Cancer cell – fibroblast crosstalk via HB-EGF/EGFR/MEK signaling promotes macrophage recruitment in squamous cell carcinoma

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. Tumour 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 derive gene expression signatures indicative of signaling between stromal fibroblasts and cancer cells, and demonstrate their prognostic significance in multiple and independent squamous cell carcinoma cohorts. By leveraging information within the signatures, we discover that the HB-EGF/EGFR/MEK axis represents a hub of tumor – stroma crosstalk, promoting the expression of CSF2 and LIF and favoring the recruitment of macrophages. Together these analyses demonstrate the utility of our approach for interrogating the extent and consequences of TME crosstalk.

Keywords

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

46 ## Abbreviations

47 Activator protein 1 (AP-1)
48 Cancer associated fibroblasts (CAF)
49 Cervical squamous cell carcinoma (CESC)
50 Clear cell Renal Cell Carcinoma (ccRCC)
51 Chemokine (C-C motif) ligand (CCL)
52 Confidence Interval (CI)
53 Conditioned medium (CM)
54 False discovery rate (FDR)
55 Fibroblast growth factor (FGF)
56 Gene-set enrichment analysis (GSEA)
57 Hazard ratio (HR)
58 Head and neck squamous cell carcinoma (HNSCC)
59 Heparin-binding epidermal growth factor-like growth factor (HB-EGF)
60 Human papillomavirus (HPV)
61 Interleukin (IL)
62 Lung squamous cell carcinoma (LUSC)
63 Overall Survival (OS)
64 Mean fluorescent intensity (MFI)
65 Normalized enrichment score (NES)
66 Non-treated (NT)
67 Pancreatic ductal adenocarcinoma (PDAC)
68 Peripheral blood mononuclear cell (PBMC)
69 Platelet derived growth factor (PDGF)
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 Transforming Growth Factor (TGF)
77 Tumor microenvironment (TME)
78 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 through multiple signaling pathways and feedback mechanisms such as growth factors (TGFβ, PDGF, FGF), contact molecules (Notch, Ephrins), and inflammatory molecules (IL1, IL6, CCL12/CXCR4) [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]. Crosstalk between cancer cells and CAFs have been demonstrated via in multiple tumors via, like oncogenic KRAS in colorectal cancer [3], EGFR in pancreatic ductal adenocarcinoma (PDAC) [4]. In addition, CAFs are correlated with a pro-tumorigenic immune landscape, including higher number of tumor-promoting myeloid cells [5], lower numbers of tumor-infiltrating lymphocytes [6] and worse prognosis [7] [8]. 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) [8–11]. Together, these different SCC account for over 800,000 deaths per year, highlighting the need for better understanding of the disease, new therapeutic strategies, and improved tools for clinical decision making[12].

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 [9,13–16] . 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 signaling events [17,18] 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 signaling between cancer cells and CAFs. This signature is associated with worse overall survival in multiple types of SCC, pancreatic cancer, and kidney cancer. 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

112 cytokines through cross-talk with CAFs. In turn, we demonstrate that this upregulation can drive
113 attraction of macrophages, ultimately linked to worse overall survival in SCC patients (Figure S1A).

114 ## Results

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

117 To identify functionally and clinically relevant gene signatures based on cancer cell – CAF cross-talk,
118 we performed a meta-analysis of transcriptomic datasets that compare co-cultures and mono-cultures
119 of cancer cells and CAFs. The datasets were generated under similar direct co-culture conditions using
120 cells derived from different cancer types [19,20]. We applied two strategies to derive gene signatures
121 indicative of upregulated cancer cell – CAF signaling: i) selection of the most significantly enriched
122 pathways via gene set enrichment analysis (GSEA) in co-culture for each transcriptomic dataset,
123 followed by the selection of the up-regulated genes most frequently present in each enriched pathway
124 (Figure 1A); ii) selection of the most up-regulated genes in co-culture for each transcriptomic dataset
125 (Figure S1B). Using the first approach, we obtained a list of 5 genes upregulated in cancer cells and of
126 4 genes upregulated in CAFs upon co-culture, with one present in both. Therefore, this gene signature
127 comprised of 8 genes (named CoCu8) (Figure 1B). The second approach led to a list of 2 genes
128 upregulated in cancer cells and of 29 genes upregulated in CAFs upon co-culture, with one gene in
129 common. Therefore, this gene signature consisted of 30 genes (named CoCu30) (Figure S1C).
130
131 We tested CoCu8 and CoCu30 on publicly available dataset of breast cancer – fibroblast co-cultures[21]
132 confirming their relevance (Figure S1D). We also tested whether CoCu8 and CoCu30 are also
133 upregulated when cancer cells are co-cultured with other stromal cell types and for this reason, we
134 analyzed a dataset of co-culture between 1205Lu cancer cells and HUVEC endothelial cells[22]: CoCu8 is
135 neither enriched in cancer cells nor in endothelial cells when co-cultured, while CoCu30 shows only a
136 weak correlation with co-culture conditions both in 1205Lu and HUVEC cells (Figure S1E). Thus, we
137 establish new gene signatures specifically indicative of cancer cell - CAF communication.
138
139 We next sought to determine the clinical relevance of these signatures by testing the effect on patient
140 survival for frequent cancer types in The Cancer Genome Atlas (TCGA). We found that high expression
141 (Q4 vs Q1) of both gene signatures correlated with worse overall survival (OS) in most of the tested
142 datasets (Figure S2A). Among them, all tested SCC datasets presented the largest effect: cervical
143 squamous cell carcinoma (CESC, CoCu8 HR:2.79, CoCu30 HR: 2.08), HNSCC (CoCu8 Hazard Ratio (HR):
144 1.95, CoCu30 HR: 1.57) and lung squamous cell carcinoma (LUSC, CoCu8 HR: 1.85, CoCu30 HR: 1.78)
145 (Figure 1C; Figure S2B), with CoCu8 consistently showing a slightly higher hazard ratio (HR) compared
146 to CoCu30 in all three tumor types. A multivariate analysis including relevant clinical variables such as
147 age, sex and clinical stage to evaluate the co-culture signatures effect as a continuous variable,

148 confirmed the relevance of the signatures in these tumor types (Figure S3 for CoCu8, Figure S4 for
149 CoCu30,). Our signature was also associated with worse survival in pancreatic and clear cell renal cell
150 carcinoma (ccRCC) (Figure S2A). No significant link to outcome was observed in lung adenocarcinoma,
151 breast, colorectal, bladder, or prostate cancer. Of note, given the clinical and biological differences
152 between HPV positive and negative tumors, which warrant a different staging classification and
153 treatment indications [23], we stratified HNSCC patients according to HPV status, observing that the
154 strongest prognostic effect was visible in HPV positive samples both for CoCu8 (Figure 1D) and CoCu30
155 (HR: 5.47) (Figure S2C).
156
157 We validated the association between CoCu8 / CoCu30 and outcome in a second independent cohort
158 of LUSC from the TRACERx study [24] (Figure 1E, Figure S2D). The multi-regional biopsies performed in
159 the study enabled us to ask if the expression of the CoCu8 signature was uniform across tumors. Of
160 the 117 tumors analyzed, 86 (74%) showed concordant expression of CoCu8 in all regions, which is
161 significantly greater than would be expected based on chance (Figure 1E). Similar results were
162 observed with CoCu30 (Figure S2D). This indicates that cancer cell-fibroblast crosstalk is typically
163 occurring across the whole tumor. Crucially, this analysis showed that the concordant up-regulation
164 of CoCu8 or CoCu30 across tumor regions is associated with worse prognosis. Overall, these data
165 indicate that CoCu8 and CoCu30 signatures are associated with worse overall survival in all SCC
166 datasets tested, therefore we decided to focus our attention on the effect of this crosstalk signature
167 in SCC.
168

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

170 Given that CoCu8 / CoCu30 reflects cancer cell-fibroblast crosstalk, the signature might be predicted
171 to correlate with fibroblast abundance. We performed a correlation analysis of CoCu8 / CoCu30 and
172 methyl CIBERSORT signatures in TCGA datasets, revealing that there is a statistically significant
173 correlation between CAFs presence and CoCu8 signature (Figure S5A-B). Similar results were observed
174 with the CoCu30 signature (Figure S5C-D). We validated these results in the TRACERx LUSC cohort and
175 in a second independent UK_HPV positive cohort (Figure S5E-J). As methylome data was not available
176 for these cohorts, we used fibroblast subtype gene signatures defined in a pan-cancer analysis by
177 Galbo et al.[9]. Strong positive correlations were observed between CoCu8 / CoCu30 signature and all
178 of the fibroblast subtypes defined both in LUSC and UK_HPV positive HNSCC (Figure S5E-J).
179
180 We speculated that our signatures of active cancer cell-fibroblast crosstalk might have better
181 prognostic power than simply CAF abundance. To test this, we analyzed the abundance of CAFs using

182 the methyl CIBERSORT deconvolution strategy in TCGA cohorts and probed links with overall survival[13].
183 This analysis indicated worse OS for HNSCC patients with higher CAF presence, but no significant
184 differences were observed in CESC or LUSC (Figure S6A-B). Of note, both CoCu8 and CoCu30 signatures
185 out-performed the methyl CIBERSORT method for HNSCC, CESC, and LUSC, which adds credence to
186 our method of deriving a signature based on the interaction with fibroblasts, not simply their
187 abundance. Overall, these data indicate that CoCu8 / CoCu30 signatures correlates with CAF
188 abundance but have greater prognostic value than gene signatures used to infer CAF abundance.
189

190 Pathway enrichment analysis of cancer-associated fibroblast and cancer cell co-culture 191 reveals a consistent upregulation of AP-1 transcription factor genes

192 To obtain insight in the molecular basis of the cancer cell-CAF interactions, we analyzed the pathways
193 that were enriched in all the transcriptomic datasets used to generate CoCu8 and CoCu30. This
194 revealed upregulation of multiple pathways linked to immune regulation, stress response, and
195 signaling (Figure 2A). Multiple genes belonging to the AP-1 transcription factor complex: *JUNB*, *FOS*,
196 *FOSB* were strongly enriched (Figure 2B). Moreover, *PLAUR* is regulated by AP-1 factors. As our meta-
197 analysis showed the strongest impact in the stratification of OS patients from HPV positive HNSCC, we
198 decided to explore *JUNB*, *FOS*, *FOSB* expression levels in 4 different co-culture combinations of human
199 HPV positive HNSCC cell lines, SCC154 and SCC47, and human oral CAFs, OCAF1 and OCAF2. Our results
200 confirmed that these three AP-1 TFs are upregulated when cancer cells and CAFs are in direct co-
201 culture, as compared to mono-culture (pooled RNA from both cell lines) (Figure 2C). Analysis of
202 indirect co-cultures [20] separated by a 0.4µm filter indicated that CoCu8 / CoCu30 are strongly enriched
203 with direct co-culture also when compared with indirect co-culture, implying that direct contact is
204 required for increased AP-1 TF expression (Figure S7A).
205

206 Interaction between cancer cells and cancer-associated fibroblasts is linked to increased RAS 207 activity

208 We next investigated possible mechanisms underlying the upregulation of the multifunctional AP-1
209 TFs both in cancer cells and CAFs[25]. RAS signaling via MAPK is known to be a major driver of AP-1 gene
210 expression[26,27]. Accordingly, we found that RAS signaling was strongly up-regulated upon direct co-
211 culture, as indicated by the enrichment of the KRAS_SIGNALLING_UP signature (Figure 2A) and of the
212 curated RAS84 gene signature [28], in all our cancer cell-CAF datasets (Figure 3A, Figure S7B).
213

214 To interrogate further the linkage between CAFs, the CoCu8 signature, and RAS signaling in HPV
215 positive patients from the TCGA cohort, we stratified them according to RAS activity. We split the

216 patient data into three groups (RAS84_0, RAS84_1 and RAS84max) according to the levels of RAS
217 activity as performed by East et al.[28]. This analysis shows that higher RAS activity (group RAS84_max)
218 correlates with worse OS, in a similar fashion to the effect observed with CoCu8 stratification (Figure
219 3B). Indeed, we observed a strong, positive correlation between RAS84 activity and CoCu8 expression
220 in both HPV positive HNSCC cohorts (TCGA - R = 0.79, Figure 3C and UK_HPV positive cohort - R = 0.84,
221 Figure S8A). We also observed a statistically significant enrichment in the extent of fibroblasts present
222 when RAS activity was higher in both cohorts (Figure 3D, Figure S8B).
223
224 Activation of the RAS pathway upon cancer cell-CAF co-culture was experimentally validated by co-
225 culturing SCC154 – OCAF1, which resulted in a strong increase in RAS-MAPK signaling, as determined
226 by phosphorylated ERK1/2 levels (Figure 3E). To test whether the RAS-MAPK pathway is responsible
227 for the upregulation of *JUNB*, *FOS* and *FOSB* genes, we used the MEK inhibitor, trametinib, in the
228 SCC154-OCAF1 co-culture. We confirmed that MEK inhibition downregulates ERK1/2 activation upon
229 co-culture (Figure 3F) and observed a significant downregulation of *JUNB*, *FOS* and *FOSB* genes (Figure
230 3G). Thus, multiple genes belonging to the AP-1 TF complex are upregulated when cancer cells and
231 CAFs are co-cultured and this is mechanistically linked to the activation of RAS-MAPK kinase signaling.
232

233 ## HB-EGF activation is crucial to trigger RAS pathway signaling.

234 To explain why AP-1 TFs get upregulated in co-culture, we looked for possible activators of RAS-MAPK
235 signaling. We noted that *HB-EGF* was among the genes upregulated in all 6 transcriptomic datasets
236 together with *JUNB*, *FOS* and *FOSB* (Figure 2B). HB-EGF is an EGFR ligand and therefore can activate
237 the RAS-MAPK pathway [29,30]. We evaluated the expression levels of all seven EGFR ligands.
238 Importantly, *HB-EGF* showed a strong and specific activation upon cancer cell - CAF co-culture (Figure
239 4A). Moreover, *HB-EGF* expression strongly correlates with CoCu8 in HPV positive HNSCC patients'
240 data from both TCGA (R=0.6, p-value=2.2e-16) and UK_HPV positive (R=0.58, p-value=1e-8) datasets
241 (Figure 4B).
242
243 HB-EGF can activate EGFR/MAPK as an un-cleaved pro-molecule at the plasma membrane [30,31] and, as
244 such, signaling by membrane-bound HB-EGF could explain the need of direct cell contact to trigger
245 the pathway. As membrane-bound HB-EGF should be expressed at about 20-25 kDa, we evaluated the
246 cellular levels of HB-EGF in OCAF1-SCC154 co-culture and found that it was strongly upregulated at
247 the protein level at a molecular weight previously reported as un-cleaved protein [32] (Figure 4C). We
248 also observed EGFR phosphorylation was increased upon direct co-culture (Figure 4C); importantly,
249 both these effects were abrogated by MEK inhibition (Figure 4D), suggesting the presence of a positive

250 feedback loop involving HB-EGF / EGFR / MAPK / AP-1 upon direct co-culture. Furthermore, the EGFR
251 inhibitor, afatinib, blocked AP-1 activation upon co-culture of SCC154-OCAF1 (Figure 4E).
252
253 These results suggest that HB-EGF might be the link that activates EGFR upon cancer cell – CAF co-
254 culture. Therefore, we evaluated the basal expression level of EGFR and HB-EGF in SCC154 and OCAF1
255 mono-cultures. Interestingly, SCC154 expressed HB-EGF at much levels than OCAF1, while both cell
256 types expressed similar levels of EGFR (Figure 4F). This suggests that both cell types can be reactive to
257 EGF ligands, but the activation of the positive feedback loop upon direct contact requires higher levels
258 of HB-EGF, expressed at the membrane of cancer cells. In that case, both cancer cells and CAFs should
259 be responsive to HB-EGF treatment, albeit with potentially different downstream effects. To test this
260 hypothesis, we incubated OCAF1 and SCC154 in mono-cultures with different concentrations of HB-
261 EGF. Firstly, HB-EGF caused an increase in proliferation in both cell types, as shown by a higher
262 proportion of EdU positive nuclei (Figure 4G). Moreover, SCC154 – but not OCAF1 – showed a scatter-
263 like phenotype when treated with high doses of HB-EGF, suggestive of an epithelial to mesenchymal
264 transition. Consistent with this, HB-EGF treatment reduced E-Cadherin staining (Figure S8C). We then
265 tested the effect of HB-EGF treatment on known CAF markers: after treatment of OCAF1 with HB-EGF
266 for 48h we noticed a slight but significant downregulation of *ACTA2*, a CAF and myofibroblast marker
267 (Figure S8D). However, no effect was observed for other widely used CAF markers (*FAP*, *LRRC15*, *FN1*)
268 (Figure S8D).
269
270 Given these data, we asked whether HB-EGF expression in cancer cells is enough to induce
271 upregulation of AP-1 genes when cancer cells and CAFs are in co-culture: we therefore performed
272 knock down of HB-EGF in SCC154 (Figure S8E) and then co-cultured them with OCAF1. Importantly,
273 the downregulation of HB-EGF in SCC154 is enough to block the upregulation of *JUNB*, *FOS* and *FOSB*
274 when cancer cells and CAFs are co-cultured (Figure 4H). These results point to an axis involving HB-
275 EGF in cancer cells and EGFR in CAFs that activates MAPK / AP-1 inducing a positive feedback loop
276 when cancer cells and CAFs are co-cultured.
277

278 ## A paracrine HB-EGF/EGFR axis regulates cytokine expression and macrophage recruitment

279 To focus on the downstream effects of this crosstalk between cancer cells and CAFs and how HB-EGF
280 could affect CAFs functions and lead to unfavorable biology, we analyzed scRNAseq dataset of HNSCC
281 with both malignant and non-malignant samples published by Choi et al. [33]. Interestingly, by using
282 myofibroblast and inflammatory markers (Figure S8F), we found that EGFR is mainly expressed by

283 inflammatory fibroblasts (iFibroblasts), but not myofibroblastic CAFs (myoFibroblasts), while HB-EGF
284 is mainly expressed by endothelial and epithelial cells (Figure 5A).
285
286 Given that high expression of EGFR is linked to iCAFs, we performed a cytokine array of conditioned
287 medium from pooled mono-cultures and co-cultures (Figure 5B): interestingly, we found a strong
288 upregulation of macrophage attraction and differentiation factors (LIF and GM-CSF – gene name *CSF2*
289 –). We validated by qPCR that *LIF* and *CSF2* are transcriptionally upregulated in co-culture (Figure 5C)
290 and that trametinib, MEK inhibitor, treatment is enough to downregulate their expression (Figure 5D).
291 To further investigate the involvement of EGFR in this crosstalk pathway, we performed Afatinib
292 treatment in co-culture and observed that both *LIF* and *CSF2* are strongly downregulated by EGFR
293 inhibition (Figure 5E). Moreover, by reducing HB-EGF expression in cancer cells and then co-culturing
294 the cells with CAFs, we also observed strong downregulation of both *LIF* and *CSF2* expression (Figure
295 5F). Importantly, HB-EGF treatment in cancer cells and CAFs mono-cultures shows that: *LIF* is strongly
296 upregulated only by CAFs, indicating that these are the cells responsible for its production when in co-
297 culture (Figure S9A); *CSF2* was upregulated by HB-EGF treatment both in cancer cells and in CAFs
298 (Figure S9A). In line with these, transcriptomic data of HPV positive HNSCC patients from TCGA show
299 that there is a strong positive correlation between both *LIF* and *CSF2* mRNA and *HBEGF* mRNA
300 expression (Figure S9B). These data establish that HB-EGF/EGFR signaling is required for the up-
301 regulation of cytokines and that EGFR is most highly expressed in human iCAFs.
302
303 Given the established literature behind CSF2/GM-CSF and LIF involvement in macrophage biology [34,35]
304 , we isolated primary monocytes from peripheral blood mononuclear cells (PBMCs) from healthy
305 donors, differentiated them into macrophages and then performed a migration assay using
306 conditioned medium (CM) to ask if cancer cell - CAF direct CM was sufficient to increase macrophage
307 attraction. Importantly, CM derived from the co-culture of OCAF1 and SCC154 increased the numbers
308 of migrating macrophages, compared with pooled CM derived from each cell in monoculture (Figure
309 5G). We next tested if the attraction of macrophages depended on the activation of EGFR or MEK
310 upon cancer cell-fibroblast interaction. When cancer cells and CAFs are co-cultured in the presence of
311 trametinib, there was stark decrease in the number of migrating macrophages attracted by the CM
312 (Figure 5H). Crucially, this was not the case when MEK inhibitor is freshly added to CM after it is
313 harvested from the cancer cell-CAF co-culture, indicating that any residual inhibitor in the CM is not
314 the cause of reduced macrophage attraction (Figure 5H). Moreover, blockade of EGFR using the
315 inhibitor afatinib during the co-culture phase significantly reduced the attraction of macrophages
316 (Figure 5H).

317

318 These data suggest that cancer cell – CAF crosstalk is sufficient to promote macrophage recruitment,
319 therefore we asked whether CoCu8 high patients showed higher levels of macrophages. Of note, when
320 TCGA patients' data are separated according to CoCu8 expression, we observe a strong enrichment
321 for cells defined by a CD14-related methylation signature (monocytes/macrophages) (Figure 5I). We
322 found a similar pattern when patients were separated by fibroblast abundance and by RAS activity
323 (Figure S9C). Importantly, we also observed enrichment for monocyte/macrophage lineages in our
324 second cohort of 84 HPV positive HNSCC patients when separated for CoCu8 expression levels (Figure
325 5I).
326
327 To conclude, we have demonstrated that cancer cell – CAF cross-talk increases expression of different
328 cytokines that, in turn, recruit higher numbers of macrophages. This loop is established by HB-EGF
329 expression in cancer cells that induces a paracrine cross-talk with CAFs via EGFR dependent by RAS /
330 MAPK activity. Activation of this pathway in both CAFs and cancer cells is needed to increase the
331 expression of both LIF and GM-CSF. MEK inhibitor and EGFR inhibitor are sufficient to reduce the
332 macrophage attraction.

333 ## Discussion

334 The presence of CAFs in tumors correlates with worse patient survival and an immune suppressive
335 TME in multiple tumor types [20,36–40], with recent studies linking different CAF subpopulations to
336 prognosis [41]. However, analysis based on the presence or absence of CAFs does not account for
337 variability in the extent of functional crosstalk between cancer cells and CAFs. The approach we
338 develop here is based on the selection of genes that are commonly upregulated in both cancer cells
339 and CAFs upon direct cell-to-cell contact, thus focusing on the functional cancer cell-CAF interactions,
340 rather than just on the presence of CAFs in the tumor. We applied two different strategies to select
341 genes indicative of cancer cell-CAF interactions. The approach to define CoCu30 enriches for genes
342 that are strongly up-regulated, which has been employed previously [9]. To define the CoCu8 signature,
343 we used a new approach based on the selection of a coherent set of genes linked by function.
344 Strikingly, this method generates a signature with prognostic power in all types of SCC investigated,
345 pancreatic, and ccRCC, with particularly strong links to outcome in HPV positive SCC. Our signature did
346 not signify poor prognosis in breast, colorectal, or prostate cancer. We speculate that the different
347 relevance of the signature in cancer arising in different tissues might reflect varying roles for
348 fibroblasts in the tissue in coordinating wound healing responses, including engagement with myeloid
349 cells.
350
351 Comparative analysis of CoCu8 and CoCu30 with annotated gene sets (KRAS SIGNALLING UP and
352 RAS84 signature [28]) suggested a mechanism of cross-talk between cancer cells and CAFs based on the
353 activation of MAPK / AP-1 pathway (Figure S10). Consistent with this, the upregulation of CoCu30
354 genes – FOS, FOSB, JUNB, and HBEGF – required MEK activity. These data extend previous literature
355 showing that KRAS mutation is associated with higher stromal presence [42] and with higher cancer cell
356 – stromal interaction [43]. We hypothesize that our signatures are highly prognostic in HPV positive
357 HNSCC because it lacks oncogenic activation of EGFR or RAS, which frequently occurs in HPV negative
358 disease [44]. Thus, in a subset of HPV positive disease, RAS pathway activation and unfavorable
359 downstream biology are triggered by cancer cell – fibroblast interaction.
360
361 Our analyses indicate that HB-EGF is central to the activation of MAPK signaling upon cancer cell – CAF
362 contact. HB-EGF is the only EGF ligand to be consistently upregulated in co-culture across diverse
363 models. Accordingly, EGFR activation is upregulated in co-culture, suggesting the presence of a
364 positive feedback loop, including HB-EGF / EGFR / MAPK / AP-1. HB-EGF stimulation allowed us to
365 decipher the cell type-dependent consequences of the activation of the pathway. While both cancer
366 cells and CAFs respond to HB-EGF by activating MAPK and inducing changes in AP-1 TF expression, we

observed different downstream activation mechanisms depending by the cell type. These data are consistent with the finding that AP-1 activation leads to diverse molecular and phenotypic consequences depending on the cell type studied [45]. The modest level of HB-EGF expression by cancer cells is not sufficient to initiate these events. We propose the presence of CAFs acts as a mechanism to amplify the expression of HB-EGF, enabling a threshold for productive signaling to be exceeded. CAFs also up-regulate inflammatory cytokines more strongly than cancer cells, meaning that co-culture is required for HB-EGF to drive high levels of expression and subsequent macrophage recruitment. The mechanism through which HB-EGF is upregulated could be associated with proteolytic processing of HB-EGF at the interface between cancer cell and CAFs and will be interesting to test in future studies.

Ultimately, we link increased MAPK and EGFR activity to the chemo-attraction of macrophages. Our data provide insights into the molecular mechanism behind the correlation of CAFs and macrophages in tumors and, more generally, for links between CAFs and a pro-tumorigenic and immune-suppressive milieu [37]. Indeed, CSF2/GM-CSF is known to be associated with macrophage enrichment and chronic inflammation and in cancer [34] and LIF can promote macrophage recruitment and induce a more pro-tumorigenic polarization to alter immune response during anti PD-1 therapy [35]. Of note, CSF2/GM-CSF is produced both by cancer cells and CAFs when stimulated with HB-EGF, while LIF is specifically produced by CAFs. Distinct from our work, Mucciolo and colleagues reported that in pancreatic cancer stromal EGFR activated by AREG is involved in acquisition of pro-tumorigenic properties that favor cancer cells via myofibroblast activation[4]. This difference may reflect either difference between SCC, which is the experimental model in our work, and pancreatic cancer, or that AREG and HB-EGF may trigger different patterns of gene expression. Thus, EGFR is a critical determinant of CAF functions, with further studies required to disentangle tissue- and ligand-specific biology.

Our findings have clinical implications for patient stratification and treatment. Although HPV positive HNSCC patients typically have a better prognosis than HPV negative HNSCC patients, about 25% of these patients still have poor overall survival [46,47].CoCu8 / CoCu30 signatures and CAF abundance could help stratify those patients with worse prognosis within the HPV positive SCC. This improved patient stratification would be especially relevant in the context of the recent unsatisfactory efforts to de-escalate and de-intensify treatment for patients with HPV positive tumors [48] and could help reduce toxicity without compromising outcomes. Moreover, our results suggest that this subset of patients could benefit from a targeted approach, for example re-purposing the use of MEK or EGFR inhibitors. Indeed, trametinib – MEK inhibitor – is already used in the treatment of melanoma [49] and non-small-

401 cell lung cancer [50] and it has been tested in phase I / II oral cavity SCC patients, showing some reduction
402 in RAS / MAPK activity as neoadjuvant treatment [51]. Our data argue that trametinib or EGFR inhibitors
403 may be beneficial for HPV positive HNSCC patients with high stromal content.
404
405 In conclusion, our results demonstrate a new approach to detect biologically meaningful stromal
406 signatures. We show that signatures based on signaling in the TME have the potential to both improve
407 patient stratification and to identify new mechanisms of cross-talk between cancer cells and CAFs.

408 # Materials and methods

409 ## Cell lines and reagents

410 OCAF1 and OCAF2 human fibroblasts were isolated from patient tissues of oral cancer and
411 immortalized with lentiviral HTERT as described in [52]. These patient samples were collected under the
412 ethical approval REC reference 06/Q0403/125).
413 CRUK0764 were derived from patients with lung adenocarcinoma. These fibroblasts were established
414 from the tumor tissue. The primary CRUK0764 was immortalized by the following infection with
415 retroviruses expressing human telomerase reverse transcriptase.
416 PC9 was obtained from the Crick Institute Central Cell Services facility. PC9 were stably transfected
417 with Lipofectamine 2000 Reagent (Thermo Fisher Scientific) according to the manufacturer's
418 instructions. Briefly, PC9 cell line was seeded at 50-70 % confluence in a six-well plate and transfected
419 2 µg of Piggybac transposase (pPBase-piggyBac) and 2 µg of mEGFP (pPBbsr2-mEGFP) plasmid DNAs.
420 After 24h of incubation, the medium with Lipofectamine/plasmid DNA mix was replaced with a fresh
421 medium. Cells were selected using 2 µg ml -1 blasticidin.
422 SCC154 (UPCI-SCC154) and SCC47 (UM-SCC47) were purchased from ATCC.
423 OCAF1, OCAF2, SCC154, SCC47 and CRUK0764 cells were cultured in DMEM (ThermoFisher,
424 #41966052) containing 10% fetal bovine serum (Gibco, #10270-106), 1% penicillin/streptomycin
425 (Invitrogen, #15140122), 1% insulin–transferrin–selenium (Invitrogen, #41400045) and kept at 37°C
426 and 5% CO2.
427 PC9 were cultured in RPMI-1640 (Thermo Fisher Scientific, Rockford, IL) supplemented with 10% fetal
428 bovine serum (Gibco, #10270-106), 1% penicillin/streptomycin (Invitrogen, #15140122) and kept at
429 37 °C in 5% CO2.
430 Cells were not allowed to reach more than 90% confluency for routine cell culture cultivation. Cell
431 lines that are not commercially obtainable are available from the authors upon reasonable request.
432 Routine screening for *Mycoplasma* testing was performed for all cell lines with negative results. STR
433 profiles of human non-commercially available cell lines are included in Supplementary Table 1.
434

435 ## Cell cultures conditions and treatments

436 Co-cultures and mono-cultures were performed with a ratio of 1:2, typically plating 5.5 x 10^5 CAFs
437 and 2.75 x 10^5 cancer cells for a single well of a 6 well plate for the specified time point. When co-
438 cultures were compared to pooled mono-cultures, for the mono-culture condition, same number of
439 cells was plated but in two separated wells and then lysed together (pooled condition).
440 When cancer cells and CAFs mono-culture were compared among themselves, 1 x 10^6 OCAF1 and 1
441 x 10^6 SCC154 cells were plated in a 10cm dish.

442 For PC9 and CRUK0764 cells monocultures and co-cultures used for RNAseq, following 24 h co-
443 cultures, the culture media was replaced with fresh medium with DMSO, then harvested after an
444 additional 24 h. PC9 – CRUK0764 co-cultures were performed in a mixture of RPMI-1640 and DMEM
445 (1:1) containing 1% fetal bovine serum (Gibco, #10270-106).
446 For macrophage cultivation, please see "Macrophage migration assay" section.
447 For cell culture treatments: drugs / factors were added when cells were plated and then added fresh
448 after 24h. Drugs / factors used: trametinib (Selleckchem, #GSK1120212), afatinib (Selleckchem,
449 #BIBW2992), human recombinant HB-EGF (Peprotech, #100-47).
450 For trametinib treatment to collect conditioned medium (CM), in order to control the effect of the
451 drug presence regardless of its effect on secreted factors, we added fresh trametinib treatment to
452 DMSO co-culture CM at the same concentration used for the cell co-culture treatment.
453 All concentrations used are specified in the figures.
454 RNA interference was performed with Lipofectamine RNAimax reagent from Invitrogen, according to
455 the manufacturer's instructions. For transient knock down of HB-EGF, cells were subjected to reverse
456 transfection with 20 nM RNAi oligos plus forward transfection the day after, then analyzed 4 days after
457 reverse transfection. The following RNAi oligo (Dharmacon) was used: siHB-EGF A (Cat # D-019624-
458 02), siHB-EGF B (Cat # D-019624-03), as control the following non-targeting siRNA oligo (All Stars
459 Negative, Quiagen, Cat # 1027281).
460

461 ## Fluorescence-activated cell sorting

462 For PC9 – CRUK0764 RNAseq experiment, CRUK0764 were labelled with CellVue® Red Mini Kit for
463 Membrane Labeling (Polysciences, 25567-1) according to the manufacturer's instructions. Briefly, 1 ×
464 10^7 cells of CRUK0764 were resuspended in the Diluent C and mixed with CellVue® Red working dye
465 solutions (final concentration: 5 × 10^6 cells/mL, 2 × 10^6 M dye) and then incubated for 5 min at RT.
466 Cells were washed twice with DMEM, 10% FBS medium to ensure removal of unbound fluorescence
467 dye.
468 For fluorescence-activated cell sorting, cells were sorted using a flow cytometer–cell sorter BD
469 FACSAria™ II. PC9-GFP and CRUK0764 -CellVue Red were sorted by FACS 48 h after seeding them in
470 monoculture or direct co-culture. The cells were then trypsinised and resuspended in 3% FBS in PBS,
471 1 mM EDTA in preparation for sorting. Cells were separated into two populations: PC9-GFP and
472 CRUK0764 with CellVue Red using a 488 nm laser with collection filter 530 nm/30 nm for GFP and 561
473 nm laser with collection filter 582 nm/20 nm for CellVue Red. Gates were designed on the basis of
474 negative and single-color controls. All cell populations were tested for purity, and data were analyzed
475 using FlowJo software.

476

477 ## RNA sequencing analysis for co-cultures

478 PC9 and CRUK0764 cells were immediately centrifuged at 300 × g for 4 min to remove supernatant
479 and add 350 µl RLT buffer (Qiagen, 79216) containing 1% β-mercaptoethanol (Sigma, M6250) and total
480 RNA was extracted using the RNAeasy Mini kit (Qiagen, 74104; n = 3 independent experiments). Prior
481 to library construction, the quality of total RNA was assessed by Bioanalyzer 2100 (Agilent
482 Technologies Inc).
483 For RNAseq analysis: biological replicates libraries were prepared using the polyA KAPA mRNA
484 HyperPrep Kit and sequenced on the Illumina HiSeq 4000 platform generating ~28 million 75bp single
485 end reads per sample. FASTQ_files were quality trimmed and adaptor removed using Trimmomatic
486 (version 0.36) [53]. The RSEM package (version 1.3.30) [54] in conjunction with the STAR alignment
487 software (version 2.5.2a) [55] was used for the mapping and subsequent gene level counting of the
488 mapped reads with respect to the Ensembl human GRCh38 (release 89) transcriptome. Normalization
489 of raw count data was performed with the DESeq2 package (version 1. 18.1) [56]. All the analysis was
490 done (version 1. 18.1) [56] within the R programming environment (version 3. 4. 3).
491 To check the purity of the samples, we analyzed the resulting transcriptomic data for the expression
492 of 'lineage markers'. CDH1, EPCAM, CD24, and KRT genes were used as markers of carcinoma cells and
493 for fibroblasts we used COL1A1, COL1A2, DCN, CD248, and PDGFR genes. This revealed high sample
494 purity for all transcriptomic data, except in the PC9 – CRUK0764 experiment that had variable purity
495 between samples. Therefore, we estimated the impurity in each sample based on the expression of
496 the lineage marker genes and calculated the expected level of transcript if the two mono-cultures
497 (cancer cells alone and CAFs alone) were mixed in proportion with the impurity estimate. The
498 observed transcript in the co-culture condition was then normalized to account for the effect of
499 contamination.

500

501 ## EdU proliferation assay

502 The Click-iT Plus EdU Imaging Kit (Invitrogen #c10640) was used to perform the assay. Briefly, 48h after
503 mono-cultures of OCAF1 and SCC154 were seeding, a solution with Edu 20µM was prepared and then
504 diluted 1:1 with cell media to add EdU 10µM final concentration. After 90 minutes incubation, cells
505 were washes twice in PBS, then fixed for 15 minutes with paraformaldehyde 3.7% and then washed
506 twice in BSA 3%. Following this step, cells were incubated for 20 minutes with 0.5% Triton X-100 in
507 PBS. After two BSA 3% washes, the Click-iT reaction buffer was added for 30 minutes, followed by one
508 wash in BSA 3% and one wash in PBS. Subsequently, nuclei were stained with Hoechst 33342 at 5
509 µg/mL in PBS incubation for 30 minutes, followed by two PBS washes.

510 Samples were imaged with Zeiss 980 microscope.

511

512 Immunofluorescence assay

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

520

521 Peripheral Blood Mononuclear Cell extraction and monocytes selection

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

527

528 Cytokine array

529 Cytokine array used is "Proteome Profiler Human XL Cytokine Array Kit" (R&D Systems, # ARY022B)
530 following manufacturer's instruction. Briefly, conditioned media was isolated and filtered through a
531 0.4µm low protein binding PVDF Miltex syringe-driven filter (Millipore #SLHV033RS) to remove cellular
532 debris. Media was then concentrated to 4X using Amicon® Ultra-15 and used for subsequent
533 incubation with array.

534

535 Macrophage migration assay

536 Monocytes were plated into 12-well plates ($1x10^5$ cells / well) in RPMI 1640 media (ThermoFisher
537 #12633-012) containing 10% FBS, 1% streptomycin/penicillin and 50ng/mL of M-CSF (Peprotech #300-
538 25) and kept at 37°C and 5% CO2 for 5 days to allow macrophage differentiation. During incubation
539 period, OCAF1 – SCC154 mono- and co-cultures were set-up for 48h. Conditioned media was isolated
540 and filtered through a 0.4µm low protein binding PVDF Miltex syringe-driven filter (Millipore
541 #SLHV033RS) to remove cellular debris. Media was then concentrated to 4X using Amicon® Ultra-15
542 centrifugal filter units (Millipore #UFC901024) and frozen into aliquots until needed. Conditioned

543 media was added to 24-well plates, 8µm hanging cell culture inserts (Millipore #MCEP12H48) were
544 placed on top of each well. The now differentiated macrophages were seeded inside the hanging cell
545 culture insert and left to settle for 10 minutes before topping up media. Plates were left in the
546 incubator for 5 hours to allow macrophages to migrate through membrane pores. After this time, the
547 inserts were removed and the macrophages sat on top of the membrane were wiped off with a cotton
548 bud, leaving behind the migrated macrophages at the bottom. Inserts were stained with 0.05% crystal
549 violet for 30 minutes before washing and then imaged. Inserts were imaged using Zeiss Observer Z1
550 mounted with a QImaging Color camera. Quantification of crystal violet staining was carried out using
551 ImageJ through 'Cell Counter' function.
552

553 ## Gene Set Enrichment Analysis

554 Gene set enrichment analysis was performed with GSEA software v4.1.0. The dataset used to perform
555 the comparative analysis are: RAS84 derived from [28], CoCu8 derived from our own analysis, Hallmarks
556 (h.all.v7.5.symbols.gmt) for all the other analysis. All the parameters have been used as defaults
557 except: permutation type (gene set) and metric for ranking genes (Student's t-test). Gene signatures
558 with a false discovery rate < 0.05 were considered as statistically significant.
559

560 ## Transcriptomic data

561 The transcriptomic data used are: microarray data of A431 / VCAF2b under conditions of mono-
562 cultures, co-cultures in direct contact and indirect contact are available at the Gene Expression
563 Omnibus under record GSE121058. The microarray data of MAF2 under conditions of mono-culture
564 and co-culture in direct contact is available at the Gene Expression Omnibus under record GSE63160.
565 The microarray data of 5555 under conditions of mono-culture and co-culture in direct contact will be
566 submitted at the Gene Expression Omnibus.
567 The RNAseq data of PC9 / CRUK0764 under conditions of mono-cultures and co-cultures in direct
568 contact will be submitted to the European Genome-Phenome Archive before publication.
569 The microarray data used for HUVEC – 1205Lu analysis is available at Gene Expression Omnibus under
570 record GSE8699.
571 The microarray data used for breast cancer cell lines co-culture with fibroblasts analysis is available at
572 Gene Expression Omnibus under record GSE41678.
573

574 ## Co-culture gene signature generation

575 CoCu8 gene signature generation: the A431/VCAF2b, 5555/MAF2, PC9/CRUK0764 co-cultures vs
576 mono-cultures transcriptional datasets have been analyzed with GSEA (see Gene Set Enrichment
577 Analysis method) to obtain a list of enriched pathways in co-culture with FDR < 0.05 for each condition.
578 For each cell type, all the genes statistically upregulated have been pulled together. From this list,
579 genes that were present in 20% or more of the enriched pathways have been selected. The results
580 obtained for each sample have been merged according to the cell type: the three cancer cells in co-
581 culture have been pulled together, same for the three CAFs. To select the final list, only genes present
582 in the three different cancer cells or in the three different CAFs have been selected to generate CoCu8.
583

584 CoCu30 gene signature generation: the genes with a fold change upregulation of 1.5 or higher have
585 been selected for each cell type upon co-culture. The results of the three cancer cells in co-culture
586 have been pulled together, same for the three CAFs in co-culture. To select the final list, only genes
587 present in the three different cancer cells or in the three different CAFs have been selected to generate
588 CoCu30.
589

590 ## TCGA analysis

591 Clinical data, RSEM (RNA-Seq by Expectation-Maximization) normalized expression data (Illumina
592 RNASEQ platform) and Methylation data (Illumina Human Methylation 450 platform) for TCGA cohorts
593 were downloaded from the Firebrowse website hosed by Broad Institute of MIT and Harvard.
594 [http://firebrowse.org/]. Data downloads were all version 2016012800.0.0.
595 De-convolution strategies:
596 MethylCIBERSORT: signature matrix and mixture files were obtained using MethylCIBERSORT R
597 package, hosted on Zenodo. The detailed origin of the signatures and the procedure to create the
598 deconvolution strategy is explained in Chakravarthy et al. [13].
599 Absolute-CIBERSORT: To calculate the immune infiltrate per sample, the library 'CIBERSORT' (version
600 1.04 [57]) was run within R version 3.4.3 on the RSEM normalized data and the LM22 signature using the
601 parameters absolute=TRUE and abs_method ="no.sumto1".
602

603 ## RNA extraction and RT-qPCR

604 Cells were collected and lysed with RLT buffer and total RNA was extracted using the RNeasy Mini kit
605 (Qiagen, #74104), according to the manufacturer's protocol.

606 The cDNA was prepared using M-MLV reverse transcriptase (Promega, #M3682), and quantitative PCR
607 was performed using PowerUp™ SYBR™ Green Master Mix (ThermoFisher, #A25778), using the
608 QuantStudio 3 and 7 Real-Time PCR systems (Applied Biosystems).
609 Custom primers were acquired from Sigma; sequences are available in Supplementary Table 2. RNA
610 levels were normalized using three house-keeping genes using the ΔΔC method and reported as
611 relative fold change compared with Ctr/not treated cells/mono-culture. For each sample, technical
612 triplicates were obtained performed and, if one of the three technical replicates was an outlier, it has
613 been excluded. Samples with expression levels below 37 or undetected have been considered as not
614 expressed and – in order to perform statistics – a Ct value of 40 has been assigned.
615

616 ## scRNAseq analysis

617 scRNAseq data from Choi et al. [33] was downloaded from GEO (GSE181919) and analyzed using Seurat
618 package (version 4) [58].
619

620 ## Protein extraction, quantification and Western Blot analysis

621 Cells were lysed in RIPA buffer (50 mM TrisHCl, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 1% sodium
622 deoxycholate, 0.1% SDS), supplemented with a protease and phosphatase inhibitors (PhosSTOP tablet
623 Roche #04906837001, cOmplete EDTA-free Roche #11873580001, 50 mM NaF). Lysis was performed
624 directly in the cell culture plates using a cell scraper, lysates were kept for 10min on ice and then
625 clarified by centrifugation at 16,000 g for 30 min at 4 °C.
626 Total protein was quantified using the bicinchoninic acid method in accordance with manufacturer's
627 instructions (ThermoFisher, 23225). Following protein quantification, 20 µg of sample was loaded on
628 a 4–15% gradient Mini-PROTEAN TGX Gels (Biorad, #4561084) and transferred to a Trans-Blot Turbo
629 Mini 0.2 µm PVDF membrane (Biorad, 1704156) for blotting. The membrane was blocked for 1 h in
630 5% BSA or 5% milk in TBST and then incubated overnight at 4°C or 1 h at room temperature with
631 antibodies. The membrane was then washed before adding the horseradish peroxidase-conjugated
632 secondary antibody (ThermoFisher), and incubating for 1 h at room temperature. The membrane was
633 washed again before developing with Luminata Classico Western HRP substrate (Millipore,
634 #WBLUR0100) Luminata Classico Western HRP substrate (Millipore, # WBLUF0100) and imaging.
635 Antibody information are listed in Supplementary Table 3. All original blots are provided as source
636 data.
637

638 ## Software and visualization

639 Graphs were generated with Prism software (Graphpad Software v9.4.0) and R (version 4.2.1) using
640 package 'ggplot' except for correlation plot in Figure S10B that was generated with cBioportal [59].
641 scRNAseq data were analyzed with Seurat package (version 4).
642

643 ## Statistics

644 Statistical analysis was performed using Prism software (Graphpad Software v9.4.0), Excel software
645 (Microsoft Corporation v16.0) and R (version 4.2.1).
646 All Student's t-tests have been performed with two tailed strategy.
647 P-value information: * is p-value<0.05; ** is p-value<0.01, *** is p-value<0.001, **** is p-
648 value<0.0001.
649 For GSEA, we used FDR with a threshold below 0.05 to definite the significance.
650 Kaplan-Meier, Log-Rank and Cox regression on survival data was calculated using the R package
651 'survminer' using univariable analysis. Correlations were calculated using the Spearman method in R
652 and the package 'corplot' was used to generate the graphs.
653

654 ## UK_HPV positive cohort

655 FFPE tumor samples from 2 studies formed this cohort:
656 - INOVATE (MR/R015589/1 ISRCTN32335415), a prospective sample collection study in patients
657 with T1-T2/N1-3 or T3-T4/N0-3 oropharyngeal cancer (AJCC TNM classification 7.0) receiving
658 treatment with radical radiotherapy with or without additional platin-based chemotherapy.
659 - INSIGHT-2 (C7224/A23275 NCT04242459), a prospective study of optimizing radiation therapy
660 in head and neck cancers using functional image-guided radiotherapy and novel biomarkers.
661 INOVATE was approved by the London -Bloomsbury Research Ethics Committee (19/LO/1558) and
662 INSIGHT-2 was approved by London – Queen Square Research Ethics Committee (19/LO/0638).
663 Written informed consent was obtained from all participants prior to any study procedure.
664

665 ## UK_HPV positive cohort RNAseq and data analysis

666 Baseline diagnostic biopsies embedded in paraffin blocks were obtained from the above-mentioned
667 cohort. Relevant tumor sections were selected and RNA was extracted from 3-4 slides using the
668 Qiagen AllPrep® DNA/RNA FFPE kit (#80234). Ribosomal RNA was depleted using QIAGEN FastSelect
669 rRNA H/M/R kit (#334375). RNA sequencing libraries were prepared using the NEBNext Ultra II RNA
670 Library Prep Kit (#E770) for Illumina following manufacturer's instructions. The sequencing libraries

671 were multiplexed and loaded on the flow cell on the Illumina NovaSeq 6000 instrument according to
672 manufacturer's instructions. The samples were sequenced using a 2x150 Pair-End (PE) configuration
673 v1.5 for an estimated output of ~50M paired-end reads per sample. Image analysis and base calling
674 were conducted by the NovaSeq Control Software v1.7 on the NovaSeq instrument. Raw sequence
675 data (.bcl files) generated from Illumina NovaSeq was converted into fastq files and de-multiplexed
676 using Illumina bcl2fastq program version 2.20. One mismatch was allowed for index sequence
677 identification. Sample adequacy was confirmed using FASTQC, low quality bases and reads were
678 trimmed using Trimmomatic, we run Hisat2-Stringtie for alignment.
679 RNAseq was performed on 103 patient samples, of which 7 were from the INSIGHT2 and 96 from
680 INOVATE. The data from RNAseq was analyzed to identify samples with presence of HPV (by aligning
681 the unmapped sequences to the whole HPV16 genome sequence obtained from GEO using HISAT2
682 and StingTie) and these samples were classed as HPV positive. 84 samples (77 INOVATE and 7
683 INSIGHT2) were classified at HPV positive and RNAseq data from these was used for analysis in this
684 study.
685

686 ### TRACERx cohort

687 Tumor samples used in this study were collected from LUSC patients enrolled as a part of TRACERx
688 study (accession code: NCT01888601) which is sponsored by University College London (UCL/12/0279)
689 and has been approved by an independent research ethics committee (13/LO/1546). Multiple regions
690 were sampled per tumor and processed as described by Frankell et al. [24] yielding whole-RNA
691 sequencing data for 295 regions from 117 LUSC patients. Expression count and transcript per million
692 (TPM) were quantified by the RSEM package [54]. Genes with expression level of at least 1 TPM in at
693 least 20% of the samples were included. A variance stabilizing transformation (VST) was then applied
694 to filtered count using the DESeq2 package [56].
695

696 ## Supplementary tables

697 Supplementary Tables 1, 2 and 3 are provided with this article.
698

699 ## Data accessibility

700 The transcriptomic data for UK_HPV positive cohort is part of ongoing clinical trials, therefore the data
701 cannot be deposited in a public repository until the trial is finalised. Data can be shared upon
702 reasonable request following corresponding Ethical Research Committee approval following the ICR-
703 Clinical Trials and Statistics Unit policy.

Authors contributions

G.G. and A.R. conceptualized and designed the research, with guidance from E.S. GG performed the experimental analysis and identified the cancer cell – CAF crosstalk gene signature strategy. A.R. performed the clinical data analysis on TCGA and UK_HPV positive cohorts and performed the deconvolution strategies. G.G., A.R., S.H., E.H., Y.N., P.N. and P.C. acquired and analyzed the data, except for TRACERx data that were analyzed 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 and reagents. G.G., A.R. and E.S. wrote the manuscript. All authors reviewed and edited the manuscript. All authors authorized the final version.

Declaration of Interests and acknowledgments

We thank the Crick healthy donors' program and donators for blood donations from the Francis Crick Institute. We thank Sophie de Carne and Phil East for useful discussions around RAS signature and RAS signaling.

G.G. is funded by Merck Sharp & Dohme Corp, New Jersey, USA (LKR190557). A.R. acknowledges funding from the Spanish Society for Medical Oncology (Beca Fundación SEOM), a CRUK accelerator grant to The Francis Crick Institute, and the Royal Marsden NIHR/BRC Bridge funding program. E.H. is supported by the Japan Society for the Promotion of Science (JSPS) Kakenhi Grant [20H03510]. Y.N was supported by a Grant-in-aid for JSPS Overseas Research Fellowship (No. 201860634).

K.H. and A.M. acknowledge funding by the Wellcome Trust, ICR/RM NIHR Biomedical Research Centre, The Institute of Cancer Research/Royal Marsden Hospital Centre for Translational Immunotherapy, CRUK Head and Neck Programme Grant (C7224/A23275) and ICR/RM CRUK RadNet Centre of Excellence (C7224/A28724). Disclosures: Honoraria: Arch Oncology (Inst), AstraZeneca (Inst), BMS (Inst), Boehringer Ingelheim (Inst), Codiak Biosciences (Inst), F-Star Therapeutics (Inst), Inzen Therapeutics (Inst), Merck Serono (Inst), MSD (Inst), Oncolys Biopharma (Inst), Pfizer (Inst), Replimune (Inst), VacV Biotherapeutics (Inst); Consulting or Advisory Role: Arch Oncology (Inst), AstraZeneca (Inst), BMS (Inst), Boehringer Ingelheim (Inst), Inzen Therapeutics (Inst), Merck Serono (Inst), MSD (Inst), Oncolys BioPharma (Inst), Replimune (Inst); Speakers' Bureau: BMS (Inst), Merck Serono (Inst), MSD (Inst); Research Funding: AstraZeneca (Inst), Boehringer Ingelheim (Inst), Merck Sharp & Dohme (Inst), Replimune (Inst).

S.B. declares no conflict of Interest and funding support from ICR/RM NIHR Biomedical Research Centre, CRUK Head and Neck Programme Grant (C7224/A23275), ICR/RM CRUK RadNet Centre of Excellence (C7224/A28724) and the Medical Research Council Developmental Pathway Funding Scheme [MR/R015589/1].

738 D.B. was supported by funding from a Cancer Research UK (CRUK) Early Detection and Diagnosis
739 Project award, the Idea to Innovation (i2i) Crick translation scheme supported by the Medical Research
740 Council, the National Institute for Health Research Biomedical Research Centre and the Breast Cancer
741 Research Foundation (BCRF). D.B. reports personal fees from NanoString and AstraZeneca, and has a
742 patent PCT/GB2020/050221 issued on methods for cancer prognostication.
743 C.S. acknowledges grant support from AstraZeneca, Boehringer-Ingelheim, Bristol Myers Squibb,
744 Pfizer, Roche-Ventana, Invitae (previously Archer Dx Inc - collaboration in minimal residual disease
745 sequencing technologies), and Ono Pharmaceutical. He is an AstraZeneca Advisory Board member and
746 Chief Investigator for the AZ MeRmaiD 1 and 2 clinical trials and is also Co-Chief Investigator of the
747 NHS Galleri trial funded by GRAIL and a paid member of GRAIL's Scientific Advisory Board. He receives
748 consultant fees from Achilles Therapeutics (also SAB member), Bicycle Therapeutics (also a SAB
749 member), Genentech, Medicxi, Roche Innovation Centre – Shanghai, Metabomed (until July 2022),
750 and the Sarah Cannon Research Institute. C.S has received honoraria from Amgen, AstraZeneca, Pfizer,
751 Novartis, GlaxoSmithKline, MSD, Bristol Myers Squibb, Illumina, and Roche-Ventana. C.S. had stock
752 options in Apogen Biotechnologies and GRAIL until June 2021, and currently has stock options in Epic
753 Bioscience, Bicycle Therapeutics, and has stock options and is co-founder of Achilles Therapeutics.
754 C.S. holds patents relating to assay technology to detect tumor recurrence (PCT/GB2017/053289); to
755 targeting neoantigens (PCT/EP2016/059401), identifying patent response to immune checkpoint
756 blockade (PCT/EP2016/071471), determining HLA LOH (PCT/GB2018/052004), predicting survival
757 rates of patients with cancer (PCT/GB2020/050221), identifying patients who respond to cancer
758 treatment (PCT/GB2018/051912), US patent relating to detecting tumor mutations
759 (PCT/US2017/28013), methods for lung cancer detection (US20190106751A1) and both a European
760 and US patent related to identifying insertion/deletion mutation targets (PCT/GB2018/051892). C.S.
761 is a Royal Society Napier Research Professor (RSRP\R\210001). This work was supported by the Francis
762 Crick Institute that receives its core funding from Cancer Research UK (CC2041), the UK Medical
763 Research Council (CC2041), and the Wellcome Trust (CC2041). For the purpose of Open Access, the
764 author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising
765 from this submission. C.S. is funded by Cancer Research UK (TRACERx (C11496/A17786), PEACE
766 (C416/A21999) and CRUK Cancer Immunotherapy Catalyst Network); Cancer Research UK Lung Cancer
767 Centre of Excellence (C11496/A30025); the Rosetrees Trust, Butterfield and Stoneygate Trusts;
768 NovoNordisk Foundation (ID16584); Royal Society Professorship Enhancement Award
769 (RP/EA/180007); National Institute for Health Research (NIHR) University College London Hospitals
770 Biomedical Research Centre; the Cancer Research UK-University College London Centre; Experimental
771 Cancer Medicine Centre; the Breast Cancer Research Foundation (US) BCRF-22-157; Cancer Research

772 UK Early Detection an Diagnosis Primer Award (Grant EDDPMA-Nov21/100034); and The Mark
773 Foundation for Cancer Research Aspire Award (Grant 21-029-ASP). This work was supported by a
774 Stand Up To Cancer-LUNGevity-American Lung Association Lung Cancer Interception Dream Team
775 Translational Research Grant (Grant Number: SU2C-AACR-DT23-17 to S.M. Dubinett and A.E. Spira).
776 Stand Up To Cancer is a division of the Entertainment Industry Foundation. Research grants are
777 administered by the American Association for Cancer Research, the Scientific Partner of SU2C. C.S. is
778 in receipt of an ERC Advanced Grant (PROTEUS) from the European Research Council under the
779 European Union's Horizon 2020 research and innovation programme (grant agreement no. 835297).
780 E.S. is supported by the Francis Crick Institute which receives its core funding from Cancer Research
781 UK (CC2040), the UK Medical Research Council (CC2040), and the Wellcome Trust (CC2040). E.S. is
782 additionally funded by the European Research Council ERCAdG CAN_ORGANISE 101019366. E.S.
783 declares research funding from Merck Sharp & Dohme and Astrazeneca and consultancy work for
784 Phenomic AI and Theolytics. For the purpose of Open Access, the author has applied a CC BY public
785 copyright license to any Author Accepted Manuscript version arising from this submission.
786 RNA sequencing of the UK_HPV positive cohort was funded by the NIHR through the RMH BRC Pump
787 Priming theme (grant B140).
788 The INOVATE study was funded by the Medical Research Council Developmental Pathway Funding
789 Scheme [MR/R015589/1]. The INSIGHT-2 trial was supported by CRUK Head and Neck Programme
790 Grant (C7224/A23275).
791 This study represents independent research supported by the National Institute for Health Research
792 (NIHR) Biomedical Research Centre at the Royal Marsden National Health Service Foundation Trust
793 and the Institute of Cancer Research, London, UK.

794

795 ## References

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

799 2. Sahai E, Astsaturov I, Cukierman E, et al. A framework for advancing our understanding of
800 cancer-associated fibroblasts. *Nat Rev Cancer*. 2020;20(3):174-186. doi:10.1038/s41568-019-
801 0238-1

802 3. Hsu WH, Labella KA, Lin Y, et al. Oncogenic KRAS Drives Lipofibrogenesis to Promote
803 Angiogenesis and Colon Cancer Progression. *Cancer Discov*. 2023;13(12):2652-2673.
804 doi:10.1158/2159-8290.CD-22-1467

805 4. Mucciolo G, Araos Henríquez J, Jihad M, et al. EGFR-activated myofibroblasts promote
806 metastasis of pancreatic cancer. *Cancer Cell*. 2024;42(1):101-118.e11.
807 doi:10.1016/J.CCELL.2023.12.002

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

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

814 7. Grout JA, Sirven P, Leader AM, et al. Spatial Positioning and Matrix Programs of Cancer-
815 Associated Fibroblasts Promote T-cell Exclusion in Human Lung Tumors. *Cancer Discov*.
816 2022;12(11):2606-2625. doi:10.1158/2159-8290.CD-21-1714/708811/AM/SPATIAL-
817 POSITIONING-AND-MATRIX-PROGRAMS-OF-CANCER

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

821 9. Galbo PM, Zang X, Zheng D. Molecular features of cancer-associated fibroblast subtypes and
822 their implication on cancer pathogenesis, prognosis, and immunotherapy resistance. *Clinical*
823 *Cancer Research*. 2021;27(9):2636-2647. doi:10.1158/1078-0432.CCR-20-
824 4226/79067/AM/IMPLICATION-OF-CANCER-ASSOCIATED-FIBROBLAST

825 10. Kürten CHL, Kulkarni A, Cillo AR, et al. Investigating immune and non-immune cell
826 interactions in head and neck tumors by single-cell RNA sequencing. *Nature Communications*
827 *2021 12:1*. 2021;12(1):1-16. doi:10.1038/s41467-021-27619-4

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

830 12. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of
831 Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*.
832 2021;71(3). doi:10.3322/caac.21660

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

836 14. Craven KE, Gökmen-Polar Y, Badve SS. CIBERSORT analysis of TCGA and METABRIC identifies
837 subgroups with better outcomes in triple negative breast cancer. *Scientific Reports 2021 11:1*.
838 2021;11(1):1-19. doi:10.1038/s41598-021-83913-7

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

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

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

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

852 19. Hirata E, Girotti MR, Viros A, et al. Intravital imaging reveals how BRAF inhibition generates
853 drug-tolerant microenvironments with high integrin β1/FAK Signaling. *Cancer Cell*.
854 2015;27(4):574-588. doi:10.1016/j.ccell.2015.03.008

855 20. Arwert EN, Milford EL, Rullan A, et al. STING and IRF3 in stromal fibroblasts enable sensing of
856 genomic stress in cancer cells to undermine oncolytic viral therapy. *Nat Cell Biol*.
857 2020;22(7):758-766. doi:10.1038/s41556-020-0527-7

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

861 22. Stine MJ, Wang CJ, Moriarty WF, et al. Integration of genotypic and phenotypic screening
862 reveals molecular mediators of melanoma-stromal interaction. *Cancer Res*. 2011;71(7):2433-
863 2444. doi:10.1158/0008-5472.CAN-10-1875

864 23. Ruffin AT, Li H, Vujanovic L, Zandberg DP, Ferris RL, Bruno TC. Improving head and neck
865 cancer therapies by immunomodulation of the tumour microenvironment. *Nat Rev Cancer*.
866 2023;23(3):173-188. doi:10.1038/s41568-022-00531-9

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

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

873 26. Westwick JK, Cox AD, Der CJ, et al. Oncogenic Ras activates c-Jun via a separate pathway from
874 the activation of extracellular signal-regulated kinases. *Proc Natl Acad Sci U S A*.
875 1994;91(13):6030-6034. doi:10.1073/PNAS.91.13.6030

876 27. Eferl R, Wagner EF. AP-1: a double-edged sword in tumorigenesis. *Nature Reviews Cancer*
877 *2003 3:11*. 2003;3(11):859-868. doi:10.1038/nrc1209

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

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

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

887 31. Prince RN, Schreiter ER, Zou P, et al. The heparin-binding domain of HB-EGF mediates
888 localization to sites of cell-cell contact and prevents HB-EGF proteolytic release. *J Cell Sci*.
889 2010;123(13):2308-2318. doi:10.1242/JCS.058321

890 32. Provenzano AP, Besner GE, James PF, Harding PA. Heparin-binding EGF-like growth factor
891 (HB-EGF) overexpression in transgenic mice downregulates insulin-like growth factor binding
892 protein (IGFBP)-3 and -4 mRNA. *http://dx.doi.org/101080/08977140512331344012*.
893 2009;23(1):19-31. doi:10.1080/08977140512331344012

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

897 34. Dougan M, Dranoff G, Dougan SK. GM-CSF, IL-3, and IL-5 Family of Cytokines: Regulators of
898 Inflammation. *Immunity*. 2019;50(4):796-811. doi:10.1016/J.IMMUNI.2019.03.022

899 35. Pascual-García M, Bonfill-Teixidor E, Planas-Rigol E, et al. LIF regulates CXCL9 in tumor-
900 associated macrophages and prevents CD8+ T cell tumor-infiltration impairing anti-PD1
901 therapy. *Nature Communications 2019 10:1*. 2019;10(1):1-11. doi:10.1038/s41467-019-
902 10369-9

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

906 37. Qi J, Sun H, Zhang Y, et al. Single-cell and spatial analysis reveal interaction of FAP+ fibroblasts
907 and SPP1+ macrophages in colorectal cancer. *Nature Communications 2022 13:1*.
908 2022;13(1):1-20. doi:10.1038/s41467-022-29366-6

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

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

915 40. Dominguez CX, Müller S, Keerthivasan S, et al. Single-cell RNA sequencing reveals stromal
916 evolution into LRRC15+ myofibroblasts as a determinant of patient response to cancer
917 immunotherapy. *Cancer Discov*. 2020;10(2):232-253. doi:10.1158/2159-8290.CD-19-
918 0644/333443/AM/SINGLE-CELL-RNA-SEQUENCING-REVEALS-STROMAL

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

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

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

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

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

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

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

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

941 49. Dummer R, Hauschild A, Santinami M, et al. Five-Year Analysis of Adjuvant Dabrafenib plus
942 Trametinib in Stage III Melanoma. *New England Journal of Medicine*. 2020;383(12):1139-
943 1148. doi:10.1056/NEJMOA2005493/SUPPL_FILE/NEJMOA2005493_DATA-SHARING.PDF

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

947 51. Uppaluri R, Winkler AE, Lin T, et al. Biomarker and tumor responses of oral cavity squamous
948 cell carcinoma to trametinib: A phase II neoadjuvant window-of-opportunity clinical trial.
949 *Clinical Cancer Research*. 2017;23(9):2186-2194. doi:10.1158/1078-0432.CCR-16-
950 1469/116024/AM/BIOMARKER-AND-TUMOR-RESPONSES-OF-ORAL-CAVITY

951 52. Gaggioli C, Hooper S, Hidalgo-Carcedo C, et al. Fibroblast-led collective invasion of carcinoma
952 cells with differing roles for RhoGTPases in leading and following cells. *Nature Cell Biology*
953 *2007 9:12*. 2007;9(12):1392-1400. doi:10.1038/ncb1658

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

956 54. Li B, Dewey CN. RSEM: Accurate transcript quantification from RNA-Seq data with or without
957 a reference genome. *BMC Bioinformatics*. 2011;12(1):1-16. doi:10.1186/1471-2105-12-
958 323/TABLES/6

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

961 56. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-
962 seq data with DESeq2. *Genome Biol*. 2014;15(12):1-21. doi:10.1186/S13059-014-0550-
963 8/FIGURES/9

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

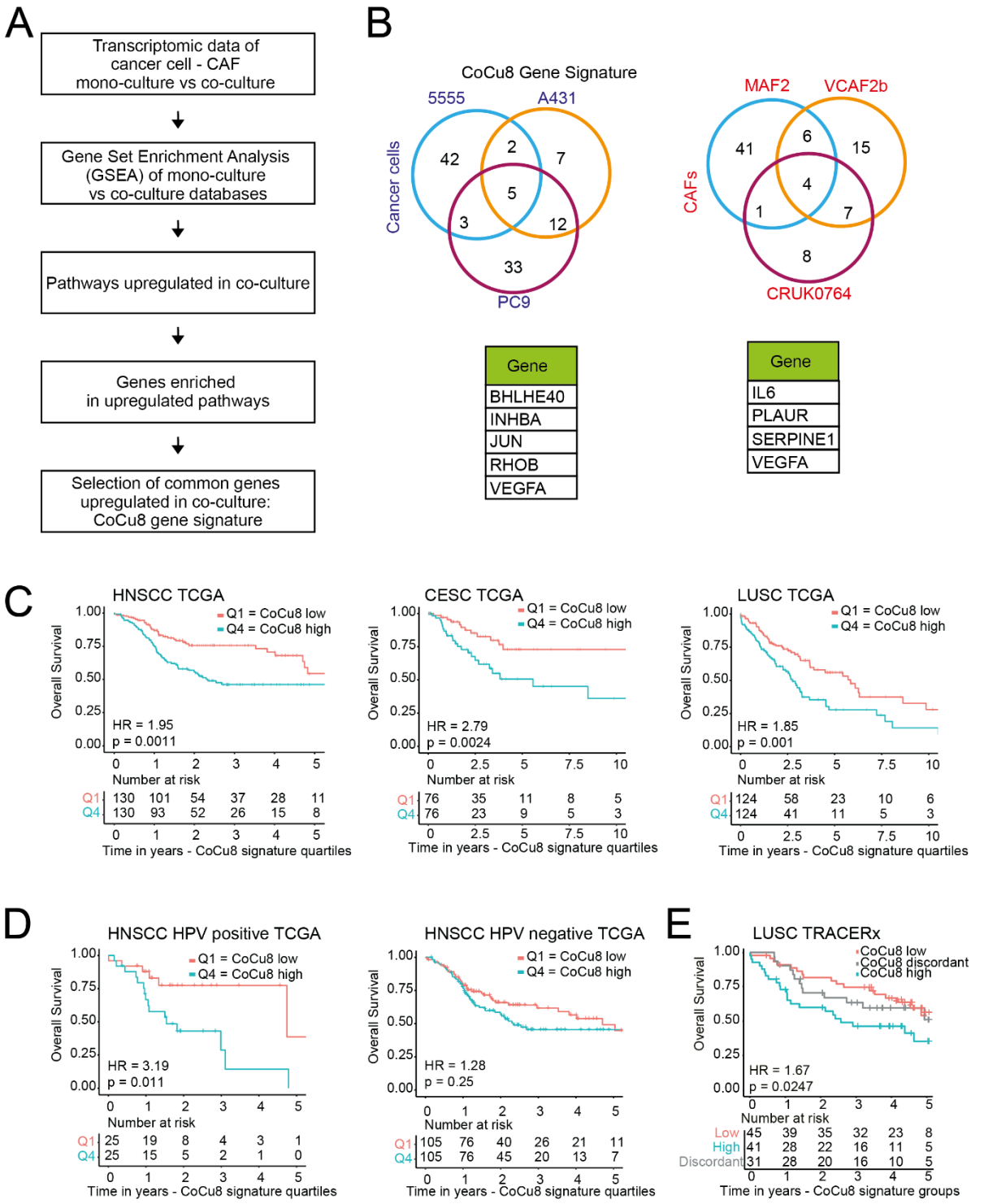
966 58. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data
967 across different conditions, technologies, and species. *Nature Biotechnology 2018 36:5*.
968 2018;36(5):411-420. doi:10.1038/nbt.4096

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

972 60. Biswas D, Birkbak NJ, Rosenthal R, et al. A clonal expression biomarker associates with lung
973 cancer mortality. *Nature Medicine 2019 25:10*. 2019;25(10):1540-1548. doi:10.1038/s41591-
974 019-0595-z

975

976 # Figures



977 Figure 1

978 **Figure 1:** Cancer cell / CAF co-culture gene signature CoCu8 is associated with worse overall survival
979 in multiple squamous cell carcinoma datasets. **A)** Strategy used to obtain CoCu8 gene signature. **B)**
980 Venn diagrams of the genes upregulated in the different datasets (top) and summary table of the
981 genes upregulated in all the datasets (bottom) for cancer cells (right) and CAFs (left). **C)** Kaplan-Meier
982 overall survival analysis of HNSCC (right), CESC (center), LUSC (left) TCGA datasets stratified for CoCu8
983 first vs last quartile. Numbers at risk shown in tables below graphs. HNSCC HR=1.95 (95% Confidence
984 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
985 HR=1.85 (95%CI 1.28-2.70), p-value=0.001. HR and CI were calculated using Cox regression. p-value
986 was calculated using logRank test. **D)** Kaplan-Meier overall survival analysis of HNSCC HPV positive
987 (right) and negative (left) TCGA datasets stratified for CoCu8 first vs last quartile. Numbers at risk
988 shown in tables below graphs. HNSCC HPV positive HR=3.19 (95%CI 1.24-8.18), p-value=0.011. HPV
989 negative HR=1.28 (95%CI 0.84-1.93), p-value=0.25. HR and CI were calculated using Cox regression. p-
990 value was calculated using logRank test. **E)** Kaplan-Meier overall survival analysis of LUSC TRACERx
991 dataset. Individual tumors stratified as high-, discordant or low-risk according to expression profile of
992 CoCu8 signature across multiple regions, as previously described and stratified according to Biswas et
993 al. [60]. Briefly, patients were classified as discordant when different tumor regions from the same
994 patient presented not unique signature levels. Below are shown the numbers at risk in years. HR=1.67
995 (95%CI 1.13-2.47), p-value=0.0247. HR and CI calculated using Cox regression and are referred to
996 CoCu8 low vs CoCu8 high. p-value was calculated using logRank test.

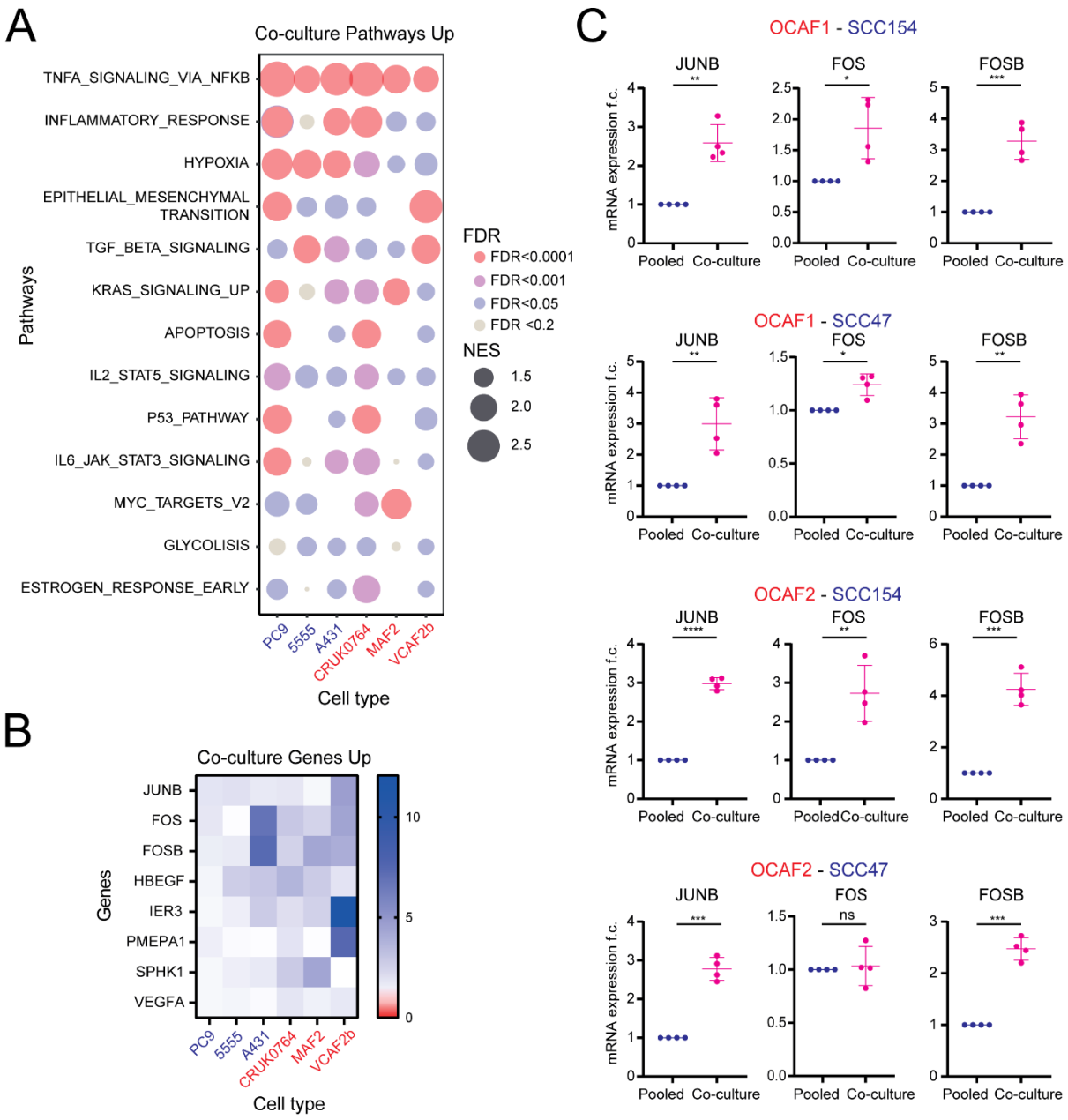

A
Co-culture Pathways Up
TNFA_SIGNALING_VIA_NFKB
INFLAMMATORY_RESPONSE
HYPOXIA
EPITHELIAL_MESENCHYMAL TRANSITION
TGF_BETA_SIGNALING
KRAS_SIGNALING_UP
APOPTOSIS
IL2_STAT5_SIGNALING
P53_PATHWAY
IL6_JAK_STAT3_SIGNALING
MYC_TARGETS_V2
GLYCOLISIS
ESTROGEN_RESPONSE_EARLY
Pathways
PC9
5555
A431
CRUK0764
MAF2
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
OCAF1 - SCC47
OCAF2 - SCC154
OCAF2 - SCC47
JUNB
FOS
FOSB
mRNA expression f.c.
Pooled
Co-culture
ns

Figure 2

999 **Figure 2:** Cancer cell / CAFs co-culture upregulates AP-1 TF genes. **A)** Bubble plot of the Hallmarks
1000 pathways upregulated by cancer cells / CAFs culture. Normalized enrichment score (NES) is depicted
1001 as bubble size; False discovery rate (FDR) is depicted as color intensity. **B)** Heatmap of expression of
1002 the genes commonly upregulated upon co-culture in all the tested conditions from the
1003 TNFA_SIGNALLING_VIA_NKFB pathway. **C)** qPCR analysis of *JUNB*, *FOS* and *FOSB* genes in
1004 OCAF1/OCAF2 with SCC154/SCC47 pooled mono-cultures and co-culture after 24h. mRNA expression
1005 is reported as mean ± standard deviation (SD) fold change difference over pooled mono-culture.
1006 Genes have been normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 4
1007 independent experiments. Two tailed paired Student's t-test.

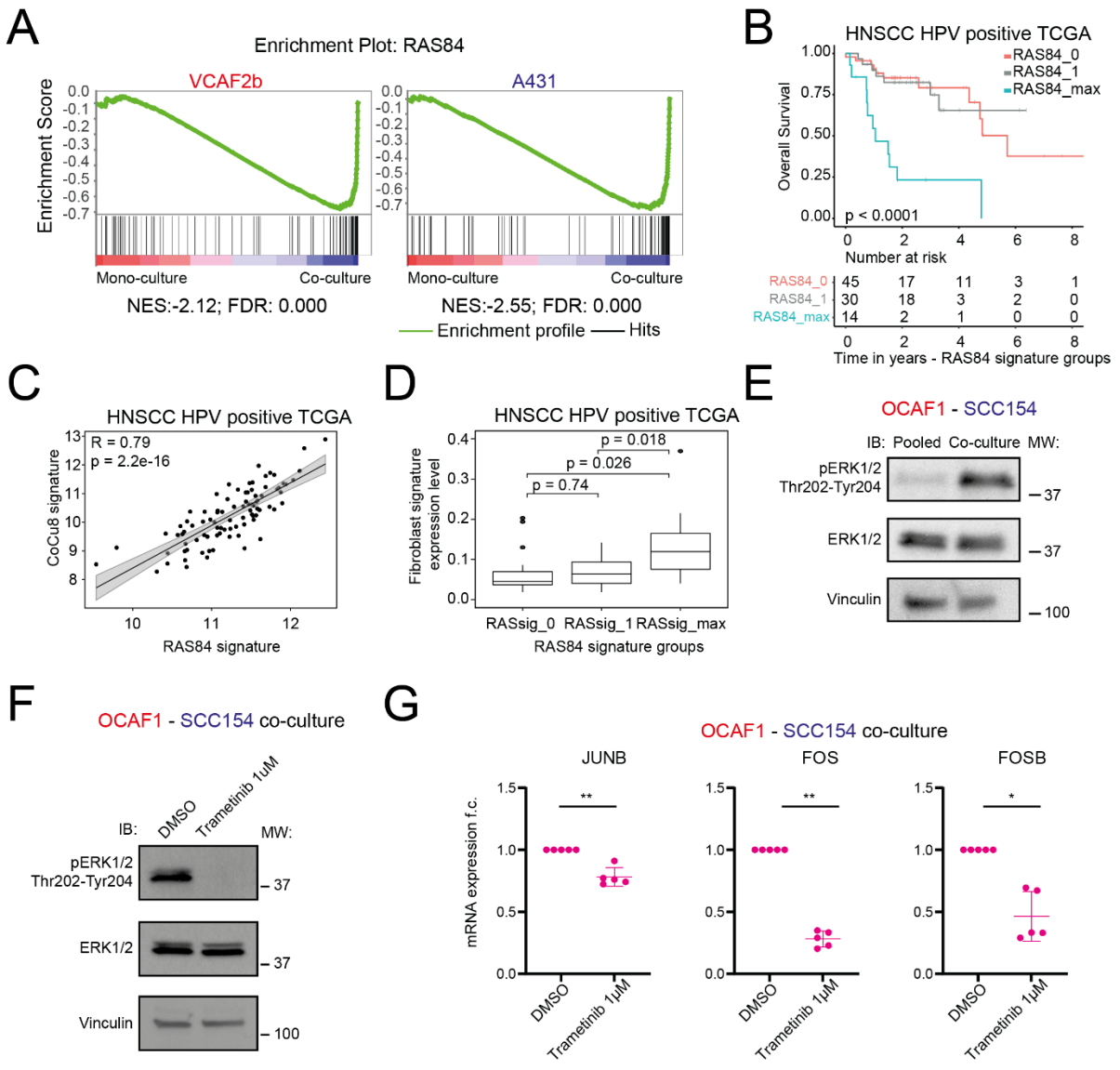

A
Enrichment Plot: RAS84
VCAF2b
A431
Enrichment Score
Mono-culture
Co-culture
NES:-2.12; FDR: 0.000
NES:-2.55; FDR: 0.000
Enrichment profile
Hits
B
HNSCC HPV positive TCGA
RAS84_0
RAS84_1
RAS84_max
Overall Survival
p < 0.0001
Number at risk
Time in years - RAS84 signature groups
C
HNSCC HPV positive 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

1010 **Figure 3:** RAS activity is upregulated in cancer cell and CAFs upon co-culture. **A)** Gene set enrichment
1011 analysis (GSEA) plot of RAS84 gene signature in mono-culture and co-culture. NES and FDR are
1012 specified below each plot. **B)** Kaplan-Meier overall survival analysis of HNSCC HPV positive TCGA
1013 dataset stratified for Ras84 activity according to [28]. Below are shown the numbers at risk in years.
1014 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
1015 (95%CI 2.46-13.9) p-value<0.001. HR and CI calculated using Cox regression. **C)** Correlation plot of
1016 RAS84 expression level and CoCu8 expression level in HNSCC HPV positive TCGA dataset. R is
1017 Spearman correlation coefficient. n=97. **D)** Box plot analysis of fibroblast abundance via Methyl
1018 CIBERSORT deconvolution strategy in HNSCC HPV positive TCGA dataset according to RAS84 activity.
1019 Independent Student's t-test. **E)** Western blot analysis of OCAF1 – SCC154 pooled mono-culture vs co-
1020 culture for 48h showing the indicated antibodies. Vinculin is used as loading control. **F)** Western blot
1021 analysis of OCAF1 – SCC154 co-cultures for 48h at the indicated conditions showing the indicated
1022 antibodies. Vinculin is used as loading control. **G)** qPCR analysis of *JUNB*, *FOS* and *FOSB* genes in OCAF1
1023 - SCC154 co-cultures for the indicated treatments after 48h. mRNA expression is reported as
1024 mean ± standard deviation (SD) fold change difference over co-culture DMSO. Genes have been
1025 normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 4 independent
1026 experiments. Two tailed paired Student's t-test.

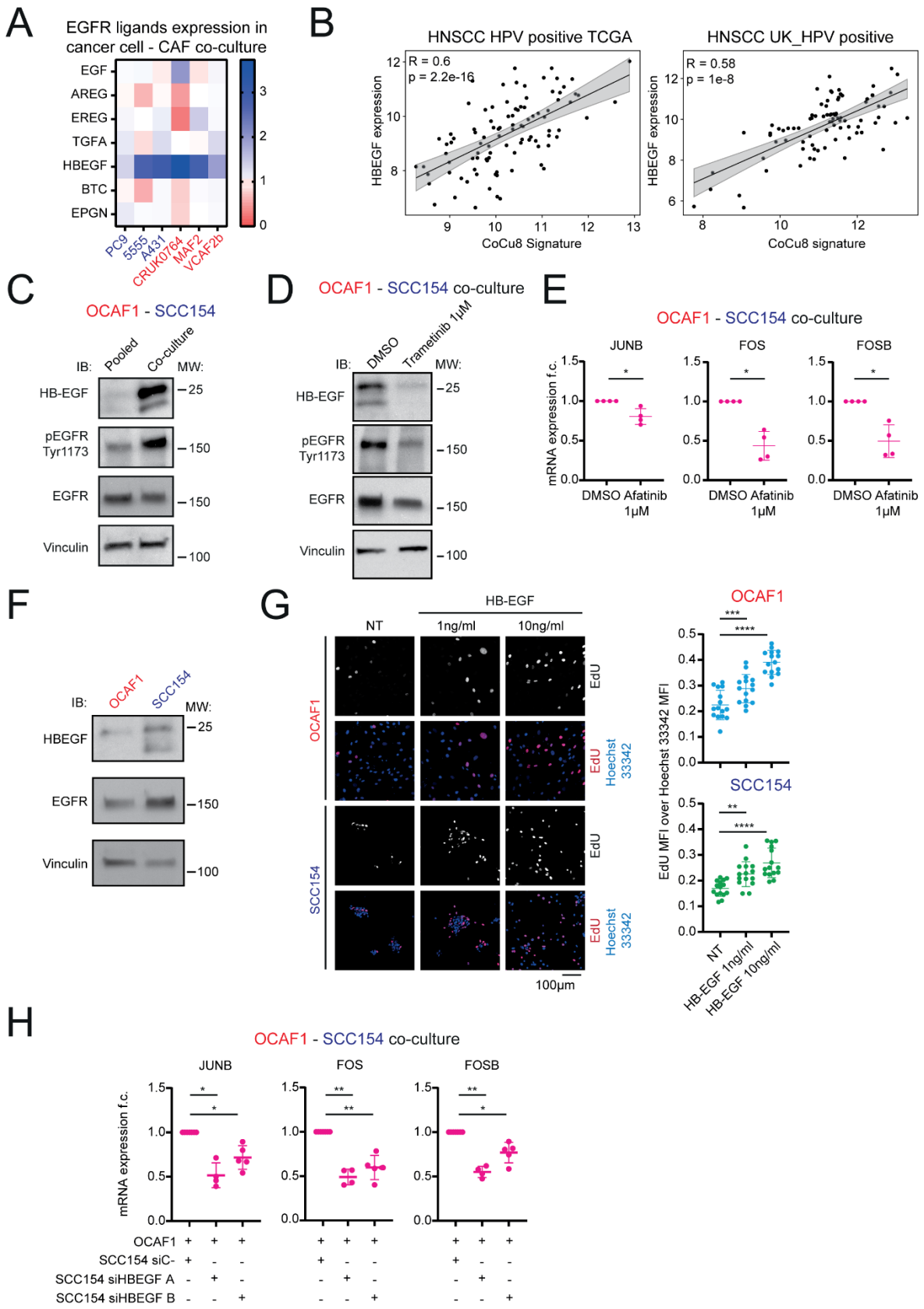

A
EGFR ligands expression in cancer cell - CAF co-culture
EGF
AREG
EREG
TGFA
HBEGF
BTC
EPGN
PC9
5555
A431
CRUK0764
MAF2
VCAF2b
B
HNSCC HPV positive TCGA
R = 0.6
p = 2.2e-16
HBEGF expression
CoCu8 Signature
HNSCC UK_HPV positive
R = 0.58
p = 1e-8
HBEGF expression
CoCu8 signature
C
OCAF1 - SCC154
IB:
Pooled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

Figure 4

1029 **Figure 4:** HB-EGF / EGFR axis activates AP-1 TF genes in cancer cells and CAFs upon co-culture via RAS
1030 pathway. **A)** Pattern of expression of the 7 EGFR ligands in all the tested transcriptomic datasets. **B)**
1031 Correlation plot of *HB-EGF* expression level and CoCu8 expression level in HNSCC HPV positive TCGA
1032 dataset (left, n=97) and UK_HPV positive cohort (right, n=84). R is Spearman correlation coefficient.
1033 **C)** Western blot analysis of OCAF1 – SCC154 pooled mono-culture vs co-culture for 48h showing the
1034 indicated antibodies. Vinculin is used as loading control. **D)** Western blot analysis of OCAF1 – SCC154
1035 co-cultures for 48h at the indicated conditions showing the indicated antibodies. Vinculin is used as
1036 loading control. **E)** qPCR analysis of *JUNB*, *FOS* and *FOSB* genes in OCAF1 - SCC154 co-cultures for the
1037 indicated treatments after 48h. mRNA expression is reported as mean ± standard deviation (SD) fold
1038 change difference over co-culture DMSO. Genes have been normalized over the average of *GAPDH*,
1039 *ACTB* and *RPLP0* housekeeping genes. n = 4 independent experiments. The DMSO treated sample is
1040 the same used for Figure 3G. Two tailed paired Student's t-test. **F)** Western blot analysis of OCAF1 and
1041 SCC154 mono-cultures after 48h showing the indicated antibodies. Vinculin is used as loading control.
1042 **G)** Proliferation assay of OCAF1 and SCC154 mono-cultures for the indicated treatments after 48h and
1043 stained with EdU and Hoechst 33342. On the left, a representative image is shown with bar graph. On
1044 the right, dot plot of mean fluorescent intensity (MFI) of EdU over Hoechst 33342 with
1045 mean ± standard deviation (SD) highlighted. Each dot is a field of view. n = 3 independent experiments.
1046 Two tailed Student's t-test. **H)** qPCR analysis of *JUNB*, *FOS* and *FOSB* genes in OCAF1 - SCC154 co-
1047 cultures after 48h with SCC154 pre-treated with the indicated conditions. mRNA expression is
1048 reported as mean ± standard deviation (SD) fold change difference over siC- condition. Genes has been
1049 normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n ≥ 4 independent
1050 experiments. Two tailed paired Student's t-test.

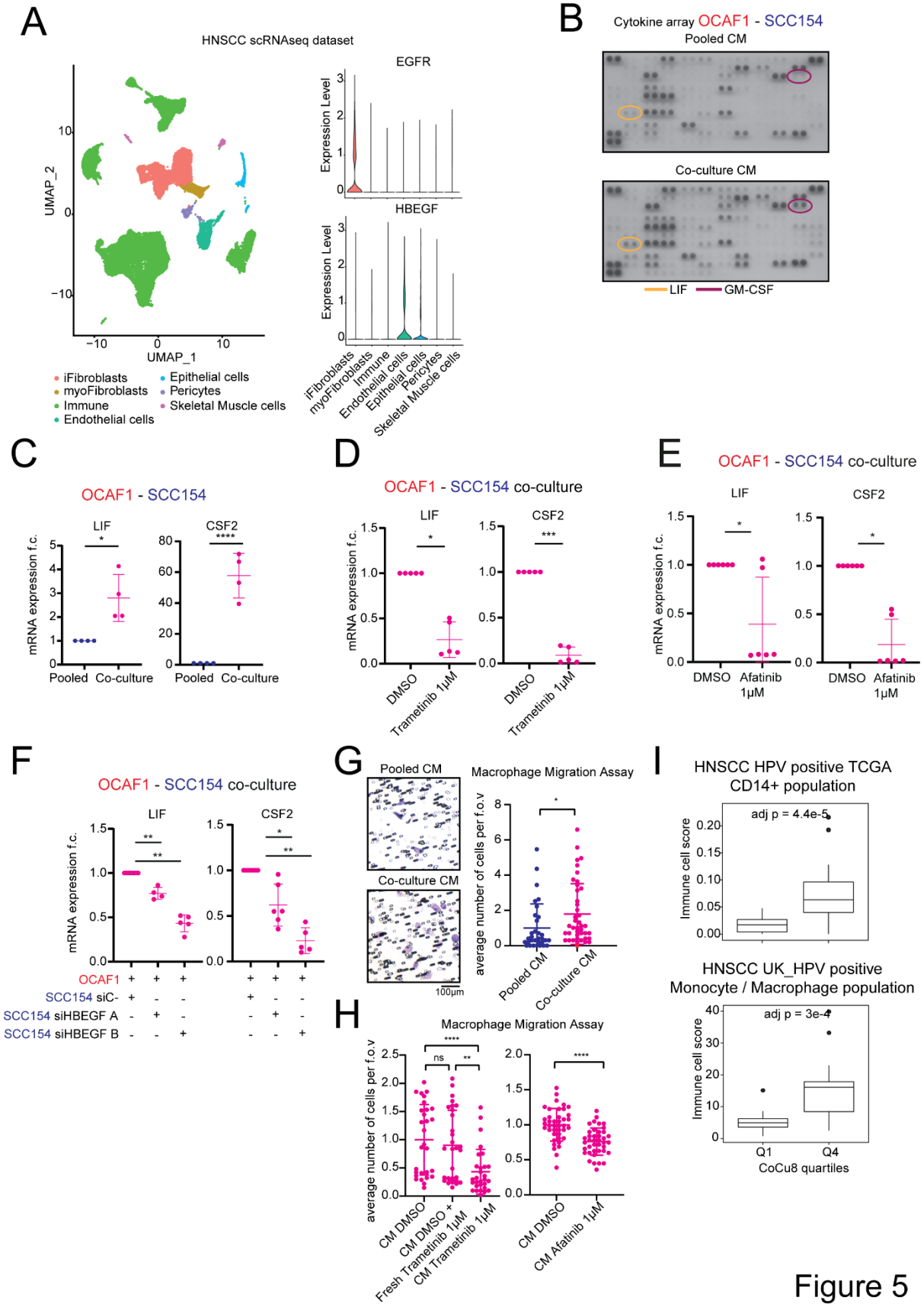


Figure 5

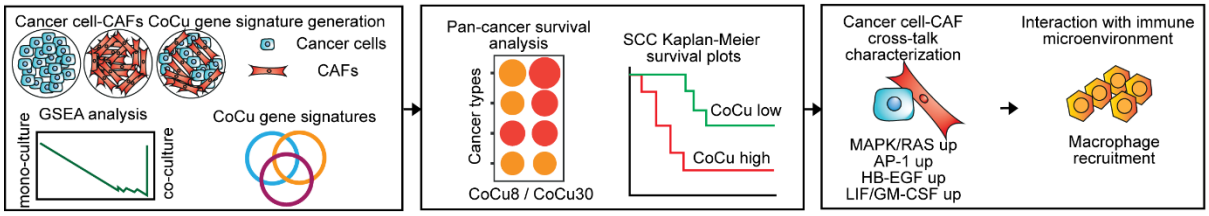
1051

1052

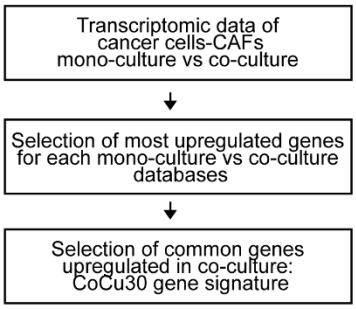
1053 **Figure 5:** Cancer cells – CAFs co-culture induces production of specific cytokines to attract
1054 macrophages. **A)** On the right is shown scRNAseq UMAP analysis from Choi et al. [33]. On the left is
1055 shown violin plot of EGFR and HBEGF mRNA expression levels in the indicated clusters. **B)** Cytokine
1056 array of conditioned medium from pooled mono-culture and co-culture of OCAF1 and SCC154.
1057 Highlighted relevant cytokines. n = 2 independent experiments. **C)** qPCR analysis of *LIF* and *CSF2* genes
1058 in OCAF1 – SCC154 pooled mono-culture vs co-culture after 24h. mRNA expression is reported as
1059 mean ± standard deviation (SD) fold change difference over pooled mono-culture. Genes have been
1060 normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 4 independent
1061 experiments. Two tailed paired Student's t-test. **D)** qPCR analysis of *LIF* and *CSF2* genes in OCAF1 -
1062 SCC154 co-cultures for the indicated treatments after 48h. mRNA expression is reported as
1063 mean ± standard deviation (SD) fold change difference over co-culture DMSO. Genes have been
1064 normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 5 independent
1065 experiments. Paired t-test. **E)** qPCR analysis of *LIF* and *CSF2* genes in OCAF1 - SCC154 co-cultures for
1066 the indicated treatments after 48h. mRNA expression is reported as mean ± standard deviation (SD)
1067 fold change difference over co-culture DMSO. Genes have been normalized over the average of
1068 *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 6 independent experiments. The DMSO treated
1069 sample is the same used for Figure 5D. Two tailed paired Student's t-test. **F)** qPCR analysis of *LIF* and
1070 *CSF2* genes in OCAF1 - SCC154 co-cultures after 48h with SCC154 pre-treated with the indicated
1071 conditions. mRNA expression is reported as mean ± standard deviation (SD) fold change difference
1072 over siC- condition. Gene has been normalized over the average of *GAPDH*, *ACTB* and *RPLP0*
1073 housekeeping genes. n ≥ 4 independent experiments. Two tailed paired Student's t-test. **G)** Migration
1074 assay of macrophages plated in transwells with conditioned medium (CM) from OCAF1-SCC154 pooled
1075 mono-culture or co-culture. CM have been obtained after 48h culture. On the left, a representative
1076 field of view is shown with bar graph. On the right, dot plot of number of cells per field of view as
1077 mean ± standard deviation (SD). Each dot is a field of view normalized by the average of the pooled
1078 mono-culture CM sample. n = 4 different donors. Two tailed Student's t-test. **H)** Migration assay of
1079 macrophages plated in transwells with CM from OCAF1-SCC154 co-culture with the indicated
1080 treatments. CM have been obtained after 48h culture and, for the fresh trametinib sample the drug
1081 has been added after the CM was collected. Dot plot of number of cells per field of view as
1082 mean ± standard deviation (SD) is shown. Each dot is a field of view normalized by the average of the
1083 co-culture DMSO CM sample. n = 3 different donors for trametinib effect and 4 donors for afatinib
1084 effect. Two tailed Student's t-test. **I)** At the top, box plot analysis of CD14+ monocytic/macrophage
1085 lineage immune cell absolute score via Methyl CIBERSORT deconvolution strategy in HNSCC HPV
1086 positive TCGA dataset separate by first and last quartile of CoCu8 expression. Independent Student's

1087 t-test, Bonferroni correction for multiple comparisons. At the bottom, box plot analysis of monocyte
1088 and macrophage immune cell score using Absolute CIBERSORT deconvolution strategy in HNSCC HPV
1089 positive UK_HPV positive dataset separate by first and last quartile of CoCu8 expression. Independent
1090 Student's t-test.
1091

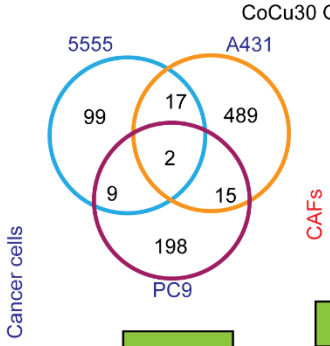
A



B



C



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

D

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.

E

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 S1

1094 **Figure S1:** Generation and validation of CoCu8 / CoCu30 gene signatures. **A)** Flow chart description of
1095 the manuscript is provided. **B)** Strategy used to obtain CoCu30 gene signature. **C)** Venn diagram of the
1096 genes upregulated in the different datasets (top) and a table to summarize the genes constantly
1097 upregulated in the datasets (bottom) for cancer cells (left) and CAFs (right). **D)** Table with Normalized
1098 Enrichment Score (NES) values of different combinations of cancer cells and CAFs breast cancer cell
1099 lines from Rajaram et al. [21]. Negative values represent enrichment towards co-culture condition. Color
1100 legend is shown. **E)** Table with NES values of CoCu8 and CoCu30 gene signatures for cancer cells and
1101 endothelial cells mono-culture vs co-culture available from Stine et al. [22]. Negative values represent
1102 enrichment towards co-culture condition. Color legend is shown.
1103
1104

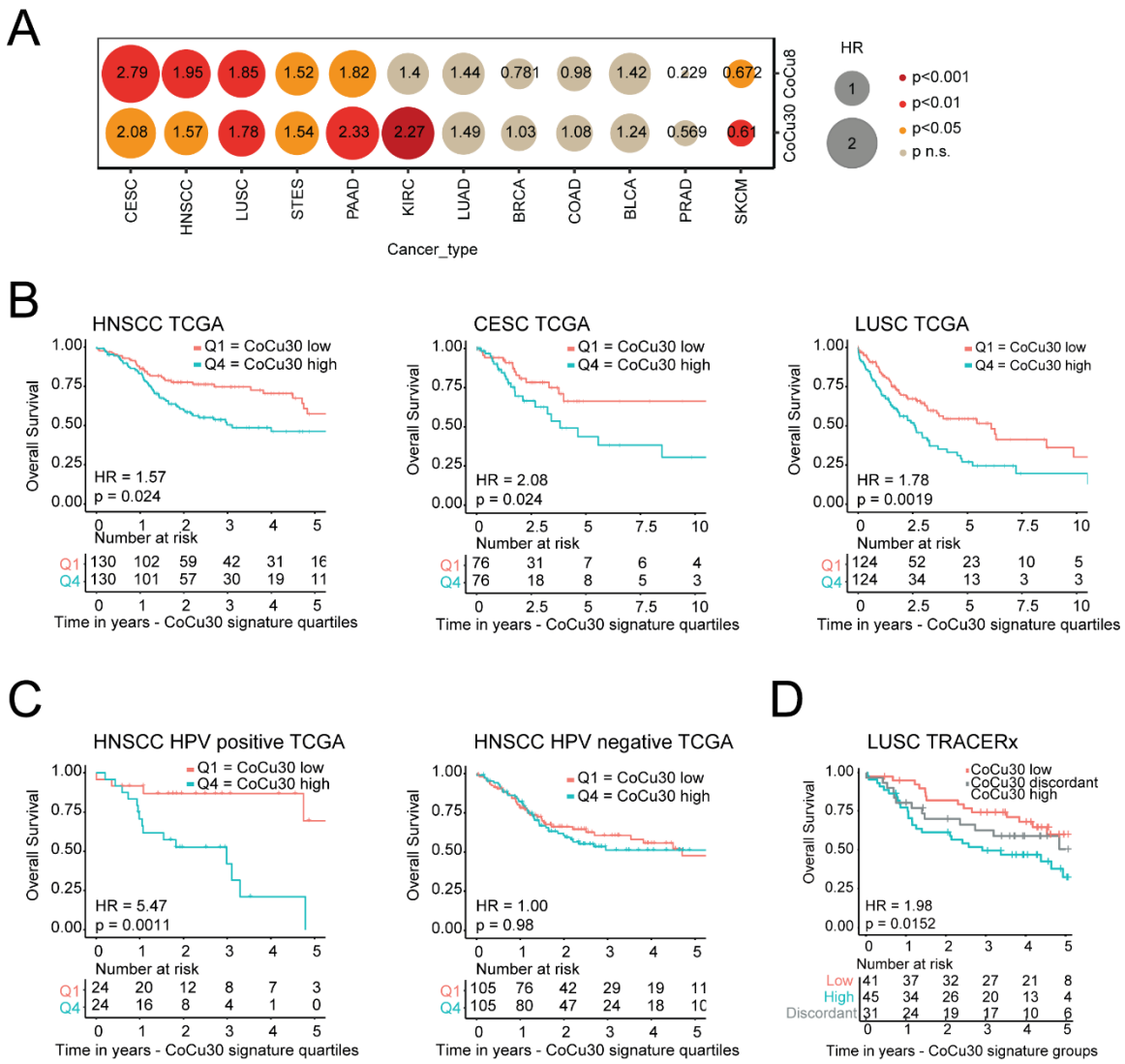

A
2.79 1.95 1.85 1.52 1.82 1.4 1.44 0.781 0.98 1.42 0.229 0.672
2.08 1.57 1.78 1.54 2.33 2.27 1.49 1.03 1.08 1.24 0.569 0.61
CoCu30 CoCu8
CESC HNSCC LUSC STES PAAD KIRC LUAD BRCA COAD BLCA PRAD SKCM
Cancer_type
HR
1
2
p<0.001
p<0.01
p<0.05
p n.s.
B
HNSCC TCGA
Q1 = CoCu30 low
Q4 = CoCu30 high
HR = 1.57
p = 0.024
Overall Survival
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
Q1 = CoCu30 low
Q4 = CoCu30 high
HR = 2.08
p = 0.024
Number at risk
Q1 76 31 7 6 4
Q4 76 18 8 5 3
Time in years - CoCu30 signature quartiles
LUSC TCGA
Q1 = CoCu30 low
Q4 = CoCu30 high
HR = 1.78
p = 0.0019
Number at risk
Q1 124 52 23 10 5
Q4 124 34 13 3 3
Time in years - CoCu30 signature quartiles
C
HNSCC HPV positive TCGA
Q1 = CoCu30 low
Q4 = CoCu30 high
HR = 5.47
p = 0.0011
Number at risk
Q1 24 20 12 8 7 3
Q4 24 16 8 4 1 0
Time in years - CoCu30 signature quartiles
HNSCC HPV negative TCGA
Q1 = CoCu30 low
Q4 = CoCu30 high
HR = 1.00
p = 0.98
Number at risk
Q1 105 76 42 29 19 11
Q4 105 80 47 24 18 10
Time in years - CoCu30 signature quartiles
D
LUSC TRACERx
CoCu30 low
CoCu30 discordant
CoCu30 high
HR = 1.98
p = 0.0152
Number at risk
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 S2

1107 **Figure S2:** CoCu8 / CoCu30 gene signatures are associated with worse overall survival in multiple
1108 squamous cell carcinoma datasets. **A)** Bubble plot hazard ratios and p- values for overall survival for
1109 multiple TCGA cancer types using both CoCu8 and CoCu30 signatures. CESC - cervical squamous cell
1110 carcinoma; HNSCC - head and neck squamous cell carcinoma; LUSC – lung squamous cell carcinoma;
1111 PRAD - prostate adenocarcinoma; PAAD - pancreatic adenocarcinoma; LUAD - lung adenocarcinoma;
1112 KIRC - kidney clear cell carcinoma; COAD - colorectal adenocarcinoma; BLCA - bladder urothelial
1113 carcinoma; STES - esophagogastric carcinoma; SKCM – melanoma; BRCA - breast cancer. **B)** Kaplan-
1114 Meier overall survival analysis of HNSCC (left), CESC (center), LUSC (right) TCGA datasets stratified for
1115 CoCu30 first vs last quartile. Below each analysis are shown the corresponding numbers at risk, time
1116 in years. HNSCC HR=1.57 (95%CI 1.06-2.35), p-value=0.024. CESC HR=2.08 (95%CI 1.09-4.01), p-
1117 value=0.024. LUSC HR=1.78 (95%CI 1.23-2.59), p-value=0.0019. HR and CI were calculated using Cox
1118 regression. p-value was calculated using logRank test. **C)** Kaplan-Meier overall survival analysis of
1119 HNSCC HPV positive (left) and HNSCC HPV positive (right) TCGA datasets stratified for CoCu30 first vs
1120 last quartile. Below each analysis are shown the corresponding numbers at risk, time in years. HPV
1121 positive HR=5.47 (95%CI 1.76-17.0), p-value=0.0011. HPV negative HR=1.00 (95%CI 0.67-1.51), p-
1122 value=0.98. HR and CI were calculated using Cox regression. p-value was calculated using logRank test.
1123 **D)** Kaplan-Meier overall survival analysis of LUSC TRACERx dataset. Individual tumors stratified as high-
1124 , discordant or low-risk according to expression profile of CoCu30 signature across multiple regions,
1125 as previously described and stratified according to Biswas et al. [60]. Below are shown the numbers at
1126 risk in years. HR=1.98 (95% CI 1.09-3.6), p-value=0.0152. HR and CI calculated using Cox regression
1127 and are referred to CoCu30 low vs CoCu30 high. p-value was calculated using logRank test.

CoCu8 Overall Survival

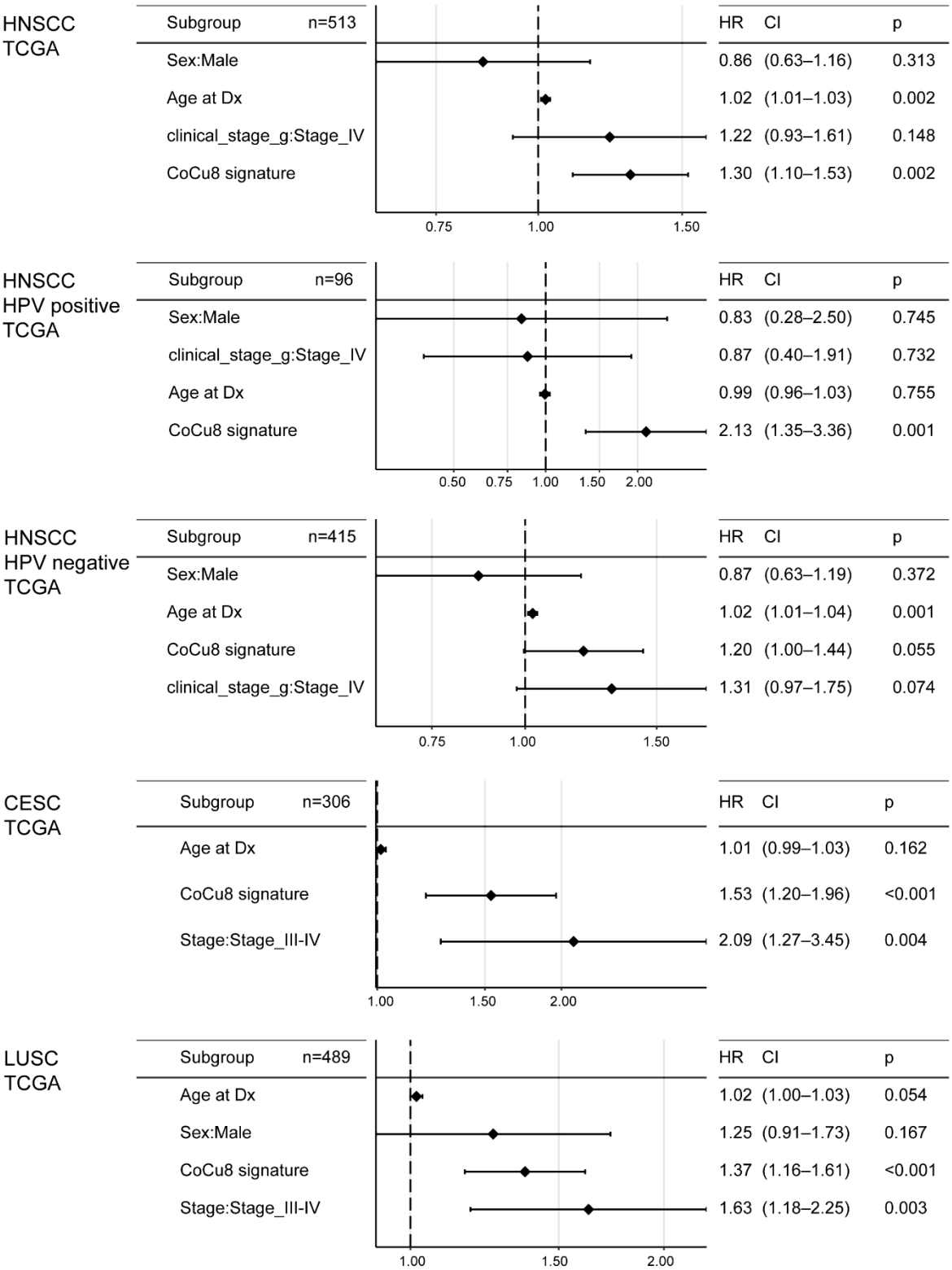


Figure S3

1129 **Figure S3:** Multivariate analysis of CoCu8 overall survival. Forest plot showing Hazard Ratios, 95%
1130 confidence interval and p value calculated using multivariate Cox regression from patients with HNSCC
1131 (HPV positive and negative), LUSC and CESC from the TCGA cohort. Variables include: age (continuous,
1132 years), sex (male vs female, except for CESC as all patients were female), clinical stage (categorical)
1133 and the CoCu8 signature (continuous variable).

CoCu30 Overall Survival

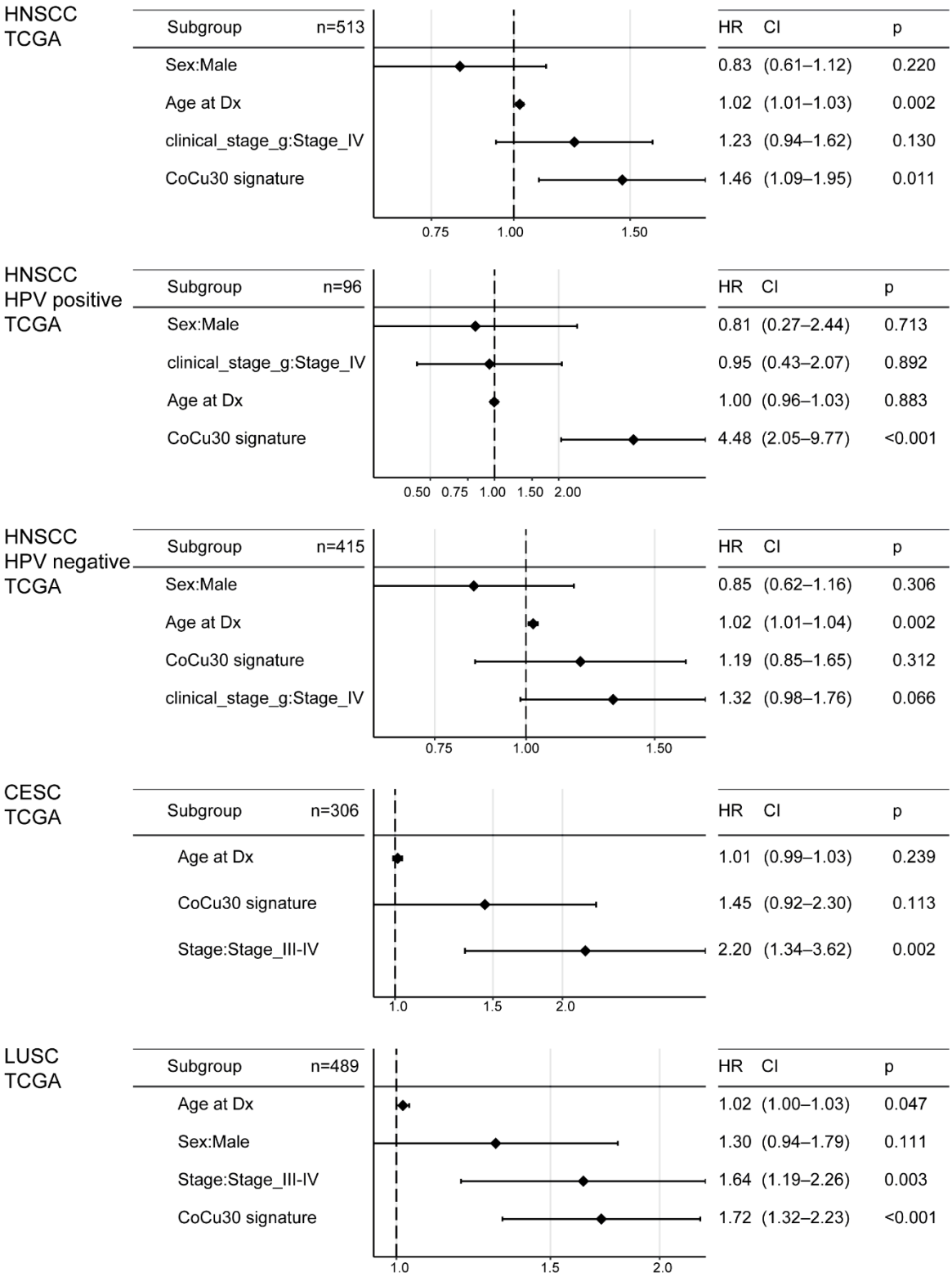


Figure S4

1135 **Figure S4:** Multivariate analysis of CoCu30 overall survival. Forest plot showing Hazard Ratios, 95%
1136 confidence interval and p value calculated using multivariate Cox regression from patients with HNSCC
1137 (HPV positive and negative), LUSC and CESC from the TCGA cohort. Variables include: age (continuous,
1138 years), sex (male vs female, except for CESC as all patients were female), clinical stage (categorical)
1139 and the CoCu30 signature (continuous variable).

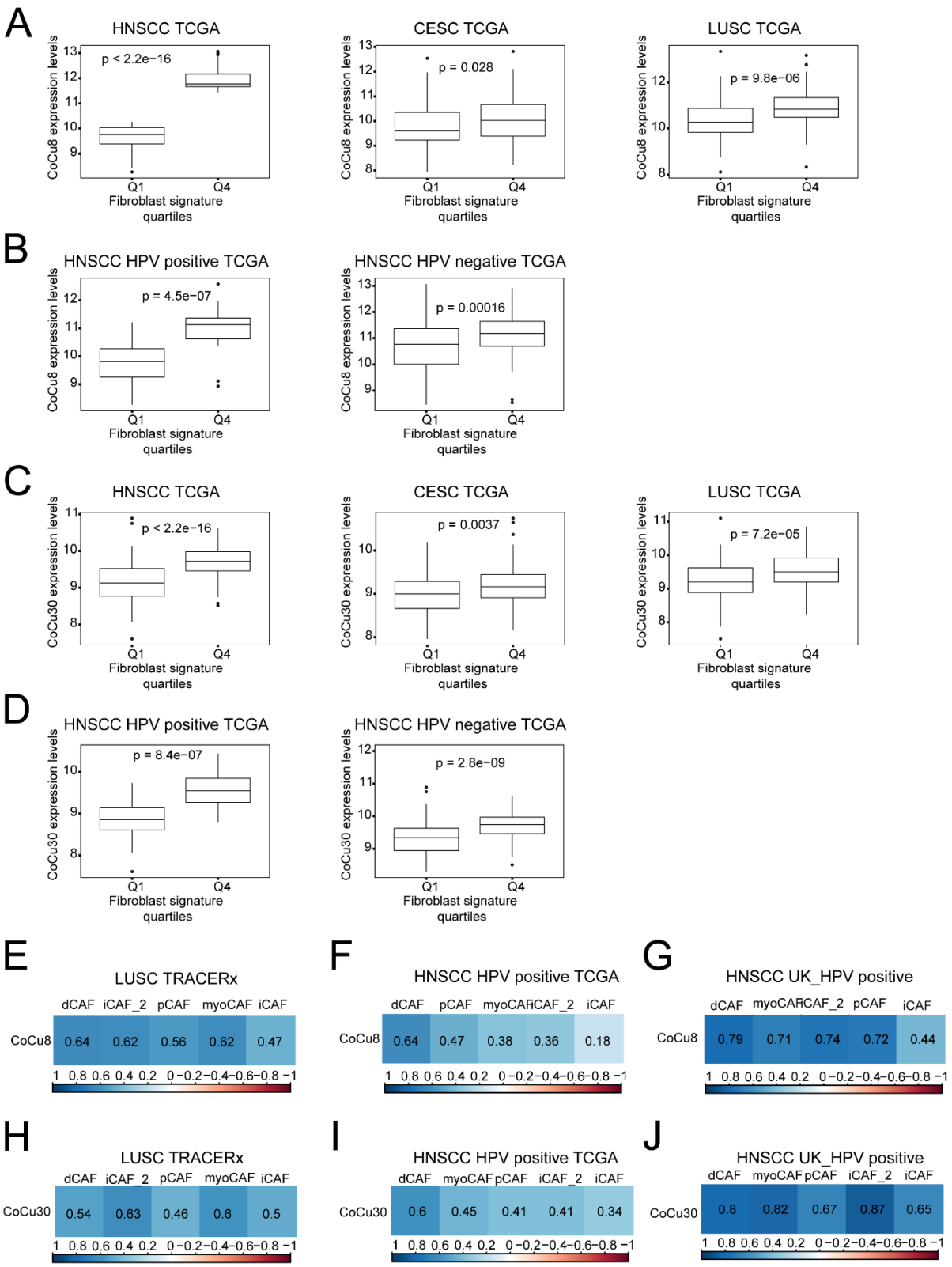

A
HNSCC TCGA
p < 2.2e−16
CESC TCGA
p = 0.028
LUSC TCGA
p = 9.8e−06
CoCu8 expression levels
Q1
Q4
Fibroblast signature quartiles
B
HNSCC HPV positive TCGA
p = 4.5e−07
HNSCC HPV negative TCGA
p = 0.00016
C
HNSCC TCGA
p < 2.2e−16
CESC TCGA
p = 0.0037
LUSC TCGA
p = 7.2e−05
CoCu30 expression levels
D
HNSCC 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

1142 **Figure S5:** Fibroblast abundance correlates with CoCu8 / CoCu30 gene signature in different squamous
1143 cell carcinoma datasets. **A)** Box plot analysis of CoCu8 expression in HNSCC, CESC and LUSC TCGA
1144 separated by first and last quartile of fibroblast abundance via Methyl CIBERSORT deconvolution
1145 strategy. Independent Student's t-test. **B)** Box plot analysis of CoCu8 expression in HPV positive and
1146 negative TCGA separated by first and last quartile of fibroblast abundance via Methyl CIBERSORT
1147 deconvolution strategy. Independent Student's t-test. **C)** Box plot analysis of CoCu30 expression in
1148 HNSCC, CESC and LUSC TCGA separated by first and last quartile of fibroblast abundance via Methyl
1149 CIBERSORT deconvolution strategy. Independent Student's t-test. **D)** Box plot analysis of CoCu30
1150 expression in HPV positive and negative TCGA separated by first and last quartile of fibroblast
1151 abundance via Methyl CIBERSORT deconvolution strategy. Independent Student's t-test. **E-J)**
1152 Correlation plot of different fibroblast subpopulations derived from Galbo et al. [9] with CoCu8 gene
1153 signature in LUSC TRACERx **(E)**, HPV positive HNSCC TCGA **(F)**, UK_HPV positive HNSCC **(G)** patients
1154 and with CoCu30 gene signature in LUSC TRACERx **(H)**, HPV positive HNSCC TCGA **(I)**, UK_HPV positive
1155 HNSCC **(J)** patients. The number inside the square represents the R, Spearman correlation coefficient.
1156 The color legend is shown at the bottom. All correlations are significant p-value<0.05.

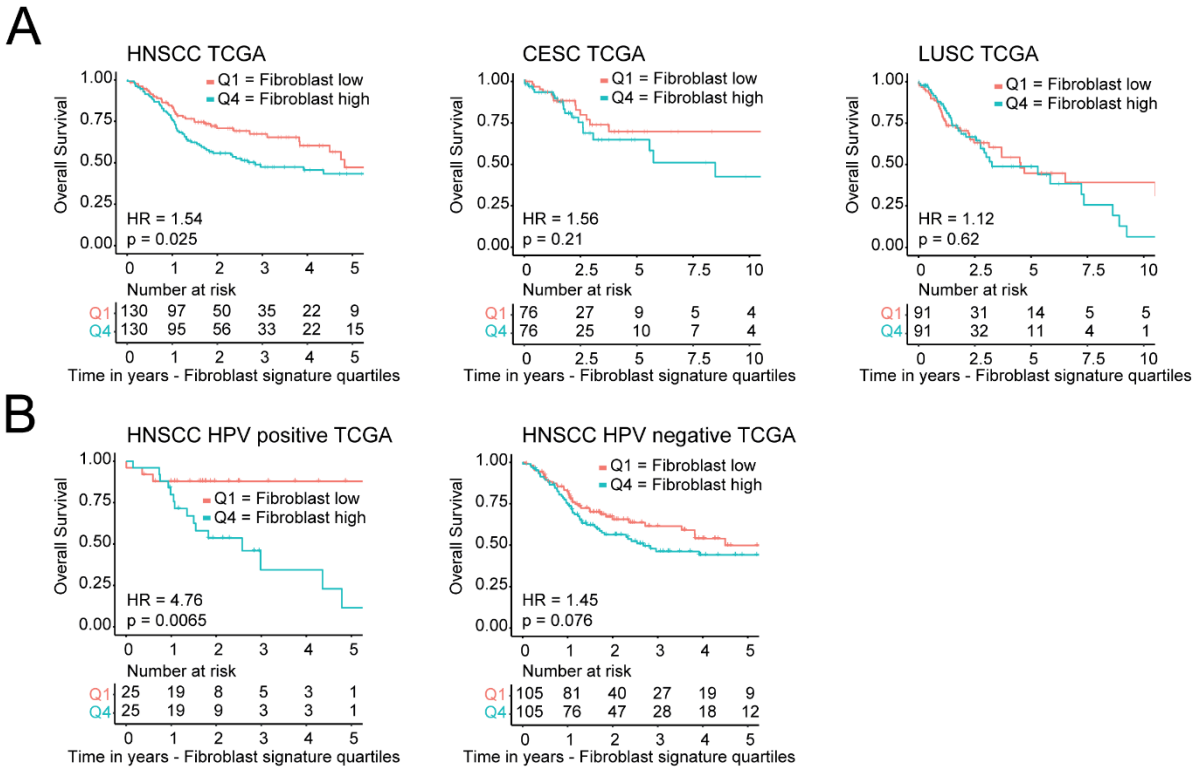
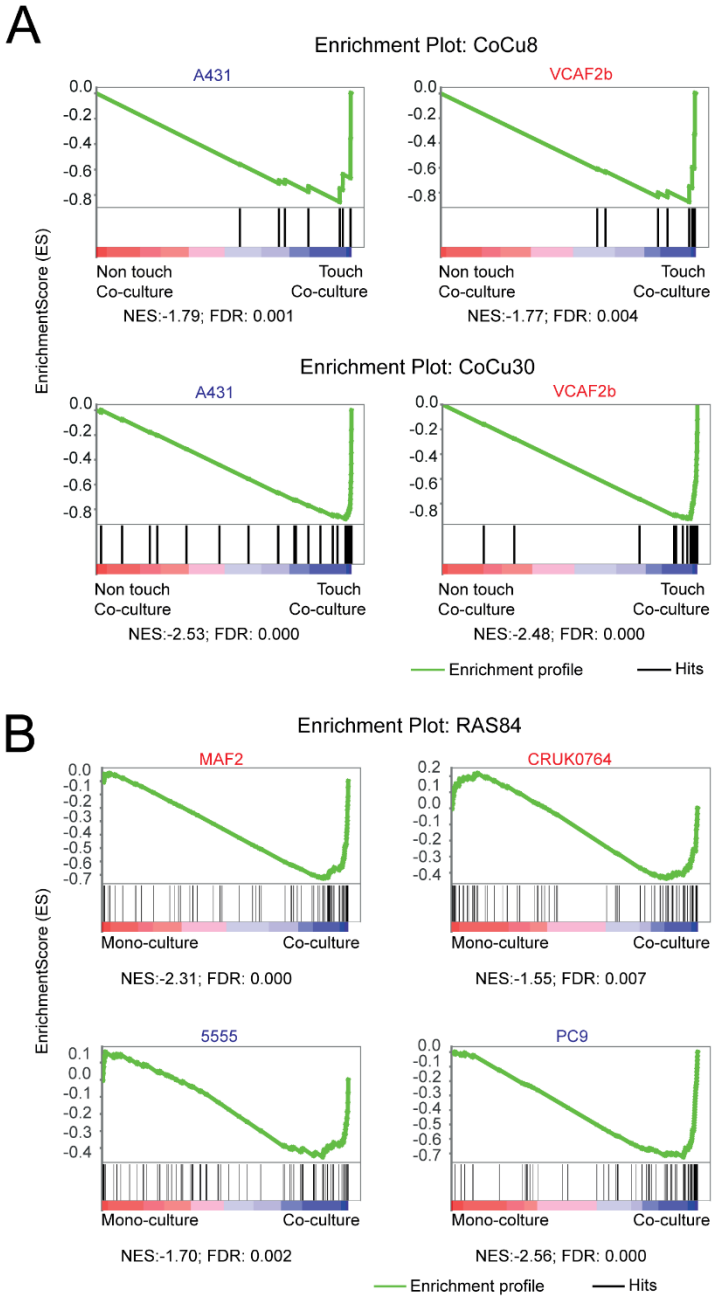

A
HNSCC TCGA
Q1 = Fibroblast low
Q4 = Fibroblast high
Overall Survival
HR = 1.54
p = 0.025
Number at risk
Q1 130 97 50 35 22 9
Q4 130 95 56 33 22 15
Time in years - Fibroblast signature quartiles
CESC TCGA
HR = 1.56
p = 0.21
Q1 76 27 9 5 4
Q4 76 25 10 7 4
LUSC TCGA
HR = 1.12
p = 0.62
Q1 91 31 14 5 5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

Figure S6

1158 **Figure S6:** Fibroblast abundance meta-analysis on different squamous cell carcinoma datasets. **A)**
1159 Kaplan-Meier overall survival analysis of HNSCC (left), CESC (center), LUSC (right) TCGA datasets
1160 stratified for fibroblast abundance first vs last quartile. Below each analysis are shown the
1161 corresponding numbers at risk, time in years. HNSCC HR=1.54 (95%CI 1.05-2.25), p-value=0.025. CESC
1162 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
1163 were calculated using Cox regression. p-value was calculated using logRank test. **B)** Kaplan-Meier
1164 overall survival analysis of HNSCC HPV positive (left) and HNSCC HPV positive (right) TCGA datasets
1165 stratified for fibroblast abundance first vs last quartile. Below each analysis are shown the
1166 corresponding numbers at risk, time in years. HPV positive HR=4.76 (95%CI 1.36-16.5), p-
1167 value=0.0065. HPV negative HR=1.45 (95%CI 0.96-2.18), p-value=0.076. HR and CI were calculated
1168 using Cox regression. p-value was calculated using logRank test.

1169



1170

Figure S7

1171 **Figure S7:** Enrichment of CoCu8, CoCu30 and RAS84 gene signatures in different co-culture conditions.
1172 **A)** Gene set enrichment analysis (GSEA) plot of CoCu8 gene signature (top) and CoCu30 (bottom) in
1173 co-culture indirect vs direct condition in A431 / VCAF2b transcriptomic dataset. NES and FDR are
1174 specified below each plot. **B)** Gene set enrichment analysis (GSEA) plot of RAS84 gene signature [28] in
1175 mono-culture and co-culture. NES and FDR are specified below each plot.

1176

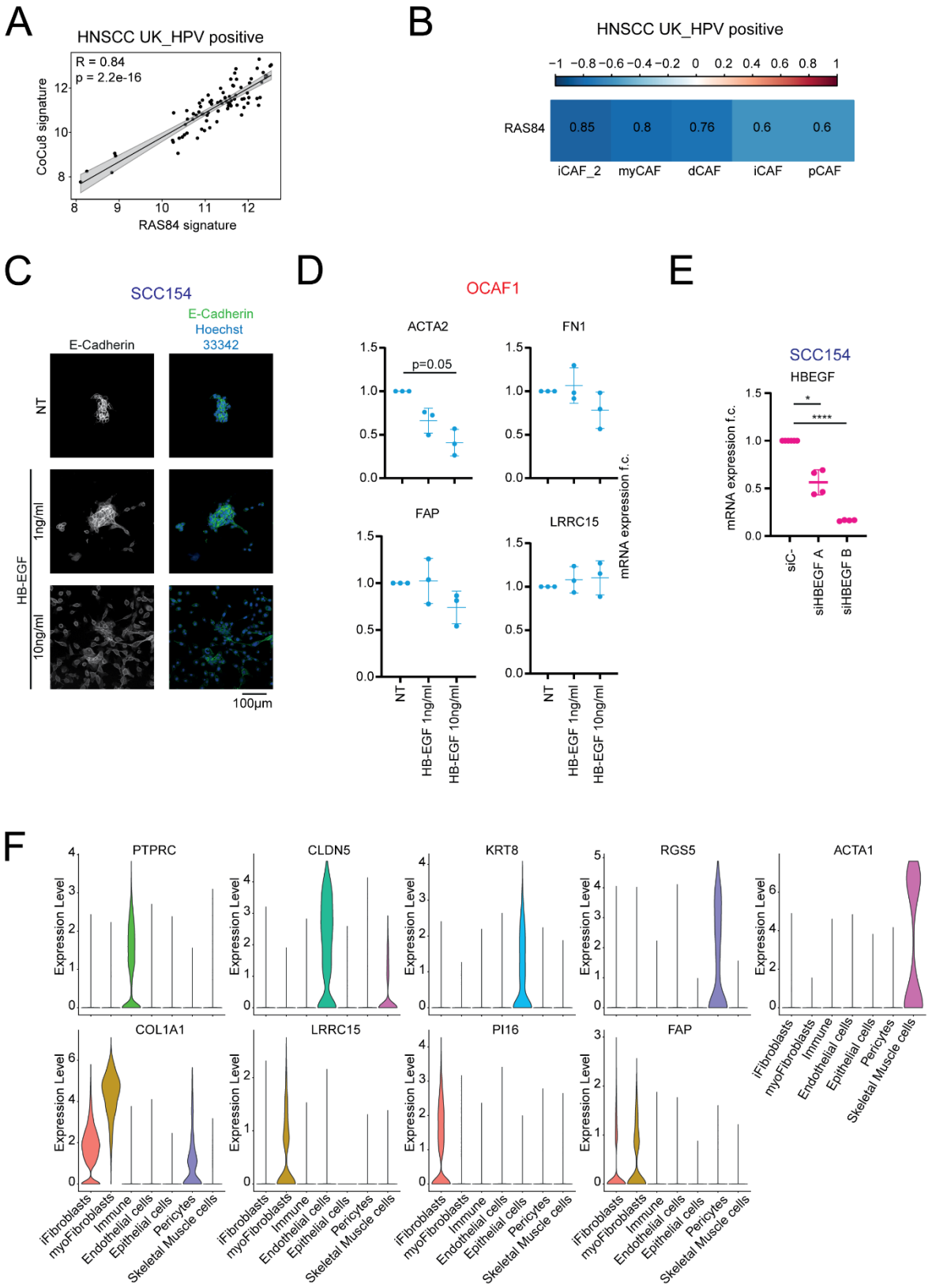


Figure S8

1177

1178 **Figure S8:** HB-EGF/RAS/MAPK activity in cancer cells and CAFs mono-culture vs co-culture. **A)**
1179 Correlation plot of CoCu8 and RAS84 expression levels in HNSCC HPV positive UK_HPV positive cohort.
1180 R is Spearman correlation coefficient. **B)** Correlation table of different fibroblast subpopulations
1181 derived from Galbo et al. [9] with RAS84 gene signature in the independent cohort UK_HPV positive of
1182 HNSCC HPV positive patients. The number inside the square represents the Spearman R, correlation
1183 coefficient. The color legend is shown on the top. All correlations are significant at p-value<0.001.
1184 n=97. **C)** Immunofluorescence staining of E-Cadherin and Hoechst 33342 for SCC154 mono-culture for
1185 the indicated treatments after 48h. **D)** qPCR analysis of *ACTA2*, *FAP*, *FN1* and *LRRC15* genes in OCAF1
1186 mono-culture for the indicated treatments after 48h. mRNA expression is reported as
1187 mean ± standard deviation (SD) fold change difference over non-treated (NT) condition. Genes have
1188 been normalized over the average of *GAPDH*, *ACTB* and *RPLP0* housekeeping genes. n = 3 independent
1189 experiments. **E)** qPCR analysis of *HBEGF* gene in SCC154 after 96h of treatment with the indicated
1190 conditions. mRNA expression is reported as mean ± standard deviation (SD) fold change difference
1191 over siC- condition. Gene has been normalized over the average of *GAPDH*, *ACTB* and *RPLP0*
1192 housekeeping genes. n ≥ 4 independent experiments. Two tailed paired Student's t-test. **F)** Violin plot
1193 of mRNA expression levels of the indicated genes for each cluster from Choi et al. [33].

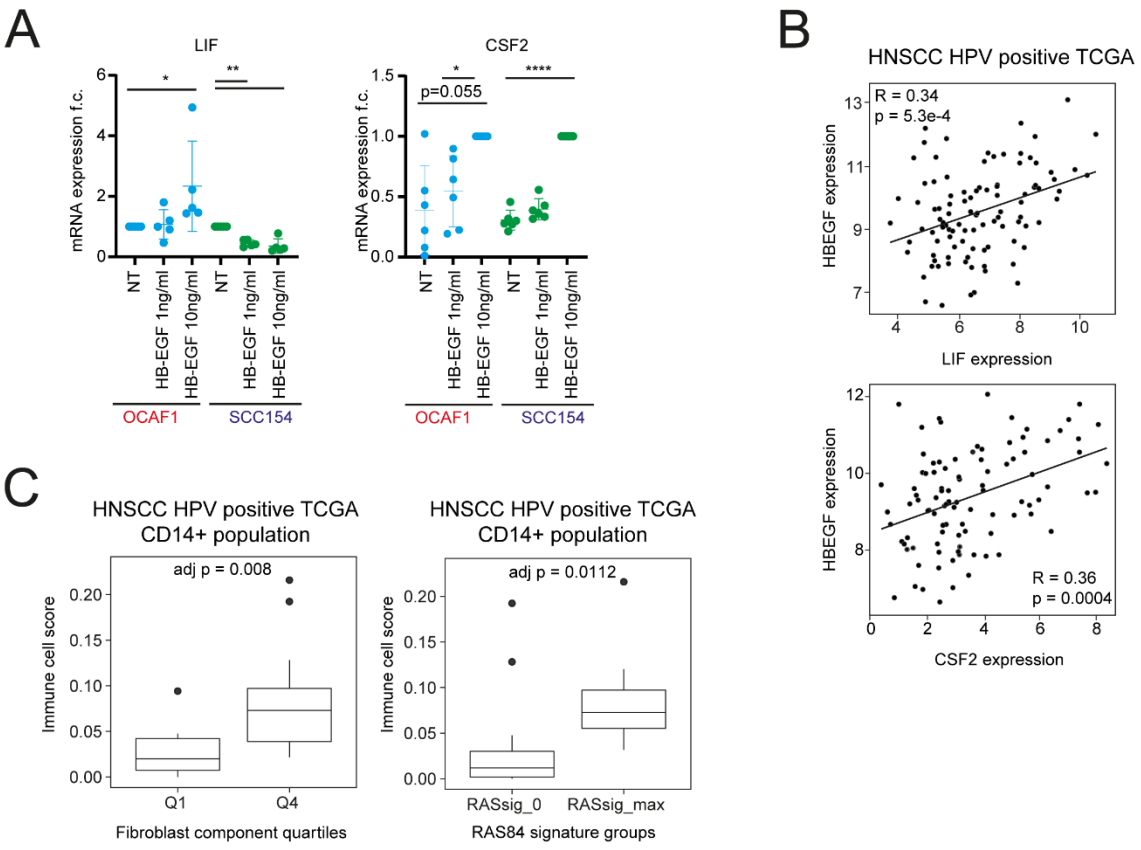

A
LIF
CSF2
mRNA expression f.c.
p=0.055
NT
HB-EGF 1ng/ml
HB-EGF 10ng/ml
OCAF1
SCC154
B
HNSCC HPV positive TCGA
R = 0.34
p = 5.3e-4
HBEGF expression
LIF expression
R = 0.36
p = 0.0004
CSF2 expression
C
HNSCC HPV positive TCGA
CD14+ population
adj p = 0.008
adj p = 0.0112
Immune cell score
Q1
Q4
Fibroblast component quartiles
RASsig_0
RASsig_max
RAS84 signature groups

Figure S9

1195 **Figure S9:** HB-EGF effect on cytokine production and monocyte/macrophage enrichment. **A)** qPCR
1196 analysis of *LIF* and *CSF2* genes in OCAF1 and SCC154 mono-cultures for the indicated treatments after
1197 48h. mRNA expression is reported as mean ± standard deviation (SD) fold change difference over non-
1198 treated (NT) condition for each cell type with *LIF*, while for *CSF2* fold change difference is reported
1199 over Hb-EGF 10ng/ml treated sample. Genes have been normalized over the average of *GAPDH*, *ACTB*
1200 and *RPLP0* housekeeping genes. n ≥ 5 independent experiments. Two tailed paired Student's t-test. **B)**
1201 Correlation plot of *HBEGF* with *LIF* (left) and *HBEGF* with *CSF2* (right) expression levels in HNSCC HPV
1202 positive TCGA dataset. R is Spearman correlation coefficient. **C)** (Right) Box plot analysis of immune
1203 cell absolute score via Methyl CIBERSORT deconvolution strategy in HNSCC HPV positive separated by
1204 first and last quartile of fibroblast abundance. Independent Student's t-test. Bonferroni correction for
1205 multiple comparisons. (Left) Box plot analysis of immune cell absolute score via Methyl CIBERSORT
1206 deconvolution strategy in HNSCC HPV positive separated by RAS84_0 and RAS84_max. Student's t-
1207 test Bonferroni correction for multiple comparisons.

1208

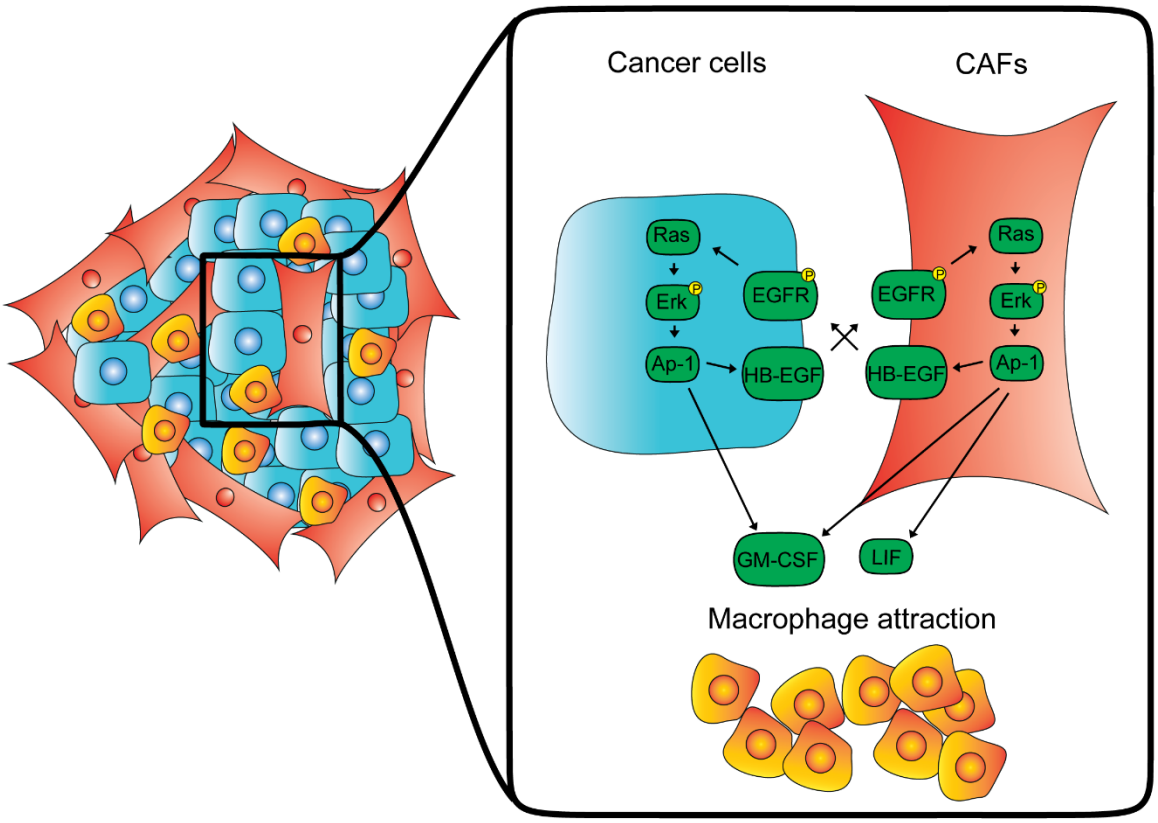


Figure S10

1209

1210

1211 **Figure S10:** Schematic representation of the current cross-talk model of cancer cells – CAFs co-culture.