for the correct structure of the lnc-FANCI-2
594 locus and its 14 isoforms.

595 Downstream analysis and visualization were performed within the NIH Integrated
596 Analysis Portal (NIDAP) using R programs developed on the Foundry platform (Palantir

597 Technologies). Briefly, raw counts data produced by RSEM were imported into the NIDAP
598 platform, genes were filtered for low counts (<1 cpm) and the voom algorithm [97] from the
599 Limma R package (version 3.40.6) [98] was used for quantile normalization and calculation of
600 differentially expressed genes. Pre-ranked GSEA was performed using the Molecular Signatures
601 Database version 6.2 [99, 100] and the fgsea package [101]. Raw data and the analyzed RNA-seq
602 data supporting the findings in this study have been deposited in the NCBI GEO database (the
603 accession#: GSE190904).

604 **Human Soluble Receptor Array**

605 To detect the changes of soluble receptors expressed and released from parental WT CaSki cells
606 and lnc-FANCI-2 KO cells, Proteome Profiler Human sReceptor (Soluble Receptor) Array
607 (R&D Systems, Minneapolis, MN, # ARY012) was performed according to a manufacturer
608 protocol. Briefly, 7×10^6 WT CaSki cells and ΔPr-A9 cells were cultured in 15 ml of DMEM
609 supplemented with 10% FBS in a T75 flask for 24 h. Cell culture supernatants were then
610 collected and centrifuged at 4,000 × g for 15 min to remove cell debris. Cells were rinsed by 10
611 ml of PBS and solubilized in Lysis Buffer 17 at 4°C for 30 min. The cell lysis was then
612 centrifuged at 14,000 × g for 5 min to remove cell debris. Protein concentration of cell lysis was
613 measured by Micro BCA Protein Assay (PIERCE, Waltham, MA, #23235). Next, two sets of N
614 and C membranes were blocking by Buffer 8/1 in 4-Well Multi-dish for 1 h on a rocking
615 platform. 500 µl of culture supernatant or 100 µg cell lysate generated from the WT CaSki cells
616 or ΔPr-A9 cells were diluted in 3 ml of Buffer 8/1, and then applied to each membrane after
617 three washes with 1× Wash Buffer. The membranes were incubated overnight at 4 °C on a
618 rocking platform and washed three times to remove unbound materials followed by incubation
619 with their specific cocktail of biotinylated detection antibodies for 2 h at room temperature on a

620 rocking platform. 2 ml of diluted Streptavidin-HRP was then applied to each membrane for 30
621 min at room temperature on a rocking platform after three times of wash. 1 ml of the prepared
622 Chemi Reagent Mix was added onto the membranes after three times of wash. The signal was
623 then captured by ChemiDoc Touch Imaging System (Bio-Rad). Quantification of the relative
624 protein expression in ΔPr-A9 cells compared to the WT cells was determined in Image Lab
625 software (Bio-Rad). The average signal of duplicate spots subtracting background value from
626 negative control spots represented the levels of each protein. The relative change in protein
627 levels was then determined by comparing ΔPr-A9 cells to the WT cells.

628 **Antibodies and Immunoblot**

629 Immunoblot analysis was performed for individual proteins by using the following antibodies:
630 anti-MCAM (#17564-1-AP), anti-E2F1 (#12171-1-AP), anti-MAP4K4 (#55247-1-AP), and anti-
631 p53 (#10442-1-AP) antibodies were from ProteinTech (Rosemont, IL). Anti-tubulin (#T5201)
632 antibodies were from Sigma-Aldrich (St. Louis, MO). Anti-hnRNP C1/C2 (#Ab10294) and anti-
633 RAC3 (#ab129062) antibodies were from Abcam (Cambridge, United Kingdom). Anti- CCND2
634 (Cyclin D2, #3741), anti-Akt (#9272), anti-phospho-Akt (#9271), anti-Erk1/2 (p44/42 MAPK,
635 #4695), anti-phospho-Erk1/2 (#4370), and anti-IGFBP3 (D1U9C, #25864) were from Cell
636 Signaling Technology (Danvers, MA). Anti-VIM (#MA5-11883) was from Thermo Fisher
637 Scientific (Waltham, MA). Anti- PODXL2 (#AF1524), anti-ECM1 (#MAB39371), anti-TIMP2
638 (#AF971) and anti-ADAM8 (#AF1031) were from R&D Systems (Minneapolis, MN). Anti-
639 SRSF3 (#NBP2-76892) was from Novus Biologicals (Centennial, CO). Rabbit polyclonal anti-
640 HPV16 E7 (#GTX133411) was from GeneTex (Irvine, CA).

641 **RAS GTPase Chemi ELISA**

642 RAS GTPase activity in cell lysate of the WT CaSki or ΔPr-A9 cells was determined by RAS
643 GTPase Chemi ELISA Kit (Active Motif, Carlsbad, CA, # 52097) according to the manufacturer
644 protocol. Briefly, 2×10^7 cells were solubilized in 500 µl of Complete Lysis/Binding buffer at
645 4°C for 15 min. Protein concentration of cell extract was then measured by Micro BCA Protein
646 Assay (PIERCE, #23235) after cell debris was removed by centrifuged at 14,000 × g for 10 min.
647 2 µg of GST-Raf-RBD in 50 µl of Complete Lysis/Binding buffer was coated in each well of a
648 96-well plate by incubating for 1 h at 4°C with 100 rpm agitation. After three times of wash, 50
649 µg of extract in 50 µl Complete Lysis/Binding buffer was then added in each well. HeLa (EGF
650 treated) extract was served as a positive control. The plate was incubated for 1 h at room
651 temperature with 100 rpm agitation. The active RAS protein (GST-Raf-RBD binding) in each
652 well was then detected by incubating with primary RAS antibody and three times of wash, and
653 then, HRP-conjugated secondary antibody and three times of wash. Chemiluminescence in each
654 well was read in luminometer by adding 50 µl room-temperature Chemiluminescent Working
655 Solution.

656 **RNA *in situ* hybridization (RNA-ISH)**

657 Endogenous lnc-FANCI-2 in CaSki cells was examined by RNAScope Multiplex Fluorescent V2
658 Assay (Advanced Cell Diagnostics, Minneapolis, MN, #323100) as described previously [8]. A
659 custom-designed probe targeting to nt 359-1713 of GenBank Acc. No. MT669800.1 transcript
660 for lnc-FANCI-2a-PA2 was utilized. Dual staining of lnc-FANCI-2 and IGFBP RNA was
661 performed according to RNAscope Multplex Fluorescent v2 Manual part 2 (#323100), and the
662 channel one probe for IGFBP3 (#310351) and channel three probe for lnc-FANCI-2 (#509061-
663 C3) were applied.

664 For RNA-ISH combined IF staining of MCAM protein, 3×10^5 of CaSki cells were
665 grown on a glass coverslip in a 6-well plate for 24 h before transfection. Plasmid pHBL05
666 containing a long isoform of lnc-FANCI-2a-PA2 (GenBank Acc. No. MT669800.1) were
667 transfected into cells with LipoD 293 DNA in vitro transfection reagent. Cultured Adherent Cell
668 Sample Preparation for the RNAscope Multplex Fluorescent v2 was performed according to a
669 manufacturer protocol (Advanced Cell Diagnostics, Minneapolis, MN, MK-50-010). Briefly, the
670 cells were washed with PBS and fixed by 10% neutral buffered formalin at room temperature for
671 30 min. The cells were washed with PBS three times and dehydrated by 50% ethanol for 5 min,
672 70% ethanol for 5 min, 100% ethanol for 5 min, and 100% ethanol for 10 min. The cells were
673 then rehydrated by 70% ethanol for 2 min, 50% ethanol for 2 min, and PBS for 10 min. The cells
674 were then applied to hydrogen peroxide at RT for 10 min and washed with distilled water for
675 three times. The cells were then applied to Protease III (1:15 dilution with PBS) at RT for 10
676 min. The cells were ready for RNAscope Multplex Fluorescent v2 Manual part 2 (#323100). IF
677 was performed after RNA-ISH but before counterstaining with DAPI. The cells were washed
678 with PBS, and then were blocked with 2% bovine serum albumin (BSA) in PBS for 1 h at 37°C
679 or overnight at 4°C. The cells were incubated with the MCAM primary antibody (ProteinTech,
680 Rosemont, IL, #17564-1-AP, diluted 1:200 in 2% BSA blocking buffer) for 2 h at 37°C. An
681 Alexa Fluor-conjugated secondary antibody (1: 500, ThermoFisher Scientific) was diluted in a
682 blocking solution and incubated for 1 h at 37˚C. The slides were washed with PBS for three
683 times. DAPI was used for nuclei counterstaining before mounting in a Prolong Gold Antifade
684 mounting medium (Thermo Fisher Scientific, #P36934).

685 Confocal images were collected using a Zeiss LSM710 laser-scanning microscope equipped with
686 a 63 × Plan-Apochromat (N.A. 1.4) objective lens. Three dimensional distributions of lnc-
687 FANCI-2 were generated by Z stacks using ZEN 2.3 software (Zeiss).

688 **Isolation of lnc-FANCI-2 RNA-protein complex by RNA Purification (IRPCRP) and LC-**
689 **MS/MS analysis**

690 To identify the lnc-FANCI-2 RNA-associated proteins, we pulled down lnc-FANCI-2 RNA from
691 WT CaSki cells using IRPCRP, a modified ChIRP protocol from the published method [102],
692 followed by mass spectrometry [70]. Briefly, a total of 30 lnc-FANCI-2 anti-sense oligo probes
693 (oLLY496-oLLY525), each with biotinTEG at 3' end, across entire lnc-FANCI-2 RNA
694 (GenBank Acc. No MT669800.1) were designed using the online probe designer
695 (singlemoleculefish.com). Two probe pools were used in IRPCRP assays: the probe pool 1 was
696 the mixed oligos in even numbers (oLLY496
697 /498/500/502/504/506/508/510/512/514/516/518/520/522/524) and the probe pool 2 mixed with
698 the oligos in odd numbers
699 (oLLY497/499/501/503/505/507/509/511/513/515/517/519/521/523/525) (Table S5). CaSki
700 cells were seeded and harvested at 24 h and the cell lysates from 200 million cells were used for
701 each IRPCRP reaction. Instead of formaldehyde crosslink in the standard protocol [102], UV
702 cross-linking (254 nm, energy at 480 mJ/cm^2) was applied in our study to minimize non-specific
703 binding. After UV cross-linking, the cells were lysed on ice in 1× radioimmunoprecipitation
704 assay (RIPA) buffer (Boston Bioproducts, Ashland, MA, #BP-115) (50 mM Tris base/Tris-HCl
705 [pH 7.4], 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate [SDS], 1%
706 NP-40) supplemented with protease inhibitors (complete mini-EDTA-free protease inhibitor
707 cocktail, Millipore Sigma, Burlington, MA, #469315900) and RNase inhibitor (Thermo Fisher

708 Scientific, #AM2694) for 30 min, followed by a brief sonication. The obtained cell lysates were
709 spin for 10 min of centrifugation at 20,000 × g at 4°C and the collected cell lysate supernatants
710 were pre-absorbed with C-1 magnetic beads (Thermo Fisher Scientific, #65002) for 1 h at room
711 temperature before proceeding to oligo probe hybridization.

712 Two different hybridization protocols were used to pull down lnc-FANCI-2 RNA. One
713 had the pre-absorbed cell lysate incubated with two separate probe pools in hybridization buffer
714 (750 mM NaCl, 0.1% SDS, 50 mM Tris-Cl pH 7.0, 1 mM EDTA, 15% formamide) at room
715 temperature overnight and the pre-washed C-1 magnetic beads was then added and incubated for
716 additional 4 h; the other had two separate oligo probe pools immobilized onto pre-washed C-1
717 magnetic beads first at room temperature for 1 h and then mixed with the pre-absorbed cell
718 lysates in hybridization buffer overnight at room temperature, along with the negative control-
719 beads only IRPCRP reaction without oligo probes serving as a control in the modified protocol.
720 After hybridization, the beads were washed with 1× wash buffer (2× NaCl and sodium citrate,
721 0.1% SDS) for 5 times and resuspend in 1 ml lysis buffer. 100 µl beads from each IRPCRP were
722 set aside for RNA isolation of lnc-FANCI-2 pulldown efficiency and the leftover beads were sent
723 for mass spectrometry analysis. The input and IRPCRP RNA were isolated using a MIRNeasy
724 mini kit (Qiagen, #217004) following the standard protocol after proteinase K (Millipore Sigma,
725 #71049) digestion.

726 For mass spectrometry analysis, the IRPCRP beads were processed for trypsin/LysC
727 digestion before submitted to Thermo Scientific Orbitrap Exploris 240 Mass Spectrometer and a
728 Thermo Dionex UltiMate 3000 RSLCnano System for Proteinomics. Peptides from trypsin
729 digestion were loaded onto a peptide trap cartridge at a flow rate of 5 µl/min. The trapped
730 peptides were eluted onto a reversed-phase Easy-Spray Column PepMap RSLC, C18, 2 µM,

731 100A, 75 µm × 250 mm (Thermo Scientific) using a linear gradient of acetonitrile (3-36%) in
732 0.1% formic acid. The elution duration was 110 min at a flow rate of 0.3 µl/min. Eluted peptides
733 from the Easy-Spray column were ionized and sprayed into the mass spectrometer, using a Nano
734 Easy-Spray Ion Source (Thermo Fisher Scientific) under the following settings: spray voltage,
735 1.6 kV, Capillary temperature, 275 °C. Other settings were empirically determined. Raw data
736 files were searched against human protein sequences database using the Proteome Discoverer 2.4
737 software (Thermo Fisher Scientific, San Jose, CA) based on the SEQUEST algorithm. All the
738 protein peptides identified in the IRPCRP pulldowns were summarized in Table S4.

739 **ACKNOWLEDGEMENTS**

740 We thank Louise T. Chow of the University of Alabama at Birmingham for her critical reading
741 of the manuscript. We thank Craig Meyers of Penn State University Hershey Medical Center for
742 providing HPV16- and HPV18- infected raft cultures, Xing Xie and Yang Li of Zhejiang
743 University Women's Hospital of China for cervical tissue sections, and Johannes G. Schweizer
744 from Arbor Vita corporation for anti-HPV16 E6-specific antibodies. This study was supported
745 by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer
746 Research.

747 **DATA AVAILABILITY**

748 NCBI GEO database Acc. No. GSE190904 for CaSki cell RNA-seq.

749 **AUTHORS CONTRIBUTIONS:** H. L., L. Y., and Z.M.Z. designed the study. H. L. and L. Y.
750 performed the experiments. H. L., L. Y., V.M., T. M., M. Y., P. J., M.C., D. R. L., and Z.M.Z.
751 analyzed data and participated in the discussion and interpretation of the results. H. L., L. Y., and

752 Z.M.Z drafted the manuscript. H. L., L. Y., and Z.M.Z. revised the manuscript. All authors read
753 and approved the final manuscript.

754 **DECLARATION OF INTERESTS:** The authors declare no competing interests.

755

756 **REFERENCES**

757

758 1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A: **Global cancer statistics, 2012.** *CA*
759 *Cancer J Clin* 2015, **65:**87-108.
760 2. Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV: **The causal relation between human**
761 **papillomavirus and cervical cancer.** *J Clin Pathol* 2002, **55:**244-265.
762 3. Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV, Snijders PJ, Meijer CJ,
763 International Agency for Research on Cancer Multicenter Cervical Cancer Study G:
764 **Epidemiologic classification of human papillomavirus types associated with cervical cancer.** *N*
765 *Engl J Med* 2003, **348:**518-527.
766 4. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J,
767 Meijer CJ, Munoz N: **Human papillomavirus is a necessary cause of invasive cervical cancer**
768 **worldwide.** *J Pathol* 1999, **189:**12-19.
769 5. McCredie MR, Sharples KJ, Paul C, Baranyai J, Medley G, Jones RW, Skegg DC: **Natural history of**
770 **cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia**
771 **3: a retrospective cohort study.** *Lancet Oncol* 2008, **9:**425-434.
772 6. Niraj J, Farkkila A, D'Andrea AD: **The Fanconi Anemia Pathway in Cancer.** *Annu Rev Cancer Biol*
773 2019, **3:**457-478.
774 7. Hoskins EE, Morris TA, Higginbotham JM, Spardy N, Cha E, Kelly P, Williams DA, Wikenheiser-
775 Brokamp KA, Duensing S, Wells SI: **Fanconi anemia deficiency stimulates HPV-associated**
776 **hyperplastic growth in organotypic epithelial raft culture.** *Oncogene* 2009, **28:**674-685.
777 8. Liu H, Xu J, Yang Y, Wang X, Wu E, Majerciak V, Zhang T, Steenbergen RDM, Wang HK, Banerjee
778 NS, et al: **Oncogenic HPV promotes the expression of the long noncoding RNA lnc-FANCI-2**
779 **through E7 and YY1.** *Proc Natl Acad Sci U S A* 2021, **118**.
780 9. Santegoets LA, van Baars R, Terlou A, Heijmans-Antonissen C, Swagemakers SM, van der Spek PJ,
781 Ewing PC, van Beurden M, van der Meijden WI, Helmerhorst TJ, Blok LJ: **Different DNA damage**
782 **and cell cycle checkpoint control in low- and high-risk human papillomavirus infections of the**
783 **vulva.** *Int J Cancer* 2012, **130:**2874-2885.
784 10. Hoskins EE, Morreale RJ, Werner SP, Higginbotham JM, Laimins LA, Lambert PF, Brown DR,
785 Gillison ML, Nuovo GJ, Witte DP, et al: **The fanconi anemia pathway limits human**
786 **papillomavirus replication.** *J Virol* 2012, **86:**8131-8138.
787 11. Park JW, Pitot HC, Strati K, Spardy N, Duensing S, Grompe M, Lambert PF: **Deficiencies in the**
788 **Fanconi anemia DNA damage response pathway increase sensitivity to HPV-associated head**
789 **and neck cancer.** *Cancer Res* 2010, **70:**9959-9968.
790 12. Park JW, Shin MK, Lambert PF: **High incidence of female reproductive tract cancers in FA-**
791 **deficient HPV16-transgenic mice correlates with E7's induction of DNA damage response, an**
792 **activity mediated by E7's inactivation of pocket proteins.** *Oncogene* 2014, **33:**3383-3391.

793 13. Hietanen S, Lain S, Krausz E, Blattner C, Lane DP: **Activation of p53 in cervical carcinoma cells by**
794 **small molecules.** *Proc Natl Acad Sci U S A* 2000, **97:**8501-8506.
795 14. Tommasino M, Accardi R, Caldeira S, Dong W, Malanchi I, Smet A, Zehbe I: **The role of TP53 in**
796 **Cervical carcinogenesis.** *Hum Mutat* 2003, **21:**307-312.
797 15. Prior IA, Lewis PD, Mattos C: **A comprehensive survey of Ras mutations in cancer.** *Cancer Res*
798 2012, **72:**2457-2467.
799 16. Fernández-Medarde A, Santos E: **Ras in cancer and developmental diseases.** *Genes Cancer*
800 2011, **2:**344-358.
801 17. Roman A, Munger K: **The papillomavirus E7 proteins.** *Virology* 2013, **445:**138-168.
802 18. Vande Pol SB, Klingelhutz AJ: **Papillomavirus E6 oncoproteins.** *Virology* 2013, **445:**115-137.
803 19. Liu X, Dakic A, Zhang Y, Dai Y, Chen R, Schlegel R: **HPV E6 protein interacts physically and**
804 **functionally with the cellular telomerase complex.** *Proc Natl Acad Sci U S A* 2009, **106:**18780-
805 18785.
806 20. Klingelhutz AJ, Foster SA, McDougall JK: **Telomerase activation by the E6 gene product of**
807 **human papillomavirus type 16.** *Nature* 1996, **380:**79-82.
808 21. Storey A, Pim D, Murray A, Osborn K, Banks L, Crawford L: **Comparison of the in vitro**
809 **transforming activities of human papillomavirus types.** *EMBO J* 1988, **7:**1815-1820.
810 22. Barbosa MS, Schlegel R: **The E6 and E7 genes of HPV-18 are sufficient for inducing two-stage in**
811 **vitro transformation of human keratinocytes.** *Oncogene* 1989, **4:**1529-1532.
812 23. Hawley-Nelson P, Vousden KH, Hubbert NL, Lowy DR, Schiller JT: **HPV16 E6 and E7 proteins**
813 **cooperate to immortalize human foreskin keratinocytes.** *EMBO J* 1989, **8:**3905-3910.
814 24. Munger K, Phelps WC, Bubb V, Howley PM, Schlegel R: **The E6 and E7 genes of the human**
815 **papillomavirus type 16 together are necessary and sufficient for transformation of primary**
816 **human keratinocytes.** *J Virol* 1989, **63:**4417-4421.
817 25. Halbert CL, Demers GW, Galloway DA: **The E7 gene of human papillomavirus type 16 is**
818 **sufficient for immortalization of human epithelial cells.** *J Virol* 1991, **65:**473-478.
819 26. Phelps WC, Yee CL, Munger K, Howley PM: **The human papillomavirus type 16 E7 gene encodes**
820 **transactivation and transformation functions similar to those of adenovirus E1A.** *Cell* 1988,
821 **53:**539-547.
822 27. DiPaolo JA, Woodworth CD, Popescu NC, Notario V, Doniger J: **Induction of human cervical**
823 **squamous cell carcinoma by sequential transfection with human papillomavirus 16 DNA and**
824 **viral Harvey ras.** *Oncogene* 1989, **4:**395-399.
825 28. Schreiber K, Cannon RE, Karrison T, Beck-Engeser G, Huo D, Tennant RW, Jensen H, Kast WM,
826 Krausz T, Meredith SC, et al: **Strong synergy between mutant ras and HPV16 E6/E7 in the**
827 **development of primary tumors.** *Oncogene* 2004, **23:**3972-3979.
828 29. Narisawa-Saito M, Yoshimatsu Y, Ohno S, Yugawa T, Egawa N, Fujita M, Hirohashi S, Kiyono T: **An**
829 **in vitro multistep carcinogenesis model for human cervical cancer.** *Cancer Res* 2008, **68:**5699-
830 5705.
831 30. Conover CA, Bale LK, Durham SK, Powell DR: **Insulin-like growth factor (IGF) binding protein-3**
832 **potentiation of IGF action is mediated through the phosphatidylinositol-3-kinase pathway and**
833 **is associated with alteration in protein kinase B/AKT sensitivity.** *Endocrinology* 2000, **141:**3098-
834 3103.
835 31. Allard JB, Duan C: **IGF-Binding Proteins: Why Do They Exist and Why Are There So Many?** *Front*
836 *Endocrinol (Lausanne)* 2018, **9:**117.
837 32. Bach LA: **IGF-binding proteins.** *J Mol Endocrinol* 2018, **61:**T11-t28.
838 33. Song S, Pitot HC, Lambert PF: **The human papillomavirus type 16 E6 gene alone is sufficient to**
839 **induce carcinomas in transgenic animals.** *J Virol* 1999, **73:**5887-5893.

840 34. Herber R, Liem A, Pitot H, Lambert PF: **Squamous epithelial hyperplasia and carcinoma in mice**
841 **transgenic for the human papillomavirus type 16 E7 oncogene.** *J Virol* 1996, **70:**1873-1881.
842 35. Song S, Liem A, Miller JA, Lambert PF: **Human papillomavirus types 16 E6 and E7 contribute**
843 **differently to carcinogenesis.** *Virology* 2000, **267:**141-150.
844 36. Riley RR, Duensing S, Brake T, Munger K, Lambert PF, Arbeit JM: **Dissection of human**
845 **papillomavirus E6 and E7 function in transgenic mouse models of cervical carcinogenesis.**
846 *Cancer Res* 2003, **63:**4862-4871.
847 37. Brake T, Lambert PF: **Estrogen contributes to the onset, persistence, and malignant**
848 **progression of cervical cancer in a human papillomavirus-transgenic mouse model.** *Proc Natl*
849 *Acad Sci U S A* 2005, **102:**2490-2495.
850 38. Kato S, Endoh H, Masuhiro Y, Kitamoto T, Uchiyama S, Sasaki H, Masushige S, Gotoh Y, Nishida E,
851 Kawashima H, et al: **Activation of the estrogen receptor through phosphorylation by mitogen-**
852 **activated protein kinase.** *Science* 1995, **270:**1491-1494.
853 39. Statello L, Guo CJ, Chen LL, Huarte M: **Gene regulation by long non-coding RNAs and its**
854 **biological functions.** *Nat Rev Mol Cell Biol* 2021, **22:**96-118.
855 40. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT: **Polycomb proteins targeted by a short repeat RNA to**
856 **the mouse X chromosome.** *Science* 2008, **322:**750-756.
857 41. Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM:
858 **Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by**
859 **polycomb CBX7 in transcriptional silencing of INK4a.** *Mol Cell* 2010, **38:**662-674.
860 42. Cho SW, Xu J, Sun R, Mumbach MR, Carter AC, Chen YG, Yost KE, Kim J, He J, Nevins SA, et al:
861 **Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element.** *Cell* 2018,
862 **173:**1398-1412 e1322.
863 43. Xiang JF, Yin QF, Chen T, Zhang Y, Zhang XO, Wu Z, Zhang S, Wang HB, Ge J, Lu X, et al: **Human**
864 **colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the**
865 **MYC locus.** *Cell Res* 2014, **24:**513-531.
866 44. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya
867 PA, et al: **The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by**
868 **modulating SR splicing factor phosphorylation.** *Mol Cell* 2010, **39:**925-938.
869 45. Yap K, Mukhina S, Zhang G, Tan JSC, Ong HS, Makeyev EV: **A Short Tandem Repeat-Enriched**
870 **RNA Assembles a Nuclear Compartment to Control Alternative Splicing and Promote Cell**
871 **Survival.** *Mol Cell* 2018, **72:**525-540 e513.
872 46. Nishizawa M, Ikeya Y, Okumura T, Kimura T: **Post-transcriptional inducible gene regulation by**
873 **natural antisense RNA.** *Front Biosci (Landmark Ed)* 2015, **20:**1-36.
874 47. Lee S, Kopp F, Chang TC, Sataluri A, Chen B, Sivakumar S, Yu H, Xie Y, Mendell JT: **Noncoding**
875 **RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins.** *Cell* 2016,
876 **164:**69-80.
877 48. Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP: **A coding-independent**
878 **function of gene and pseudogene mRNAs regulates tumour biology.** *Nature* 2010, **465:**1033-
879 1038.
880 49. Grelet S, Link LA, Howley B, Obellianne C, Palanisamy V, Gangaraju VK, Diehl JA, Howe PH: **A**
881 **regulated PNUTS mRNA to lncRNA splice switch mediates EMT and tumour progression.** *Nat*
882 *Cell Biol* 2017, **19:**1105-1115.
883 50. Aznaourova M, Janga H, Sefried S, Kaufmann A, Dorna J, Volkers SM, Georg P, Lechner M, Hoppe
884 J, Dokel S, et al: **Noncoding RNA MaIL1 is an integral component of the TLR4-TRIF pathway.**
885 *Proc Natl Acad Sci U S A* 2020, **117:**9042-9053.

886 51. Guo CJ, Ma XK, Xing YH, Zheng CC, Xu YF, Shan L, Zhang J, Wang S, Wang Y, Carmichael GG, et al:
887 **Distinct Processing of lncRNAs Contributes to Non-conserved Functions in Stem Cells.** *Cell*
888 2020, **181:**621-636 e622.
889 52. Scheffner M, Münger K, Byrne JC, Howley PM: **The state of the p53 and retinoblastoma genes**
890 **in human cervical carcinoma cell lines.** *Proceedings of the National Academy of Sciences* 1991,
891 **88:**5523-5527.
892 53. BeltCappellino A, Majerciak V, Lobanov A, Lack J, Cam M, Zheng ZM: **CRISPR/Cas9-Mediated**
893 **Knockout and In Situ Inversion of the ORF57 Gene from All Copies of the Kaposi's Sarcoma-**
894 **Associated Herpesvirus Genome in BCBL-1 Cells.** *J Virol* 2019, **93**.
895 54. Schlomann U, Wildeboer D, Webster A, Antropova O, Zeuschner D, Knight CG, Docherty AJ,
896 Lambert M, Skelton L, Jockusch H, Bartsch JW: **The metalloprotease disintegrin ADAM8.**
897 **Processing by autocatalysis is required for proteolytic activity and cell adhesion.** *J Biol Chem*
898 2002, **277:**48210-48219.
899 55. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P: **The Molecular**
900 **Signatures Database (MSigDB) hallmark gene set collection.** *Cell Syst* 2015, **1:**417-425.
901 56. Downward J: **Targeting RAS signalling pathways in cancer therapy.** *Nat Rev Cancer* 2003, **3:**11-
902 22.
903 57. Simanshu DK, Nissley DV, McCormick F: **RAS Proteins and Their Regulators in Human Disease.**
904 *Cell* 2017, **170:**17-33.
905 58. Castellano E, Downward J: **RAS Interaction with PI3K: More Than Just Another Effector**
906 **Pathway.** *Genes Cancer* 2011, **2:**261-274.
907 59. Asati V, Mahapatra DK, Bharti SK: **PI3K/Akt/mTOR and Ras/Raf/MEK/ERK signaling pathways**
908 **inhibitors as anticancer agents: Structural and pharmacological perspectives.** *Eur J Med Chem*
909 2016, **109:**314-341.
910 60. Cuesta C, Arevalo-Alameda C, Castellano E: **The Importance of Being PI3K in the RAS Signaling**
911 **Network.** *Genes (Basel)* 2021, **12**.
912 61. Vanhaesebroeck B, Perry MWD, Brown JR, Andre F, Okkenhaug K: **PI3K inhibitors are finally**
913 **coming of age.** *Nat Rev Drug Discov* 2021, **20:**741-769.
914 62. Favata MF, Horiuchi KY, Manos EJ, Daulerio AJ, Stradley DA, Feeser WS, Van Dyk DE, Pitts WJ,
915 Earl RA, Hobbs F, et al: **Identification of a novel inhibitor of mitogen-activated protein kinase**
916 **kinase.** *J Biol Chem* 1998, **273:**18623-18632.
917 63. Serrano ML, Sánchez-Gómez M, Bravo MM: **Insulin-like growth factor system gene expression**
918 **in cervical scrapes from women with squamous intraepithelial lesions and cervical cancer.**
919 *Growth Horm IGF Res* 2007, **17:**492-499.
920 64. Wang Z, Xu Q, Zhang N, Du X, Xu G, Yan X: **CD146, from a melanoma cell adhesion molecule to**
921 **a signaling receptor.** *Signal Transduct Target Ther* 2020, **5:**148.
922 65. Kebir A, Harhouri K, Guillet B, Liu JW, Foucault-Bertaud A, Lamy E, Kaspi E, Elganfoud N, Vely F,
923 Sabatier F, et al: **CD146 short isoform increases the proangiogenic potential of endothelial**
924 **progenitor cells in vitro and in vivo.** *Circ Res* 2010, **107:**66-75.
925 66. Stalin J, Harhouri K, Hubert L, Garrigue P, Nollet M, Essaadi A, Muller A, Foucault-Bertaud A,
926 Bachelier R, Sabatier F, et al: **Soluble CD146 boosts therapeutic effect of endothelial**
927 **progenitors through proteolytic processing of short CD146 isoform.** *Cardiovasc Res* 2016,
928 **111:**240-251.
929 67. Joshkon A, Heim X, Dubrou C, Bachelier R, Traboulsi W, Stalin J, Fayyad-Kazan H, Badran B,
930 Foucault-Bertaud A, Leroyer AS, et al: **Role of CD146 (MCAM) in Physiological and Pathological**
931 **Angiogenesis-Contribution of New Antibodies for Therapy.** *Biomedicines* 2020, **8**.
932 68. Baxter RC: **Insulin-like growth factor binding protein-3 (IGFBP-3): Novel ligands mediate**
933 **unexpected functions.** *J Cell Commun Signal* 2013, **7:**179-189.

934 69. Varma Shrivastav S, Bhardwaj A, Pathak KA, Shrivastav A: **Insulin-Like Growth Factor Binding**
935 **Protein-3 (IGFBP-3): Unraveling the Role in Mediating IGF-Independent Effects Within the Cell.**
936 *Front Cell Dev Biol* 2020, **8:**286.
937 70. Chu C, Chang HY: **ChIRP-MS: RNA-Directed Proteomic Discovery.** *Methods Mol Biol* 2018,
938 **1861:**37-45.
939 71. Gao X, Gao C, Liu G, Hu J: **MAP4K4: an emerging therapeutic target in cancer.** *Cell Biosci* 2016,
940 **6:**56.
941 72. Gao X, Chen G, Gao C, Zhang DH, Kuan SF, Stabile LP, Liu G, Hu J: **MAP4K4 is a novel MAPK/ERK**
942 **pathway regulator required for lung adenocarcinoma maintenance.** *Mol Oncol* 2017, **11:**628-
943 639.
944 73. Gonzalez-Montero J, Rojas CI, Burotto M: **MAP4K4 and cancer: ready for the main stage?** *Front*
945 *Oncol* 2023, **13:**1162835.
946 74. Patterson V, Ullah F, Bryant L, Griffin JN, Sidhu A, Saliganan S, Blaile M, Saenz MS, Smith R,
947 Ellingwood S, et al: **Abrogation of MAP4K4 protein function causes congenital anomalies in**
948 **humans and zebrafish.** *Sci Adv* 2023, **9:**eade0631.
949 75. Roth Flach RJ, Guo CA, Danai LV, Yawe JC, Gujja S, Edwards YJ, Czech MP: **Endothelial Mitogen-**
950 **Activated Protein Kinase Kinase Kinase Kinase 4 Is Critical for Lymphatic Vascular**
951 **Development and Function.** *Mol Cell Biol* 2016, **36:**1740-1749.
952 76. Crook T, Storey A, Almond N, Osborn K, Crawford L: **Human papillomavirus type 16 cooperates**
953 **with activated ras and fos oncogenes in the hormone-dependent transformation of primary**
954 **mouse cells.** *Proc Natl Acad Sci U S A* 1988, **85:**8820-8824.
955 77. Matlashewski G, Schneider J, Banks L, Jones N, Murray A, Crawford L: **Human papillomavirus**
956 **type 16 DNA cooperates with activated ras in transforming primary cells.** *EMBO J* 1987,
957 **6:**1741-1746.
958 78. Greenhalgh DA, Wang XJ, Rothnagel JA, Eckhardt JN, Quintanilla MI, Barber JL, Bundman DS,
959 Longley MA, Schlegel R, Roop DR: **Transgenic mice expressing targeted HPV-18 E6 and E7**
960 **oncogenes in the epidermis develop verrucous lesions and spontaneous, rasHa-activated**
961 **papillomas.** *Cell Growth Differ* 1994, **5:**667-675.
962 79. Riou G, Barrois M, Sheng ZM, Duvillard P, Lhomme C: **Somatic deletions and mutations of c-Ha-**
963 **ras gene in human cervical cancers.** *Oncogene* 1988, **3:**329-333.
964 80. Eiben GL, Velders MP, Schreiber H, Cassetti MC, Pullen JK, Smith LR, Kast WM: **Establishment of**
965 **an HLA-A*0201 human papillomavirus type 16 tumor model to determine the efficacy of**
966 **vaccination strategies in HLA-A*0201 transgenic mice.** *Cancer Res* 2002, **62:**5792-5799.
967 81. Golijow CD, Mouron SA, Gomez MA, Dulout FN: **Differences in K-ras codon 12 mutation**
968 **frequency between "high-risk" and "low-risk" HPV-infected samples.** *Gynecol Oncol* 1999,
969 **75:**108-112.
970 82. Prokopakis P, Sourvinos G, Koumantaki Y, Koumantakis E, Spandidos DA: **K-ras mutations and**
971 **HPV infection in cervicitis and intraepithelial neoplasias of the cervix.** *Oncol Rep* 2002, **9:**129-
972 133.
973 83. Strickland SW, Vande Pol S: **The Human Papillomavirus 16 E7 Oncoprotein Attenuates AKT**
974 **Signaling To Promote Internal Ribosome Entry Site-Dependent Translation and Expression of**
975 **c-MYC.** *J Virol* 2016, **90:**5611-5621.
976 84. Yang J, Nie J, Ma X, Wei Y, Peng Y, Wei X: **Targeting PI3K in cancer: mechanisms and advances**
977 **in clinical trials.** *Mol Cancer* 2019, **18:**26.
978 85. Saliani M, Mirzaiebadizi A, Javadmanesh A, Siavoshi A, Ahmadian MR: **KRAS-related long**
979 **noncoding RNAs in human cancers.** *Cancer Gene Ther* 2021.

980 86. Martin JL, Weenink SM, Baxter RC: **Insulin-like growth factor-binding protein-3 potentiates**
981 **epidermal growth factor action in MCF-10A mammary epithelial cells. Involvement of p44/42**
982 **and p38 mitogen-activated protein kinases.** *J Biol Chem* 2003, **278:**2969-2976.
983 87. Józefiak A, Larska M, Pomorska-Mól M, Ruszkowski JJ: **The IGF-1 Signaling Pathway in Viral**
984 **Infections.** *Viruses* 2021, **13**.
985 88. Jozefiak A, Pacholska-Bogalska J, Myga-Nowak M, Kedzia W, Kwasniewska A, Luczak M, Kedzia H,
986 Gozdzicka-Jozefiak A: **Serum and tissue levels of insulin-like growth factor-I in women with**
987 **dysplasia and HPV-positive cervical cancer.** *Mol Med Rep* 2008, **1:**231-237.
988 89. Serrano ML, Romero A, Cendales R, Sánchez-Gómez M, Bravo MM: **Serum levels of insulin-like**
989 **growth factor-I and -II and insulin-like growth factor binding protein 3 in women with**
990 **squamous intraepithelial lesions and cervical cancer.** *Biomedica* 2006, **26:**258-268.
991 90. Anfosso F, Bardin N, Frances V, Vivier E, Camoin-Jau L, Sampol J, Dignat-George F: **Activation of**
992 **human endothelial cells via S-endo-1 antigen (CD146) stimulates the tyrosine phosphorylation**
993 **of focal adhesion kinase p125(FAK).** *J Biol Chem* 1998, **273:**26852-26856.
994 91. Dye DE, Karlen S, Rohrbach B, Staub O, Braathen LR, Eidne KA, Coombe DR: **hShroom1 links a**
995 **membrane bound protein to the actin cytoskeleton.** *Cell Mol Life Sci* 2009, **66:**681-696.
996 92. Bardin N, Frances V, Combes V, Sampol J, Dignat-George F: **CD146: biosynthesis and production**
997 **of a soluble form in human cultured endothelial cells.** *FEBS Lett* 1998, **421:**12-14.
998 93. Yu L, Zheng ZM: **Human Papillomavirus Type 16 Circular RNA Is Barely Detectable for the**
999 **Claimed Biological Activity.** *mBio* 2022, **13:**e0359421.
1000 94. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR:
1001 **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics* 2013, **29:**15-21.
1002 95. Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-Seq data with or without a**
1003 **reference genome.** *BMC Bioinformatics* 2011, **12:**323.
1004 96. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D,
1005 Zadissa A, Searle S, et al: **GENCODE: the reference human genome annotation for The ENCODE**
1006 **Project.** *Genome Res* 2012, **22:**1760-1774.
1007 97. Law CW, Chen Y, Shi W, Smyth GK: **voom: Precision weights unlock linear model analysis tools**
1008 **for RNA-seq read counts.** *Genome Biol* 2014, **15:**R29.
1009 98. Smyth GK: **Linear models and empirical bayes methods for assessing differential expression in**
1010 **microarray experiments.** *Stat Appl Genet Mol Biol* 2004, **3:**Article3.
1011 99. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP: **Molecular**
1012 **signatures database (MSigDB) 3.0.** *Bioinformatics* 2011, **27:**1739-1740.
1013 100. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A,
1014 Pomeroy SL, Golub TR, Lander ES, Mesirov JP: **Gene set enrichment analysis: a knowledge-**
1015 **based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci U S A*
1016 2005, **102:**15545-15550.
1017 101. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A: **Fast gene set**
1018 **enrichment analysis.** *bioRxiv* 2021.
1019 102. Chu C, Quinn J, Chang HY: **Chromatin isolation by RNA purification (ChIRP).** *J Vis Exp* 2012.

1020

1021 **Figure legends**

1022

1023 **Fig. 1.** Increased expression of lnc-FANCI-2 in HPV16-infected cervical cancer tissues and
1024 HPV16- and HPV18-infected raft culture tissues. (A) Expression lnc-FANCI-2 in HPV16+
1025 cervical cancer tissues were examined by RNAscope RNA-ISH analysis. Nuclei were stained
1026 with DAPI (blue). (B) Human foreskin keratinocyte (HFK)-derived raft cultures without (HFK)
1027 or with HPV16 or HPV18 infections (HFK-HPV16 or HFK-HPV18) were examined at day 10
1028 for lnc-FANCI-2 RNA by RNAscope RNA-ISH analysis. Scale bars: 25 μm in the original
1029 figures and 10 μm in the zoomed insets.

1030 **Fig. 2.** Knockout (KO) of lnc-FANCI-2 in CaSki cells affects cell proliferation, colony formation
1031 and migration. (A) Diagram and KO strategies of lnc-FANCI-2 gene. On the figure top is the
1032 lnc-FANCI-2 gene structure and its alternative transcription start sites (TSS) and polyadenylation
1033 sites (pA). TSS2 or pA2 with a heavier arrow are predominately used for lnc-FANCI-2
1034 expression. The lower figure part shows KO strategies. Red slashes represent gRNA-targeted
1035 sites to create genomic DNA deletions by CRISPR-Cas9 technology. Deletion of YY1-binding
1036 motifs (ΔYY1) led to a deletion of 86-bp DNA fragment containing two YY1-binding motifs.
1037 Deletion of lnc-FANCI-2 promoter (ΔPr) led to delete a ~3.3-kb promoter region. Primers (P1-
1038 P4) used for PCR screening were shown as black arrows. (B) PCR screening of single cell clones
1039 with homozygous lnc-FANCI-2 KO, with indicated primer sets diagramed (A). Single cell clones
1040 selected from the cells transfected with an empty vector served as control (Ctrl) cells. (C)
1041 Evaluation of lnc-FANCI-2 KO efficiency in the selected single cell clones (B) by RT-qPCR.
1042 (D) Northern blot validation of lnc-FANCI-2 KO efficiency from the individual single cell
1043 clones or the parental (WT) CaSki cells on polyA$^+$ RNA (enriched from 100 μg of total cell
1044 RNA, lanes 1-4) or 10 μg of total cell RNA (lanes 5-9). Total cell RNA (2 μg) of HEK293T cells

1045 with ectopic expression of two isoforms (short, S or long, L) of lnc-FANCI-2 cDNA served as a
1046 control. Antisense oligo probe P5 (A) labeled with ^{32}P was used for hybridization. GAPDH RNA
1047 served as a loading control and hybridized with a GAPDH-specific oligo probe. (E-F) RNA in
1048 situ hybridization (RNA-ISH) validation of lnc-FANCI-2 KO efficiency by RNAscope
1049 technology. Two single cell clones were examined with red color for lnc-FANCI-2 and blue
1050 color for the nucleus (E). Bar graphs show the copy number of lnc-FANCI-2 per cell in the WT,
1051 ΔYY1-D5 or ΔPr-A9 CaSki cells (each averaged from 200 cells) (F). **, P<0.01; ***, P<0.001
1052 by two tailed Student *t* test.

1053 **Fig. 3.** KO of lnc-FANCI-2 in CaSki cells affects expression of cellular soluble receptors. (A)
1054 Dot blots shows differentially expressed soluble receptors in ΔPr-A9 cells when compared to
1055 those in the WT CaSki cells. A Proteome Profiler Human sReceptor Array was used to examine
1056 105 cellular soluble receptors. Label in red, upregulated soluble receptors; label in blue,
1057 downregulated soluble receptors. (B) Quantification of differentially expressed and/or released
1058 soluble receptors (A) in ΔPr-A9 cells when compared to those in the WT CaSki cells. Error bar
1059 represents standard deviation of replicates in each proteome array. The dash line indicates the
1060 threshold of fold change (FC) with a score above or below the threshold (-/+ 0.3 FC) being
1061 determined as differentially expressed and/or released soluble receptors. (C) Validation of
1062 several differentially expressed soluble receptors by immunoblot analysis using specific
1063 antibodies. Tubulin served as an internal loading control. * non-specific band in MCAM
1064 immunoblot.

1065 **Fig. 4.** Transcriptomic effect of lnc-FANCI-2 KO in CaSki cells by RNA-seq analysis. (A)
1066 RNA-seq reads-coverage maps by IGV showing the expression levels of lnc-FANCI-2 from two
1067 KO cell clones, ΔPr-A9 and ΔYY1-D5, to the WT CaSki cells. one representative coverage

1068 profile of four in each type of cells is shown by IGV. Red slashes represent the deleted genomic
1069 region, with validated fourteen lnc-FANCI-2 RNA isoforms shown below [8]. (B) Similarity in
1070 heatmap comparison among ΔPr-A9, ΔYY1-D5 and WT cells was generated using Limma-
1071 normalized counts from each sample by Pearson complete linkage method. (C) Volcano plot
1072 visualization of differentially expressed genes (DEGs) in ΔPr-A9 and ΔYY1-D5. The genes with
1073 the most significant change in the increased (red) or decreased (blue) expression are indicated.
1074 FC, fold change. (D) Venn diagram depicting the overlapped DEGs between ΔPr-A9 and ΔYY1-
1075 D5 cells over the WT CaSki cells. (E) Heatmap shows overlapped DEGs with FPKM ≥7 in ΔPr-
1076 A9 and ΔYY1-D5 cells when compared with the WT CaSki cells. (F) Validation of selective 12
1077 upregulated or downregulated DEGs shown in the heatmap by RT-PCR with gene-specific
1078 primers in the presence (+) or absence (-) of reverse transcriptase (RT). GAPDH served as an
1079 internal RNA control. Relative expression level of each gene was calculated based on band
1080 density after normalizing to GAPDH RNA band, with the expression level in the WT CaSki cells
1081 setting as 1. (G) Validation of the upregulated expression of IGFBP3 in ΔPr-A9 cells over the
1082 WT CaSki cells by immunoblot analysis.

1083 **Fig. 5.** Pathway analyses of DEGs identified by RNA-seq. (A) Top 3 upregulated and top 4
1084 downregulated pathways in ΔPr-A9 and ΔYY1-D5 cells over the WT CaSki cells. Data were
1085 generated by Gene set enrichment analysis (GSEA) performed with Hallmark gene sets. NES
1086 stands for normalized enrichment score. (B-E) Each GSEA Enrichment plot shows the
1087 enrichment score and gene hits enriched in ΔPr-A9 cells using Hallmark gene sets,
1088 KRAS_SIGNALING_UP (B), EPITHELIAL_MESENCHYMAL_TRANSITION (C),
1089 INTERFERON_GAMMA_ RESPONSE (D) or INTERFERON_ALPHA_RESPONSE (E). padj,
1090 adjusted p value; Zero crossed 7342, the middle number of total genes in the GSEA at ranking.

1091 The heatmaps below the enrichment plots (B and C) visualize the genes enriched in respective
1092 pathways of KRAS_SIGNALING_UP (B) and
1093 EPITHELIAL_MESENCHYMAL_TRANSITION (C) in ΔPr-A9 cells when compared to the
1094 WT CaSki cells.

1095 **Fig. 6.** CaSki cells with lnc-FANCI-2 KO exhibit activation of RAS signaling pathway. (A)
1096 CaSki cells with lnc-FANCI-2 KO (top, ΔPr-A9 cells) or knockdown (KD, lower) by lnc-
1097 FANCI-2-specific siRNAs display increased RAS GTPase activity when compared to the WT
1098 CaSki cells (top) or the WT cells treated with a nonspecific control siRNA (siNS). The relative
1099 RAS GTPase activity was measured by a RAS GTPase Chemi ELISA assay. *, $P< 0.05$; ***,
1100 $P<0.001$ by two tailed Student *t* test. (B and C) Selective validation of increased expression of
1101 RAS signaling-related downstream genes in lnc-FANCI-2 KO ΔPr-A9 cells (B) or lnc-FANCI-2
1102 KD CaSki cells (C) over the WT CaSki cells. Relative expression of the indicated genes in ΔPr-
1103 A9 cells in comparison to the WT cells (B) or in the WT CaSki cells treated for 48 h and 96 h
1104 with a nonspecific control siRNA (siNS) or a siRNA specifically targeting the lnc-FANCI-2
1105 exon 3 (C) were examined by immunoblot analyses using corresponding antibodies as indicated.
1106 Tubulin served as an internal control for each blot. The level of p-Akt or p-Erk was calculated by
1107 normalizing to total Akt or Erk protein and the other proteins by normalizing to tubulin. The
1108 protein level in the WT cells or WT cells treated with siNS set as 1. The KD efficiency of lnc-
1109 FANCI-2 RNA (C, top bar graphs) was examined by RT-qPCR. (D) The effect of blocking RAS
1110 signaling on the expression of MCAM and VIM using PI3K inhibitor LY294002 (20 μM) and
1111 MEK inhibitor U0126 (10 μM). The protein at each time point was examined by immunoblot
1112 analysis with a corresponding antibody. The level of MCAM or VIM was calculated after
1113 normalizing to tubulin. The protein level in ΔPr-A9 cells without the inhibitor was set as 100%.

1114 *, nonspecific protein band. (E) The time dependent cell viability of the WT CaSki and ΔPr-A9
1115 cells in the presence of 20 μM LY294002 or 10 μM U0126. Data was obtained in each time point
1116 after normalizing to the cells treated with DMSO. The mean ± SD at each data point was
1117 calculated from 6 samples combined from two independent experiments.

1118 **Fig. 7.** lnc-FANCI-2 is suppressive to the expression of IGFBP3 and MCAM. (A and B)
1119 Transient rescue expression of lnc-FANCI-2 in Δpr-A9 cells inhibits expression of IGFBP3.
1120 ΔPr-A9 cells were transfected with a major isoform lnc-FANCI-2a-PA2 (GenBank ACC. No.
1121 MT669800.1) cDNA plasmid. lnc-FANCI-2 in red and IGFBP3 RNA in green were detected by
1122 RNAscope RNA-ISH at 24h post transfection using each specific antisense RNA probe and
1123 imaged by confocal microscopy (A). Expression levels of lnc-FANCI-2 RNA and IGFBP3 RNA
1124 in four neighboring cells (B) were measured by signal intensity of a line crossing over the stained
1125 cells in A (white line arrow). (C) Validation of differential expression of MCAM RNA in ΔPr-
1126 A9 cells by RT-PCR in the presence (+) or absence (-) of reverse transcriptase (RT). One pair of
1127 primers with one primer at the exons 11 and the other at exon 13 of MCAM RNA (NM_006500)
1128 were used. GAPDH served as a loading control. (D) The major isofrom lnc-FANCI-2a-PA2
1129 repressed the expression of MCAM. ΔPr-A9 cells were transfected with lnc-FANCI-2a-PA2
1130 cDNA plasmid. lnc-FANCI-2 RNA in red was detected by RNAscope ISH and MCAM protein
1131 in green was detected by IF with an anti-MCAM antibody. (E) Expression levels of lnc-FANCI-
1132 2 RNA and MCAM protein in three neighboring cells were measured by signal intensity of a line
1133 crossing over the stained cells (D, white line arrow). (F) Calculation of the expression levels of
1134 MCAM in lnc-FANCI-2 positive cells (n=35) and lnc-FANCI-2 negative cells (n=54) by
1135 fluorescent intensity. (G) Subcellular MCAM distributions in the nucleus (N) and cytoplasm (C)
1136 of WT CaSki cells and ΔPr-A9 by immunoblot analysis. Fractionation efficiency and sample

1137 loading were controlled by cytoplasmic (represented by tubulin) and nuclear (represented by
1138 hnRNP C1/C2) proteins. (H-I) Correlation and survival analysis of lnc-FANCI-2 and MCAM
1139 expression with cervical squamous cell carcinoma (CESC) cases from the cancer genome atlas
1140 (TCGA) datasets by GEPIA web server (http://gepia.cancer-pku.cn/). The negative correlation at
1141 R=-0.321 (with p-value=1.02e-08) of Inc-FANCI-2 with MCAM in cervical cancer patients was
1142 obtained from the RNA-seq data from the TCGA for CESC tumor type downloaded from the
1143 TCGA data portal (https://portal.gdc.cancer.gov/). Only primary solid tumor samples (n=304
1144 after exclusion of 2 metastatic samples and 3 normal samples) were subjected to analysis, with
1145 the data showing as a Scatter plot (H). Kaplan-Meier plot with log rank test p-value at 5.74e-05
1146 (I) shows MCAM a biomarker for poor prognosis of cervical cancer survival with a lower
1147 quartile group cutoff. RNA-seq and survival data are derived from the TCGA CESC cancer
1148 patients. (J/K) AKT and ERK phosphorylation is partially regulated by MCAM and IGFBP3 in
1149 CaSki cells. KD of MCAM (J) and IGFBP3 (K) expression in WT parental CaSki (left) or ΔPr-
1150 A9 (right) cells was performed by treatment of MCAM siRNA or IGFBP3 siRNA along with or
1151 without lnc-FANCI-2 siRNA for 48 h. Expression levels of individual proteins as indicated were
1152 examined by immunoblot analyses using the corresponding antibodies. The level of MCAM, p-
1153 Akt, or p-Erk1/2 was calculated after normalizing to tubulin. The protein level in a non-targeting
1154 siRNA (siNS) control cells was set as 1.

1155 **Fig. 8.** lnc-FANCI-2 interacts with host factors to regulate RAS signaling pathway. (A) The lnc-
1156 FANCI-2-associated proteins in the WT CaSki cells were identified by isolation of RNA-protein
1157 complexes using RNA purification (IRPCRP)-Mass spectrometry technology. (B) lnc-FANCI-2
1158 RNA in the IRPCRP-1 and IRPCRP-2 pulldowns had the pooled antisense biotinylated oligos
1159 (pool 1 with oligos in even numbers and pool 2 with oligos in odd numbers) immobilized to

1160 avidin-beads first before mixed with cell lysates, while the IRPCRP-3 and IRPCRP-4 pulldowns
1161 had the oligos pool 1 and 2 separately mixed with cell lysates first before addition to the avidin-
1162 beads for the RNA pull-downs. RT-PCR in the absence (-) or presence (+) of reverse
1163 transcriptase (RT) was carried out using the RNA isolated from the individual IRPCRP
1164 experiments using a primer pair of oHBL5 and oHBL12 (Table S5) specific for lnc-FANCI-2
1165 RNA detection. Beads only (no oligos) IRPCRP experiments served as a negative control. Total
1166 RNA from the WT CaSki cells after sonication was used as an input control. Arrow indicates the
1167 detected lnc-FANCI-2 RNA. (C) Proteins associated with lnc-FANCI-2 RNA identified from
1168 lnc-FANCI-2 IRPCRP pulldowns. A total of 32 proteins were specifically pulled down from lnc-
1169 FANCI-2 IRPCRP reactions 1-4 (PSM≥2 from two separate pulldowns, Table S4), with top 10
1170 proteins binding lnc-FANCI-2 shown in the order by the number of identified PSM. (D)
1171 Expression of p-Akt and p-Erk from CaSki WT cells 48 h after siRNA-mediated KD of
1172 MAP4K4 or lnc-FANCI-2 was immunoblotted by the corresponding antibodies. GAPDH served
1173 as a sample loading control. Fold change of the indicated proteins in the cells with KD of
1174 MAP4K4 or lnc-FANCI-2 over the cells treated by a non-targeting siRNA (siNS) was calculated
1175 after normalizing to GAPDH. (E) A proposed model illustrates how lnc-FANCI-2-protein
1176 complexes inhibits the RAS signaling pathway to control Akt/ErK phosphorylation and
1177 expression of host genes. In the absence of lnc-FANCI-2 in CaSki cells, RAS signaling can be
1178 triggered by extracellular stimuli, such as IGFBP3, MCAM ligands, etc. As a result,
1179 phosphorylation of AKT and Erk leads to cascaded responses of transcription factors (TFs) to
1180 regulate the expression of RAS signaling responder genes, such as IGFBP3, MCAM, etc.

1181

1182 **Supporting Tables**

1183 **Table S1.** Genome-wide Effect of lnc-FANCI-2 KO in ΔPr-A9 and ΔYY1 cells on gene
1184 expression when compared with parental WT CaSki cells.

1185 **Table S2**. The genes with altered expression found in both ΔPr-A9 and ΔYY1-D5 cells when
1186 compared with parental WT CaSki cells.

1187 **Table S3.** The genes with increased expression in KRAS signaling and EMT and with decreased
1188 expression in IFN-gamma and IFN-alpha responses in lnc-FANCI-2 KO cells (ΔPr-A9 and
1189 ΔYY1-D5) over the parental WT CaSki cells.

1190 **Table S4.** CaSki proteins and their PSMs identified by lnc-FANCI-2 IRPCRP-LC-MS/MS

1191 **Table S5**. Sequences of oligonucleotides used in this study.

1192

1193 **Supporting Figures**

1194 **Figure S1**. Selective expression of lnc-FANCI-2 RNA in a subset of cervical cancer cell lines.

1195 **Figure S2.** Characterization of lnc-FANCI-2 KO on CaSki cell growth and viral oncoprotein
1196 expression.

1197 **Figure S3.** GSEA Enrichment plots show enrichment score and gene hits enriched in ΔYY1-D5
1198 cells using Hallmark gene sets.

1199 **Figure S4.** Pathway map of RAS Initiative with highlighted differentially expressed genes
1200 (DEGs) in ΔPr-A9 cells vs parental wild type cells.

1201 **Figure S5**. Knockout or Knockdown of lnc-FANCI-2 affects RAS signaling.

1202 **Figure S6**. MCAM and IGFBP3 expression and lnc-FANCI-2.

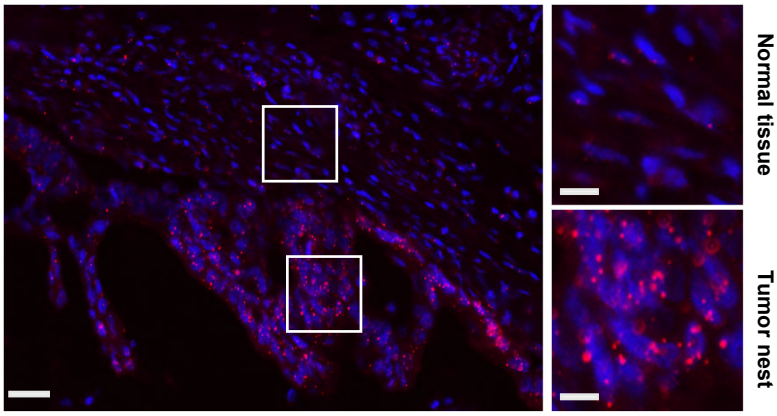
1203 **Figure S7.** Expression levels of lnc-FANCI-2 and MCAM are significantly associated with

1204 TCGA CESC cancer patients' survival outcome.

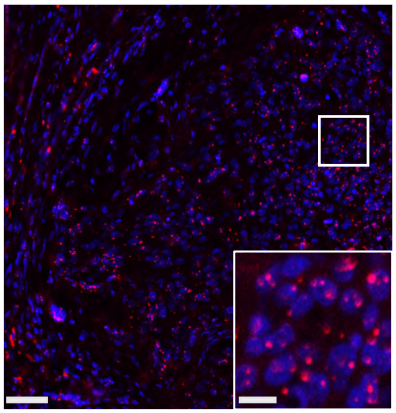
Fig. 1

A

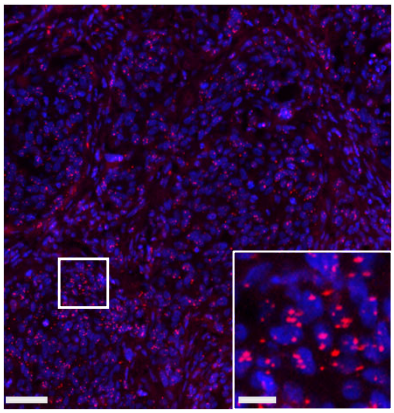
Cervical cancer tissue-tumor nest and its adjacent normal tissue area



Cervical cancer tissue-Case 1

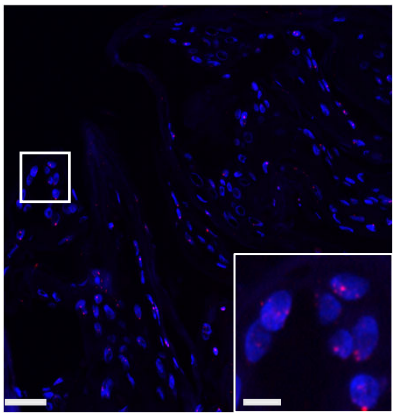


Cervical cancer tissue-Case 2

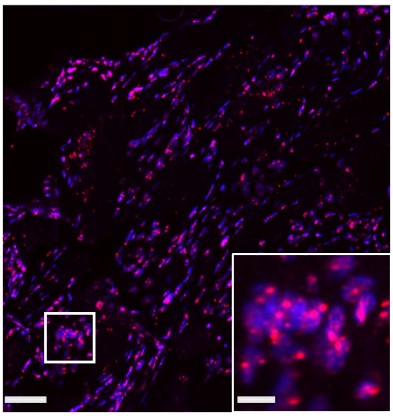


B

HFK



HFK-HPV16



HFK-HPV18

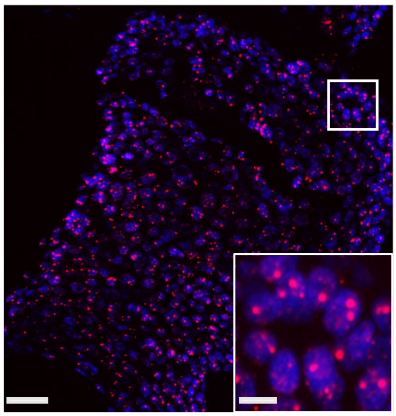


Fig. 2

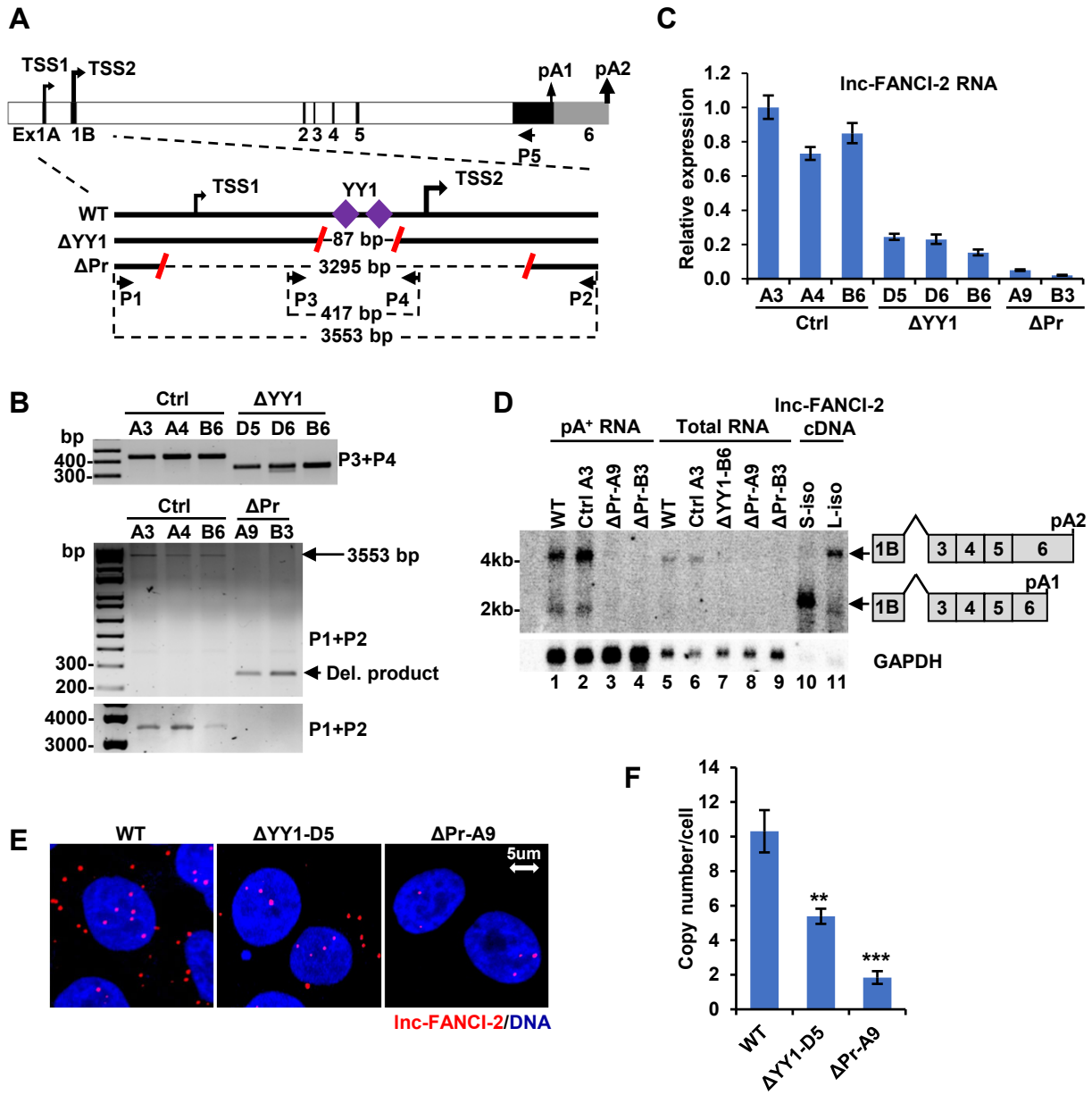

A
TSS1
TSS2
pA1
pA2
Ex1A
1B
2 3 4 5
6
P5
TSS1
YY1
TSS2
WT
ΔYY1
87 bp
ΔPr
3295 bp
P1
P3
P4
P2
417 bp
3553 bp
C
Inc-FANCI-2 RNA
Relative expression
1.2
1.0
0.8
0.6
0.4
0.2
0.0
A3 A4 B6 D5 D6 B6 A9 B3
Ctrl
ΔYY1
ΔPr
B
Ctrl
ΔYY1
A3 A4 B6 D5 D6 B6
bp
400
300
P3+P4
Ctrl
ΔPr
A3 A4 B6 A9 B3
bp
3553 bp
P1+P2
300
200
Del. product
4000
3000
P1+P2
D
Inc-FANCI-2
pA+ RNA
Total RNA
cDNA
WT
Ctrl A3
ΔPr-A9
ΔPr-B3
WT
Ctrl A3
ΔYY1-B6
ΔPr-A9
ΔPr-B3
S-iso
L-iso
4kb
2kb
pA2
1B
3 4 5 6
pA1
1B
3 4 5 6
GAPDH
1 2 3 4 5 6 7 8 9 10 11
E
WT
ΔYY1-D5
ΔPr-A9
5um
Inc-FANCI-2/DNA
F
Copy number/cell
14
12
10
8
6
4
2
0
**
***
WT
ΔYY1-D5
ΔPr-A9

Fig. 3

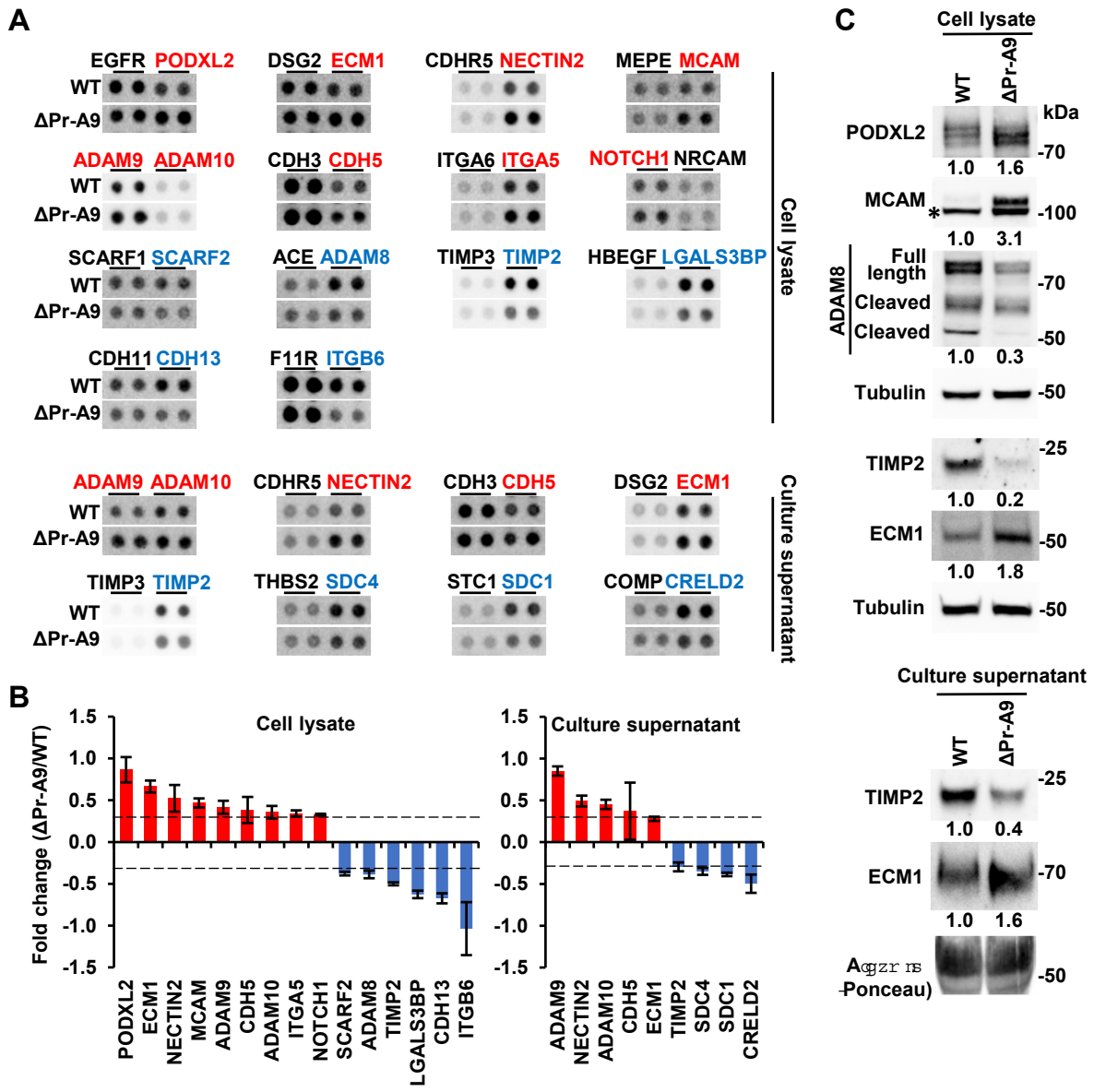

A
EGFR PODXL2
DSG2 ECM1
CDHR5 NECTIN2
MEPE MCAM
ADAM9 ADAM10
CDH3 CDH5
ITGA6 ITGA5
NOTCH1 NRCAM
SCARF1 SCARF2
ACE ADAM8
TIMP3 TIMP2
HBEGF LGALS3BP
CDH11 CDH13
F11R ITGB6
WT
ΔPr-A9
Cell lysate
ADAM9 ADAM10
CDHR5 NECTIN2
CDH3 CDH5
DSG2 ECM1
TIMP3 TIMP2
THBS2 SDC4
STC1 SDC1
COMP CRELD2
Culture supernatant
B
Cell lysate
Culture supernatant
Fold change (ΔPr-A9/WT)
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
PODXL2
ECM1
NECTIN2
MCAM
ADAM9
CDH5
ADAM10
ITGA5
NOTCH1
SCARF2
ADAM8
TIMP2
LGALS3BP
CDH13
ITGB6
ADAM9
NECTIN2
ADAM10
CDH5
ECM1
TIMP2
SDC4
SDC1
CRELD2
C
Cell lysate
WT
ΔPr-A9
kDa
PODXL2
-70
1.0 1.6
MCAM
*
-100
1.0 3.1
ADAM8
Full length
Cleaved
Cleaved
-70
-50
1.0 0.3
Tubulin
-50
TIMP2
-25
1.0 0.2
ECM1
-50
1.0 1.8
Tubulin
-50
Culture supernatant
WT
ΔPr-A9
TIMP2
-25
1.0 0.4
ECM1
-70
1.0 1.6
Ponceau)
-50

Fig. 4

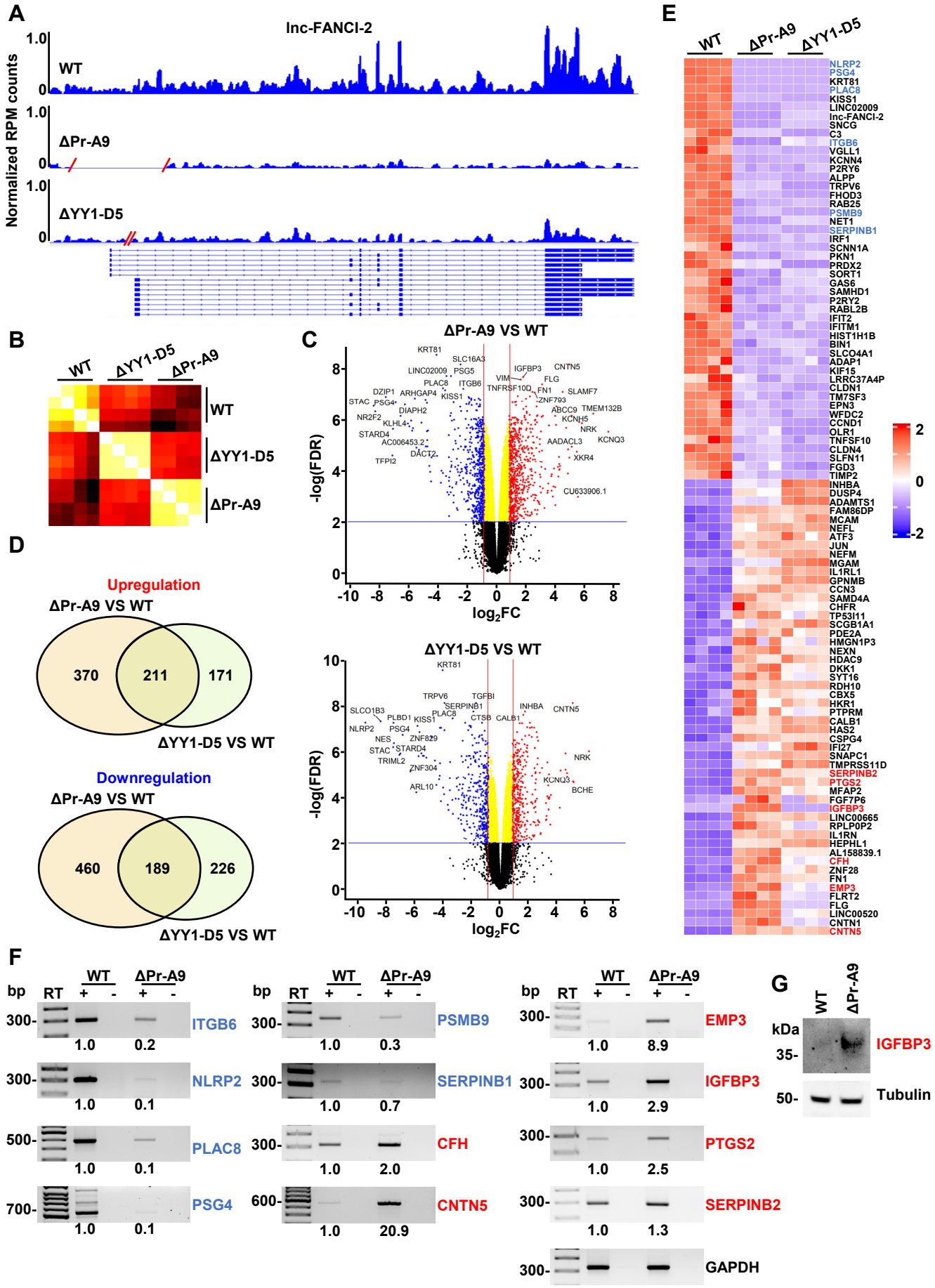

A
lnc-FANCI-2
Normalized RPM counts
WT
ΔPr-A9
ΔYY1-D5
B
WT
ΔYY1-D5
ΔPr-A9
C
ΔPr-A9 VS WT
-log(FDR)
log2FC
ΔYY1-D5 VS WT
D
Upregulation
ΔPr-A9 VS WT
370
211
171
ΔYY1-D5 VS WT
Downregulation
ΔPr-A9 VS WT
460
189
226
ΔYY1-D5 VS WT
E
WT
ΔPr-A9
ΔYY1-D5
F
bp
RT
WT
ΔPr-A9
ITGB6
NLRP2
PLAC8
PSG4
PSMB9
SERPINB1
CFH
CNTN5
EMP3
IGFBP3
PTGS2
SERPINB2
GAPDH
G
WT
ΔPr-A9
kDa
IGFBP3
Tubulin

Fig. 5

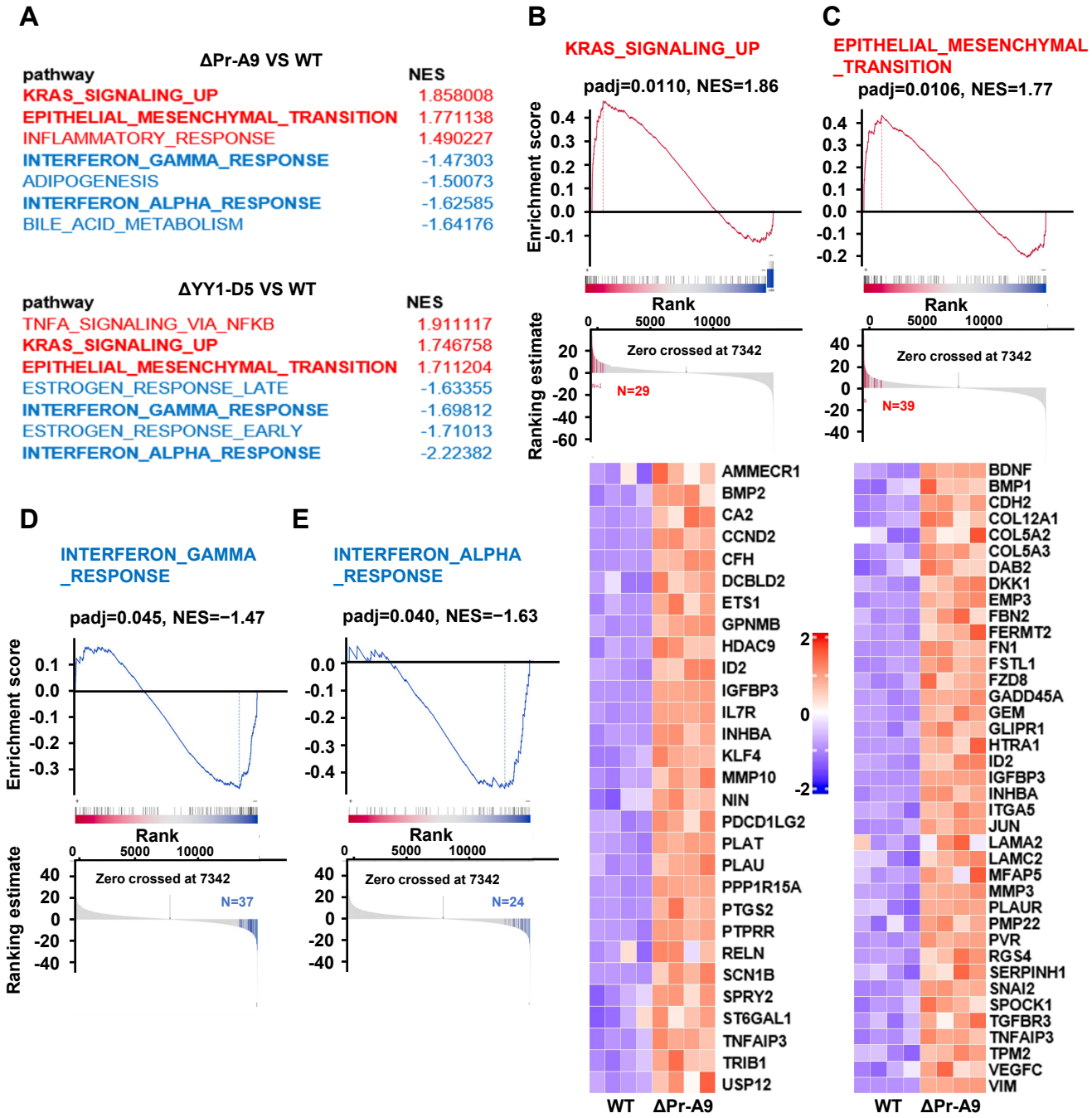

A
ΔPr-A9 VS WT
pathway NES
KRAS_SIGNALING_UP 1.858008
EPITHELIAL_MESENCHYMAL_TRANSITION 1.771138
INFLAMMATORY_RESPONSE 1.490227
INTERFERON_GAMMA_RESPONSE -1.47303
ADIPOGENESIS -1.50073
INTERFERON_ALPHA_RESPONSE -1.62585
BILE_ACID_METABOLISM -1.64176
ΔYY1-D5 VS WT
pathway NES
TNFA_SIGNALING_VIA_NFKB 1.911117
KRAS_SIGNALING_UP 1.746758
EPITHELIAL_MESENCHYMAL_TRANSITION 1.711204
ESTROGEN_RESPONSE_LATE -1.63355
INTERFERON_GAMMA_RESPONSE -1.69812
ESTROGEN_RESPONSE_EARLY -1.71013
INTERFERON_ALPHA_RESPONSE -2.22382
B
KRAS_SIGNALING_UP
padj=0.0110, NES=1.86
Enrichment score
Rank
Ranking estimate
Zero crossed at 7342
N=29
C
EPITHELIAL_MESENCHYMAL_TRANSITION
padj=0.0106, NES=1.77
Zero crossed at 7342
N=39
AMMECR1
BMP2
CA2
CCND2
CFH
DCBLD2
ETS1
GPNMB
HDAC9
ID2
IGFBP3
IL7R
INHBA
KLF4
MMP10
NIN
PDCD1LG2
PLAT
PLAU
PPP1R15A
PTGS2
PTPRR
RELN
SCN1B
SPRY2
ST6GAL1
TNFAIP3
TRIB1
USP12
WT ΔPr-A9
BDNF
BMP1
CDH2
COL12A1
COL5A2
COL5A3
DAB2
DKK1
EMP3
FBN2
FERMT2
FN1
FSTL1
FZD8
GADD45A
GEM
GLIPR1
HTRA1
ID2
IGFBP3
INHBA
ITGA5
JUN
LAMA2
LAMC2
MFAP5
MMP3
PLAUR
PMP22
PVR
RGS4
SERPINH1
SNAI2
SPOCK1
TGFBR3
TNFAIP3
TPM2
VEGFC
VIM
WT ΔPr-A9
D
INTERFERON_GAMMA_RESPONSE
padj=0.045, NES=−1.47
Enrichment score
Rank
Ranking estimate
Zero crossed at 7342
N=37
E
INTERFERON_ALPHA_RESPONSE
padj=0.040, NES=−1.63
Rank
Zero crossed at 7342
N=24

Fig. 6

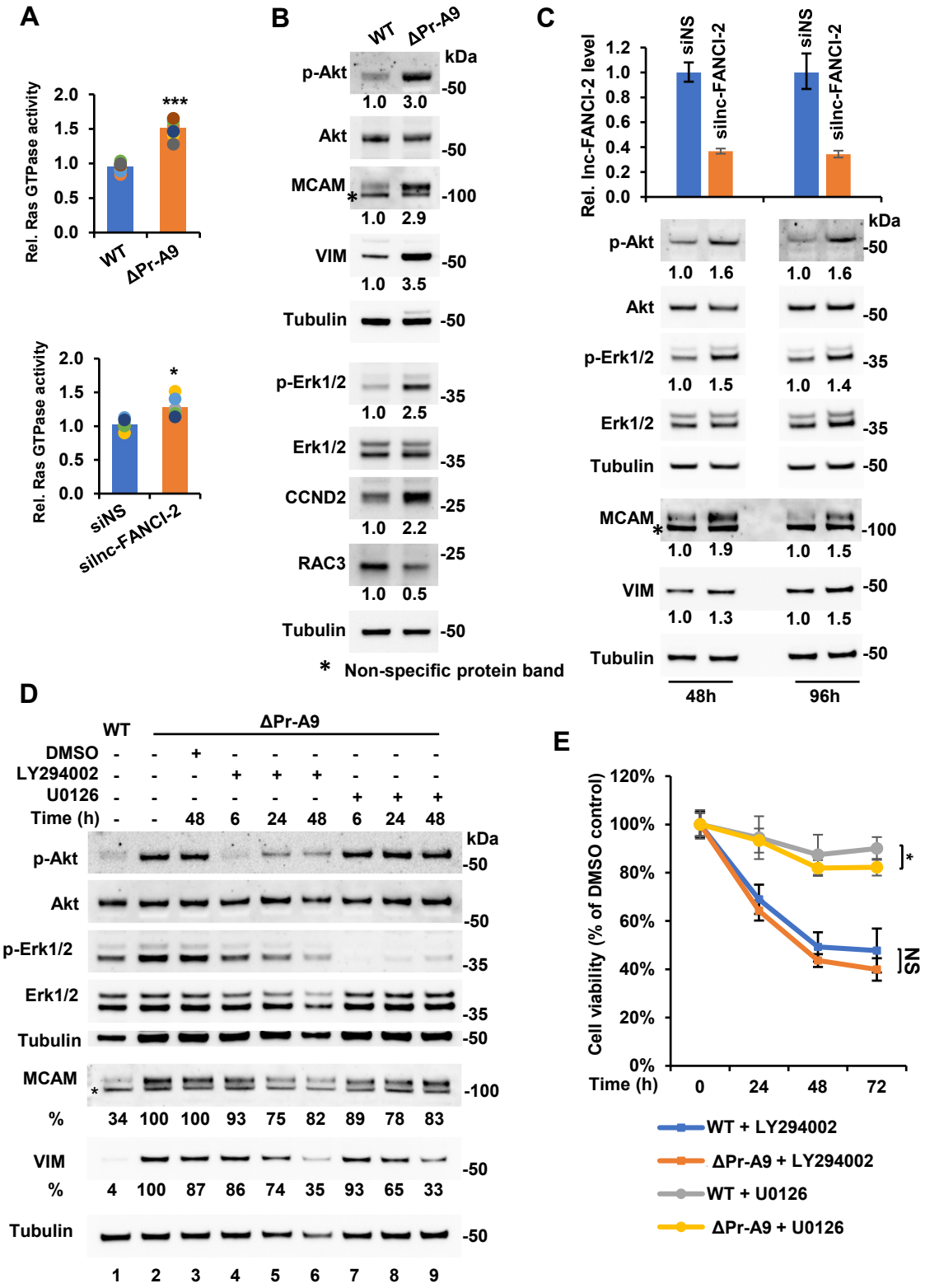

A
Rel. Ras GTPase activity
2.0
1.5
1.0
0.5
0.0
***
WT
ΔPr-A9
Rel. Ras GTPase activity
2.0
1.5
1.0
0.5
0.0
*
siNS
siInc-FANCI-2
B
WT
ΔPr-A9
kDa
p-Akt
-50
1.0 3.0
Akt
-50
MCAM
*
-100
1.0 2.9
VIM
-50
1.0 3.5
Tubulin
-50
p-Erk1/2
-35
1.0 2.5
Erk1/2
-35
CCND2
-25
1.0 2.2
RAC3
-25
1.0 0.5
Tubulin
-50
* Non-specific protein band
C
Rel. lnc-FANCI-2 level
1.2
1.0
0.8
0.6
0.4
0.2
0.0
siNS
siInc-FANCI-2
siNS
siInc-FANCI-2
kDa
p-Akt
-50
1.0 1.6
1.0 1.6
Akt
-50
p-Erk1/2
-35
1.0 1.5
1.0 1.4
Erk1/2
-35
Tubulin
-50
MCAM
*
-100
1.0 1.9
1.0 1.5
VIM
-50
1.0 1.3
1.0 1.5
Tubulin
-50
48h
96h
D
WT
ΔPr-A9
DMSO - - + - - - - - -
LY294002 - - - + + + - - -
U0126 - - - - - - + + +
Time (h) - - 48 6 24 48 6 24 48
kDa
p-Akt
-50
Akt
-50
p-Erk1/2
-35
Erk1/2
-35
Tubulin
-50
MCAM
*
-100
% 34 100 100 93 75 82 89 78 83
VIM
-50
% 4 100 87 86 74 35 93 65 33
Tubulin
-50
1 2 3 4 5 6 7 8 9
E
Cell viability (% of DMSO control)
120%
100%
80%
60%
40%
20%
0%
Time (h)
0
24
48
72
]*
NS
WT + LY294002
ΔPr-A9 + LY294002
WT + U0126
ΔPr-A9 + U0126

Fig. 7

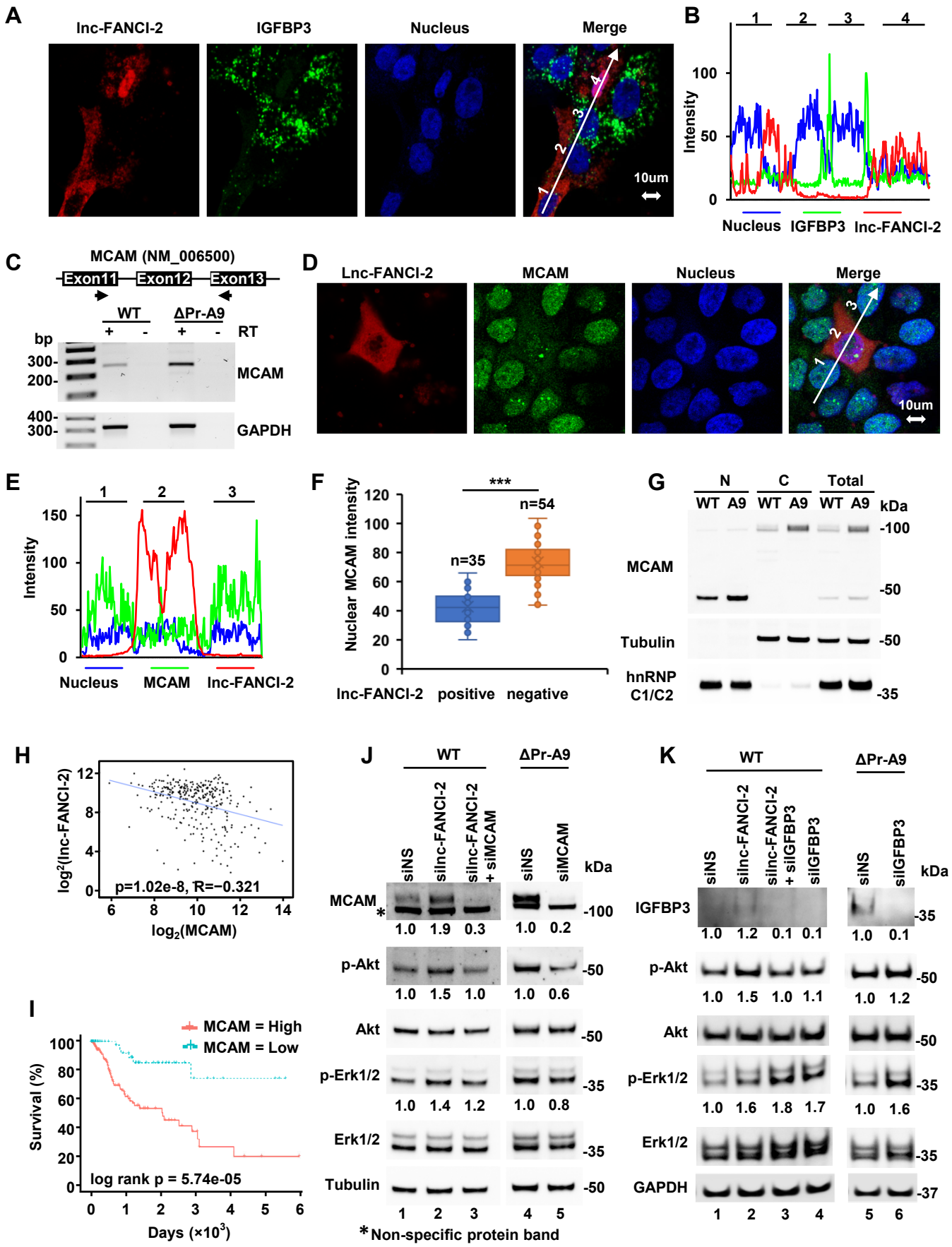
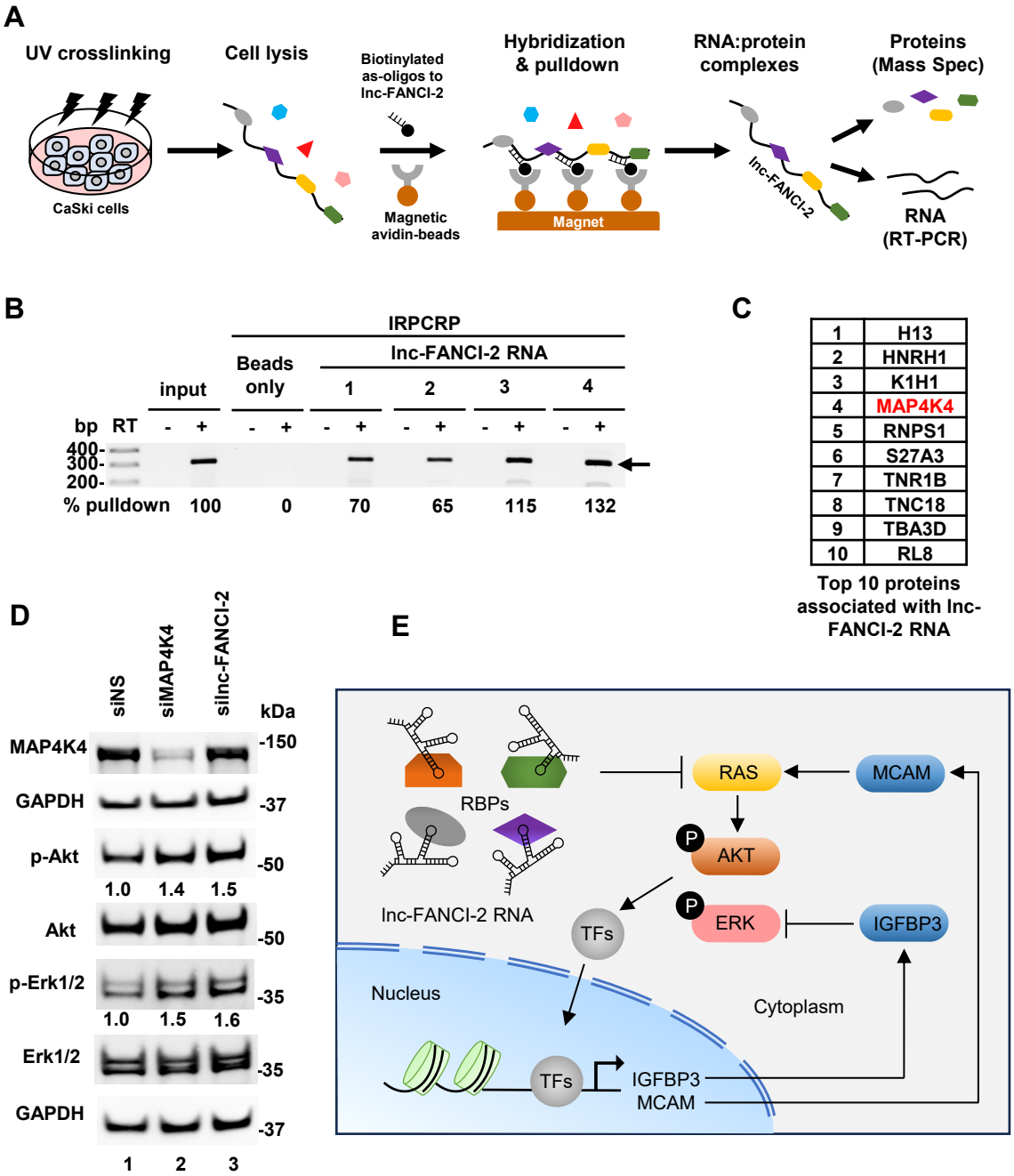

A
lnc-FANCI-2
IGFBP3
Nucleus
Merge
10um
B
Intensity
Nucleus
IGFBP3
lnc-FANCI-2
C
MCAM (NM_006500)
Exon11
Exon12
Exon13
WT
ΔPr-A9
RT
bp
MCAM
GAPDH
D
Lnc-FANCI-2
MCAM
Nucleus
Merge
10um
E
Intensity
Nucleus
MCAM
lnc-FANCI-2
F
Nuclear MCAM intensity
n=35
n=54
lnc-FANCI-2
positive
negative
G
N
C
Total
WT
A9
kDa
MCAM
Tubulin
hnRNP C1/C2
H
log2(lnc-FANCI-2)
log2(MCAM)
p=1.02e-8, R=−0.321
I
MCAM = High
MCAM = Low
Survival (%)
log rank p = 5.74e-05
Days (×10³)
J
WT
ΔPr-A9
siNS
siInc-FANCI-2
siInc-FANCI-2 + siMCAM
siMCAM
kDa
MCAM
p-Akt
Akt
p-Erk1/2
Erk1/2
Tubulin
*Non-specific protein band
K
WT
ΔPr-A9
siNS
siInc-FANCI-2
siInc-FANCI-2 + siIGFBP3
siIGFBP3
kDa
IGFBP3
p-Akt
Akt
p-Erk1/2
Erk1/2
GAPDH

Fig. 8


A
UV crosslinking
CaSki cells
Cell lysis
Biotinylated as-oligos to lnc-FANCI-2
Magnetic avidin-beads
Hybridization & pulldown
Magnet
RNA:protein complexes
lnc-FANCI-2
Proteins (Mass Spec)
RNA (RT-PCR)
B
IRPCRP
Beads only
lnc-FANCI-2 RNA
input
1
2
3
4
bp RT - + - + - + - + - + - +
400-
300-
200-
% pulldown 100 0 70 65 115 132
C
1 H13
2 HNRH1
3 K1H1
4 MAP4K4
5 RNPS1
6 S27A3
7 TNR1B
8 TNC18
9 TBA3D
10 RL8
Top 10 proteins associated with lnc-FANCI-2 RNA
D
siNS
siMAP4K4
siInc-FANCI-2
kDa
MAP4K4 -150
GAPDH -37
p-Akt -50
1.0 1.4 1.5
Akt -50
p-Erk1/2 -35
1.0 1.5 1.6
Erk1/2 -35
GAPDH -37
1 2 3
E
RBPs
lnc-FANCI-2 RNA
RAS
MCAM
P
AKT
P
ERK
IGFBP3
TFs
Nucleus
Cytoplasm
TFs
IGFBP3
MCAM

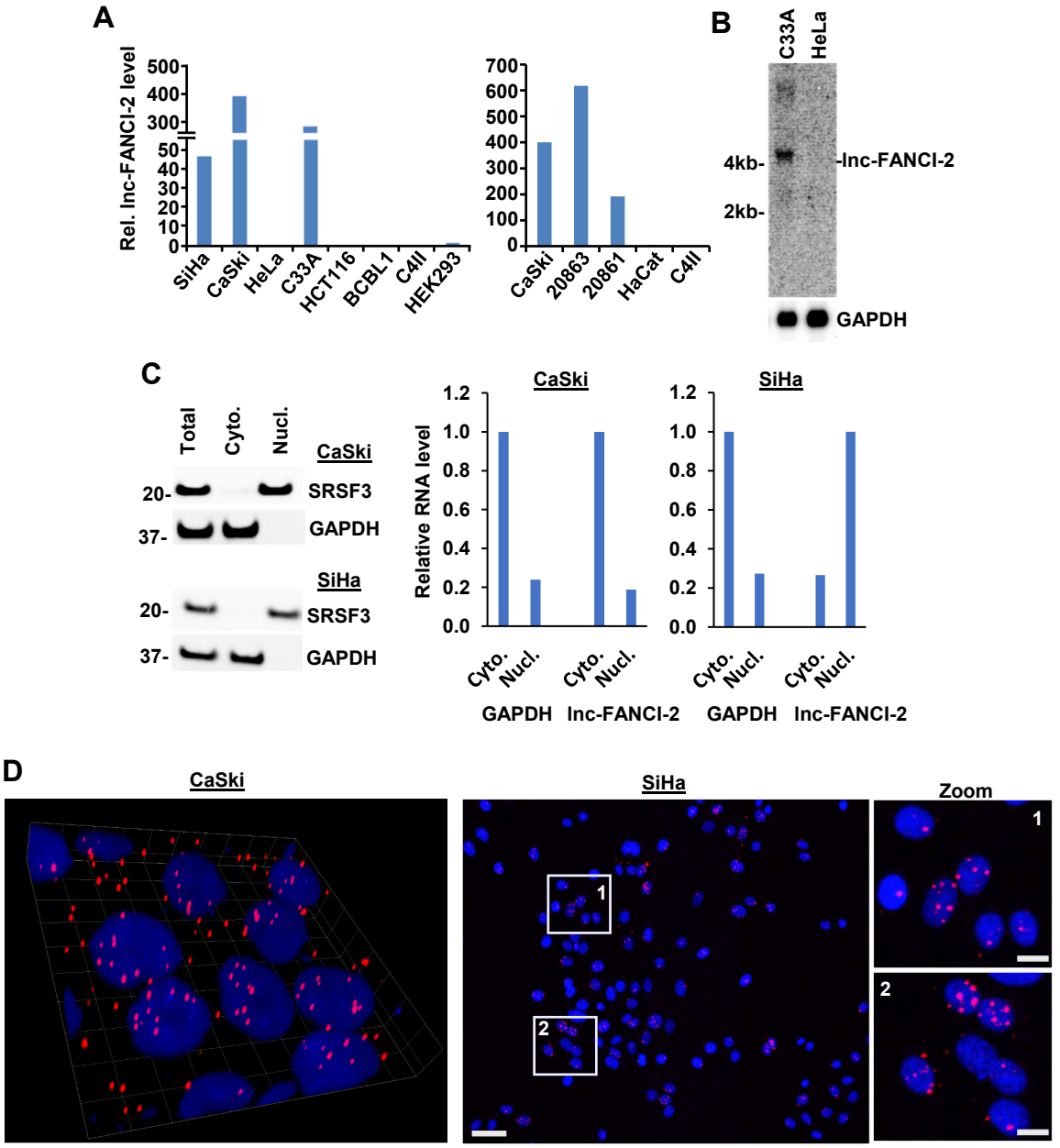


Fig S1. Selective expression of lnc-FANCI-2 RNA in a subset of cervical cancer cell lines. (**A**) RT-qPCR detection of lnc-FANCI-2 expression in HPV-positive cervical cancer cell lines: SiHa (HPV16+), CaSki (HPV16+), HeLa (HPV18+), C4II (HPV18+), and W12 subclones 20861 (HPV16+, integrated) and 20863 (HPV16+, episomal) and HPV-negative cell lines: C33A (cervical cancer cells with mutations of p53 and pRb), HCT116 (colorectal cancer cells), BCBL-1 (body cavity B lymphoma cells), HEK293 (Ad5 E1/E2-immortalized human kidney cells), and HaCaT (spontaneously immortalized human epidermal cells). (**B**) HeLa cells express no lnc-FANCI-2 when compared with C33A cells as determined by Northern blot. (**C**) Subcellular lnc-FANCI-2 by fractionation determines lnc-FANCI-2 being predominantly cytoplasmic in CaSki but nuclear in SiHa cells. Cytoplasmic and nuclear fractionation efficiency was determined by Western blot analyses of nuclear protein SRSF3 (serine- and arginine-rich splicing factor 3) and cytoplasmic GAPDH protein. Total fractionated cytoplasmic and nuclear RNAs from CaSki and SiHa cells were quantified for lnc-FANCI-2 by RT-qPCR, with cytoplasmic GAPDH RNA serving as internal controls for RNA fractionation efficiency. (**D**) Subcellular lnc-FANCI-2 (red) localization in CaSki and SiHa cells was examined by RNAscope RNA ISH analysis. Nuclei were stained with DAPI (blue). Scale bars: 25 μm in the left and 10 μm in the zoom.

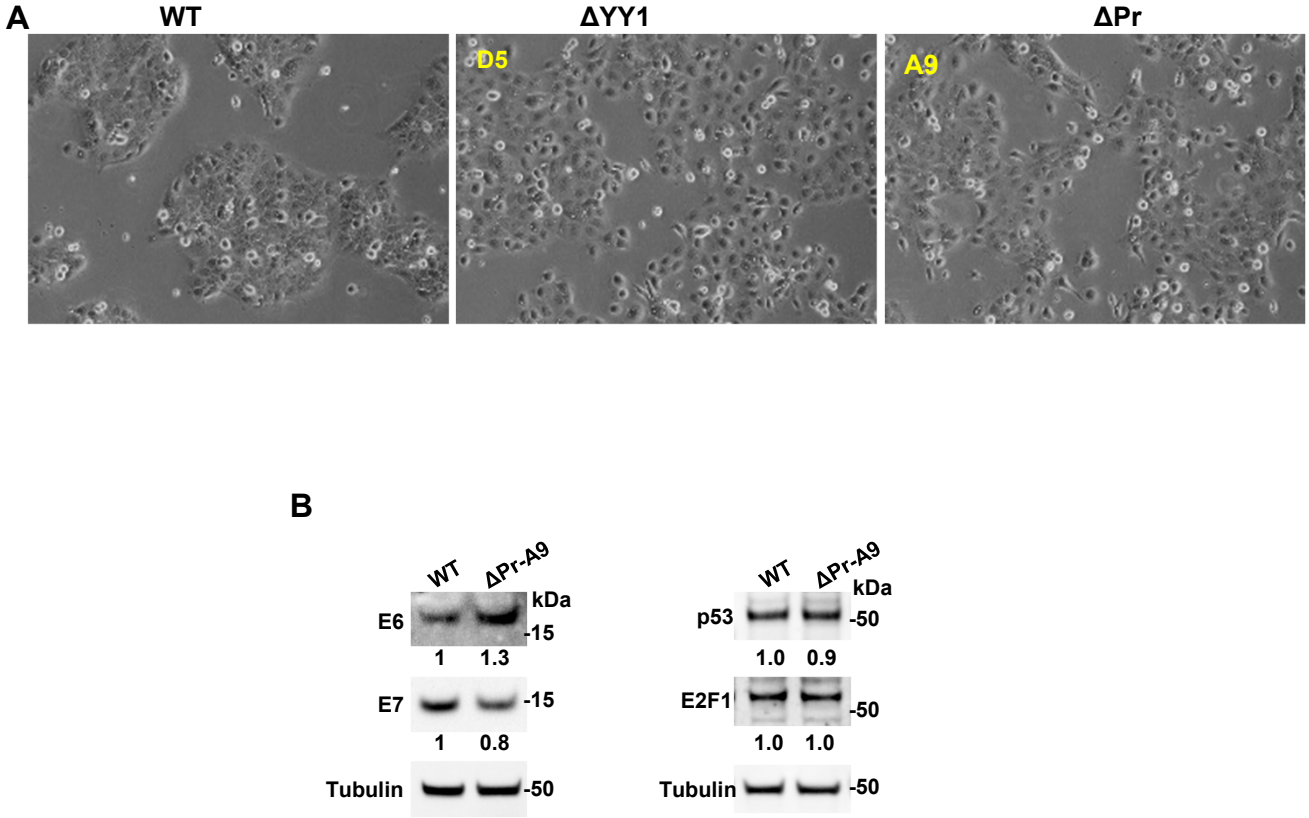


Fig. S2. Characterization of lnc-FANCI-2 KO on CaSki cell growth and viral oncoprotein expression. (A) The characteristic cell growth behavior and morphology of the wild type (WT) and single cell clones with deletion of ΔYY1 motifs (ΔYY1, clone # D5) and ΔPromoter motif (ΔPr, clone # A9) were imaged at 24 h after spreading. (B) Effect of lnc-FANCI-2 KO in CaSki cells on the expression of HPV16 E6 and E7 and their downstream targets. Total cell extracts from parental CaSki cells and ΔPr-A9 cells were examined by immunoblotting with corresponding antibodies. Tubulin served as a protein loading control. The relative protein levels of E6, E7, p53, and E2F1 were calculated after normalizing to tubulin.

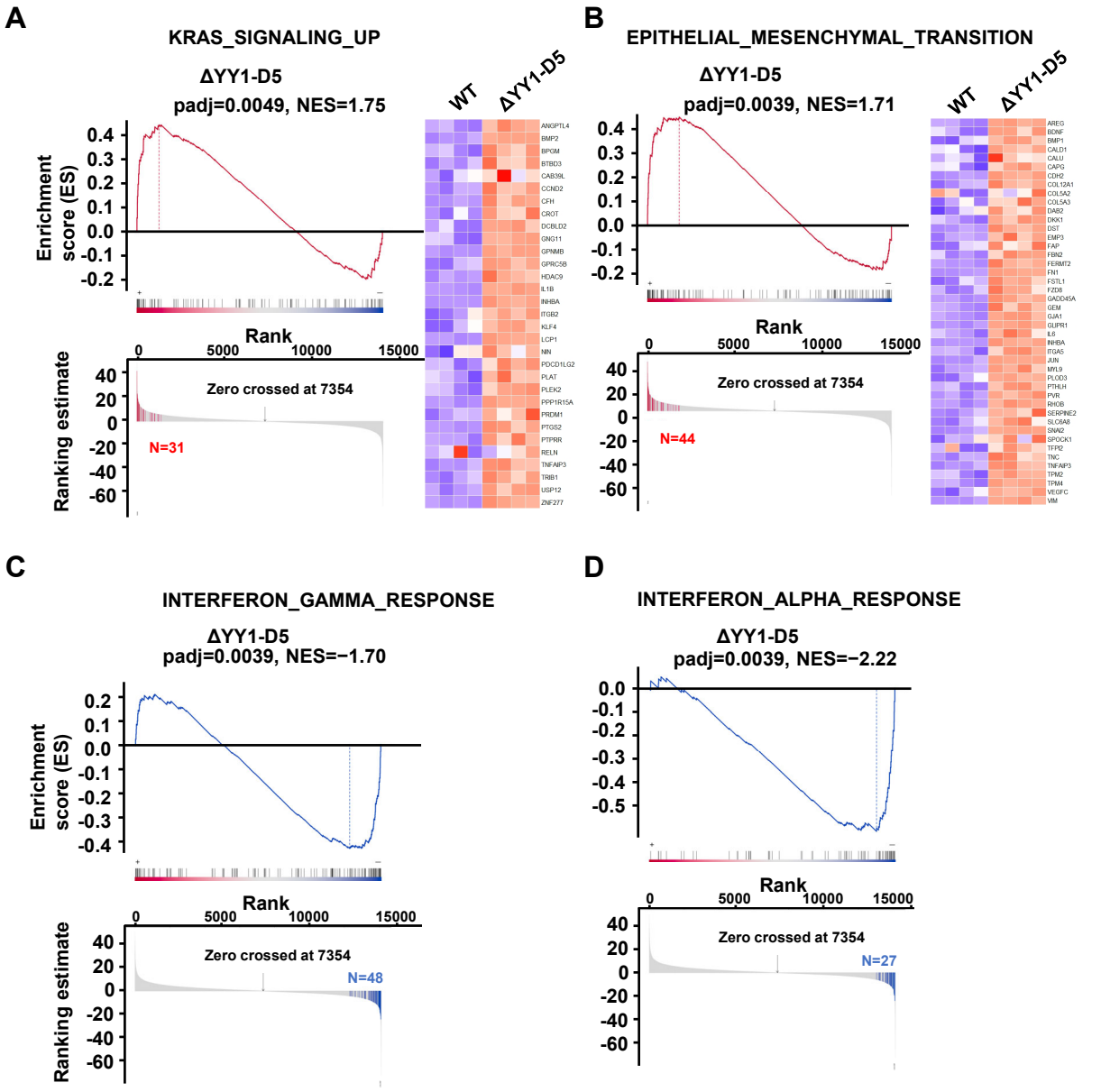


Fig S3. GSEA Enrichment plots show enrichment score and gene hits enriched in ΔYY1-D5 cells using Hallmark gene sets, (A) KRAS_SIGNALING_UP. (B) EPITHELIAL_MESENCHYMAL_TRANSITION. (C) INTERFERON_ALPHA_RESPONSE. (D) INTERFERON_GAMMA_RESPONSE. The heatmap on the right side of Enrichment plot in each panel visualize the genes with differentiated expression enriched in each pathway in ΔYY1-D5 cells.

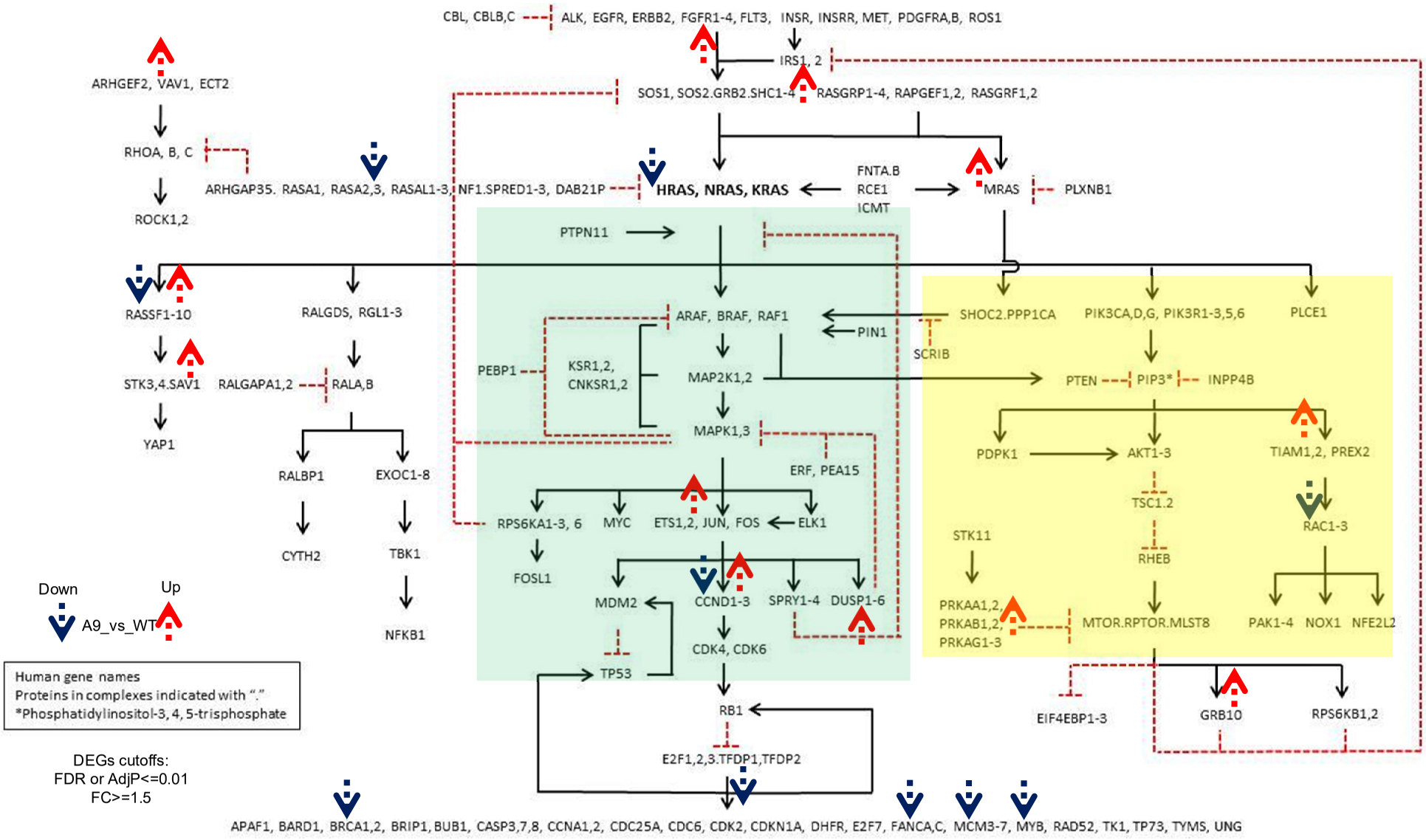

CBL, CBLB,C
ALK, EGFR, ERBB2, FGFR1-4, FLT3, INSR, INSRR, MET, PDGFRA,B, ROS1
IRS1, 2
ARHGEF2, VAV1, ECT2
SOS1, SOS2.GRB2.SHC1-4
RASGRP1-4, RAPGEF1,2, RASGRF1,2
RHOA, B, C
ARHGAP35. RASA1, RASA2,3, RASAL1-3, NF1.SPRED1-3, DAB21P
HRAS, NRAS, KRAS
FNTA.B
RCE1
ICMT
MRAS
PLXNB1
ROCK1,2
PTPN11
RASSF1-10
RALGDS, RGL1-3
ARAF, BRAF, RAF1
PIN1
SHOC2.PPP1CA
PIK3CA,D,G, PIK3R1-3,5,6
PLCE1
SCRIB
PEBP1
KSR1,2,
CNKSR1,2
MAP2K1,2
STK3,4.SAV1
RALGAPA1,2
RALA,B
PTEN
PIP3*
INPP4B
YAP1
MAPK1,3
PDPK1
AKT1-3
TIAM1,2, PREX2
RALBP1
EXOC1-8
ERF, PEA15
TSC1.2
RAC1-3
RPS6KA1-3, 6
MYC
ETS1,2, JUN, FOS
ELK1
STK11
CYTH2
TBK1
RHEB
FOSL1
MDM2
CCND1-3
SPRY1-4
DUSP1-6
PRKAA1,2,
PRKAB1,2,
PRKAG1-3
MTOR.RPTOR.MLST8
PAK1-4
NOX1
NFE2L2
NFKB1
CDK4, CDK6
TP53
RB1
EIF4EBP1-3
GRB10
RPS6KB1,2
E2F1,2,3.TFDP1,TFDP2
APAF1, BARD1, BRCA1,2, BRIP1, BUB1, CASP3,7,8, CCNA1,2, CDC25A, CDC6, CDK2, CDKN1A, DHFR, E2F7, FANCA,C, MCM3-7, MYB, RAD52, TK1, TP73, TYMS, UNG
Down
Up
A9_vs_WT
Human gene names
Proteins in complexes indicated with "."
*Phosphatidylinositol-3, 4, 5-trisphosphate
DEGs cutoffs:
FDR or AdjP<=0.01
FC>=1.5

Fig. S4. Pathway map of RAS Initiative with highlighted differentially expressed genes (DEGs) in ΔPr-A9 cells vs parental wild type cells. DEGs are mainly distributed and clustered in two main branches of RAS pathway: MAPK signaling branch and PI3K/AKT branches. DEGs were derived from RNA-seq data at cutoff of adjusted p-value <=0.01 and fold change>1.5 for both up- and down-regulated genes (indicated by red and blue arrows respectively) between A9 cells vs WT CaSki cells using three common analysis methods (DESeq2, edgeR, and limma-voom) and in-house BRB analysis pipeline from NCI. Highlighted genes in the pathway map are derived as DEGs in at least three of the analysis methods, although in most cases, these DEGs behaved consistently across all 4 methods. MAPK signaling branch is highlighted in transparent green box and PI3K/AKT branch is highlighted in transparent yellow box within the pathway map of RAS signaling pathway, which was collectively collated from community inputs that were organized and stimulated by RAS Initiative, and maintained as a common knowledge basis at URL below: https://www.cancer.gov/research/key-initiatives/ras/ras-central/blog/2015/ras-pathway-v2, which was also described in recent review (Figure 1 in Nissley and McCormick, 2022).

Reference (where the Ras pathway map has been described):

Nissley, D. V. and McCormick, F. (2022). RAS at 40: update from the RAS Initiative. *Cancer Discov.* https://doi.org/10.1158/2159-8290.CD-21-1554

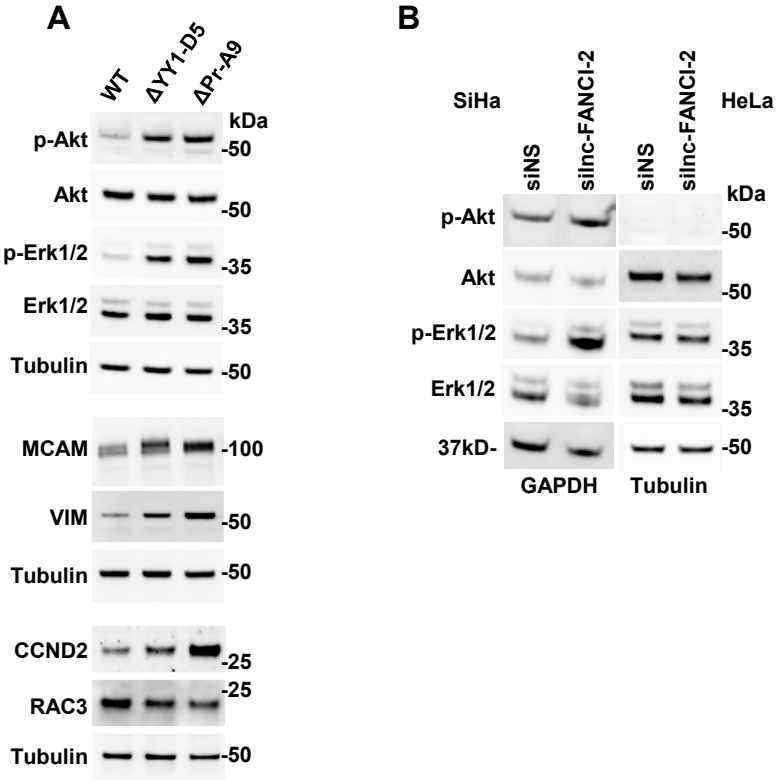


Fig S5. Knockout or Knockdown of lnc-FANCI-2 affects RAS signaling. (A) Selective validation of increased expression of RAS signaling-related downstream genes in ΔPr-A9 and ΔYY1-D5 cells. (B) KD of lnc-FANCI-2 expression in SiHa cells enhanced phosphorylation of Akt and Erk1/2 but did not in HeLa cells expressing no lnc-FANCI-2.

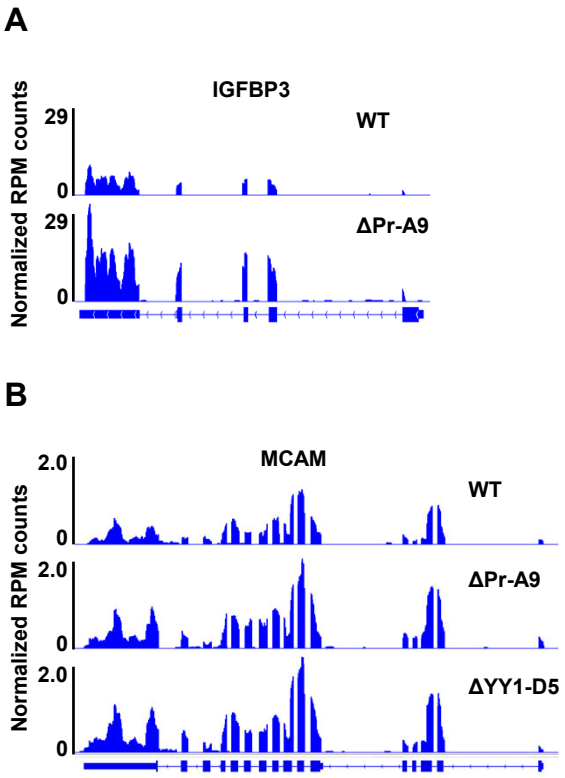


Fig. S6. IGFBP3 and MCAM expression and lnc-FANCI-2. (A-B) RNA-seq reads-coverage of IGFBP3 and MCAM by IGV illustrates the increased expression of IGFBP3 and MCAM in lnc-FANCI-2 KO cells ΔPr-A9 and/or ΔYY1-D5 when compared to the parental WT CaSki cells.

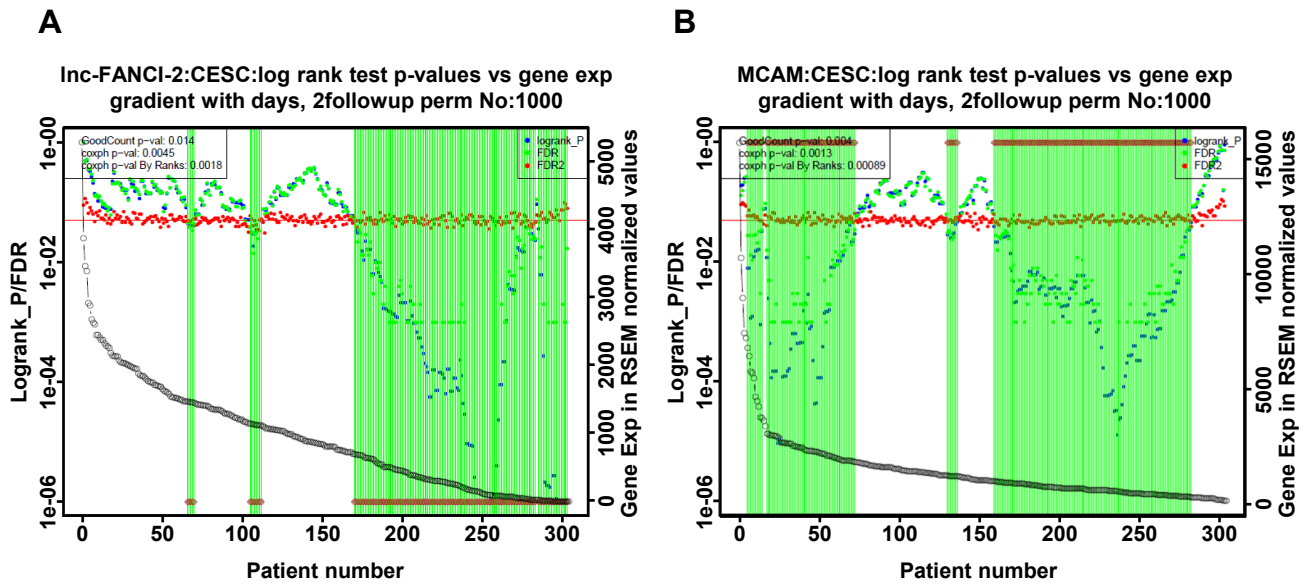


Fig. S7. Expression levels of lnc-FANCI-2 and MCAM are significantly associated with TCGA CESC cancer patients' survival outcome. The expression of lnc-FANCI-2 (A) and of MCAM (B) along with survival of the same group of 304 cervical cancer patients were analyzed by our in-house GradientScanSurv pipeline as described (Yi et al 2018, PLoS One 13(12):e0207590). The **open grey circles** in each plot are ranked expression levels from high (left) to low (right) of lnc-FANCI-2 (A) and MCAM (B) from all patient samples as indicated by y-axis at the right side of the plot (in RSEM normalized value of RNAseq data). At each breaking point denoted by patient number (x-axis), the log rank test (y-axis at the left side of the plot) was performed between the higher expression group at the left side of the cut-point vs the lower expression group at the right side of the cut-point and the indicated log rank p-values (**blue dots**) are shown as at each cut-point at x-axis. The **horizontal red line** is the p=0.05 cutoff line for the log rank test p-values. If a log rank p-value (**blue dot**) at a cut-point is below the cutoff (p<0.05), there will be a **vertical green line** shown at the corresponding cut-point, indicative of a significant log rank test p-value, and also a **brown** diamond plotted either at the bottom part of the **vertical green line** (A) if higher expression of lnc-FANCI-2 led to less severe outcome (dying slower than the lower expressors), or at the top part of the **vertical green line** (B) if higher expression of MCAM led to more severe outcome (dying faster than the lower expressors).

The original GoodCount was defined as the number of significant log rank tests (the default significance level was set as 0.05) across all possible cut-points for the original dataset. We employed the bootstrap approach to permutate the data for 1000 times in this case, and for each time, a similar procedure was performed like the original dataset, and a permutated GoodCount was obtained for each permutated dataset. The GoodCount p-value was derived as the proportion of the times that the GoodCounts of permutated datasets were no less than that of the original dataset (see Yi et al 2018, PLoS One 13(12):e0207590) and shown at the left top corner of each plot, 0.014 for lnc-FANCI-2 in panel A and 0.004 for MCAM in panel B. A significant GoodCount p-value (**blue dots** in both panels A and B) is indicative of significant association of the expression levels with the corresponding survival outcomes of the analyzed patient samples. Derived p-values from univariant expression-based cox regression models (coxph p-value: derived by original gene expression values; or coxph p-value By Rank: derived by corresponding ranks of gene expression values) were shown at the top left corner of each plot. The cut-point-specific FDR (**green dots**) and FDR2 (**red dots**) were shown at the top right corner of each plot. FDR for each cut-point was defined as the portion of cases of permutated datasets that have log rank test p-values no higher than that of original dataset. FDR2 for each cut-point was defined as the portion of cases of permutated datasets that have log rank test p-values no higher than default setting of significance as 0.05.