Cancer cell – fibroblast crosstalk signature promotes macrophage recruitment via HB-EGF and stromal EGFR/MEK signalling

Giovanni Giangreco#€[1] (ORCID: 0000-0002-5464-0199), Antonio Rullan#[1,2,3] (ORCID: 0000-0002-7365-9629), Yutaka Naito[1,4], Dhruva Biswas[5,6,7], Yun-Hsin Liu[6], Steven Hooper[1], Pablo Nenclares[2,3] (ORCID: 0000-0002-1750-3179), Shreerang Bhide[2,3], Maggie Chon U Cheang[2,3], Probir Chakravarty[8], Eishu Hirata[1,9] (ORCID: 0000-0001-7479-7460), Charles Swanton[5,6,10] (ORCID: 0000-0002-4299-3018), Alan Melcher[2,3], Kevin Harrington[2,3], Erik Sahai€[1] (ORCID: 0000-0002-3932-5086)

1. Tumor Cell Biology Laboratory, The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, United Kingdom.
2. Department of Radiotherapy and Imaging. The Institute of Cancer Research, 237 Fulham Road, London, SW3 6JB, United Kingdom.
3. Head and Neck Unit. The Royal Marsden Hospital, 203 Fulham Road, London, SW3 6JJ, United Kingdom.
4. Current affiliation: Department of Bioregulation, Institute for Advanced Medical Sciences, Nippon Medical School, Tokyo, Japan.
5. Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, 72 Huntley Street, London WC1E 6DD, United Kingdom.
6. Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, United Kingdom.
7. Bill Lyons Informatics Centre, University College London Cancer Institute, 72 Huntley Street, London WC1E 6DD, United Kingdom.
8. Bioinformatics Platform, Francis Crick Institute, 1 Midland Road, London NW1 1AT, United Kingdom.
9. Current affiliation: Division of Tumor Cell Biology and Bioimaging, Cancer Research Institute of Kanazawa University, Kanazawa, Japan.
10. Department of Oncology, University College London Hospitals, London, UK

# These authors contributed equally
€ Corresponding authors: giovanni.giangreco@crick.ac.uk ; erik.sahai@crick.ac.uk

Summary

Interactions between cells in the tumor microenvironment (TME) shape cancer progression and patient outcomes. To gain new insights into how the TME influences cancer outcomes, we adopt a distinct approach that exploits patterns of gene expression indicative of signalling between stromal fibroblasts and cancer cells, and demonstrate the prognostic significance of these gene signatures in multiple and independent squamous cell carcinoma cohorts. We discover that HB-EGF in cancer cells and EGFR in cancer associated fibroblasts represent a hub of tumor – stroma crosstalk through MAPK signalling and AP-1 transcription factor. We find divergent consequences for pathway activation in cancer cells compared to stromal fibroblasts that favours macrophages recruitment. Together these analyses demonstrate the utility of this new approach to interrogating the extent and consequences of TME crosstalk. By focusing on the transcriptional consequences of cancer cell-fibroblast interactions we derive prognostic signatures and uncover molecular mechanisms promoting fibroblast to macrophage communication.

Keywords

Fibroblasts; Cancer-associated fibroblasts; CAFs; EGFR; HB-EGF; stroma; Macrophages; Immune Microenvironment; gene signature; RAS

49 ## Abbreviations

50 Activator protein 1 (AP-1)
51 Cancer associated fibroblasts (CAF)
52 Cervical squamous cell carcinoma (CESC)
53 Chemokine (C-C motif) ligand (CCL)
54 Confidence Interval (CI)
55 Conditioned medium (CM)
56 False discovery rate (FDR)
57 Gene-set enrichment analysis (GSEA)
58 Hazard ratio (HR)
59 Head and neck squamous cell carcinoma (HNSCC)
60 Heparin-binding epidermal growth factor-like growth factor (HB-EGF)
61 Human papillomavirus (HPV)
62 Interleukin (IL)
63 Lung squamous cell carcinoma (LUSC)
64 Overall Survival (OS)
65 Mean fluorescent intensity (MFI)
66 Normalized enrichment score (NES)
67 Non-treated (NT)
68 Pancreatic ductal adenocarcinoma (PDAC)
69 Peripheral blood mononuclear cell (PBMC)
70 Single-cell RNA sequencing (scRNAseq)
71 Standard deviation (SD)
72 Short tandem repeats (STR)
73 Squamous cell carcinoma (SCC)
74 The Cancer Genome Atlas (TCGA)
75 Transcription factor (TF)
76 Tumor microenvironment (TME)
77 Tumor necrosis factor (TNF)

Introduction

Cross-talk between cancer cells and non-malignant cells in the tumor microenvironment (TME) influences tumor growth, metastasis and therapy resistance [1]. Cancer-associated fibroblasts (CAFs) promote the invasion of cancer cells, reduce the efficacy of both targeted and cytotoxic therapies and modulate immune cell recruitment and functionality [2]. In addition, CAFs are correlated with a pro-tumorigenic immune landscape, including higher number of tumor-promoting myeloid cells [3], lower numbers of tumor-infiltrating lymphocytes [4] and worse prognosis [5 6]. Of note, CAFs are linked to poor outcomes in squamous cell carcinoma (SCC) arising at multiple anatomical locations, including lungs (LUSC), cervix (CESC), and head and neck (HNSCC) [6–9].

Indeed, the development of advanced sequencing techniques allows multiple inferences about the type and abundance of different TME components, including CAFS, both from bulk transcriptome and genomic methylation data [7,10–13] . Both methods rely on the identification of cell type specific genes and the application of deconvolution strategies ultimately to infer the abundance of a particular population in a bulk dataset. However, these methods struggle to identify the functionally relevant interactions between cell types, such as signalling events [14,15] and the biological mechanisms associated with cell type crosstalk to be linked to patient outcomes remain incompletely understood.

Here, we propose an alternative approach to identify key players involved in tumor-stroma interaction. Instead of focusing on the abundance of CAFs or specific CAF subpopulations, we identify a signature indicative of signalling between cancer cells and CAFs. This signature is statistically associated with worse overall survival in multiple SCC subtypes. Moreover, we leverage information within the signature to identify a novel mechanism of interaction between cancer cells and CAFs. In co-culture, the RAS / MAPK pathway is strongly activated in both cell types, converging on the upregulation of Activator Protein 1 (AP-1) transcription factor (TF) components. We identify heparin-binding epidermal growth factor-like growth factor (HB-EGF) as a key mediator of cancer cell – CAF cross-talk, primarily expressed by cancer cells and able to upregulate the expression cytokines in CAFs. In turn, we demonstrate that this upregulation can drive attraction of macrophages, ultimately linked to worse overall survival in SCC patients.

107 ## Results

108 ### Meta-analysis of transcriptomic data of cancer cell and cancer-associated fibroblast co-
109 cultures identifies gene signatures with prognostic value

110 To identify functionally and clinically relevant gene signatures based on cancer cell – CAF cross-talk,
111 we performed a meta-analysis of transcriptomic datasets that compare co-cultures and mono-cultures
112 of cancer cells and CAFs. The datasets were generated under similar direct co-culture conditions and
113 employed cells derived from different models and cancer types [16,17]. We applied two strategies to
114 derive gene signatures indicative of upregulated cancer cell – CAF signalling: i) selection of the most
115 significantly enriched pathways via gene set enrichment analysis (GSEA) in co-culture for each
116 transcriptomic dataset, followed by the selection of the up-regulated genes most frequently present
117 in each enriched pathway (Figure 1A); ii) selection of the most up-regulated genes in co-culture for
118 each transcriptomic dataset (Figure S1A). Using the first approach, we obtained a list of 5 genes
119 upregulated in cancer cells and of 4 genes upregulated in CAFs upon co-culture, with one present in
120 both. Therefore, this gene signature comprised of 8 genes (named CoCu8) (Figure 1B). The second
121 approach led to a list of 2 genes upregulated in cancer cells and of 29 genes upregulated in CAFs upon
122 co-culture, with one gene in common. Therefore, this gene signature consisted of 30 genes (named
123 CoCu30) (Figure S1B).
124
125 We tested CoCu8 and CoCu30 on publicly available dataset of breast cancer – fibroblast co-cultures[18]
126 confirming their relevance (Figure S2A). We also tested whether CoCu8 and CoCu30 are also
127 upregulated when cancer cells are co-cultured with other stromal cell types and for this reason, we
128 analysed a dataset of co-culture between 1205Lu cancer cells and HUVEC endothelial cells[19]: CoCu8 is
129 neither enriched in cancer cells nor in endothelial cells when co-cultured, while CoCu30 shows only a
130 weak correlation with co-culture conditions both in 1205Lu and HUVEC cells (Figure S2B). Thus, we
131 establish new gene signatures specifically indicative of cancer cell - CAF communication.
132
133 We next sought to determine the clinical relevance of these signatures by testing in The Cancer
134 Genome Atlas (TCGA). We found that high expression (Q4 vs Q1) of both gene signatures correlated
135 with worse overall survival (OS) in multiple SCC datasets: HNSCC (CoCu8 Hazard Ratio (HR): 1.95,
136 CoCu30 HR: 1.57), cervical squamous cell carcinoma (CESC, CoCu8 HR:2.79, CoCu30 HR: 2.08) and lung
137 squamous cell carcinoma (LUSC, CoCu8 HR: 1.85, CoCu30 HR: 1.78) (Figure 1C; Figure S1C), with CoCu8
138 consistently showing a slightly higher hazard ratio (HR) compared to CoCu30 in all three tumor types.
139 A multivariate analysis including relevant clinical variables such as age, sex and clinical stage to

140 evaluate the co-culture signatures effect as a continuous variable, confirmed the relevance of the
141 signatures (Figure S3 for CoCu8, Figure S4 for CoCu30,). Of note, given the clinical and biological
142 differences between HPV positive and negative tumors, we stratified HNSCC patients according to HPV
143 status, observing that the strongest prognostic effect was visible in HPV positive samples both for
144 CoCu8 (Figure 1D) and CoCu30 (HR: 5.47) (Figure S1D).
145
146 We validated the association between CoCu8 / CoCu30 and outcome in a second independent cohort
147 of LUSC from the TRACERx study [20] (Figure 1E, Figure S1E). The multi-regional biopsies performed in
148 the study enabled us to ask if the expression of the CoCu8 signature was uniform across tumors. Of
149 the 117 tumors analysed, 86 (74%) showed concordant expression of CoCu8 in all regions, which is
150 significantly greater than would be expected based on chance (Figure 1E). Similar results were
151 observed with CoCu30 (Figure S1E). This indicates that cancer cell-fibroblast crosstalk is typically
152 occurring across the whole tumor. Crucially, this analysis showed that the concordant up-regulation
153 of CoCu8 / CoCu30 across tumour regions is associated with worse prognosis.
154

155 ## The crosstalk gene signature has greater prognostic power than fibroblast abundance

156 Given that CoCu8 / CoCu30 reflects cancer cell-fibroblast crosstalk, the signature might be predicted
157 to correlate with fibroblast abundance. We performed a correlation analysis of CoCu8 / CoCu30 and
158 methyl CIBERSORT signatures in TCGA datasets, showing that there is a statistically significant
159 correlation between CAFs presence and CoCu8 signature (Figure S5A-B). Similar results were observed
160 with the CoCu30 signature (Figure S5C-D). Then, we sought to validate these results in the TRACERx
161 LUSC cohort and in a second independent UK_HPV positive cohort (Figure S5E-J). As methylome data
162 was not available for these cohorts, we used fibroblast subtype gene signatures defined in a pan-
163 cancer analysis by Galbo et al.[7]. We observed strong positive correlations between CoCu8 / CoCu30
164 signature and all of the fibroblast subtypes defined both in LUSC and UK_HPV positive HNSCC (Figure
165 S5E-J).
166
167 We speculated that our signatures of active cancer cell-fibroblast crosstalk might have better
168 prognostic power than simply CAF abundance. To test this, we analysed the abundance of CAFs using
169 the methyl CIBERSORT deconvolution strategy in TCGA cohorts and probed links with overall survival[10].
170 This analysis indicated worse OS for HNSCC patients with higher CAF presence, but no significant
171 differences were observed in CESC or LUSC (Figure S6A-B). Of note, both CoCu8 and CoCu30 signatures
172 out-performed the methyl CIBERSORT method for HNSCC, CESC, and LUSC, which adds credence to
173 our method of deriving a signature based on the interaction with fibroblasts, not simply their

174 abundance. Overall, these data indicate that CoCu8 / CoCu30 signatures correlates with CAF
175 abundance but have greater prognostic value than gene signatures used to infer CAF abundance.
176

177 Pathway enrichment analysis of cancer-associated fibroblast and cancer cell co-culture 178 reveals a consistent upregulation of AP-1 transcription factor genes

179 To obtain insight in the molecular basis of the cancer cell-CAF interactions, we analysed the pathways
180 that were enriched in all the transcriptomic datasets used to generate CoCu8 and CoCu30. This
181 revealed upregulation of multiple pathways linked to immune regulation, stress response, and
182 signalling (Figure 2A). Multiple genes belonging to the AP-1 transcription factor complex: *JUNB*, *FOS*,
183 *FOSB* were strongly enriched (Figure 2B). Moreover, *PLAUR* is regulated by AP-1 factors. As our meta-
184 analysis showed the strongest impact in the stratification of OS patients from HPV positive HNSCC, we
185 decided to explore *JUNB*, *FOS*, *FOSB* expression levels in 4 different co-culture combinations of human
186 HPV positive HNSCC cell lines, SCC154 and SCC47, and human oral CAFs, OCAF1 and OCAF2. Our results
187 confirmed that these three AP-1 TFs are upregulated when cancer cells and CAFs are in direct co-
188 culture, as compared to mono-culture (pooled RNA from both cell lines) (Figure 2C). Analysis of
189 indirect co-cultures [17] separated by a 0.4µm filter indicated that CoCu8 / CoCu30 are strongly enriched
190 with direct co-culture also when compared with indirect co-culture, implying that direct contact is
191 required for increased AP-1 TF expression (Figure S7A).
192

193 Interaction between cancer cells and cancer-associated fibroblasts is linked to increased RAS 194 activity

195 We next investigated possible mechanisms underlying the upregulation of the multifunctional AP-1
196 TFs both in cancer cells and CAFs[21]. RAS signalling via MAPK is known to be a major driver of AP-1 gene
197 expression[22,23]. Accordingly, we found that RAS signalling was strongly up-regulated upon direct co-
198 culture, as indicated by the enrichment of the KRAS_SIGNALLING_UP signature (Figure 2A) and of the
199 curated RAS84 gene signature [24], in all our cancer cell-CAF datasets (Figure 3A, Figure S7B).
200

201 To interrogate further the linkage between CAFs, the CoCu8 signature, and RAS signalling in HPV
202 positive patients from the TCGA cohort, we stratified them according to RAS activity. We split the
203 patient data into three groups (RAS84_0, RAS84_1 and RAS84max) according to the levels of RAS
204 activity as performed by East et al.[24]. This analysis shows that higher RAS activity (group RAS84_max)
205 correlates with worse OS, in a similar fashion to the effect observed with CoCu8 stratification (Figure

206 3B). Indeed, we observed a strong, positive correlation between RAS84 activity and CoCu8 expression
207 in both HPV positive HNSCC cohorts (TCGA - R = 0.79, Figure 3C and UK_HPV positive cohort - R = 0.84,
208 Figure S8A). We also observed a statistically significant enrichment in the extent of fibroblasts present
209 when RAS activity was higher in both cohorts (Figure 3D, Figure S8B).
210
211 Activation of the RAS pathway upon cancer cell-CAF co-culture was experimentally validated by co-
212 culturing SCC154 – OCAF1, which resulted in a strong increase of phosphorylated ERK1/2 (Figure 3E).
213 To test whether the RAS-MAPK pathway is responsible for the upregulation of *JUNB*, *FOS* and *FOSB*
214 genes, we used the MEK inhibitor, trametinib, in the SCC154-OCAF1 co-culture. We confirmed that
215 MEK inhibition downregulates ERK1/2 activation upon co-culture (Figure 3F) and observed a
216 significant downregulation of *JUNB*, *FOS* and *FOSB* genes (Figure 3G). Thus, multiple genes belonging
217 to the AP-1 TF complex are upregulated when cancer cells and CAFs are co-cultured and this is
218 mechanistically linked to the activation of RAS-MAPK kinase signalling.
219

220 ## HB-EGF activation is crucial to trigger RAS pathway signalling.

221 To explain why AP-1 TFs get upregulated in co-culture, we looked for possible activators of RAS-MAPK
222 signalling. We noted that *HB-EGF* was among the genes upregulated in all 6 transcriptomic datasets
223 together with *JUNB*, *FOS* and *FOSB* (Figure 2B). HB-EGF is an EGFR ligand and therefore can activate
224 MAPK pathway [25,26]. We evaluated the expression levels of all seven EGFR ligands. Importantly, *HB-*
225 *EGF* showed a strong and specific activation upon cancer cell - CAF co-culture (Figure 4A). Moreover,
226 *HB-EGF* expression strongly correlates with CoCu8 in HPV positive HNSCC patients' data from both
227 TCGA (R=0.6, p-value=2.2e-16) and UK_HPV positive (R=0.58, p-value=1e-8) datasets (Figure 4B).
228
229 HB-EGF can activate EGFR/MAPK as an un-cleaved pro-molecule at the plasma membrane [26,27] and, as
230 such, signalling by membrane-bound HB-EGF could explain the need of direct cell contact to trigger
231 the pathway. As membrane-bound HB-EGF should be expressed at about 20-25 kDa, we evaluated the
232 cellular levels of HB-EGF in OCAF1-SCC154 co-culture and found that it was strongly upregulated at
233 the protein level at a molecular weight previously reported as un-cleaved protein [28] (Figure 4C). We
234 also observed EGFR phosphorylation was increased upon direct co-culture (Figure 4C); importantly,
235 both these effects were abrogated by MEK inhibition (Figure 4D), suggesting the presence of a positive
236 feedback loop involving HB-EGF / EGFR / MAPK / AP-1 upon direct co-culture. Furthermore, the EGFR
237 inhibitor, afatinib, blocked AP-1 activation upon co-culture of SCC154-OCAF1 (Figure 4E).
238

239 These results suggest that HB-EGF might be the link that activates EGFR upon cancer cell – CAF co-
240 culture. Therefore, we evaluated the basal expression level of EGFR and HB-EGF in SCC154 and OCAF1
241 mono-cultures. Interestingly, SCC154 expressed HB-EGF at much higher levels than OCAF1, while both
242 cell types expressed similar levels of EGFR (Figure 4F). This suggests that both cell types can be reactive
243 to EGF ligands, but the activation of the positive feedback loop upon direct contact requires higher
244 levels of HB-EGF, expressed at the membrane of cancer cells. In that case, both cancer cells and CAFs
245 should be responsive to HB-EGF treatment, albeit with potentially different downstream effects. To
246 test this hypothesis, we incubated OCAF1 and SCC154 in mono-cultures with different concentrations
247 of HB-EGF. Firstly, HB-EGF caused an increase in proliferation in both cell types, as shown by a higher
248 proportion of EdU positive nuclei (Figure 4G). Moreover, SCC154 – but not OCAF1 – showed a scatter-
249 like phenotype when treated with high doses of HB-EGF, as visible from E-Cadherin staining (Figure
250 S8C). We then tested the effect of HB-EGF treatment on known CAF markers: after treatment of OCAF1
251 with HB-EGF for 48h we noticed a slight but significant downregulation of *ACTA2*, a CAF and
252 myofibroblast marker (Figure S8D). However, no effect was observed for other widely used CAF
253 markers (*FAP*, *LRRC15*, *FN1*) (Figure S8D).
254
255 Given these data, we asked whether HB-EGF expression in cancer cells is enough to induce
256 upregulation of AP-1 genes when cancer cells and CAFs are in co-culture: we therefore performed
257 knock down of HB-EGF in SCC154 (Figure S8E) and then co-cultured them with OCAF1. Importantly,
258 the downregulation of HB-EGF in SCC154 is enough to block the upregulation of *JUNB*, *FOS* and *FOSB*
259 when cancer cells and CAFs are co-cultured (Figure 4H).
260
261 These results point to an axis involving HB-EGF in cancer cells and EGFR in CAFs that activates MAPK /
262 AP-1 inducing a positive feedback loop when cancer cells and CAFs are co-cultured.
263

264 Stromal EGFR regulates cytokine expression and macrophage recruitment in tumor/stroma 265 interaction

266 To focus on the downstream effects of this crosstalk between cancer cells and CAFs and how HB-EGF
267 could affect CAFs functions and lead to unfavourable biology, we analysed scRNAseq dataset of HNSCC
268 with both malignant and non-malignant samples published by Choi et al. [29]. Interestingly, by using
269 myofibroblast and inflammatory markers (Figure S8F), we found that EGFR is mainly expressed by
270 inflammatory fibroblasts (iFibroblasts), but not myofibroblastic CAFs (myoFibroblasts), while HB-EGF
271 is mainly expressed by endothelial and epithelial cells (Figure 5A). This is in line with our preliminary

272 evidence that HB-EGF treatment in OCAF1 is likely more associated with inflammatory than
273 myofibroblast signature.
274
275 Given that high expression of EGFR is associated with inflammatory fibroblasts, we performed a
276 cytokine array of conditioned medium from pooled mono-cultures and co-cultures (Figure 5B):
277 interestingly, we found a strong upregulation of macrophage attraction and differentiation markers
278 (GM-CSF, CCL5, LIF). By cross-referencing the upregulated cytokines with genes regulated by RAS in
279 the RAS84 gene signature, we observed that LIF is both upregulated by co-culturing cancer cells and
280 CAFs and it is linked to RAS activity. We validated by qPCR that *LIF* is transcriptionally upregulated in
281 co-culture (Figure 5C) and that trametinib, MEK inhibitor, treatment is enough to downregulate its
282 expression (Figure 5D). Moreover, by reducing HB-EGF expression in cancer cells and then co-culturing
283 the cells with CAFs, we also observed strong downregulation of *LIF* expression (Figure 5E). Importantly,
284 HB-EGF treatment in cancer cells and CAFs mono-cultures shows that LIF is strongly upregulated only
285 by CAFs, indicating that these are the cells responsible for its production when in co-culture (Figure
286 5F). In line with these data, transcriptomic data of HPV positive HNSCC patients from TCGA show that
287 there is a strong positive correlation between *LIF* and *HBEGF* mRNA expression (Figure S9A).
288
289 To understand if cytokines produced upon cancer cell - CAF direct contact were sufficient to increase
290 macrophage attraction, we isolated primary monocytes from peripheral blood mononuclear cells
291 (PBMCs) from healthy donors, differentiated them into macrophages and then performed a migration
292 assay of macrophages using conditioned medium (CM). Importantly, CM derived from the co-culture
293 of OCAF1 and SCC154 increased the numbers of migrating macrophages, compared with pooled CM
294 derived from each cell in monoculture (Figure 5G). We then tested whether MEK inhibition was
295 enough to decrease the number of macrophages attracted. Indeed, when cancer cells and CAFs are
296 co-cultured in the presence of trametinib, there is a stark decrease in the number of migrating
297 macrophages attracted by the CM and this was not the case when MEK inhibitor is freshly added to
298 CM after it is harvested from the cancer cell-CAF co-culture (Figure 5H).
299
300 These data suggest that cancer cell – CAF crosstalk is consistently linked to macrophage recruitment,
301 therefore we asked whether CoCu8 high patients showed higher levels of macrophages. Of note, when
302 TCGA patients' data are separated according to CoCu8 expression, we observe a strong enrichment
303 for cells defined by a CD14-related methylation signature (monocytes/macrophages) (Figure 5I). We
304 found a similar pattern when patients were separated by fibroblast abundance (Figure S9B) and by
305 RAS activity (Figure S9C). Importantly, we also observed enrichment for monocyte/macrophage

306 lineages in our second cohort of 84 HPV positive HNSCC patients when separated for CoCu8 expression
307 levels (Figure 5I).
308
309 In summary, we have demonstrated that cancer cell – CAF cross-talk increases expression of different
310 cytokines that, in turn, recruit higher numbers of macrophages. This loop is established by HB-EGF
311 expression in cancer cells that induces a cross-talk with CAFs via stromal EGFR dependent by RAS /
312 MAPK activity. Activation of this pathway in CAFs is needed to increase the expression of LIF by the
313 stromal cells. MEK inhibitor, trametinib, is sufficient to reduce the macrophage attraction.

314 ## Discussion

315 The presence of CAFs in tumors correlates with worse patient survival and an immune suppressive
316 TME in multiple tumor types [17,30–34], with recent studies attempting to link different CAF
317 subpopulations to prognosis [35]. However, analysis based on the presence or absence of CAFs does not
318 account for variability in the extent of crosstalk between cancer cells and CAFs. The approach we
319 develop here is based on the selection of genes that are commonly upregulated in both cancer cells
320 and CAFs upon direct cell-to-cell contact, thus focusing on the functional relevance of the cancer cell-
321 CAF interaction, rather than just on the presence of CAFs in the tumor. We applied two different
322 strategies to select genes indicative of cancer cell-CAF interactions. The approach to define CoCu30
323 enriches for genes that are strongly up-regulated, which has been employed previously [7]. To define
324 the CoCu8 signature, we used a new approach based on the selection of a coherent set of genes linked
325 by function. This method produces a list of genes whose changes in expression levels may not be as
326 dramatic, but that have increased functional relevance. Importantly, CoCu8 and CoCu30 stratification
327 gave rise to similar results, strengthening the observation that cancer cell – CAF crosstalk is associated
328 with worse overall survival.
329
330 Comparative analysis of CoCu8 and CoCu30 with annotated gene sets (KRAS SIGNALLING UP from
331 Hallmark signatures and RAS84 signature [24]) suggested a mechanism of cross-talk between cancer cells
332 and CAFs based on the activation of MAPK / AP-1 pathway (Figure S10) that we confirmed through
333 experimentation. These data are consistent with previous literature showing that KRAS is associated
334 with higher stromal presence [36] and with higher cancer cell – stromal interaction [37]. Our analyses
335 indicate that HB-EGF is central to the activation of MAPK signalling upon cancer cell – CAF contact. HB-
336 EGF is the only EGF ligand to be consistently upregulated in co-culture across diverse models.
337 Accordingly, EGFR activation is upregulated in co-culture, suggesting the presence of a positive
338 feedback loop, including HB-EGF / EGFR / MAPK / AP-1. In the future, it will be interesting to explore
339 if proteolytic processing of HB-EGF at the interface between cancer cell and CAFs also plays a role in
340 activation of this feedback loop.
341
342 Treatment with HB-EGF allowed us to decipher the cell type-dependent consequences of the
343 activation of the pathway. Both cancer cells and CAFs respond to HB-EGF by activating MAPK and
344 inducing changes in AP-1 TF expression. However, cancer cells exhibited increased proliferation and
345 migration, whereas CAFs showed increased proliferation and a pronounced up-regulation of the
346 MAPK-dependent cytokine LIF. These data are consistent with the finding that AP-1 activation leads
347 to diverse molecular and phenotypic consequences depending on the cell type studied [38]. Ultimately,

348 we link increased MAPK activity to the chemo-attraction of macrophages. These data provide a
349 molecular explanation specifically for the correlation of CAFs and macrophages in tumors and, more
350 generally, for links between CAF and a pro-tumorigenic and immune-suppressive milieu [31]. Notably,
351 LIF can promote macrophage recruitment and induce a more pro-tumorigenic polarization to alter
352 immune response during anti PD-1 therapy [39].
353
354 CoCu8 and CoCu30 signatures are associated with worse overall survival in multiple SCC datasets from
355 independent cohorts, suggesting that the mechanism we have identified is broadly applicable. Of note,
356 the most striking results were observed in HPV positive HNSCC. Although HPV positive HNSCC patients
357 typically have a better prognosis than HPV negative HNSCC patients, about 25% of these patients still
358 have poor overall survival [40,41].
359
360 We hypothesize that CoCu8 / CoCu30 are particularly effective in HPV positive HNSCC compared to
361 HPV negative because HPV negative has more frequent mutation and amplification in EGF / RAS /
362 MAPK pathway [42], hence its worse prognosis and the smaller contribution of CAFs to poor prognosis
363 (CoCu8 HR is 1.28 in HPV negative HNSCC and 1.85 LUSC compared to 3.19 in HPV positive HNSCC and
364 2.79 in CESC). In HPV positive disease, the same signalling pathway is driven as the result of paracrine
365 interactions with CAFs, which occur to varying degrees in different tumors. These findings have clinical
366 implications for patient stratification and treatment. Our data suggest that the CoCu8 / CoCu30
367 signatures and CAF abundance could help stratify those patients with worse prognosis within the HPV
368 positive SCC. This improved patient stratification would be especially relevant in the context of the
369 recent unsatisfactory efforts to de-escalate and de-intensify treatment for patients with HPV positive
370 tumors [43] and could help reduce toxicity without compromising outcomes. Moreover, our results
371 suggest that this subset of patients could benefit from a targeted approach, for example re-purposing
372 the use of MEK or EGFR inhibitors. Indeed, trametinib – MEK inhibitor – is already used in the
373 treatment of melanoma [44] and non-small-cell lung cancer [45] and it has been tested in phase I / II oral
374 cavity SCC patients, showing some reduction in RAS / MAPK activity as neoadjuvant treatment [46]. Our
375 data support the hypothesis that trametinib or EGFR inhibitors may be beneficial for HPV positive
376 HNSCC patients with high tumor / stroma content.
377
378 In conclusion, our results demonstrate a new approach to detect biologically meaningful stromal
379 signatures. We show that signatures based on signalling in the TME have the potential to both improve
380 patient stratification and to identify new mechanisms of cross-talk between cancer cells and CAFs.
381

382 # Methods

383 ## Cell lines and reagents

384 OCAF1 and OCAF2 human fibroblasts were isolated from patient tissues of oral cancer and
385 immortalized with lentiviral HTERT as described in [47]. These patient samples were collected under the
386 ethical approval REC reference 06/Q0403/125).
387 CRUK0764 were derived from patients with lung adenocarcinoma. These fibroblasts were established
388 from the tumor tissue. The primary CRUK0764 was immortalised by the following infection with
389 retroviruses expressing human telomerase reverse transcriptase.
390 PC9 was obtained from the Crick Institute Central Cell Services facility. PC9 were stably transfected
391 with Lipofectamine 2000 Reagent (Thermo Fisher Scientific) according to the manufacturer's
392 instructions. Briefly, PC9 cell line was seeded at 50-70 % confluence in a six-well plate and transfected
393 2 µg of Piggybac transposase (pPBase-piggyBac) and 2 µg of mEGFP (pPBbsr2-mEGFP) plasmid DNAs.
394 After 24h of incubation, the medium with Lipofectamine/plasmid DNA mix was replaced with a fresh
395 medium. Cells were selected using 2 µg ml -1 blasticidin.
396 SCC154 (UPCI-SCC154) and SCC47 (UM-SCC47) were purchased from ATCC.
397 OCAF1, OCAF2, SCC154, SCC47 and CRUK0764 cells were cultured in DMEM (ThermoFisher,
398 #41966052) containing 10% fetal bovine serum (Gibco, #10270-106), 1% penicillin/streptomycin
399 (Invitrogen, #15140122), 1% insulin–transferrin–selenium (Invitrogen, #41400045) and kept at 37°C
400 and 5% CO2.
401 PC9 were cultured in RPMI-1640 (Thermo Fisher Scientific, Rockford, IL) supplemented with 10% fetal
402 bovine serum (Gibco, #10270-106), 1% penicillin/streptomycin (Invitrogen, #15140122) and kept at
403 37 °C in 5% CO2.
404 Cells were not allowed to reach more than 90% confluency for routine cell culture cultivation. Cell
405 lines that are not commercially obtainable are available from the authors upon reasonable request.
406 Routine screening for *Mycoplasma* testing was performed for all cell lines with negative results. STR
407 profiles of human non-commercially available cell lines are included in Supplementary Table 1.
408

409 ## Cell cultures conditions and treatments

410 Co-cultures and mono-cultures were performed with a ratio of 1:2, typically plating 5.5 x 10^5 CAFs
411 and 2.75 x 10^5 cancer cells for a single well of a 6 well plate for the specified time point. When co-
412 cultures were compared to pooled mono-cultures, for the mono-culture condition, same number of
413 cells was plated but in two separated wells and then lysed together (pooled condition).

414 When cancer cells and CAFs mono-culture were compared among themselves, 1 x 10^6 OCAF1 and 1
415 x 10^6 SCC154 cells were plated in a 10cm dish.
416 For PC9 and CRUK0764 cells monocultures and co-cultures used for RNAseq, following 24 h co-
417 cultures, the culture media was replaced with fresh medium with DMSO, then harvested after an
418 additional 24 h. PC9 – CRUK0764 co-cultures were performed in a mixture of RPMI-1640 and DMEM
419 (1:1) containing 1% fetal bovine serum (Gibco, #10270-106).
420 For macrophage cultivation, please see "Macrophage migration assay" section.
421 For cell culture treatments: drugs / factors were added when cells were plated and then added fresh
422 after 24h. Drugs / factors used: trametinib (Selleckchem, #GSK1120212), afatinib (Selleckchem,
423 #BIBW2992), human recombinant HB-EGF (Peprotech, #100-47).
424 For trametinib treatment to collect conditioned medium (CM), in order to control the effect of the
425 drug presence regardless of its effect on secreted factors, we added fresh trametinib treatment to
426 DMSO co-culture CM at the same concentration used for the cell co-culture treatment.
427 All concentrations used are specified in the figures.
428 RNA interference was performed with Lipofectamine RNAimax reagent from Invitrogen, according to
429 the manufacturer's instructions. For transient knock down of HB-EGF, cells were subjected to reverse
430 transfection with 20 nM RNAi oligos plus forward transfection the day after, then analyzed 4 days after
431 reverse transfection. The following RNAi oligo (Dharmacon) was used: siHB-EGF A (Cat # D-019624-
432 02), siHB-EGF B (Cat # D-019624-03), as control the following non-targeting siRNA oligo (All Stars
433 Negative, Quiagen, Cat # 1027281).
434

435 ## Fluorescence-activated cell sorting

436 For PC9 – CRUK0764 RNAseq experiment, CRUK0764 were labelled with CellVue® Red Mini Kit for
437 Membrane Labeling (Polysciences, 25567-1) according to the manufacturer's instructions. Briefly, 1 ×
438 10^7 cells of CRUK0764 were resuspended in the Diluent C and mixed with CellVue® Red working dye
439 solutions (final concentration: 5 × 10^6 cells/mL, 2 × 10^6 M dye) and then incubated for 5 min at RT.
440 Cells were washed twice with DMEM, 10% FBS medium to ensure removal of unbound fluorescence
441 dye.
442 For fluorescence-activated cell sorting, cells were sorted using a flow cytometer–cell sorter BD
443 FACSAria™ II. PC9-GFP and CRUK0764 -CellVue Red were sorted by FACS 48 h after seeding them in
444 monoculture or direct co-culture. The cells were then trypsinised and resuspended in 3% FBS in PBS,
445 1 mM EDTA in preparation for sorting. Cells were separated into two populations: PC9-GFP and
446 CRUK0764 with CellVue Red using a 488 nm laser with collection filter 530 nm/30 nm for GFP and 561
447 nm laser with collection filter 582 nm/20 nm for CellVue Red. Gates were designed on the basis of

448 negative and single-colour controls. All cell populations were tested for purity, and data were analysed
449 using FlowJo software.
450

451 ## RNA sequencing analysis for co-cultures

452 PC9 and CRUK0764 cells were immediately centrifuged at 300 × g for 4 min to remove supernatant
453 and add 350 µl RLT buffer (Qiagen, 79216) containing 1% β-mercaptoethanol (Sigma, M6250) and total
454 RNA was extracted using the RNAeasy Mini kit (Qiagen, 74104; n = 3 independent experiments). Prior
455 to library construction, the quality of total RNA was assessed by Bioanalyzer 2100 (Agilent
456 Technologies Inc).
457 For RNAseq analysis: biological replicates libraries were prepared using the polyA KAPA mRNA
458 HyperPrep Kit and sequenced on the Illumina HiSeq 4000 platform generating ~28 million 75bp single
459 end reads per sample. FASTQ_files were quality trimmed and adaptor removed using Trimmomatic
460 (version 0.36) [48]. The RSEM package (version 1.3.30) [49] in conjunction with the STAR alignment
461 software (version 2.5.2a) [50] was used for the mapping and subsequent gene level counting of the
462 mapped reads with respect to the Ensembl human GRCh38 (release 89) transcriptome. Normalisation
463 of raw count data was performed with the DESeq2 package (version 1. 18.1) [51]. All the analysis was
464 done (version 1. 18.1) [51] within the R programming environment (version 3. 4. 3).
465 To check the purity of the samples, we analyzed the resulting transcriptomic data for the expression
466 of 'lineage markers'. CDH1, EPCAM, CD24, and KRT genes were used as markers of carcinoma cells and
467 for fibroblasts we used COL1A1, COL1A2, DCN, CD248, and PDGFR genes. This revealed high sample
468 purity for all transcriptomic data, except in the PC9 – CRUK0764 experiment that had variable purity
469 between samples. Therefore, we estimated the impurity in each sample based on the expression of
470 the lineage marker genes and calculated the expected level of transcript if the two mono-cultures
471 (cancer cells alone and CAFs alone) were mixed in proportion with the impurity estimate. The
472 observed transcript in the co-culture condition was then normalized to account for the effect of
473 contamination.

474

475 ## EdU proliferation assay

476 The Click-iT Plus EdU Imaging Kit (Invitrogen #c10640) was used to perform the assay. Briefly, 48h after
477 mono-cultures of OCAF1 and SCC154 were seeding, a solution with Edu 20µM was prepared and then
478 diluted 1:1 with cell media to add EdU 10µM final concentration. After 90 minutes incubation, cells
479 were washes twice in PBS, then fixed for 15 minutes with paraformaldehyde 3.7% and then washed
480 twice in BSA 3%. Following this step, cells were incubated for 20 minutes with 0.5% Triton X-100 in

481 PBS. After two BSA 3% washes, the Click-iT reaction buffer was added for 30 minutes, followed by one
482 wash in BSA 3% and one wash in PBS. Subsequently, nuclei were stained with Hoechst 33342 at 5
483 µg/mL in PBS incubation for 30 minutes, followed by two PBS washes.
484 Samples were imaged with Zeiss 980 microscope. Mean fluorescent intensity was analysed using
485 ImageJ.
486

487 ## Immunofluorescence assay

488 The samples used to perform EdU proliferation assay have been then stained for E-Cadherin. Briefly,
489 samples were washes twice in PBS, followed by incubation for 30 minutes in BSA 3%. Then samples
490 were incubated over night at 4°C. After two washes in BSA 3% of 5 minutes each, samples were
491 incubated with secondary antibody Alexa Fluor 555 in BSA 3% for 45 min. Following this step, samples
492 were washed with PBS twice. Subsequently, nuclei were stained with Hoechst 33342 at 5 µg/mL in
493 PBS incubation for 30 minutes, followed by two PBS washes. Samples were imaged with Zeiss 980
494 microscope.
495

496 ## Peripheral Blood Mononuclear Cell extraction and monocytes selection

497 Donations of healthy blood donors were received from the Francis Crick Institute. PBMCs were
498 isolated from whole blood using Lymphoprep (Stemcell Technologies #7811) with SepMate™ density
499 centrifugation tubes in line with manufacturer's instructions (Stemcell Technologies #85450). Freshly
500 isolated PBMCs were then counted before isolation of monocytes (Miltenyi Biotec #130-096-537).
501 Monocytes were then counted for plated in normal plastic dishes.
502

503 ## Cytokine array

504 Cytokine array used is "Proteome Profiler Human XL Cytokine Array Kit" (R&D Systems, # ARY022B)
505 following manufacturer's instruction. Briefly, conditioned media was isolated and filtered through a
506 0.4µm low protein binding PVDF Miltex syringe-driven filter (Millipore #SLHV033RS) to remove cellular
507 debris. Media was then concentrated to 4X using Amicon® Ultra-15 and used for subsequent
508 incubation with array.
509

510 ## Macrophage migration assay

511 Monocytes were plated into 12-well plates ($1x10^5$ cells / well) in RPMI 1640 media (ThermoFisher
512 #12633-012) containing 10% FBS, 1% streptomycin/penicillin and 50ng/mL of M-CSF (Peprotech #300-

513 25) and kept at 37°C and 5% CO2 for 5 days to allow macrophage differentiation. During incubation
514 period, OCAF1 – SCC154 mono- and co-cultures were set-up for 48h. Conditioned media was isolated
515 and filtered through a 0.4µm low protein binding PVDF Miltex syringe-driven filter (Millipore
516 #SLHV033RS) to remove cellular debris. Media was then concentrated to 4X using Amicon® Ultra-15
517 centrifugal filter units (Millipore #UFC901024) and frozen into aliquots until needed. Conditioned
518 media was added to 24-well plates, 8µm hanging cell culture inserts (Millipore #MCEP12H48) were
519 placed on top of each well. The now differentiated macrophages were seeded inside the hanging cell
520 culture insert and left to settle for 10 minutes before topping up media. Plates were left in the
521 incubator for 5 hours to allow macrophages to migrate through membrane pores. After this time, the
522 inserts were removed and the macrophages sat on top of the membrane were wiped off with a cotton
523 bud, leaving behind the migrated macrophages at the bottom. Inserts were stained with 0.05% crystal
524 violet for 30 minutes before washing and then imaged. Inserts were imaged using Zeiss Observer Z1
525 mounted with a QImaging Colour camera. Quantification of crystal violet staining was carried out using
526 ImageJ through 'Cell Counter' function.
527

528 ## Gene Set Enrichment Analysis

529 Gene set enrichment analysis was performed with GSEA software v4.1.0. The dataset used to perform
530 the comparative analysis are: RAS84 derived from [24], CoCu8 derived from our own analysis, Hallmarks
531 (h.all.v7.5.symbols.gmt) for all the other analysis. All the parameters have been used as defaults
532 except: permutation type (gene set) and metric for ranking genes (t test). Gene signatures with a false
533 discovery rate < 0.05 were considered as statistically significant.
534

535 ## Transcriptomic data

536 The transcriptomic data used are: microarray data of A431 / VCAF2b under conditions of mono-
537 cultures, co-cultures in direct contact and indirect contact are available at the Gene Expression
538 Omnibus under record GSE121058. The microarray data of MAF2 under conditions of mono-culture
539 and co-culture in direct contact is available at the Gene Expression Omnibus under record GSE63160.
540 The microarray data of 5555 under conditions of mono-culture and co-culture in direct contact will be
541 submitted at the Gene Expression Omnibus.
542 The RNAseq data of PC9 / CRUK0764 under conditions of mono-cultures and co-cultures in direct
543 contact will be submitted to the European Genome-Phenome Archive before publication.
544 The microarray data used for HUVEC – 1205Lu analysis is available at Gene Expression Omnibus under
545 record GSE8699.

546 The microarray data used for breast cancer cell lines co-culture with fibroblasts analysis is available at
547 Gene Expression Omnibus under record GSE41678.
548

549 ## Co-culture gene signature generation

550 CoCu8 gene signature generation: the A431/VCAF2b, 5555/MAF2, PC9/CRUK0764 co-cultures vs
551 mono-cultures transcriptional datasets have been analysed with GSEA (see Gene Set Enrichment
552 Analysis method) to obtain a list of enriched pathways in co-culture with FDR < 0.05 for each condition.
553 For each cell type, all the genes statistically upregulated have been pulled together. From this list,
554 genes that were present in 20% or more of the enriched pathways have been selected. The results
555 obtained for each sample have been merged according to the cell type: the three cancer cells in co-
556 culture have been pulled together, same for the three CAFs. To select the final list, only genes present
557 in the three different cancer cells or in the three different CAFs have been selected to generate CoCu8.
558

559 CoCu30 gene signature generation: the genes with a fold change upregulation of 1.5 or higher have
560 been selected for each cell type upon co-culture. The results of the three cancer cells in co-culture
561 have been pulled together, same for the three CAFs in co-culture. To select the final list, only genes
562 present in the three different cancer cells or in the three different CAFs have been selected to generate
563 CoCu30.
564

565 ## TCGA analysis

566 Clinical data, RSEM (RNA-Seq by Expectation-Maximization) normalised expression data (Illumina
567 RNASEQ platform) and Methylation data (Illumina Human Methylation 450 platform) for HNSCC, CESC
568 and LUSC cohorts was downloaded from the Firebrowse website hosed by Broad Institute of MIT and
569 Harvard. [http://firebrowse.org/]. Data downloads were all version 2016012800.0.0.
570 De-convolution strategies:
571 MethylCIBERSORT: signature matrix and mixture files were obtained using MethylCIBERSORT R
572 package, hosted on Zenodo. The detailed origin of the signatures and the procedure to create the
573 deconvolution strategy is explained in Chakravarthy et al. [10].
574 Absolute-CIBERSORT: To calculate the immune infiltrate per sample, the library 'CIBERSORT' (version
575 1.04 [52]) was run within R version 3.4.3 on the RSEM normalised data and the LM22 signature using the
576 parameters absolute=TRUE and abs_method ="no.sumto1".
577

578 ## RNA extraction and RT-qPCR

579 Cells were collected and lysed with RLT buffer and total RNA was extracted using the RNeasy Mini kit
580 (Qiagen, #74104), according to the manufacturer's protocol.
581 The cDNA was prepared using M-MLV reverse transcriptase (Promega, #M3682), and quantitative PCR
582 was performed using PowerUp™ SYBR™ Green Master Mix (ThermoFisher, #A25778), using the
583 QuantStudio 3 and 7 Real-Time PCR systems (Applied Biosystems).
584 Custom primers were acquired from Sigma; sequences are available in Supplementary Table 2. RNA
585 levels were normalized using three house-keeping genes using the ΔΔC method and reported as
586 relative fold change compared with Ctr/not treated cells/mono-culture.
587

588 ## scRNAseq analysis

589 scRNAseq data from Choi et al. [29] was downloaded from GEO (GSE181919) and analysed using Seurat
590 package (version 4) [53].
591

592 ## Protein extraction, quantification and Western Blot analysis

593 Cells were lysed in RIPA buffer (50 mM TrisHCl, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 1% sodium
594 deoxycholate, 0.1% SDS), supplemented with a protease and phosphatase inhibitors (PhosSTOP tablet
595 Roche #04906837001, cOmplete EDTA-free Roche #11873580001, 50 mM NaF). Lysis was performed
596 directly in the cell culture plates using a cell scraper, lysates were kept for 10min on ice and then
597 clarified by centrifugation at 16,000 g for 30 min at 4 °C.
598 Total protein was quantified using the bicinchoninic acid method in accordance with manufacturer's
599 instructions (ThermoFisher, 23225). Following protein quantification, 20 µg of sample was loaded on
600 a 4–15% gradient Mini-PROTEAN TGX Gels (Biorad, #4561084) and transferred to a Trans-Blot Turbo
601 Mini 0.2 µm PVDF membrane (Biorad, 1704156) for blotting. The membrane was blocked for 1 h in
602 5% BSA or 5% milk in TBST and then incubated overnight at 4°C or 1 h at room temperature with
603 antibodies. The membrane was then washed before adding the horseradish peroxidase-conjugated
604 secondary antibody (ThermoFisher), and incubating for 1 h at room temperature. The membrane was
605 washed again before developing with Luminata Classico Western HRP substrate (Millipore,
606 #WBLUR0100) Luminata Classico Western HRP substrate (Millipore, # WBLUF0100) and imaging.
607 Antibody information are listed in Supplementary Table 3. All original blots are provided as source
608 data.
609

610 ## Software and visualization

611 Graphs were generated with Prism software (Graphpad Software v9.4.0) and R (version 4.2.1) using
612 package 'ggplot' except for correlation plot in Figure S9A that was generated with cBioportal [54].
613 scRNAseq data were analysed with Seurat package (version 4).
614

615 ## Statistics

616 Statistical analysis was performed using Prism software (Graphpad Software v9.4.0), Excel software
617 (Microsoft Corporation v16.0) and R (version 4.2.1).
618 All Student's t tests have been performed with two tailed strategy.
619 P-value information: * is p-value<0.05; ** is p-value<0.01, *** is p-value<0.001, **** is p-
620 value<0.0001.
621 For GSEA, we used FDR with a threshold below 0.05 to definite the significance.
622 Kaplan-Meier, Log-Rank and Cox regression on survival data was calculated using the R package
623 'survminer' using univariable analysis. Correlations were calculated using the Spearman method in R
624 and the package 'corplot' was used to generate the graphs.
625

626 ## UK_HPV positive cohort

627 FFPE tumor samples from 2 studies formed this cohort:
628 - INOVATE (MR/R015589/1 ISRCTN32335415), a prospective sample collection study in patients
629 with T1-T2/N1-3 or T3-T4/N0-3 oropharyngeal cancer (AJCC TNM classification 7.0) receiving
630 treatment with radical radiotherapy with or without additional platin-based chemotherapy.
631 - INSIGHT-2 (C7224/A23275 NCT04242459), a prospective study of optimising radiation therapy
632 in head and neck cancers using functional image-guided radiotherapy and novel biomarkers.
633 INOVATE was approved by the London -Bloomsbury Research Ethics Committee (19/LO/1558) and
634 INSIGHT-2 was approved by London – Queen Square Research Ethics Committee (19/LO/0638).
635 Written informed consent was obtained from all participants prior to any study procedure.
636

637 ## UK_HPV positive cohort RNAseq and data analysis

638 Baseline diagnostic biopsies embedded in paraffin blocks were obtained from the above-mentioned
639 cohort. Relevant tumor sections were selected and RNA was extracted from 3-4 slides using the
640 Qiagen AllPrep® DNA/RNA FFPE kit (#80234). Ribosomal RNA was depleted using QIAGEN FastSelect
641 rRNA H/M/R kit (#334375). RNA sequencing libraries were prepared using the NEBNext Ultra II RNA

642 Library Prep Kit (#E770) for Illumina following manufacturer's instructions. The sequencing libraries
643 were multiplexed and loaded on the flow cell on the Illumina NovaSeq 6000 instrument according to
644 manufacturer's instructions. The samples were sequenced using a 2x150 Pair-End (PE) configuration
645 v1.5 for an estimated output of ~50M paired-end reads per sample. Image analysis and base calling
646 were conducted by the NovaSeq Control Software v1.7 on the NovaSeq instrument. Raw sequence
647 data (.bcl files) generated from Illumina NovaSeq was converted into fastq files and de-multiplexed
648 using Illumina bcl2fastq program version 2.20. One mismatch was allowed for index sequence
649 identification. Sample adequacy was confirmed using FASTQC, low quality bases and reads were
650 trimmed using Trimmomatic, we run Hisat2-Stringtie for alignment.
651 RNAseq was performed on 103 patient samples, of which 7 were from the INSIGHT2 and 96 from
652 INOVATE. The data from RNAseq was analysed to identify samples with presence of HPV (by aligning
653 the unmapped sequences to the whole HPV16 genome sequence obtained from GEO using HISAT2
654 and StingTie) and these samples were classed as HPV positive. 84 samples (77 INOVATE and 7
655 INSIGHT2) were classified at HPV positive and RNAseq data from these was used for analysis in this
656 study.
657

658 ## TRACERx cohort

659 Tumor samples used in this study were collected from LUSC patients enrolled as a part of TRACERx
660 study (accession code: NCT01888601) which is sponsored by University College London (UCL/12/0279)
661 and has been approved by an independent research ethics committee (13/LO/1546). Multiple regions
662 were sampled per tumor and processed as described by Frankell et al. [20] yielding whole-RNA
663 sequencing data for 295 regions from 117 LUSC patients. Expression count and transcript per million
664 (TPM) were quantified by the RSEM package [49]. Genes with expression level of at least 1 TPM in at
665 least 20% of the samples were included. A variance stabilising transformation (VST) was then applied
666 to filtered count using the DESeq2 package [51].
667

668 ## Supplementary tables

669 Supplementary Tables 1, 2 and 3 are provided with this article.
670

671 ## Data accessibility

672 The transcriptomic data for UK_HPV positive cohort is part of ongoing clinical trials, therefore the data
673 cannot be deposited in a public repository until the trial is finalised. Data can be shared upon

674 reasonable request following corresponding Ethical Research Committee approval following the ICR-
675 Clinical Trials and Statistics Unit policy.
676

677 ## Authors contributions

678 G.G. and A.R. conceptualised and designed the research. GG performed the experimental analysis and
679 identified the cancer cell – CAF crosstalk gene signature strategy. A.R. performed the clinical data
680 analysis on TCGA and UK_HPV positive cohorts and performed the deconvolution strategies. G.G.,
681 A.R., S.H., E.H., Y.N., P.N. and P.C. acquired and analysed the data, except for TRACERx data that were
682 analysed by D.B. and UH.L.. C.S., A.M., K.H., S.B., M.C. and E.S. provided funding and access to samples
683 and reagents. G.G., A.R. and E.S. wrote the manuscript. All authors reviewed and edited the
684 manuscript. All authors authorised the final version.
685

686 ## Declaration of Interests and acknowledgments

687 We thank the Crick healthy donors' program and donators for blood donations from the Francis Crick
688 Institute. We thank Sophie de Carne and Phil East for useful discussions around RAS signature and RAS
689 signalling.
690 G.G. is funded by Merck Sharp & Dohme Corp, New Jersey, USA (LKR190557). A.R. acknowledges
691 funding from the Spanish Society for Medical Oncology (Beca Fundación SEOM), a CRUK accelerator
692 grant to The Francis Crick Institute, and the Royal Marsden NIHR/BRC Bridge funding program. E.H. is
693 supported by the Japan Society for the Promotion of Science (JSPS) Kakenhi Grant [20H03510]. Y.N
694 was supported by a Grant-in-aid for JSPS Overseas Research Fellowship (No. 201860634).
695 K.H. and A.M. acknowledge funding by the Wellcome Trust, ICR/RM NIHR Biomedical Research Centre,
696 The Institute of Cancer Research/Royal Marsden Hospital Centre for Translational Immunotherapy,
697 CRUK Head and Neck Programme Grant (C7224/A23275) and ICR/RM CRUK RadNet Centre of
698 Excellence (C7224/A28724). Disclosures: Honoraria: Arch Oncology (Inst), AstraZeneca (Inst), BMS
699 (Inst), Boehringer Ingelheim (Inst), Codiak Biosciences (Inst), F-Star Therapeutics (Inst), Inzen
700 Therapeutics (Inst), Merck Serono (Inst), MSD (Inst), Oncolys Biopharma (Inst), Pfizer (Inst), Replimune
701 (Inst), VacV Biotherapeutics (Inst); Consulting or Advisory Role: Arch Oncology (Inst), AstraZeneca
702 (Inst), BMS (Inst), Boehringer Ingelheim (Inst), Inzen Therapeutics (Inst), Merck Serono (Inst), MSD
703 (Inst), Oncolys BioPharma (Inst), Replimune (Inst); Speakers' Bureau: BMS (Inst), Merck Serono (Inst),
704 MSD (Inst); Research Funding: AstraZeneca (Inst), Boehringer Ingelheim (Inst), Merck Sharp & Dohme
705 (Inst), Replimune (Inst).

706 S.B. declares no conflict of Interest and funding support from ICR/RM NIHR Biomedical Research
707 Centre, CRUK Head and Neck Programme Grant (C7224/A23275), ICR/RM CRUK RadNet Centre of
708 Excellence (C7224/A28724) and the Medical Research Council Developmental Pathway Funding
709 Scheme [MR/R015589/1].
710 D.B. was supported by funding from a Cancer Research UK (CRUK) Early Detection and Diagnosis
711 Project award, the Idea to Innovation (i2i) Crick translation scheme supported by the Medical Research
712 Council, the National Institute for Health Research Biomedical Research Centre and the Breast Cancer
713 Research Foundation (BCRF). D.B. reports personal fees from NanoString and AstraZeneca, and has a
714 patent PCT/GB2020/050221 issued on methods for cancer prognostication.
715 C.S. acknowledges grant support from AstraZeneca, Boehringer-Ingelheim, Bristol Myers Squibb,
716 Pfizer, Roche-Ventana, Invitae (previously Archer Dx Inc - collaboration in minimal residual disease
717 sequencing technologies), and Ono Pharmaceutical. He is an AstraZeneca Advisory Board member and
718 Chief Investigator for the AZ MeRmaiD 1 and 2 clinical trials and is also Co-Chief Investigator of the
719 NHS Galleri trial funded by GRAIL and a paid member of GRAIL's Scientific Advisory Board. He receives
720 consultant fees from Achilles Therapeutics (also SAB member), Bicycle Therapeutics (also a SAB
721 member), Genentech, Medicxi, Roche Innovation Centre – Shanghai, Metabomed (until July 2022),
722 and the Sarah Cannon Research Institute. C.S has received honoraria from Amgen, AstraZeneca, Pfizer,
723 Novartis, GlaxoSmithKline, MSD, Bristol Myers Squibb, Illumina, and Roche-Ventana. C.S. had stock
724 options in Apogen Biotechnologies and GRAIL until June 2021, and currently has stock options in Epic
725 Bioscience, Bicycle Therapeutics, and has stock options and is co-founder of Achilles Therapeutics.
726 C.S. holds patents relating to assay technology to detect tumor recurrence (PCT/GB2017/053289); to
727 targeting neoantigens (PCT/EP2016/059401), identifying patent response to immune checkpoint
728 blockade (PCT/EP2016/071471), determining HLA LOH (PCT/GB2018/052004), predicting survival
729 rates of patients with cancer (PCT/GB2020/050221), identifying patients who respond to cancer
730 treatment (PCT/GB2018/051912), US patent relating to detecting tumor mutations
731 (PCT/US2017/28013), methods for lung cancer detection (US20190106751A1) and both a European
732 and US patent related to identifying insertion/deletion mutation targets (PCT/GB2018/051892). C.S.
733 is a Royal Society Napier Research Professor (RSRP\R\210001). This work was supported by the Francis
734 Crick Institute that receives its core funding from Cancer Research UK (CC2041), the UK Medical
735 Research Council (CC2041), and the Wellcome Trust (CC2041). For the purpose of Open Access, the
736 author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising
737 from this submission. C.S. is funded by Cancer Research UK (TRACERx (C11496/A17786), PEACE
738 (C416/A21999) and CRUK Cancer Immunotherapy Catalyst Network); Cancer Research UK Lung Cancer
739 Centre of Excellence (C11496/A30025); the Rosetrees Trust, Butterfield and Stoneygate Trusts;

740 NovoNordisk Foundation (ID16584); Royal Society Professorship Enhancement Award
741 (RP/EA/180007); National Institute for Health Research (NIHR) University College London Hospitals
742 Biomedical Research Centre; the Cancer Research UK-University College London Centre; Experimental
743 Cancer Medicine Centre; the Breast Cancer Research Foundation (US) BCRF-22-157; Cancer Research
744 UK Early Detection an Diagnosis Primer Award (Grant EDDPMA-Nov21/100034); and The Mark
745 Foundation for Cancer Research Aspire Award (Grant 21-029-ASP). This work was supported by a
746 Stand Up To Cancer-LUNGevity-American Lung Association Lung Cancer Interception Dream Team
747 Translational Research Grant (Grant Number: SU2C-AACR-DT23-17 to S.M. Dubinett and A.E. Spira).
748 Stand Up To Cancer is a division of the Entertainment Industry Foundation. Research grants are
749 administered by the American Association for Cancer Research, the Scientific Partner of SU2C. C.S. is
750 in receipt of an ERC Advanced Grant (PROTEUS) from the European Research Council under the
751 European Union's Horizon 2020 research and innovation programme (grant agreement no. 835297).
752 E.S. is supported by the Francis Crick Institute which receives its core funding from Cancer Research
753 UK (CC2040), the UK Medical Research Council (CC2040), and the Wellcome Trust (CC2040). E.S. is
754 additionally funded by the European Research Council ERCAdG CAN_ORGANISE 101019366. E.S.
755 declares research funding from Merck Sharp & Dohme and Astrazeneca and consultancy work for
756 Phenomic AI and Theolytics. For the purpose of Open Access, the author has applied a CC BY public
757 copyright licence to any Author Accepted Manuscript version arising from this submission.
758 RNA sequencing of the UK_HPV positive cohort was funded by the NIHR through the RMH BRC Pump
759 Priming theme (grant B140).
760 The INOVATE study was funded by the Medical Research Council Developmental Pathway Funding
761 Scheme [MR/R015589/1]. The INSIGHT-2 trial was supported by CRUK Head and Neck Programme
762 Grant (C7224/A23275).
763 This study represents independent research supported by the National Institute for Health Research
764 (NIHR) Biomedical Research Centre at the Royal Marsden National Health Service Foundation Trust
765 and the Institute of Cancer Research, London, UK.

766

767 ## References

768 1. Maman S, Witz IP. The significant problems we face today cannot be solved at the same level
769 of thinking we were at when we created them. *Nat Rev Cancer*. Published online 2018.
770 doi:10.1038/s41568-018-0006-7

771 2. Sahai E, Astsaturov I, Cukierman E, et al. A framework for advancing our understanding of
772 cancer-associated fibroblasts. *Nat Rev Cancer*. 2020;20(3):174-186. doi:10.1038/s41568-019-
773 0238-1

774 3. Qi J, Sun H, Zhang Y, et al. Single-cell and spatial analysis reveal interaction of FAP +
775 fibroblasts and SPP1 + macrophages in colorectal cancer. doi:10.1038/s41467-022-29366-6

776 4. Ford K, Hanley CJ, Mellone M, et al. NOX4 inhibition potentiates immunotherapy by
777 overcoming cancer-associated fibroblast-mediated CD8 T-cell exclusion from tumors. *Cancer*
778 *Res*. 2020;80(9):1846-1860. doi:10.1158/0008-5472.CAN-19-3158/653979/AM/NOX4-
779 INHIBITION-POTENTIATES-IMMUNOTHERAPY-BY

780 5. Grout JA, Sirven P, Leader AM, et al. Spatial Positioning and Matrix Programs of Cancer-
781 Associated Fibroblasts Promote T-cell Exclusion in Human Lung Tumors. *Cancer Discov*.
782 2022;12(11):2606-2625. doi:10.1158/2159-8290.CD-21-1714/708811/AM/SPATIAL-
783 POSITIONING-AND-MATRIX-PROGRAMS-OF-CANCER

784 6. Li B, Pei G, Yao J, Ding Q, Jia P, Zhao Z. Cell-type deconvolution analysis identifies cancer-
785 associated myofibroblast component as a poor prognostic factor in multiple cancer types.
786 *Oncogene 2021 40:28*. 2021;40(28):4686-4694. doi:10.1038/s41388-021-01870-x

787 7. Galbo PM, Zang X, Zheng D. Molecular features of cancer-associated fibroblast subtypes and
788 their implication on cancer pathogenesis, prognosis, and immunotherapy resistance. *Clinical*
789 *Cancer Research*. 2021;27(9):2636-2647. doi:10.1158/1078-0432.CCR-20-
790 4226/79067/AM/IMPLICATION-OF-CANCER-ASSOCIATED-FIBROBLAST

791 8. Kürten CHL, Kulkarni A, Cillo AR, et al. Investigating immune and non-immune cell
792 interactions in head and neck tumors by single-cell RNA sequencing. *Nature Communications*
793 *2021 12:1*. 2021;12(1):1-16. doi:10.1038/s41467-021-27619-4

794 9. Cillo AR, Kürten CHL, Tabib T, et al. Immune Landscape of Viral- and Carcinogen-Driven Head
795 and Neck Cancer. *Immunity*. 2020;52(1):183-199.e9. doi:10.1016/J.IMMUNI.2019.11.014

796 10. Chakravarthy A, Furness A, Joshi K, et al. Pan-cancer deconvolution of tumour composition
797 using DNA methylation. *Nature Communications 2018 9:1*. 2018;9(1):1-13.
798 doi:10.1038/s41467-018-05570-1

799 11. Craven KE, Gökmen-Polar Y, Badve SS. CIBERSORT analysis of TCGA and METABRIC identifies
800 subgroups with better outcomes in triple negative breast cancer. *Scientific Reports 2021 11:1*.
801 2021;11(1):1-19. doi:10.1038/s41598-021-83913-7

802 12. Obradovic A, Graves D, Korrer M, et al. Immunostimulatory Cancer-Associated Fibroblast
803 Subpopulations can Predict Immunotherapy Response in Head and Neck
804 CancerImmunostimulatory CAF can Predict Immunotherapy Response. *Clinical Cancer*
805 *Research*. 2022;13:OF1-OF16. doi:10.1158/1078-0432.CCR-21-3570

806 13. Peng YL, Xiong L Bin, Zhou ZH, et al. Single-cell transcriptomics reveals a low CD8+ T cell
807 infiltrating state mediated by fibroblasts in recurrent renal cell carcinoma. *J Immunother*
808 *Cancer*. 2022;10(2):e004206. doi:10.1136/JITC-2021-004206

809 14. Chu T, Wang Z, Pe'er D, Danko CG. Cell type and gene expression deconvolution with
810 BayesPrism enables Bayesian integrative analysis across bulk and single-cell RNA sequencing
811 in oncology. *Nature Cancer 2022 3:4*. 2022;3(4):505-517. doi:10.1038/s43018-022-00356-3

812 15. Noureen N, Ye Z, Chen Y, Wang X, Zheng S. Signature-scoring methods developed for bulk
813 samples are not adequate for cancer single-cell RNA sequencing data. *Elife*. 2022;11.
814 doi:10.7554/ELIFE.71994

815 16. Hirata E, Girotti MR, Viros A, et al. Intravital imaging reveals how BRAF inhibition generates
816 drug-tolerant microenvironments with high integrin β1/FAK Signaling. *Cancer Cell*.
817 2015;27(4):574-588. doi:10.1016/j.ccell.2015.03.008

818 17. Arwert EN, Milford EL, Rullan A, et al. STING and IRF3 in stromal fibroblasts enable sensing of
819 genomic stress in cancer cells to undermine oncolytic viral therapy. *Nat Cell Biol*.
820 2020;22(7):758-766. doi:10.1038/s41556-020-0527-7

821 18. Rajaram M, Li J, Egeblad M, Powers RS. System-wide analysis reveals a complex network of
822 tumor-fibroblast interactions involved in tumorigenicity. *PLoS Genet*. 2013;9(9).
823 doi:10.1371/JOURNAL.PGEN.1003789

824 19. Stine MJ, Wang CJ, Moriarty WF, et al. Integration of genotypic and phenotypic screening
825 reveals molecular mediators of melanoma-stromal interaction. *Cancer Res*. 2011;71(7):2433-
826 2444. doi:10.1158/0008-5472.CAN-10-1875

827 20. Frankell AM, Dietzen M, Al Bakir M, et al. The evolution of lung cancer and impact of
828 subclonal selection in TRACERx. *Nature 2023 616:7957*. 2023;616(7957):525-533.
829 doi:10.1038/s41586-023-05783-5

830 21. Gazon H, Barbeau B, Mesnard JM, Peloponese JM. Hijacking of the AP-1 signaling pathway
831 during development of ATL. *Front Microbiol*. 2018;8(JAN):2686.
832 doi:10.3389/FMICB.2017.02686/BIBTEX

833 22. Westwick JK, Cox AD, Der CJ, et al. Oncogenic Ras activates c-Jun via a separate pathway from
834 the activation of extracellular signal-regulated kinases. *Proc Natl Acad Sci U S A*.
835 1994;91(13):6030-6034. doi:10.1073/PNAS.91.13.6030

836 23. Eferl R, Wagner EF. AP-1: a double-edged sword in tumorigenesis. *Nature Reviews Cancer*
837 *2003 3:11*. 2003;3(11):859-868. doi:10.1038/nrc1209

838 24. East P, Kelly GP, Biswas D, et al. RAS oncogenic activity predicts response to chemotherapy
839 and outcome in lung adenocarcinoma. *Nature Communications 2022 13:1*. 2022;13(1):1-17.
840 doi:10.1038/s41467-022-33290-0

841 25. Fang L, Li G, Liu G, Lee SW, Aaronson SA. p53 induction of heparin-binding EGF-like growth
842 factor counteracts p53 growth suppression through activation of MAPK and PI3K/Akt
843 signaling cascades. *EMBO J*. 2001;20(8):1931-1939. doi:10.1093/EMBOJ/20.8.1931

844 26. Dao DT, Anez-Bustillos L, Adam RM, Puder M, Bielenberg DR. Heparin-Binding Epidermal
845 Growth Factor–Like Growth Factor as a Critical Mediator of Tissue Repair and Regeneration.
846 *Am J Pathol*. 2018;188(11):2446-2456. doi:10.1016/J.AJPATH.2018.07.016

847 27. Prince RN, Schreiter ER, Zou P, et al. The heparin-binding domain of HB-EGF mediates
848 localization to sites of cell-cell contact and prevents HB-EGF proteolytic release. *J Cell Sci*.
849 2010;123(13):2308-2318. doi:10.1242/JCS.058321

850 28. Provenzano AP, Besner GE, James PF, Harding PA. Heparin-binding EGF-like growth factor
851 (HB-EGF) overexpression in transgenic mice downregulates insulin-like growth factor binding
852 protein (IGFBP)-3 and -4 mRNA. *http://dx.doi.org/101080/08977140512331344012*.
853 2009;23(1):19-31. doi:10.1080/08977140512331344012

854 29. Choi JH, Lee BS, Jang JY, et al. Single-cell transcriptome profiling of the stepwise progression
855 of head and neck cancer. *Nature Communications 2023 14:1*. 2023;14(1):1-13.
856 doi:10.1038/s41467-023-36691-x

857 30. Knops AM, South A, Rodeck U, et al. Cancer-Associated Fibroblast Density, Prognostic
858 Characteristics, and Recurrence in Head and Neck Squamous Cell Carcinoma: A Meta-
859 Analysis. *Front Oncol*. 2020;10:2621. doi:10.3389/FONC.2020.565306/BIBTEX

860 31. Qi J, Sun H, Zhang Y, et al. Single-cell and spatial analysis reveal interaction of FAP+ fibroblasts
861 and SPP1+ macrophages in colorectal cancer. *Nature Communications 2022 13:1*.
862 2022;13(1):1-20. doi:10.1038/s41467-022-29366-6

863 32. Chen S, Morine Y, Tokuda K, et al. Cancer-associated fibroblast-induced M2-polarized
864 macrophages promote hepatocellular carcinoma progression via the plasminogen activator
865 inhibitor-1 pathway. *Int J Oncol*. 2021;59(2):1-14. doi:10.3892/IJO.2021.5239/HTML

866 33. Lambrechts D, Wauters E, Boeckx B, et al. Phenotype molding of stromal cells in the lung
867 tumor microenvironment. *Nature Medicine 2018 24:8*. 2018;24(8):1277-1289.
868 doi:10.1038/s41591-018-0096-5

869 34. Dominguez CX, Müller S, Keerthivasan S, et al. Single-cell RNA sequencing reveals stromal
870 evolution into LRRC15+ myofibroblasts as a determinant of patient response to cancer
871 immunotherapy. *Cancer Discov*. 2020;10(2):232-253. doi:10.1158/2159-8290.CD-19-
872 0644/333443/AM/SINGLE-CELL-RNA-SEQUENCING-REVEALS-STROMAL

873 35. Barrett RL, Pure E. Cancer-associated fibroblasts and their influence on tumor immunity and
874 immunotherapy. *Elife*. 2020;9:1-20. doi:10.7554/ELIFE.57243

875 36. Kim JK, Marco MR, Choi SH, et al. KRAS mutant rectal cancer cells interact with surrounding
876 fibroblasts to deplete the extracellular matrix. *Mol Oncol*. 2021;15(10):2766-2781.
877 doi:10.1002/1878-0261.12960

878 37. Tape CJ, Ling S, Dimitriadi M, et al. Oncogenic KRAS Regulates Tumor Cell Signaling via
879 Stromal Reciprocation. *Cell*. 2016;165(4):910-920. doi:10.1016/J.CELL.2016.03.029

880 38. Kadur Lakshminarasimha Murthy P, Xi R, Arguijo D, et al. Epigenetic basis of oncogenic-Kras-
881 mediated epithelial-cellular proliferation and plasticity. *Dev Cell*. 2022;57(3):310-328.e9.
882 doi:10.1016/J.DEVCEL.2022.01.006

883 39. Pascual-García M, Bonfill-Teixidor E, Planas-Rigol E, et al. LIF regulates CXCL9 in tumor-
884 associated macrophages and prevents CD8+ T cell tumor-infiltration impairing anti-PD1
885 therapy. *Nature Communications 2019 10:1*. 2019;10(1):1-11. doi:10.1038/s41467-019-
886 10369-9

887 40. Ang KK, Harris J, Wheeler R, et al. Human Papillomavirus and Survival of Patients with
888 Oropharyngeal Cancer. *New England Journal of Medicine*. 2010;363(1):24-35.
889 doi:10.1056/NEJMOA0912217/SUPPL_FILE/NEJMOA0912217_DISCLOSURES.PDF

890 41. Fakhry C, Westra WH, Li S, et al. Improved Survival of Patients With Human Papillomavirus–
891 Positive Head and Neck Squamous Cell Carcinoma in a Prospective Clinical Trial. *JNCI: Journal*
892 *of the National Cancer Institute*. 2008;100(4):261-269. doi:10.1093/JNCI/DJN011

893 42. Hayes DN, Van Waes C, Seiwert TY. Genetic landscape of human papillomavirus-associated
894 head and neck cancer and comparison to tobacco-related tumors. *Journal of Clinical*
895 *Oncology*. 2015;33(29):3227-3234. doi:10.1200/JCO.2015.62.1086

896 43. Anderson CM, Kimple RJ, Lin A, Karam SD, Margalit DN, Chua MLK. De-Escalation Strategies in
897 HPV-Associated Oropharynx Cancer—Are we Putting the Cart Before the Horse? *Int J Radiat*
898 *Oncol Biol Phys*. 2019;104(4):705-709. doi:10.1016/j.ijrobp.2019.02.054

899 44. Dummer R, Hauschild A, Santinami M, et al. Five-Year Analysis of Adjuvant Dabrafenib plus
900 Trametinib in Stage III Melanoma. *New England Journal of Medicine*. 2020;383(12):1139-
901 1148. doi:10.1056/NEJMOA2005493/SUPPL_FILE/NEJMOA2005493_DATA-SHARING.PDF

902 45. Planchard D, Smit EF, Groen HJM, et al. Dabrafenib plus trametinib in patients with previously
903 untreated BRAFV600E-mutant metastatic non-small-cell lung cancer: an open-label, phase 2
904 trial. *Lancet Oncol*. 2017;18(10):1307-1316. doi:10.1016/S1470-2045(17)30679-4

905 46. Uppaluri R, Winkler AE, Lin T, et al. Biomarker and tumor responses of oral cavity squamous
906 cell carcinoma to trametinib: A phase II neoadjuvant window-of-opportunity clinical trial.
907 *Clinical Cancer Research*. 2017;23(9):2186-2194. doi:10.1158/1078-0432.CCR-16-
908 1469/116024/AM/BIOMARKER-AND-TUMOR-RESPONSES-OF-ORAL-CAVITY

909 47. Gaggioli C, Hooper S, Hidalgo-Carcedo C, et al. Fibroblast-led collective invasion of carcinoma
910 cells with differing roles for RhoGTPases in leading and following cells. *Nature Cell Biology*
911 *2007 9:12*. 2007;9(12):1392-1400. doi:10.1038/ncb1658

912 48. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
913 *Bioinformatics*. 2014;30(15):2114-2120. doi:10.1093/BIOINFORMATICS/BTU170

914 49. Li B, Dewey CN. RSEM: Accurate transcript quantification from RNA-Seq data with or without
915 a reference genome. *BMC Bioinformatics*. 2011;12(1):1-16. doi:10.1186/1471-2105-12-
916 323/TABLES/6

917 50. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner.
918 *Bioinformatics*. 2013;29(1):15-21. doi:10.1093/BIOINFORMATICS/BTS635

919 51. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-
920 seq data with DESeq2. *Genome Biol*. 2014;15(12):1-21. doi:10.1186/S13059-014-0550-
921 8/FIGURES/9

922 52. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue
923 expression profiles. *Nat Methods*. 2015;12(5):453-457. doi:10.1038/NMETH.3337

924 53. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data
925 across different conditions, technologies, and species. *Nature Biotechnology 2018 36:5*.
926 2018;36(5):411-420. doi:10.1038/nbt.4096

927 54. Cerami E, Gao J, Dogrusoz U, et al. The cBio Cancer Genomics Portal: An Open Platform for
928 Exploring Multidimensional Cancer Genomics Data. *Cancer Discov*. 2012;2(5):401-404.
929 doi:10.1158/2159-8290.CD-12-0095

930 55. Biswas D, Birkbak NJ, Rosenthal R, et al. A clonal expression biomarker associates with lung
931 cancer mortality. *Nature Medicine 2019 25:10*. 2019;25(10):1540-1548. doi:10.1038/s41591-
932 019-0595-z

933

934 ## Figure legends

935 **Figure 1:** Cancer cell / CAFs co-culture gene signature CoCu8 is associated with worse overall survival
936 in multiple squamous cell carcinoma datasets. **A)** Strategy used to obtain CoCu8 gene signature. **B)**
937 Venn diagrams of the genes upregulated in the different datasets (top) and summary table of the
938 genes upregulated in all the datasets (bottom) for cancer cells (right) and CAFs (left). **C)** Kaplan-Meier
939 overall survival analysis of HNSCC (right), CESC (centre), LUSC (left) TCGA datasets stratified for CoCu8
940 first vs last quartile. Numbers at risk shown in tables below graphs. HNSCC HR=1.95 (95% Confidence
941 Interval (CI) 1.29-2.93), p-value=0.0011. CESC HR=2.79 (95%CI 1.40-5.56), p-value=0.0024, LUSC
942 HR=1.85 (95%CI 1.28-2.70), p-value=0.001. HR and CI were calculated using Cox regression. p-value
943 was calculated using logRank test. **D)** Kaplan-Meier overall survival analysis of HNSCC HPV positive
944 (right) and negative (left) TCGA datasets stratified for CoCu8 first vs last quartile. Numbers at risk
945 shown in tables below graphs. HNSCC HPV positive HR=3.19 (95%CI 1.24-8.18), p-value=0.011. HPV
946 negative HR=1.28 (95%CI 0.84-1.93), p-value=0.25. HR and CI were calculated using Cox regression. p-
947 value was calculated using logRank test. **E)** Kaplan-Meier overall survival analysis of LUSC TRACERx
948 dataset. Individual tumors stratified as high-, discordant or low-risk according to expression profile of
949 CoCu8 signature across multiple regions, as previously described and stratified according to Biswas et
950 al. [55]. Below are shown the numbers at risk in years. HR=1.67 (95%CI 1.13-2.47), p-value=0.0247. HR
951 and CI calculated using Cox regression and are referred to CoCu8 low vs CoCu8 high. p-value was
952 calculated using logRank test.

953

954 **Figure 2:** Cancer cell / CAFs co-culture upregulates AP-1 TF genes. **A)** Bubble plot of the Hallmarks
955 pathways upregulated by cancer cells / CAFs culture. Normalised enrichment score (NES) is depicted
956 as bubble size; False discovery rate (FDR) is depicted as colour intensity. **B)** Heatmap of expression of
957 the genes commonly upregulated upon co-culture in all the tested conditions from the
958 TNFA_SIGNALLING_VIA_NKFB pathway. **C)** qPCR analysis of *JUNB*, *FOS* and *FOSB* genes in
959 OCAF1/OCAF2 with SCC154/SCC47 pooled mono-cultures and co-culture after 24h. mRNA expression
960 is reported as mean ± standard deviation (SD) fold change difference over pooled mono-culture.
961 Genes have been normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 4
962 independent experiments. Paired t-test.

963

964 **Figure 3:** RAS activity is upregulated in cancer cell and CAFs upon co-culture. **A)** Gene set enrichment
965 analysis (GSEA) plot of RAS84 gene signature in mono-culture and co-culture. NES and FDR are
966 specified below each plot. **B)** Kaplan-Meier overall survival analysis of HNSCC HPV positive TCGA

967 dataset stratified for Ras84 activity according to [24]. Below are shown the numbers at risk in years.
968 RAS84_0 vs RAS84_1 HR=1.01 (95%CI 0.387-2.62) p-value=0.98. RAS84_0 vs RAS84_max HR=5.85
969 (95%CI 2.46-13.9) p-value<0.001. HR and CI calculated using Cox regression. **C)** Correlation plot of
970 RAS84 expression level and CoCu8 expression level in HNSCC HPV positive TCGA dataset. R is
971 Spearman correlation coefficient. n=97. **D)** Box plot analysis of fibroblast abundance via Methyl
972 CIBERSORT deconvolution strategy in HNSCC HPV positive TCGA dataset according to RAS84 activity.
973 Independent t-test. **E)** Western blot analysis of OCAF1 – SCC154 pooled mono-culture vs co-culture
974 for 48h showing the indicated antibodies. Vinculin is used as loading control. **F)** Western blot analysis
975 of OCAF1 – SCC154 co-cultures for 48h at the indicated conditions showing the indicated antibodies.
976 Vinculin is used as loading control. **G)** qPCR analysis of *JUNB*, *FOS* and *FOSB* genes in OCAF1 - SCC154
977 co-cultures for the indicated treatments after 48h. mRNA expression is reported as mean ± standard
978 deviation (SD) fold change difference over co-culture DMSO. Genes have been normalized over the
979 average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 4 independent experiments. Paired t-
980 test.
981
982 **Figure 4:** HB-EGF / EGFR axis activates AP-1 TF genes in cancer cells and CAFs upon co-culture via RAS
983 pathway. **A)** Pattern of expression of the 7 EGFR ligands in all the tested transcriptomic datasets. **B)**
984 Correlation plot of *HB-EGF* expression level and CoCu8 expression level in HNSCC HPV positive TCGA
985 dataset (left, n=97) and UK_HPV positive cohort (right, n=84). R is Spearman correlation coefficient.
986 **C)** Western blot analysis of OCAF1 – SCC154 pooled mono-culture vs co-culture for 48h showing the
987 indicated antibodies. Vinculin is used as loading control. **D)** Western blot analysis of OCAF1 – SCC154
988 co-cultures for 48h at the indicated conditions showing the indicated antibodies. Vinculin is used as
989 loading control. **E)** qPCR analysis of *JUNB*, *FOS* and *FOSB* genes in OCAF1 - SCC154 co-cultures for the
990 indicated treatments after 48h. mRNA expression is reported as mean ± standard deviation (SD) fold
991 change difference over co-culture DMSO. Genes have been normalized over the average of *GAPDH*,
992 *ACTB* and *RPLP0* housekeeping genes. n = 4 independent experiments. The DMSO treated sample is
993 the same used for Figure 3G. Paired t-test. **F)** Western blot analysis of OCAF1 and SCC154 mono-
994 cultures after 48h showing the indicated antibodies. Vinculin is used as loading control. **G)**
995 Proliferation assay of OCAF1 and SCC154 mono-cultures for the indicated treatments after 48h and
996 stained with EdU and Hoechst 33342. On the left, a representative image is shown with bar graph. On
997 the right, dot plot of mean fluorescent intensity (MFI) of EdU over Hoechst 33342 with
998 mean ± standard deviation (SD) highlighted. Each dot is a field of view. n = 3 independent experiments.
999 t-test. **H)** qPCR analysis of *JUNB*, *FOS* and *FOSB* genes in OCAF1 - SCC154 co-cultures after 48h with
1000 SCC154 pre-treated with the indicated conditions. mRNA expression is reported as mean ± standard

1001 deviation (SD) fold change difference over siC- condition. Genes has been normalized over the average
1002 of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n ≥ 4 independent experiments. Paired t-test.
1003
1004 **Figure 5:** Cancer cells – CAFs co-culture induces production of specific cytokines to attract
1005 macrophages. **A)** On the right is shown scRNAseq UMAP analysis from Choi et al. [29]. On the left is
1006 shown violin plot of EGFR and HBEGF mRNA expression levels in the indicated clusters. **B)** Cytokine
1007 array of conditioned medium from pooled mono-culture and co-culture of OCAF1 and SCC154.
1008 Highlighted relevant cytokines. n = 2 independent experiments. **C)** qPCR analysis of *LIF* gene in OCAF1
1009 – SCC154 pooled mono-culture vs co-culture after 24h. mRNA expression is reported as
1010 mean ± standard deviation (SD) fold change difference over pooled mono-culture. Genes have been
1011 normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 4 independent
1012 experiments. Paired t-test. **D)** qPCR analysis of *LIF* gene in OCAF1 - SCC154 co-cultures for the indicated
1013 treatments after 48h. mRNA expression is reported as mean ± standard deviation (SD) fold change
1014 difference over co-culture DMSO. Genes have been normalized over the average of *GAPDH*, *ACTB* and
1015 *RPLP0* housekeeping genes. n = 4 independent experiments. Paired t-test. **E)** qPCR analysis of *LIF* gene
1016 in OCAF1 - SCC154 co-cultures after 48h with SCC154 pre-treated with the indicated conditions. mRNA
1017 expression is reported as mean ± standard deviation (SD) fold change difference over siC- condition.
1018 Gene has been normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n ≥ 4
1019 independent experiments. Paired t-test. **F)** qPCR analysis of *LIF* gene in OCAF1 and SCC154 mono-
1020 cultures for the indicated treatments after 48h. mRNA expression is reported as mean ± standard
1021 deviation (SD) fold change difference over non-treated (NT) condition for each cell type. Genes have
1022 been normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 5 independent
1023 experiments. Paired t-test. **G)** Migration assay of macrophages plated in transwells with conditioned
1024 medium (CM) from OCAF1-SCC154 pooled mono-culture or co-culture. CM have been obtained after
1025 48h culture. On the left, a representative field of view is shown with bar graph. On the right, dot plot
1026 of number of cells per field of view as mean ± standard deviation (SD). Each dot is a field of view
1027 normalized by the average of the pooled mono-culture CM sample. n = 4 different donors. t-test. **H)**
1028 Migration assay of macrophages plated in transwells with CM from OCAF1-SCC154 co-culture with the
1029 indicated treatments. CM have been obtained after 48h culture and for the fresh trametinib sample,
1030 the drug has been added after the CM was collected. Dot plot of number of cells per field of view as
1031 mean ± standard deviation (SD) is shown. Each dot is a field of view normalized by the average of the
1032 co-culture DMSO CM sample. n = 3 different donors. t-test. **I)** At the top, box plot analysis of CD14+
1033 monocytic/macrophage lineage immune cell absolute score via Methyl CIBERSORT deconvolution
1034 strategy in HNSCC HPV positive TCGA dataset separate by first and last quartile of CoCu8 expression.

1035 Independent t-test, Bonferroni correction for multiple comparisons. At the bottom, box plot analysis
1036 of monocyte and macrophage immune cell score using Absolute CIBERSORT deconvolution strategy in
1037 HNSCC HPV positive UK_HPV positive dataset separate by first and last quartile of CoCu8 expression.
1038 Independent t-test.
1039
1040 **Figure S1:** Cancer cell / CAFs co-culture gene signature CoCu30 is associated with worse overall survival
1041 in multiple squamous cell carcinoma datasets. **A)** Strategy used to obtain CoCu30 gene signature. **B)**
1042 Venn diagram of the genes upregulated in the different datasets (top) and a table to summarize the
1043 genes constantly upregulated in the datasets (bottom) for cancer cells (left) and CAFs (right). **C)**
1044 Kaplan-Meier overall survival analysis of HNSCC (left), CESC (centre), LUSC (right) TCGA datasets
1045 stratified for CoCu30 first vs last quartile. Below each analysis are shown the corresponding numbers
1046 at risk, time in years. HNSCC HR=1.57 (95%CI 1.06-2.35), p-value=0.024. CESC HR=2.08 (95%CI 1.09-
1047 4.01), p-value=0.024. LUSC HR=1.78 (95%CI 1.23-2.59), p-value=0.0019. HR and CI were calculated
1048 using Cox regression. p-value was calculated using logRank test. **D)** Kaplan-Meier overall survival
1049 analysis of HNSCC HPV positive (left) and HNSCC HPV positive (right) TCGA datasets stratified for
1050 CoCu30 first vs last quartile. Below each analysis are shown the corresponding numbers at risk, time
1051 in years. HPV positive HR=5.47 (95%CI 1.76-17.0), p-value=0.0011. HPV negative HR=1.00 (95%CI 0.67-
1052 1.51), p-value=0.98. HR and CI were calculated using Cox regression. p-value was calculated using
1053 logRank test. **E)** Kaplan-Meier overall survival analysis of LUSC TRACERx dataset. Individual tumors
1054 stratified as high-, discordant or low-risk according to expression profile of CoCu30 signature across
1055 multiple regions, as previously described and stratified according to Biswas et al. [55]. Below are shown
1056 the numbers at risk in years. HR=1.98 (95% CI 1.09-3.6), p-value=0.0152. HR and CI calculated using
1057 Cox regression and are referred to CoCu30 low vs CoCu30 high. p-value was calculated using logRank
1058 test.
1059
1060 **Figure S2:** Validation of CoCu8 and CoCu30 gene signatures on different co-culture combinations. **A)**
1061 Table with Normalized Enrichment Score (NES) values of different combinations of cancer cells and
1062 CAFs breast cancer cell lines from Rajaram et al. [18]. Negative values represent enrichment towards co-
1063 culture condition. Colour legend is shown. **B)** Table with NES values of CoCu8 and CoCu30 gene
1064 signatures for cancer cells and endothelial cells mono-culture vs co-culture available from Stine et al.
1065 [19]. Negative values represent enrichment towards co-culture condition. Colour legend is shown.
1066
1067 **Figure S3:** Multivariate analysis of CoCu8 overall survival. Forest plot showing Hazard Ratios, 95%
1068 confidence interval and p value calculated using multivariate Cox regression from patients with HNSCC

1069 (HPV positive and negative), LUSC and CESC from the TCGA cohort. Variables include: age (continuous,
1070 years), sex (male vs female, except for CESC as all patients were female), clinical stage (categorical)
1071 and the CoCu8 signature (continuous variable).
1072
1073 **Figure S4:** Multivariate analysis of CoCu30 overall survival. Forest plot showing Hazard Ratios, 95%
1074 confidence interval and p value calculated using multivariate Cox regression from patients with HNSCC
1075 (HPV positive and negative), LUSC and CESC from the TCGA cohort. Variables include: age (continuous,
1076 years), sex (male vs female, except for CESC as all patients were female), clinical stage (categorical)
1077 and the CoCu30 signature (continuous variable).
1078
1079 **Figure S5:** Fibroblast abundance correlates with CoCu8 / CoCu30 gene signature on different
1080 squamous cell carcinoma datasets. **A)** Box plot analysis of CoCu8 expression in HNSCC, CESC and LUSC
1081 TCGA separated by first and last quartile of fibroblast abundance via Methyl CIBERSORT deconvolution
1082 strategy. Independent t-test. **B)** Box plot analysis of CoCu8 expression in HPV positive and negative
1083 TCGA separated by first and last quartile of fibroblast abundance via Methyl CIBERSORT deconvolution
1084 strategy. Independent t-test. **C)** Box plot analysis of CoCu30 expression in HNSCC, CESC and LUSC TCGA
1085 separated by first and last quartile of fibroblast abundance via Methyl CIBERSORT deconvolution
1086 strategy. Independent t-test. **D)** Box plot analysis of CoCu30 expression in HPV positive and negative
1087 TCGA separated by first and last quartile of fibroblast abundance via Methyl CIBERSORT deconvolution
1088 strategy. Independent t-test. **E-J)** Correlation plot of different fibroblast subpopulations derived from
1089 Galbo et al. [7] with CoCu8 gene signature in LUSC TRACERx **(E)**, HPV positive HNSCC TCGA **(F)**, UK_HPV
1090 positive HNSCC **(G)** patients and with CoCu30 gene signature in LUSC TRACERx **(H)**, HPV positive HNSCC
1091 TCGA **(I)**, UK_HPV positive HNSCC **(J)** patients. The number inside the square represents the R,
1092 Spearman correlation coefficient. The colour legend is shown at the bottom. All correlations are
1093 significant p-value<0.05.
1094
1095 **Figure S6:** Fibroblast abundance meta-analysis on different squamous cell carcinoma datasets. **A)**
1096 Kaplan-Meier overall survival analysis of HNSCC (left), CESC (centre), LUSC (right) TCGA datasets
1097 stratified for fibroblast abundance first vs last quartile. Below each analysis are shown the
1098 corresponding numbers at risk, time in years. HNSCC HR=1.54 (95%CI 1.05-2.25), p-value=0.025. CESC
1099 HR=1.56 (95%CI 0.77-3.17), p-value=0.21. LUSC HR=1.12 (95%CI 0.71-1.77), p-value=0.62. HR and CI
1100 were calculated using Cox regression. p-value was calculated using logRank test. **B)** Kaplan-Meier
1101 overall survival analysis of HNSCC HPV positive (left) and HNSCC HPV positive (right) TCGA datasets
1102 stratified for fibroblast abundance first vs last quartile. Below each analysis are shown the

1103 corresponding numbers at risk, time in years. HPV positive HR=4.76 (95%CI 1.36-16.5), p-
1104 value=0.0065. HPV negative HR=1.45 (95%CI 0.96-2.18), p-value=0.076. HR and CI were calculated
1105 using Cox regression. p-value was calculated using logRank test.
1106
1107 **Figure S7:** Enrichment of CoCu8, CoCu30 and RAS84 gene signatures in different co-culture conditions.
1108 **A)** Gene set enrichment analysis (GSEA) plot of CoCu8 gene signature (top) and CoCu30 (bottom) in
1109 co-culture indirect vs direct condition in A431 / VCAF2b transcriptomic dataset. NES and FDR are
1110 specified below each plot. **B)** Gene set enrichment analysis (GSEA) plot of RAS84 gene signature [24] in
1111 mono-culture and co-culture. NES and FDR are specified below each plot.
1112
1113 **Figure S8:** HB-EGF / RAS / MAPK activity in cancer cells and CAFs mono-culture vs co-culture. **A)**
1114 Correlation plot of CoCu8 and RAS84 expression levels in HNSCC HPV positive UK_HPV positive cohort.
1115 R is Spearman correlation coefficient. **B)** Correlation table of different fibroblast subpopulations
1116 derived from Galbo et al. [7] with RAS84 gene signature in the independent cohort UK_HPV positive of
1117 HNSCC HPV positive patients. The number inside the square represents the Spearman R, correlation
1118 coefficient. The colour legend is shown on the top. All correlations are significant at p-value<0.001.
1119 n=97. **C)** Immunofluorescence staining of E-Cadherin and Hoechst 33342 for SCC154 mono-culture for
1120 the indicated treatments after 48h. **D)** qPCR analysis of *ACTA2*, *FAP*, *FN1* and *LRRC15* genes in OCAF1
1121 mono-culture for the indicated treatments after 48h. mRNA expression is reported as
1122 mean ± standard deviation (SD) fold change difference over non-treated (NT) condition. Genes have
1123 been normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 3 independent
1124 experiments. **E)** qPCR analysis of *HBEGF* gene in SCC154 after 96h of treatment with the indicated
1125 conditions. mRNA expression is reported as mean ± standard deviation (SD) fold change difference
1126 over siC- condition. Gene has been normalized over the average of *GAPDH*, *ACTB* and *RPLP0*
1127 housekeeping genes. n ≥ 4 independent experiments. Paired t-test. **F)** Violin plot of mRNA expression
1128 levels of the indicated genes for each cluster from Choi et al. [29].
1129
1130 **Figure S9:** Enrichment of monocyte/macrophage lineage in patients with abundance of cancer cells –
1131 CAFs cross-talk. **A)** Correlation plot of *HBEGF* and *LIF* expression levels in HNSCC HPV positive TCGA
1132 dataset. R is Spearman correlation coefficient. **B)** Box plot analysis of immune cell absolute score via
1133 Methyl CIBERSORT deconvolution strategy in HNSCC HPV positive separated by first and last quartile
1134 of fibroblast abundance. Independent t-test. Bonferroni correction for multiple comparisons. **C)** Box
1135 plot analysis of immune cell absolute score via Methyl CIBERSORT deconvolution strategy in HNSCC

1136 HPV positive separated by RAS84_0 and RAS84_max. t-test Bonferroni correction for multiple
1137 comparisons.
1138
1139 **Figure S10:** Schematic representation of the current understanding of cancer cells – CAFs co-culture.
1140 Graphics created with BioRender.com.

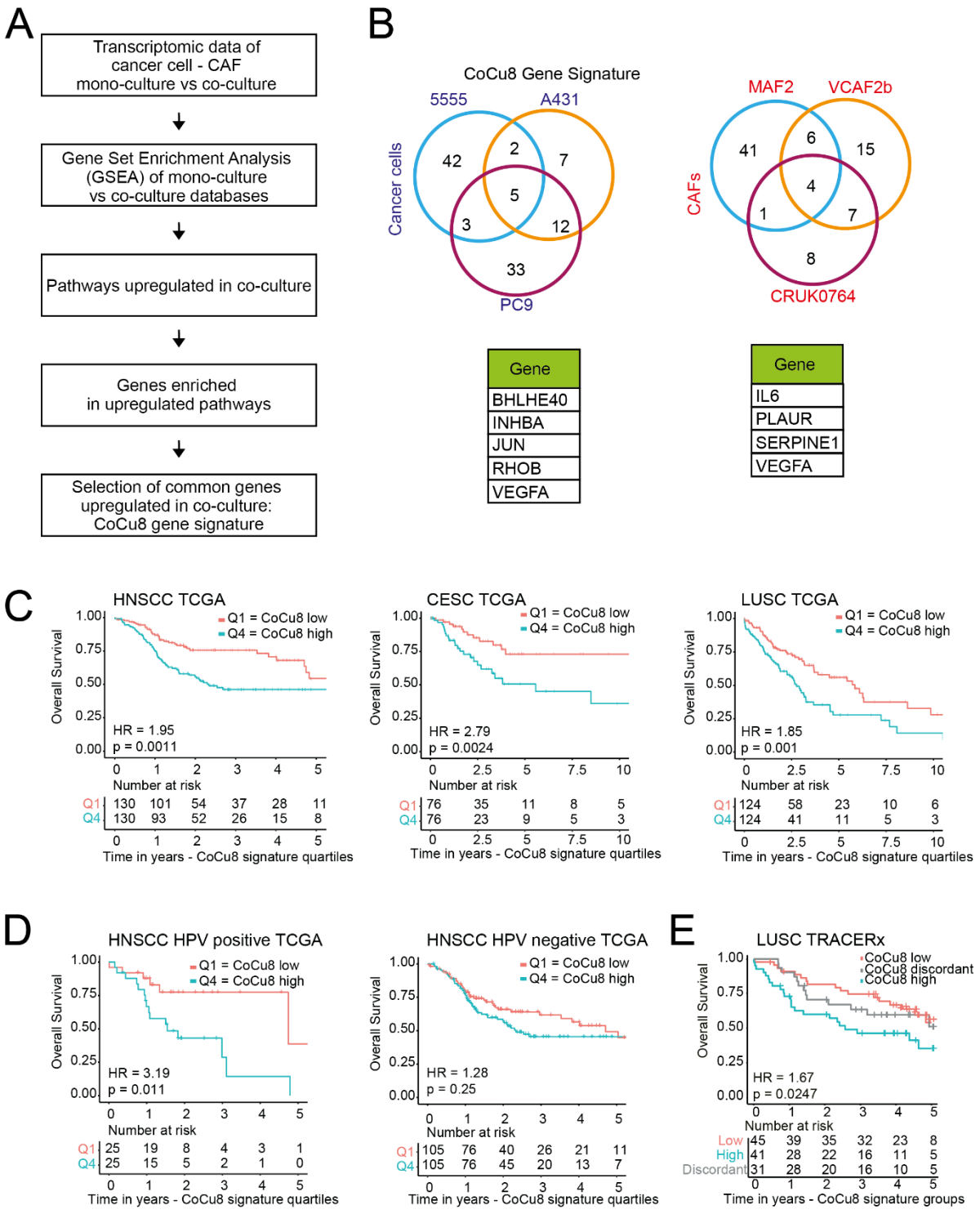
A
Transcriptomic data of cancer cell - CAF mono-culture vs co-culture
Gene Set Enrichment Analysis (GSEA) of mono-culture vs co-culture databases
Pathways upregulated in co-culture
Genes enriched in upregulated pathways
Selection of common genes upregulated in co-culture: CoCu8 gene signature
B
CoCu8 Gene Signature
5555
A431
PC9
Cancer cells
42
2
7
5
3
12
33
MAF2
VCAF2b
CRUK0764
CAFs
41
6
15
4
1
7
8
Gene
BHLHE40
INHBA
JUN
RHOB
VEGFA
Gene
IL6
PLAUR
SERPINE1
VEGFA
C
HNSCC TCGA
Q1 = CoCu8 low
Q4 = CoCu8 high
HR = 1.95
p = 0.0011
Overall Survival
Number at risk
Q1 130 101 54 37 28 11
Q4 130 93 52 26 15 8
Time in years - CoCu8 signature quartiles
CESC TCGA
HR = 2.79
p = 0.0024
Q1 76 35 11 8 5
Q4 76 23 9 5 3
LUSC TCGA
HR = 1.85
p = 0.001
Q1 124 58 23 10 6
Q4 124 41 11 5 3
D
HNSCC HPV positive TCGA
HR = 3.19
p = 0.011
Q1 25 19 8 4 3 1
Q4 25 15 5 2 1 0
HNSCC HPV negative TCGA
HR = 1.28
p = 0.25
Q1 105 76 40 26 21 11
Q4 105 76 45 20 13 7
E
LUSC TRACERx
CoCu8 low
CoCu8 discordant
CoCu8 high
HR = 1.67
p = 0.0247
Low 45 39 35 32 23 8
High 41 28 22 16 11 5
Discordant 31 28 20 16 10 5
Time in years - CoCu8 signature groups

Figure 1

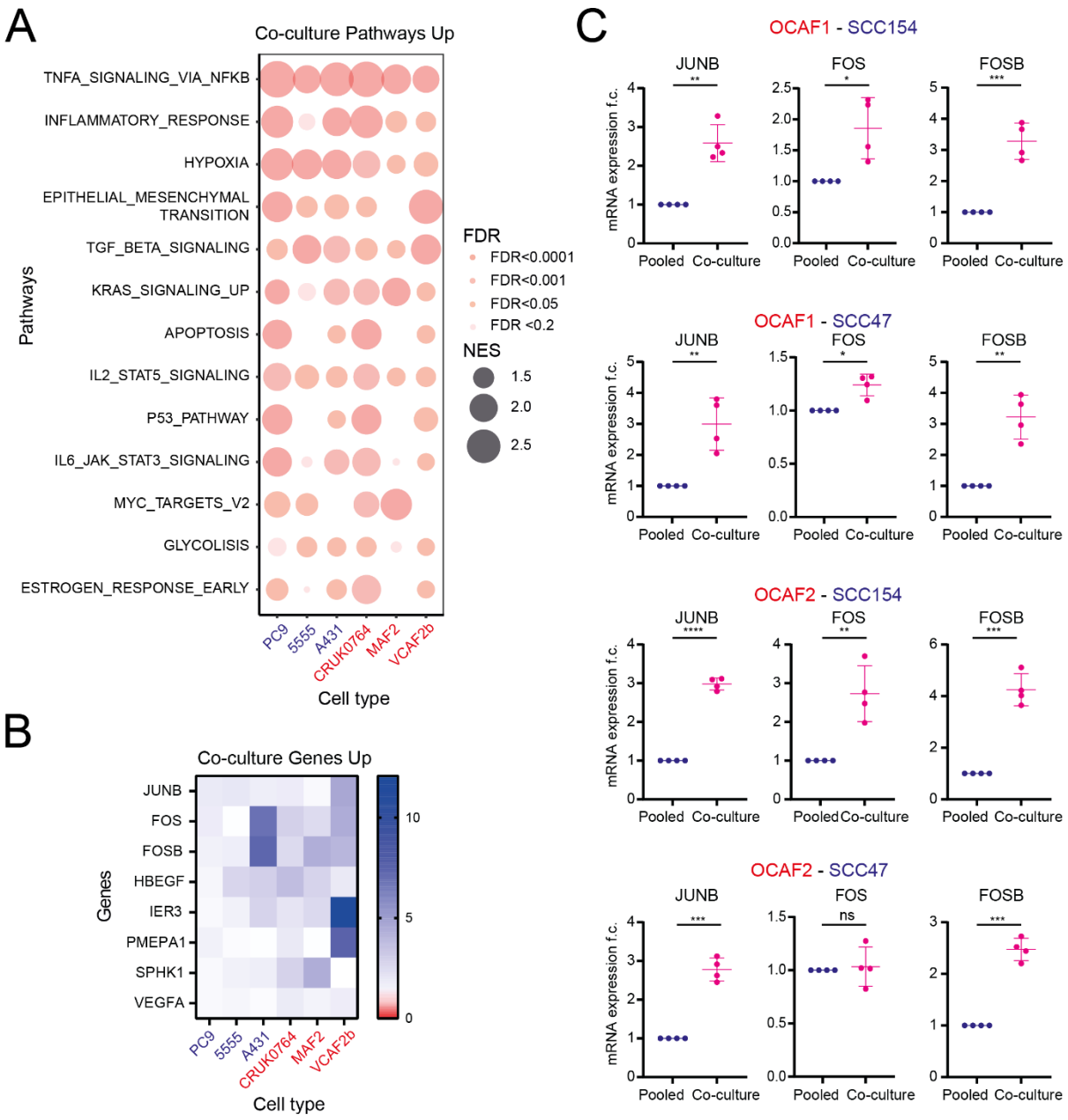
A
Co-culture Pathways Up
TNFA_SIGNALING_VIA_NFKB
INFLAMMATORY_RESPONSE
HYPOXIA
EPITHELIAL_MESENCHYMAL TRANSITION
TGF_BETA_SIGNALING
KRAS_SIGNALING_UP
APOPTOSIS
IL2_STAT5_SIGNALING
P53_PATHWAY
IL6_JAK_STAT3_SIGNALING
MYC_TARGETS_V2
GLYCOLISIS
ESTROGEN_RESPONSE_EARLY
Pathways
PC9
5555
A431
CRUK0764
MAF2
VCAF2b
Cell type
FDR
FDR<0.0001
FDR<0.001
FDR<0.05
FDR <0.2
NES
1.5
2.0
2.5
B
Co-culture Genes Up
JUNB
FOS
FOSB
HBEGF
IER3
PMEPA1
SPHK1
VEGFA
Genes
10
5
0
PC9
5555
A431
CRUK0764
MAF2
VCAF2b
Cell type
C
OCAF1 - SCC154
JUNB
FOS
FOSB
mRNA expression f.c.
Pooled
Co-culture
OCAF1 - SCC47
OCAF2 - SCC154
OCAF2 - SCC47
ns

Figure 2

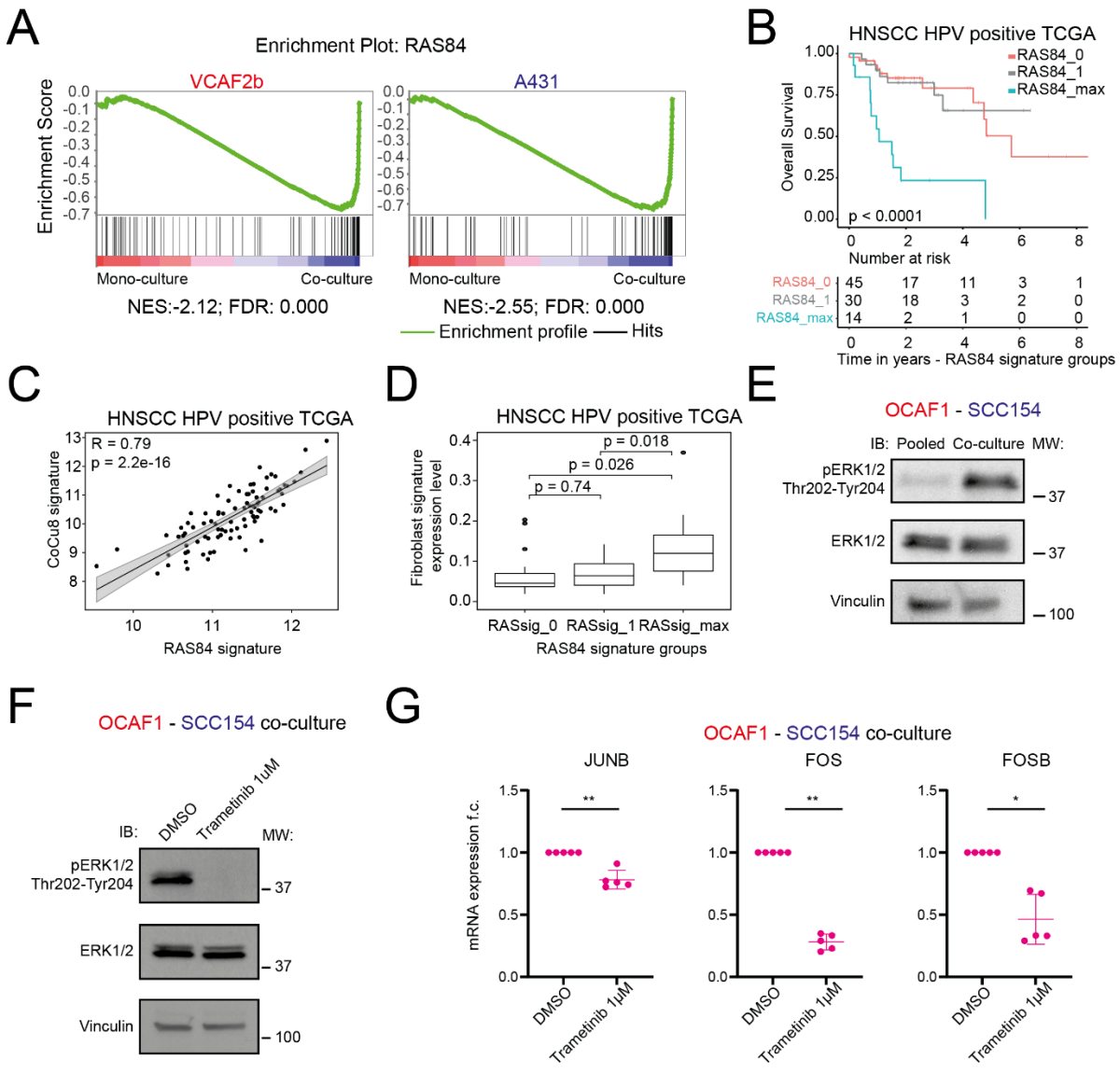
A
Enrichment Plot: RAS84
VCAF2b
A431
Enrichment Score
Mono-culture
Co-culture
NES:-2.12; FDR: 0.000
NES:-2.55; FDR: 0.000
Enrichment profile
Hits
B
HNSCC HPV positive TCGA
RAS84_0
RAS84_1
RAS84_max
Overall Survival
p < 0.0001
Number at risk
Time in years - RAS84 signature groups
C
HNSCC HPV positive TCGA
R = 0.79
p = 2.2e-16
CoCu8 signature
RAS84 signature
D
HNSCC HPV positive TCGA
p = 0.018
p = 0.026
p = 0.74
Fibroblast signature expression level
RASsig_0
RASsig_1
RASsig_max
RAS84 signature groups
E
OCAF1 - SCC154
IB: Pooled Co-culture MW:
pERK1/2 Thr202-Tyr204
ERK1/2
Vinculin
F
OCAF1 - SCC154 co-culture
IB:
DMSO
Trametinib 1uM
MW:
pERK1/2 Thr202-Tyr204
ERK1/2
Vinculin
G
OCAF1 - SCC154 co-culture
JUNB
FOS
FOSB
mRNA expression f.c.
DMSO
Trametinib 1μM

Figure 3

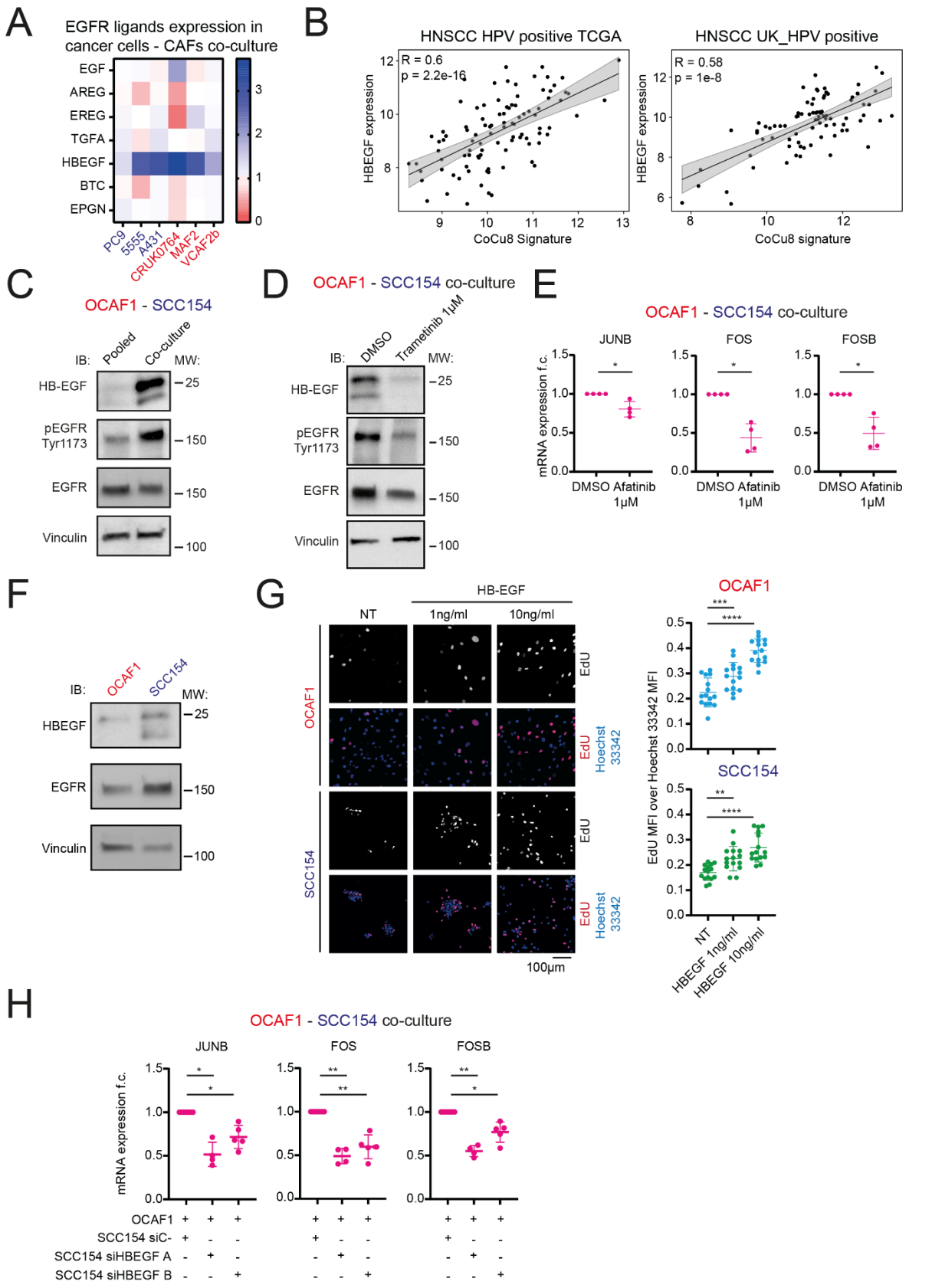
A
EGFR ligands expression in cancer cells - CAFs co-culture
EGF
AREG
EREG
TGFA
HBEGF
BTC
EPGN
PC9
5555
A431
CRUK0764
MAF2
VCAF2b
B
HNSCC HPV positive TCGA
R = 0.6
p = 2.2e-16
HBEGF expression
CoCu8 Signature
HNSCC UK_HPV positive
R = 0.58
p = 1e-8
CoCu8 signature
C
OCAF1 - SCC154
IB:
Pooled
Co-culture
MW:
HB-EGF
pEGFR Tyr1173
EGFR
Vinculin
D
OCAF1 - SCC154 co-culture
DMSO
Trametinib 1μM
E
OCAF1 - SCC154 co-culture
JUNB
FOS
FOSB
mRNA expression f.c.
DMSO
Afatinib 1μM
F
OCAF1
SCC154
HBEGF
G
HB-EGF
NT
1ng/ml
10ng/ml
EdU
Hoechst 33342
100μm
EdU MFI over Hoechst 33342 MFI
HBEGF 1ng/ml
HBEGF 10ng/ml
H
OCAF1 - SCC154 co-culture
SCC154 siC-
SCC154 siHBEGF A
SCC154 siHBEGF B

Figure 4

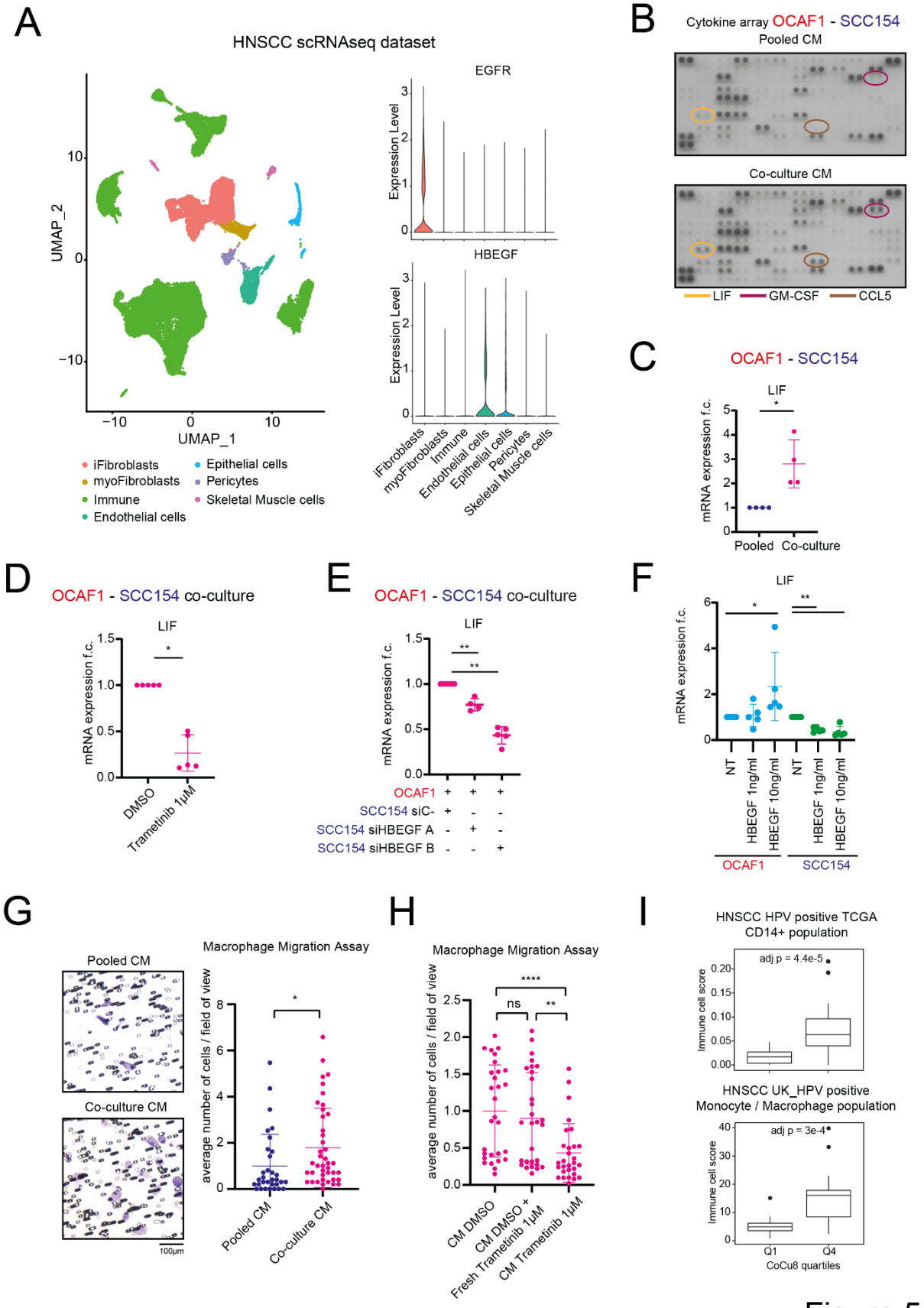
A
HNSCC scRNAseq dataset
UMAP_2
UMAP_1
iFibroblasts
myoFibroblasts
Immune
Endothelial cells
Epithelial cells
Pericytes
Skeletal Muscle cells
EGFR
HBEGF
Expression Level
B
Cytokine array OCAF1 - SCC154
Pooled CM
Co-culture CM
LIF
GM-CSF
CCL5
C
OCAF1 - SCC154
LIF
mRNA expression f.c.
Pooled
Co-culture
D
OCAF1 - SCC154 co-culture
LIF
mRNA expression f.c.
DMSO
Trametinib 1µM
E
OCAF1 - SCC154 co-culture
LIF
mRNA expression f.c.
OCAF1
SCC154 siC-
SCC154 siHBEGF A
SCC154 siHBEGF B
F
LIF
mRNA expression f.c.
NT
HBEGF 1ng/ml
HBEGF 10ng/ml
OCAF1
SCC154
G
Pooled CM
Co-culture CM
100µm
Macrophage Migration Assay
average number of cells / field of view
Pooled CM
Co-culture CM
H
Macrophage Migration Assay
average number of cells / field of view
ns
CM DMSO
CM DMSO + Fresh Trametinib 1µM
CM Trametinib 1µM
I
HNSCC HPV positive TCGA
CD14+ population
adj p = 4.4e-5
Immune cell score
HNSCC UK_HPV positive
Monocyte / Macrophage population
adj p = 3e-4
Immune cell score
Q1
Q4
CoCu8 quartiles

Figure 5

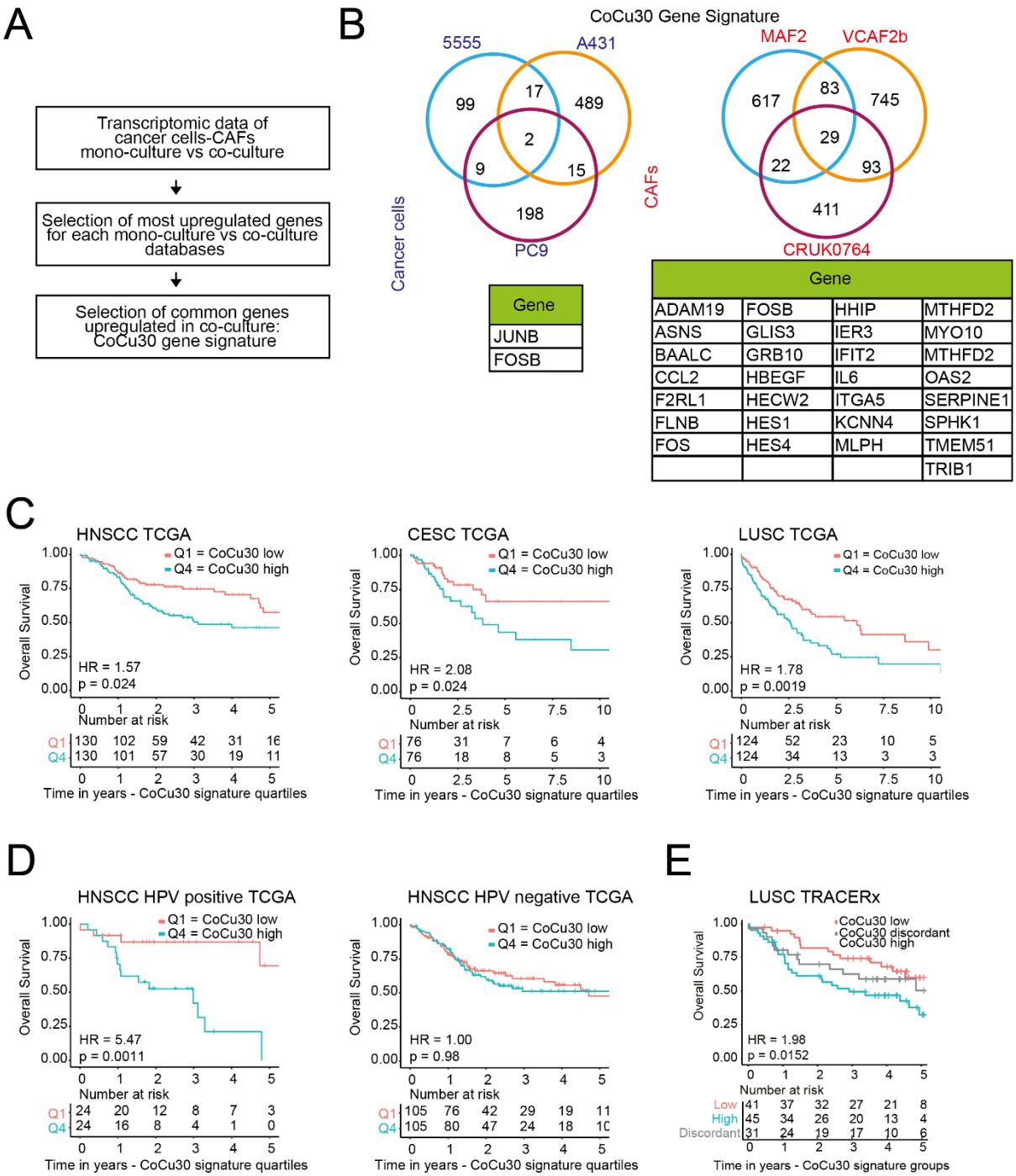
A
Transcriptomic data of cancer cells-CAFs mono-culture vs co-culture
Selection of most upregulated genes for each mono-culture vs co-culture databases
Selection of common genes upregulated in co-culture: CoCu30 gene signature
B
CoCu30 Gene Signature
5555
A431
PC9
Cancer cells
99
17
489
2
9
15
198
MAF2
VCAF2b
CRUK0764
CAFs
617
83
745
29
22
93
411
Gene
JUNB
FOSB
Gene
ADAM19 FOSB HHIP MTHFD2
ASNS GLIS3 IER3 MYO10
BAALC GRB10 IFIT2 MTHFD2
CCL2 HBEGF IL6 OAS2
F2RL1 HECW2 ITGA5 SERPINE1
FLNB HES1 KCNN4 SPHK1
FOS HES4 MLPH TMEM51
TRIB1
C
HNSCC TCGA
Q1 = CoCu30 low
Q4 = CoCu30 high
Overall Survival
HR = 1.57
p = 0.024
Number at risk
Q1 130 102 59 42 31 16
Q4 130 101 57 30 19 11
Time in years - CoCu30 signature quartiles
CESC TCGA
HR = 2.08
p = 0.024
Q1 76 31 7 6 4
Q4 76 18 8 5 3
LUSC TCGA
HR = 1.78
p = 0.0019
Q1 124 52 23 10 5
Q4 124 34 13 3 3
D
HNSCC HPV positive TCGA
HR = 5.47
p = 0.0011
Q1 24 20 12 8 7 3
Q4 24 16 8 4 1 0
HNSCC HPV negative TCGA
HR = 1.00
p = 0.98
Q1 105 76 42 29 19 11
Q4 105 80 47 24 18 10
E
LUSC TRACERx
CoCu30 low
CoCu30 discordant
CoCu30 high
HR = 1.98
p = 0.0152
Low 41 37 32 27 21 8
High 45 34 26 20 13 4
Discordant 31 24 19 17 10 6
Time in years - CoCu30 signature groups

Figure S1

A

Cancer cell / CAF cell co-culture

NES	CoCu8	CoCu30
Wi38 with MDA-MB-231	-1.99	-2.30
CDD1112Sk with MDA-MB-231	-2.00	-2.12
HFF1 with MDA-MB-231	-2.05	-2.41
HFF2 with MDA-MB-231	-1.88	-2.15
MDA-MB-231 with Wi38	-1.91	-2.39
MDA-MB-231 with CDD1112Sk	-1.26	-1.48
MDA-MB-231 with HFF1	-1.89	-2.22
MDA-MB-231 with HFF2	-1.85	-2.07
Cal51 with Wi38	-1.92	-2.52
Cal51 with CDD1112Sk	-1.36	-1.76
Cal51 with HFF1	-1.97	-2.28
Cal51 with HFF2	-1.79	-1.82
Wi38 with Cal51	-1.33	-1.16
CDD1112Sk with Cal51	-1.56	-1.61
HFF1 with Cal51	-1.68	-1.97
HFF2 with Cal51	-1.69	-1.96

FDR<0.01
FDR<0.05
FDR n.s.

B

Cancer cell / endothelial cell co-culture

NES	CoCu8	CoCu30
HUVEC with 1205Lu	+0.73	-1.46
1205Lu with HUVEC	-1.12	-1.64

FDR<0.01
FDR<0.05
FDR n.s.

Figure S2

CoCu8 Overall Survival

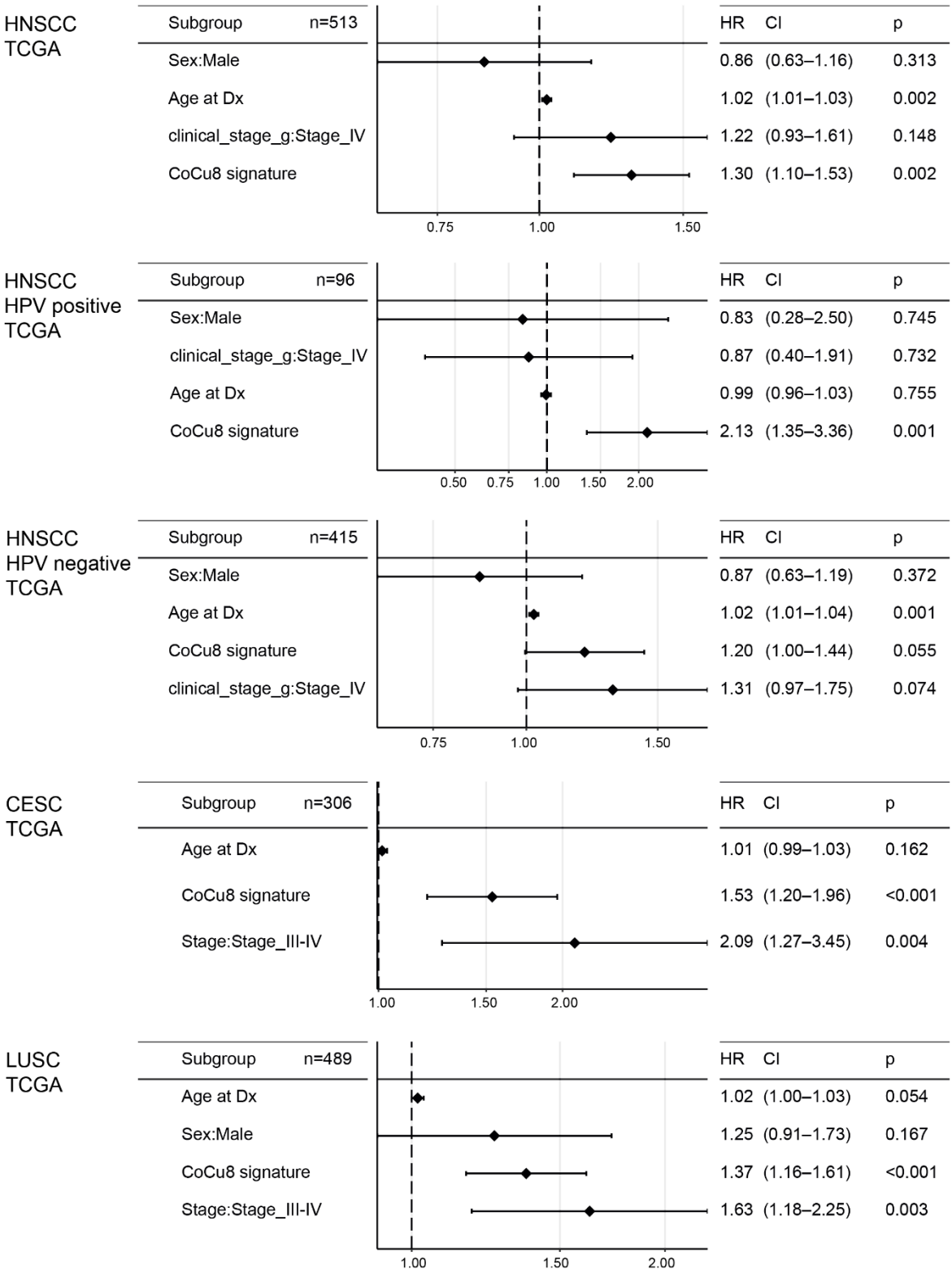

Figure S3

CoCu30 Overall Survival

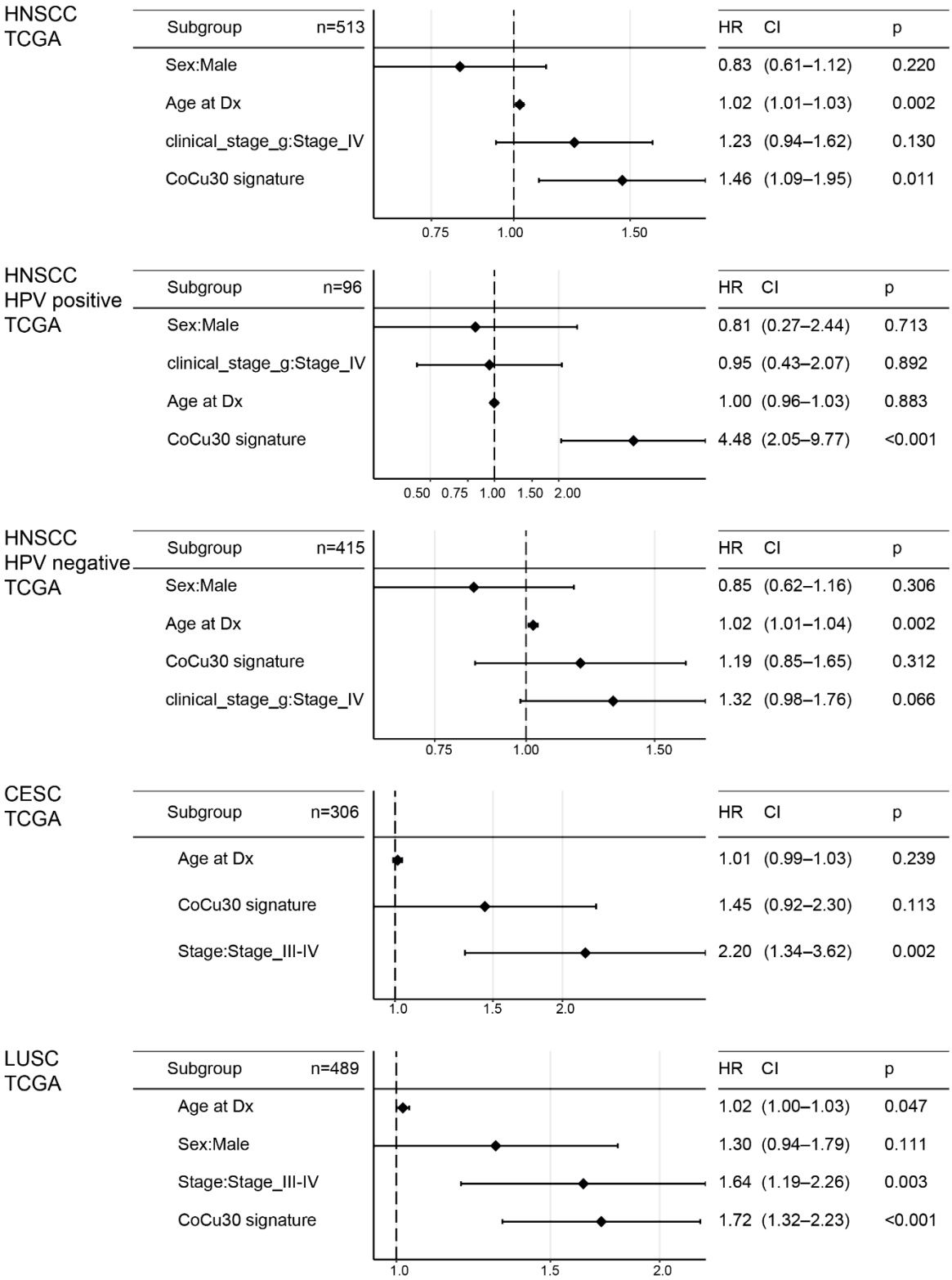

Figure S4

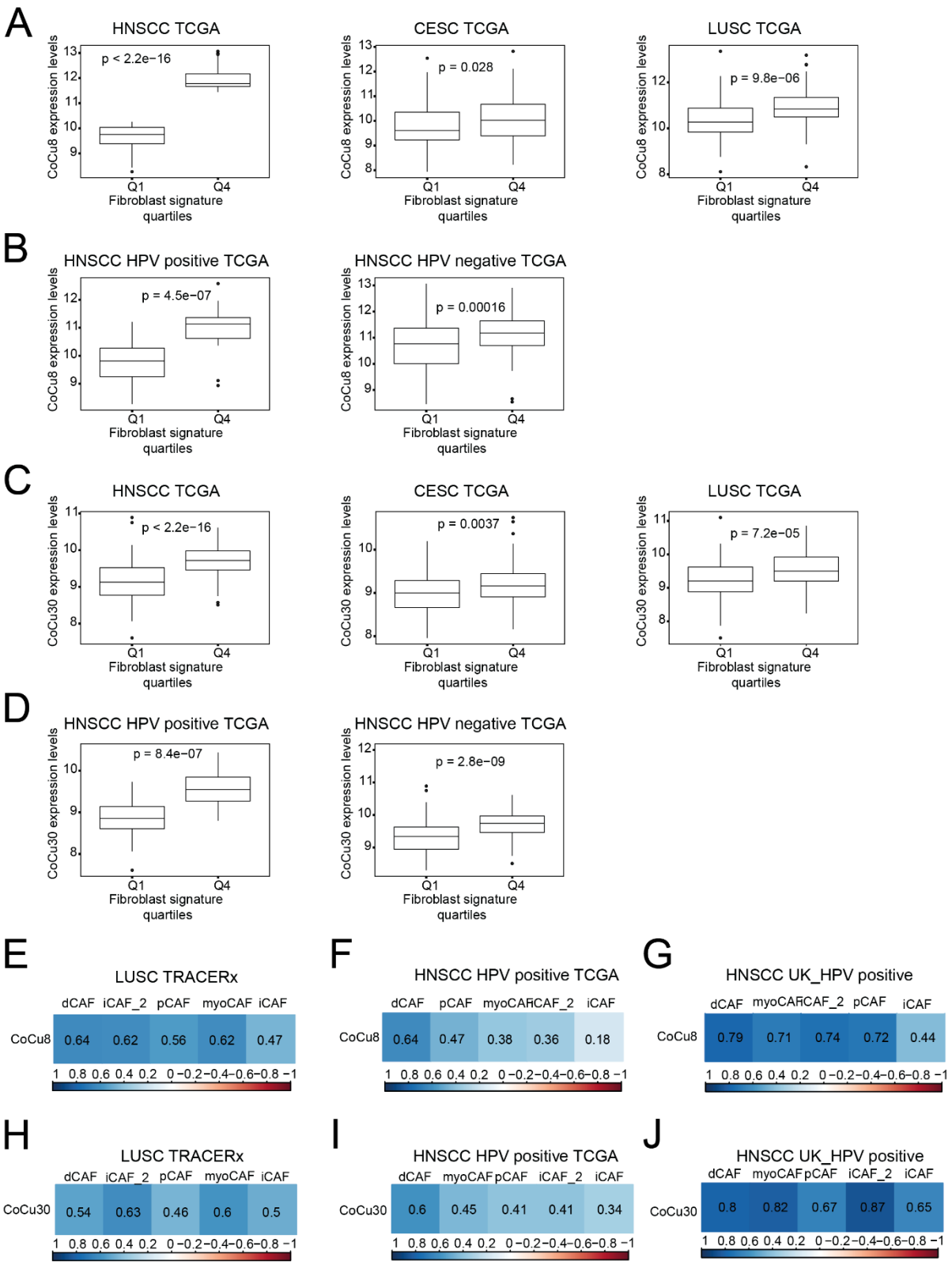
A
HNSCC TCGA
p < 2.2e−16
CESC TCGA
p = 0.028
LUSC TCGA
p = 9.8e−06
CoCu8 expression levels
Q1
Q4
Fibroblast signature quartiles
B
HNSCC HPV positive TCGA
p = 4.5e−07
HNSCC HPV negative TCGA
p = 0.00016
C
HNSCC TCGA
p < 2.2e−16
CESC TCGA
p = 0.0037
LUSC TCGA
p = 7.2e−05
CoCu30 expression levels
D
HNSCC HPV positive TCGA
p = 8.4e−07
HNSCC HPV negative TCGA
p = 2.8e−09
E
LUSC TRACERx
dCAF iCAF_2 pCAF myoCAF iCAF
CoCu8 0.64 0.62 0.56 0.62 0.47
F
HNSCC HPV positive TCGA
dCAF pCAF myoCAF iCAF_2 iCAF
CoCu8 0.64 0.47 0.38 0.36 0.18
G
HNSCC UK_HPV positive
dCAF myoCAF iCAF_2 pCAF iCAF
CoCu8 0.79 0.71 0.74 0.72 0.44
H
LUSC TRACERx
dCAF iCAF_2 pCAF myoCAF iCAF
CoCu30 0.54 0.63 0.46 0.6 0.5
I
HNSCC HPV positive TCGA
dCAF myoCAF pCAF iCAF_2 iCAF
CoCu30 0.6 0.45 0.41 0.41 0.34
J
HNSCC UK_HPV positive
dCAF myoCAF pCAF iCAF_2 iCAF
CoCu30 0.8 0.82 0.67 0.87 0.65
1 0.8 0.6 0.4 0.2 0 −0.2 −0.4 −0.6 −0.8 −1

Figure S5

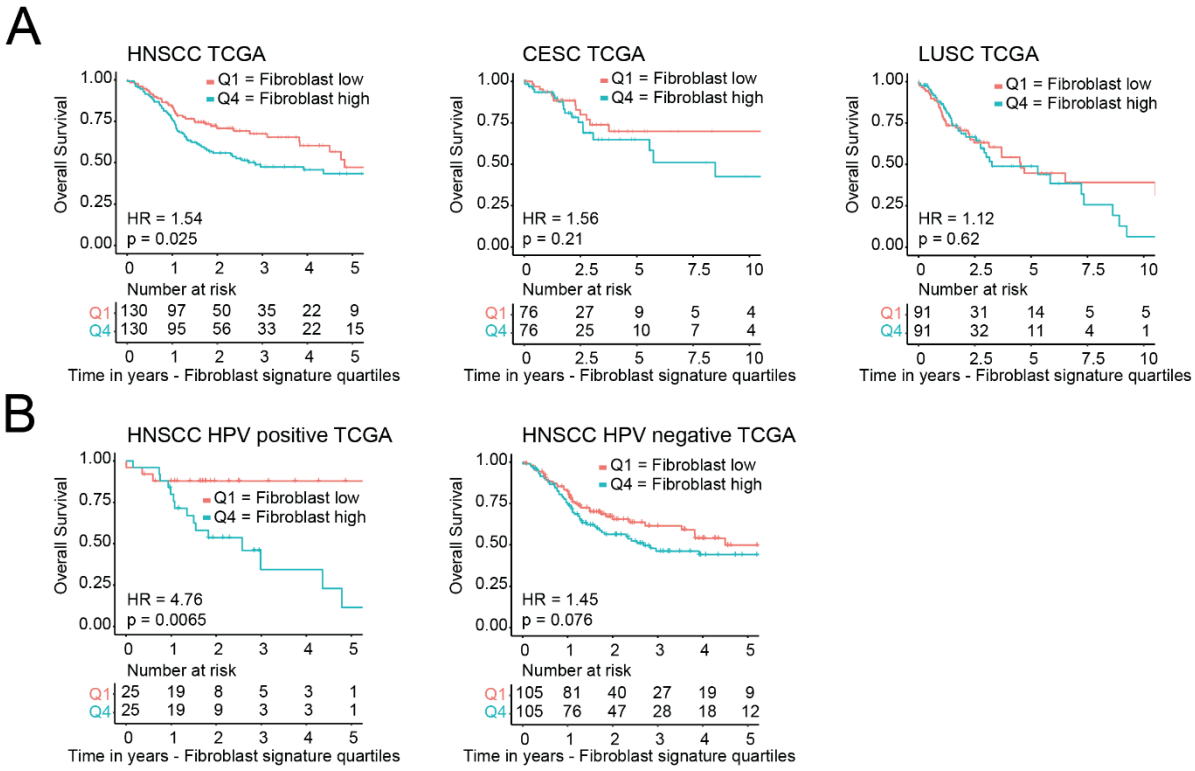
A
HNSCC TCGA
Q1 = Fibroblast low
Q4 = Fibroblast high
Overall Survival
1.00
0.75
0.50
0.25
0.00
HR = 1.54
p = 0.025
0 1 2 3 4 5
Number at risk
Q1 130 97 50 35 22 9
Q4 130 95 56 33 22 15
0 1 2 3 4 5
Time in years - Fibroblast signature quartiles
CESC TCGA
Q1 = Fibroblast low
Q4 = Fibroblast high
Overall Survival
1.00
0.75
0.50
0.25
0.00
HR = 1.56
p = 0.21
0 2.5 5 7.5 10
Number at risk
Q1 76 27 9 5 4
Q4 76 25 10 7 4
0 2.5 5 7.5 10
Time in years - Fibroblast signature quartiles
LUSC TCGA
Q1 = Fibroblast low
Q4 = Fibroblast high
Overall Survival
1.00
0.75
0.50
0.25
0.00
HR = 1.12
p = 0.62
0 2.5 5 7.5 10
Number at risk
Q1 91 31 14 5 5
Q4 91 32 11 4 1
0 2.5 5 7.5 10
Time in years - Fibroblast signature quartiles
B
HNSCC HPV positive TCGA
Q1 = Fibroblast low
Q4 = Fibroblast high
Overall Survival
1.00
0.75
0.50
0.25
0.00
HR = 4.76
p = 0.0065
0 1 2 3 4 5
Number at risk
Q1 25 19 8 5 3 1
Q4 25 19 9 3 3 1
0 1 2 3 4 5
Time in years - Fibroblast signature quartiles
HNSCC HPV negative TCGA
Q1 = Fibroblast low
Q4 = Fibroblast high
Overall Survival
1.00
0.75
0.50
0.25
0.00
HR = 1.45
p = 0.076
0 1 2 3 4 5
Number at risk
Q1 105 81 40 27 19 9
Q4 105 76 47 28 18 12
0 1 2 3 4 5
Time in years - Fibroblast signature quartiles

Figure S6

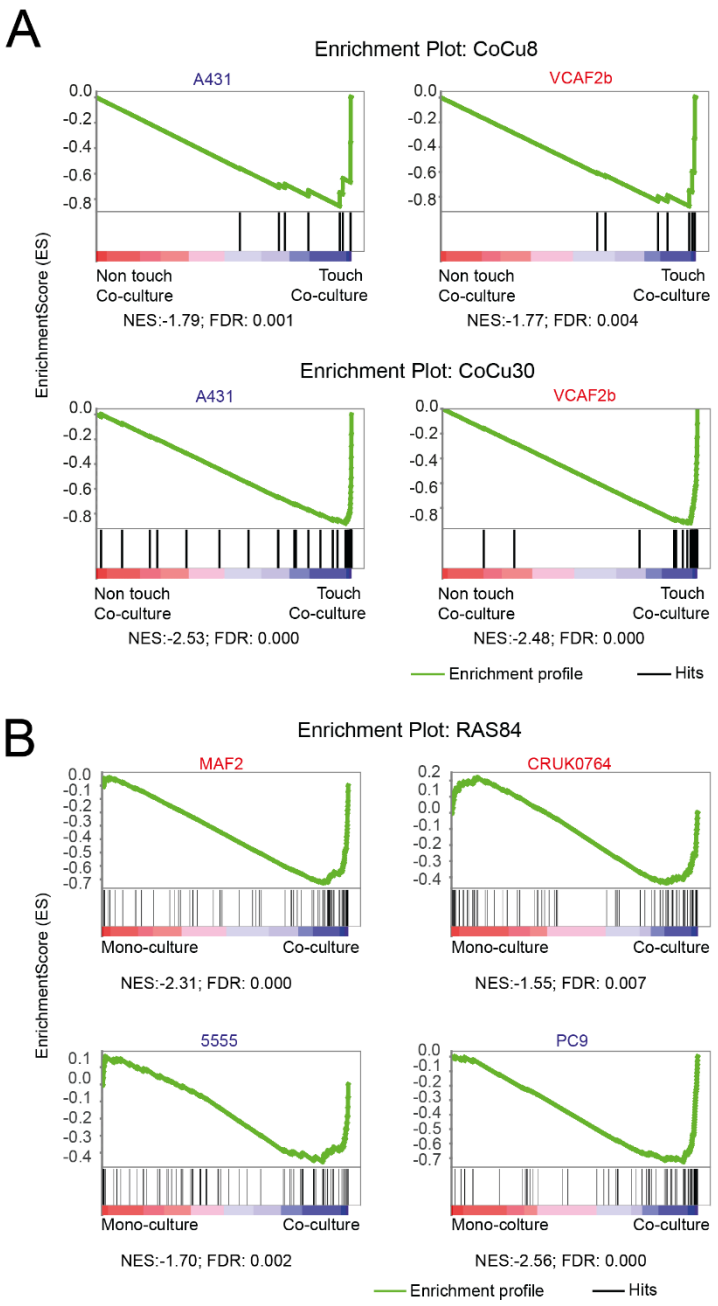
A
Enrichment Plot: CoCu8
A431
VCAF2b
EnrichmentScore (ES)
Non touch Co-culture
Touch Co-culture
NES:-1.79; FDR: 0.001
NES:-1.77; FDR: 0.004
Enrichment Plot: CoCu30
A431
VCAF2b
NES:-2.53; FDR: 0.000
NES:-2.48; FDR: 0.000
Enrichment profile
Hits
B
Enrichment Plot: RAS84
MAF2
CRUK0764
Mono-culture
Co-culture
NES:-2.31; FDR: 0.000
NES:-1.55; FDR: 0.007
5555
PC9
Mono-colture
NES:-1.70; FDR: 0.002
NES:-2.56; FDR: 0.000
Enrichment profile
Hits

Figure S7

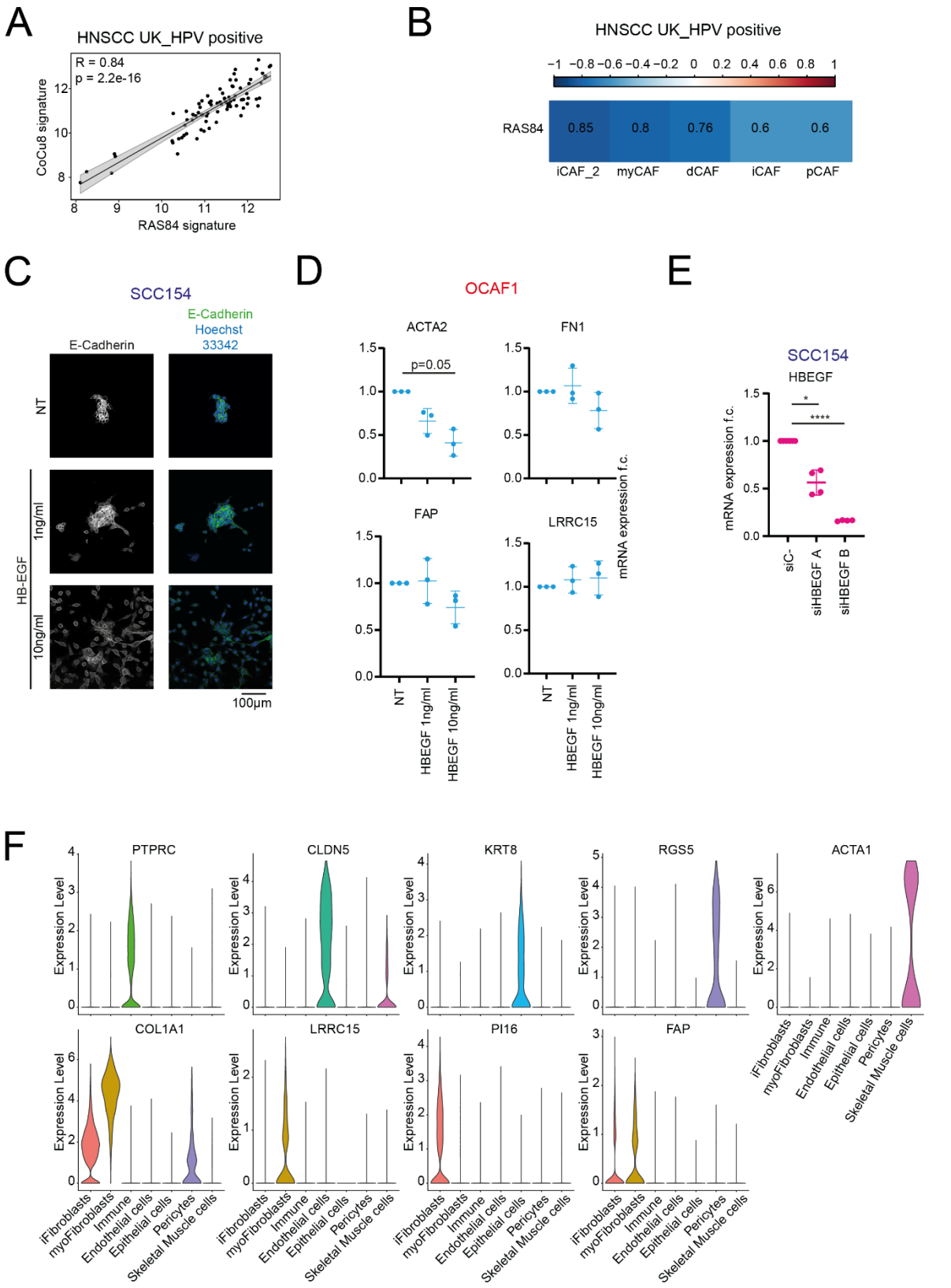
A
HNSCC UK_HPV positive
R = 0.84
p = 2.2e-16
CoCu8 signature
RAS84 signature
B
HNSCC UK_HPV positive
RAS84
0.85
0.8
0.76
0.6
0.6
iCAF_2
myCAF
dCAF
iCAF
pCAF
C
SCC154
E-Cadherin
E-Cadherin
Hoechst
33342
NT
HB-EGF
1ng/ml
10ng/ml
100µm
D
OCAF1
ACTA2
p=0.05
FN1
FAP
LRRC15
mRNA expression f.c.
NT
HBEGF 1ng/ml
HBEGF 10ng/ml
E
SCC154
HBEGF
mRNA expression f.c.
siC-
siHBEGF A
siHBEGF B
F
PTPRC
CLDN5
KRT8
RGS5
ACTA1
COL1A1
LRRC15
PI16
FAP
Expression Level
iFibroblasts
myoFibroblasts
Immune
Endothelial cells
Epithelial cells
Pericytes
Skeletal Muscle cells

Figure S8

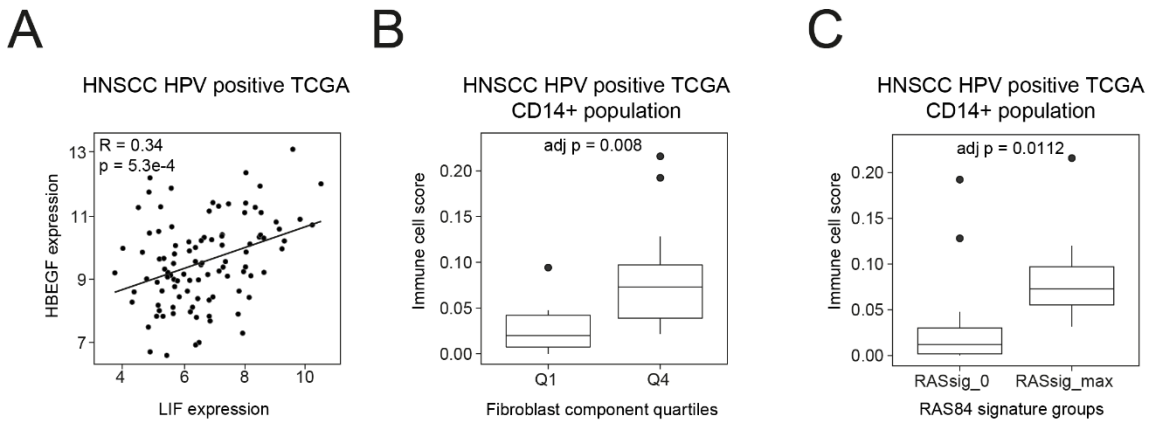
A
HNSCC HPV positive TCGA
R = 0.34
p = 5.3e-4
HBEGF expression
LIF expression
B
HNSCC HPV positive TCGA
CD14+ population
adj p = 0.008
Immune cell score
Q1
Q4
Fibroblast component quartiles
C
HNSCC HPV positive TCGA
CD14+ population
adj p = 0.0112
Immune cell score
RASsig_0
RASsig_max
RAS84 signature groups

Figure S9

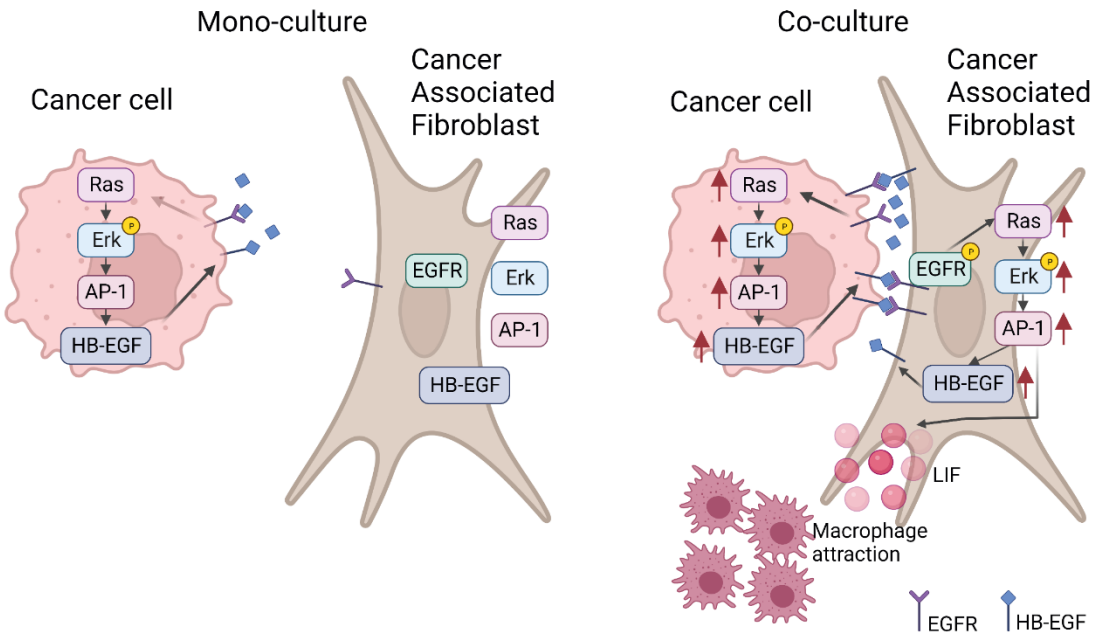
Mono-culture
Cancer cell
Cancer Associated Fibroblast
Ras
Erk
AP-1
HB-EGF
EGFR
Co-culture
LIF
Macrophage attraction
EGFR
HB-EGF

Figure S10