HHS Public Access

Author manuscript

Nat Cancer. Author manuscript; available in PMC 2022 June 13.

Published in final edited form as:

Nat Cancer. 2022 January ; 3(1): 90–107. doi:10.1038/s43018-021-00291-9.

A tumor-derived type III collagen-rich ECM niche regulates tumor cell dormancy

Julie S. Di Martino[1], **Ana Rita Nobre**[2,3,7], **Chandrani Mondal**[1], **Isra Taha**[4,5], **Eduardo F. Farias**[1], **Elana J. Fertig**[6], **Alexandra Naba**[4,5], **Julio A. Aguirre-Ghiso**[2,8], **Jose Javier Bravo-Cordero**[1,*]

[1]Division of Hematology and Oncology, Department of Medicine, The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[2]Division of Hematology and Oncology, Department of Medicine and Department of Otolaryngology, Department of Oncological Sciences, Black Family Stem Cell Institute, Precision Immunology Institute, The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[3]Abel Salazar Biomedical Sciences Institute, University of Porto, Porto, Portugal

[4]Department of Physiology and Biophysics, University of Illinois at Chicago, Chicago, IL, USA

[5]University of Illinois Cancer Center, University of Illinois at Chicago, Chicago, IL, USA

[6]Departments of Oncology, Applied Mathematics and Statistics, and Biomedical Engineering, Johns Hopkins University, Baltimore, MA, USA

[7]Current affiliation: Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[8]current affiliation: Department of Cell Biology, Cancer Dormancy and Tumor Microenvironment Institute, Gruss Lipper Biophotonics Center, Albert Einstein Cancer Center, Albert Einstein College of Medicine. Price Center for Genetic and Translational Medicine Building Room 220, 1300 Morris Park Avenue, Bronx, NY 10461

Abstract

*Corresponding author: Jose Javier Bravo-Cordero josejavier.bravo-cordero@mssm.edu.

Author Contributions Statement:

JDM designed and performed experiments, analyzed, and interpreted the data, assembled the figures, and contributed to writing and editing of the manuscript. ARN performed cell sorting experiments. CM performed lung metastasis experiments in MDA-MB-231 xenografts. EF performed the first mouse tumor surgery experiments. EJF performed the RNAseq analysis. AN and IT performed the decellularization and mass-spectrometry analysis of the tumor samples and contributed to the data interpretation. JAAG contributed to designing and interpreting experiments and provided HEp3 cellular models. JJBC coordinated the study, contributed to designing and interpretation of the experiments and to the writing of the manuscript.

Competing Interests Statement: EJF is member of the Scientific Advisory Board of Viosera Therapeutics. JAAG is a scientific co-founder of, scientific advisory board member and equity owner in HiberCell and receives financial compensation as a consultant for HiberCell, a Mount Sinai spin-off company focused on therapeutics that prevent or delay cancer recurrence. The other authors have no conflict of interest to disclose.

Peer review Information:

Nature Cancer thanks Edna Cukierman, and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

Cancer cells disseminate and seed in distant organs, where they can remain dormant for many years before forming clinically detectable metastases. Here we studied how disseminated tumor cells (DTCs) sense and remodel the extracellular matrix (ECM) to sustain dormancy. ECM proteomics revealed that dormant cancer cells assemble type III collagen-enriched ECM niche. Tumor-derived type III collagen is required to sustain tumor dormancy as its disruption restores tumor cell proliferation through DDR1-mediated STAT1 signaling. Second harmonic generation two-photon microscopy further revealed that the dormancy-to-reactivation transition is accompanied by changes in type III collagen architecture and abundance. Analysis of clinical samples revealed that type III collagen levels were increased in tumors from lymph node-negative head and neck squamous cell carcinoma (HNSCC) patients compared to lymph node-positive patients. Our data supports that the manipulation of these mechanisms could serve as a barrier to metastasis through DTCs dormancy induction.

One Sentence Summary:

‘Bravo-Cordero and colleagues demonstrate that disseminated tumor cells remodel the extracellular matrix by secreting Collagen III and generate a stromal architecture that favors dormancy through DDR1 and STAT1 signaling

Keywords

Dormancy; tumor microenvironment; extracellular matrix; matrisome; type III collagen; DDR1; multiphoton imaging

INTRODUCTION

Most cancer deaths are due to metastases [1]. Metastases can occur several years after primary tumor removal. This delay is a consequence of the activation of pathways that maintain disseminated tumor cells (DTCs) in a non-proliferative state, also called cellular dormancy [2–4]. Cellular dormancy is defined as a reversible growth arrest (G0/G1) that can be regulated by the interaction of DTCs with the microenvironment [3]. Tumor mass dormancy is another type of dormancy where metastatic outgrowth is suppressed by a balance of similar rates of proliferation and apoptosis [5–8]. In the present study, we focus on cellular dormancy.

Microenvironmental conditions, such as increased hypoxia [9] as well as upregulation of signaling factors, including many that are extracellular matrix (ECM) bound (i.e. TGFβII [10], BMP4, BMP7 [2,11,12], IFN-γ [13]), have been involved in dormancy induction in solid tumors. However, how tumor cells utilize ECM-derived cues to induce and maintain dormancy is largely unknown.

The ECM is a meshwork of proteins which can anchor tumor cells and provide signals that regulate tumor cell behavior during metastasis [14], showing dynamic reciprocity between tumor cells and the ECM [15,16]. The three-dimensional architecture of the ECM has been shown to correlate with the invasiveness and metastatic potential of tumors [17–19]. Recent proteomics studies have also profiled “the matrisome” (a collection of ECM and ECM-related proteins), of both primary tumors and metastases [20–23]. These studies highlighted

that tumor-cell-derived ECM molecules have important roles in metastatic progression [21,24–26].

Importantly, the ECM also plays a role in determining the fate of DTCs. ECM proteins such as tenascin-C [27], type I collagen [28], or fibronectin [29] act as pro-metastatic cues, in addition to being structural components of the metastatic niche. The perivascular niche contains ECM molecules that signal to maintain dormancy, including thrombospondin 1 [30]. Recently, proteolysis of the ECM protein laminin-111 by the matrix metalloproteinase MMP9 secreted by neutrophils was shown to interrupt dormancy and drive metastatic outgrowth through integrin α3β1 [31], suggesting that ECM remodeling can activate metastatic growth after a dormancy period. Although these studies suggest that ECM molecules and ECM remodeling can drive dormant cell reactivation, how specific tumor cell-derived ECM molecules affect DTCs dormancy *in vivo* is not well explored.

Tumor cells can sense the ECM through transmembrane ECM receptors. While the role of integrin receptors during tumor progression has been described [32], the role of other ECM receptors, such as DDR1 (Discoidin Domain Receptor Tyrosine Kinase) and DDR2 collagen receptors, is not as well understood. Recent work has shown that in MMTV-PyMT (mouse mammary tumor virus-polyoma middle T) mammary tumors, downregulation of *DDR1* induced aggressiveness and metastasis [33]. Interestingly, DDR1 has been involved in reactivation after dormancy [34], and has been identified as part of a tumor cell quiescence signature [35,36] that predicted for dormancy and late relapse in ER+ (Estrogen Receptor-positive) breast tumors [37], suggesting it may be involved in initiating and/or sustaining tumor cell dormancy, although this has not been explored.

The induction of cellular dormancy is regulated by several intracellular molecular pathways. The balance of ERK and p38 activation downstream of fibronectin and adhesion signaling was one of the earliest mechanisms describing cellular dormancy [29,38]. Other pathways involving transcription factors such as *NR2F1* [36], *SOX9, RARβ, SOX2* [36,39], SMAD1/5 [10], or STAT1 [13] have been linked to dormancy initiation/activation suggesting that the molecular landscape of dormancy may include coordination of multiple signaling nodes.

In this study, we address the role of tumor-derived ECM in initiating and sustaining cancer cell dormancy and how alterations in this mechanism may drive metastatic outgrowth. By using a combination of ECM proteomics, multiphoton imaging, and a cell cycle biosensor, we have defined the matrisome of dormant cells and identified type III collagen as a key ECM protein required to induce and sustain dormancy *in vivo*. Type III collagen activates a novel signaling mechanism of tumor cell dormancy involving *DDR1/STAT1* signaling to establish a type III collagen-enriched ECM niche.

RESULTS

Changes in ECM collagen orientation characterizes the dormancy-to-reactivation transition

First, we sought to determine the architectural organization of the ECM surrounding dormant and proliferative cancer cells. To do so, we used several established cellular

dormancy models and their proliferative counterparts: proliferative (T-HEp3) and dormant (D-HEp3) human head and neck squamous cell carcinomas (HNSCC) [10,35,40], and two murine models of proliferative and dormant mammary tumors (4T1 and D2.A1 tumors and respective dormant counterparts 4T07 [11,41] and D2.OR [28,31]). These models have been widely used by different laboratories as representative of proliferative and dormant states *in vivo*. When inoculated on chicken chorioallantoic membranes (CAM), a well-characterized *in vivo* model to study tumor dormancy [9,36,42], or in mice [28,36] D-HEp3 and D2.OR are indolent and do not grow tumors (Extended Data Fig. 1A, B). Instead, they form small nodules where most of the cells enter in G0 phase [43]. On the other hand, D2.A1 and T-HEp3 are highly proliferative, forming tumors *in vivo*. When injected intravenously into mice, dormant D-HEp3 and D2.OR cancer cells are capable of extravasation but remain as single cells or small clusters for many weeks in the lungs [44].

To analyze the ECM architecture surrounding indolent dormant nodules and proliferative tumors, we performed multiphoton second harmonic generation (SHG) imaging and measured collagen fiber orientation. Our results show that the ECM within dormant nodules is mainly composed of wavy collagen fibers across the three different dormancy models. Overall, there is a significantly lower degree of linear organization of collagen fibers in dormant nodules when compared to proliferative tumors (Fig. 1A–C). Interestingly, analysis of tissue microarrays by SHG multiphoton imaging reveals that ECM orientation in HNSCC tumors is more aligned than in normal tissues (Extended Data Fig. 1C, **top right**). Moreover, ECM orientation significantly discriminates between early (I to III) and advanced (IV) stages in these tumors suggesting a higher degree of ECM alignment in advanced HNSCC tumors compared to early-stage indolent ones (Extended Data Fig. 1C, **bottom right**).

Given the architectural differences observed between the ECM organization of dormant nodules and proliferative tumors, we decided to explore the ECM organization around dormant solitary cells *in vivo*. After T-HEp3 primary tumor growth and surgical removal, residual dormant cells reside in the surgical margins and give rise to local recurrences after a period of time [36]. Multiphoton SHG imaging of residual cancer cells three weeks after primary tumor surgery shows that collagen fibers have a low degree of linear orientation around individual tumor cells (Fig. 1D), similar to the phenotype described in dormant nodules (Fig. 1A–C). To assess the dormant state of these solitary cancer cells, we used a CDK2 biosensor (DHB-mVenus) that allows identification of each phase of the cell cycle using the fluorescent reporter mVenus (Fig. 1E). The fluorescent mVenus sensor shuttles between the nucleus and the cytoplasm upon CDK2 activation and consequent CDK2 phosphorylation [45]. As a result, G0/G1 cells present nuclear fluorescence, while fluorescence is equally distributed across the cell in S phase and excluded from the nucleus in the G2 phase preceding mitosis. Image analysis of residual cancer cells in the surgical margins and spontaneously DTCs in the lungs showed that solitary cells are in G0/G1, confirming that cells surrounded by an ECM with a low degree of linear orientation are dormant (Fig. 1F). On the other hand, micrometastases show a mixed population of cancer cells at different phases of the cell cycle confirming their proliferative status. The cell cycle progression of these micrometastases correlates with a shift in ECM organization toward

a more aligned ECM. These ECM architectures are similar to the one observed in the dormancy models (Fig. 1A, B, C).

Our SHG imaging analysis has demonstrated a change in the ECM between solitary cells and metastatic outgrowing clusters. This architectural rearrangement can occur swiftly as aligned ECM can be visualized in the primary site as quickly as three days after orthotopic engraftment of proliferative cancer cells by multiphoton imaging through an implanted imaging window *in vivo* (Extended Data Fig. 1D). The increase of collagen fiber alignment around growing cells suggests that ECM remodeling correlates with an escape from dormancy in local and distant sites in human HNSCC and breast cancer models (Fig. 1F, Extended Data Fig. 1D–E). Overall, these results show that collagen fibers around dormant cells are characterized by a low degree of linear orientation while a high degree of linear orientation is observed during cancer cell reactivation.

Dormant cells produce an ECM enriched in type III collagen

Transcriptomic analysis of indolent D-HEp3 dormant nodules and aggressive T-HEp3 tumors grown in mice show a significant dysregulation of the matrisome signature [46] (p=4.97e-05) (Fig. 2A). To investigate the matrisome composition of these dormant and proliferative tumors, we performed ECM-enriched mass spectrometry using decellularized T-HEp3 tumors and D-HEp3 dormant nodules grown in mice (Extended Data Fig.2A–C, Supplementary Table 1). This differential analysis revealed that the matrisome of dormant nodules is highly enriched in collagens; collagen peptides represent 55% of the matrisome in dormant nodules and only 36% of the matrisome in proliferative tumors (Fig. 2B). This enrichment in collagens was confirmed with Masson's trichrome and picrosirius red staining of D-HEp3 and T-HEp3 tissue sections and validated in the D2 mammary cancer model (Fig. 2C, Extended Data Fig. 2D). The human-in-mouse xenograft system used further allowed us to determine the relative contribution of tumor cells (human peptide sequences) and stromal cells (murine peptide sequences) [20] to the production of the matrisome (Fig. 2D). Tumor cell contribution to the ECM is as great as 6% of total ECM in dormant nodules and 17% in proliferative tumors (Extended Data Fig. 2E). Importantly, the analysis of the collagen composition of the tumor cell-derived ECM reveals that type III collagen is the most abundant collagen produced by dormant tumor cells and the only collagen enriched in dormant nodules when compared to proliferative tumors (Fig. 2E, Supplementary Table 1). Type III collagen is also enriched in the stromal compartment of dormant D-HEp3 nodules along with type VI collagen alpha 1 and 2 subunits (Fig. 2F,Supplementary Table 1). To distinguish between type I and type III collagen within tissue sections, we applied a picrosirius-polarization method (picrosirius red staining combined with polarized light detection). With this method, type I collagen is displayed in red-orange-yellow (thick fibers) and type III collagen in green (thin fibers) [47–50] (Fig. 2G). Using this method, we quantitatively validated the higher abundance of total collagen in D-HEp3 nodules compared with T-HEp3 tumors. Further analysis of polarized-light images confirms a higher abundance of type III collagen in the dormant nodules compared with proliferative tumors. Consistently, immunofluorescence staining for type III collagen of tumor sections shows an increased abundance in D-HEp3 nodules as compared to proliferative T-HEp3 tumors (Fig. 2H, S2F). A human-specific vimentin antibody was used to distinguish tumor cells from

stromal cells. Furthermore, type III collagen staining can be found surrounding single DTCs in lungs and significantly decreased around Ki67+ micrometastasis (Fig. 2I). All together, these data show that the ECM of dormant tumors is enriched in collagens, with type III collagen being the most abundant collagen ECM protein produced by dormant cancer cells.

Finally, we analyzed HNSCC tissue microarrays stained for type III collagen and found that patients with positive lymph nodes (N+) present lower concentrations of type III collagen in their primary tumor compared with patients with non-colonized lymph nodes (N0) (Fig. 2J, Extended Data Fig. 2F). These data suggest that DTCs originating from HNSCC tumors low for type III collagen expression may be able to initiate metastasis more efficiently in lymph nodes.

Type III collagen induces tumor cell dormancy

Previous studies have shown that stromal type III collagen restricts metastasis formation from mammary tumors [51], but this phenotype was not explored in the context of dormancy regulation. Given the high abundance of type III collagen in dormant D-HEp3 tumors and in samples from N0 HNSCC patients, we asked whether increasing type III collagen concentration in the microenvironment of highly proliferative tumors could drive proliferative cancer cells into dormancy and prevent their growth.

We first performed a series of *in vivo* experiments using the CAM assay (Extended Data Fig. 1A). T-HEp3 cells were inoculated on CAMs with type I, III, IV collagen or DPBS (vehicle control). We observed that T-HEp3 cells co-injected with type III collagen grew significantly smaller tumors than any other collagens tested or the vehicle (Fig. 3A). Then, we injected T-HEp3 cells in the right flank of nude mice and T-HEp3 cells with type III collagen contralaterally and observed significantly slower growth in the presence of type III collagen (Fig. 3B). Masson Trichrome staining of both tumors at the time the mice were sacrificed revealed a higher collagen content in the type III collagen co-injection site (Extended Data Fig. 3A) and a lower degree of orientation of collagen fibers in SHG (Extended Data Fig. 3B). Similar results were obtained for growth in mice orthotopically injected with 4T1 or D2A1 (Extended Data Fig. 3C–D) co-injected with type III collagen. Cancer cells stopped proliferating as demonstrated by a decrease in phospho-histone H3 protein levels (Fig. 3C). Importantly, the growth inhibition is not associated with an increase in apoptosis as shown by TUNEL assays *in vivo* (Fig. 3D) and Annexin V analysis *in vitro* (Extended Data Fig. 3E) but rather the induction of dormancy as shown by the increase in nuclear p27 dormancy marker [9,36] (Fig. 3E) and decreased proliferation as shown by live cell imaging of T-HEp3 cells expressing the CDK2 cell cycle sensor (Fig. 3F,Supplementary Movies 1 and 2). A decrease in proliferation is also observed *in vitro* using D-HEp3 on type III collagen matrices when compared with type I collagen (Extended Data Fig. 3F,SupplementaryMovies 3 and 4). Concentrations of type III collagen equal or greater than 1mg/ml inhibit cell proliferation of T-Hep3 *in vitro* (Fig. 3G). In addition, we find that type III collagen treatment has no effect on cytotoxicity, apoptosis, or viability (Extended Data Fig. 3G).

Next, we sought to determine whether treating local residual cancer cells (observed post-resection of the primary tumor) with a type III-collagen-loaded bioengineered scaffold could

prevent local relapses. To do so, we used T-HEp3 cells expressing a CDK2 biosensor and injected them orthotopically in mice. At the time of surgery, a type-III-collagen-loaded scaffold (or DPBS-loaded control) was applied to the wound area. Mice were then monitored for local relapse. Only 20% of the mice treated with type III collagen scaffolds present recurrences after tumor surgery, while 80% of tumors relapse in the control group (Fig. 3H). Intravital imaging at the time of recurrences reveals that mice treated with the type III collagen scaffold present fewer cancer cells. These few remaining cells are growth arrested in G0 (strictly nuclear localization of the CDK2 sensor). On the other hand, mice treated with the control scaffolds develop growing masses with cells at different phases of the cell cycle, confirming the growing phenotype of local relapses (Fig. 3I).

Tumor cell-derived type III collagen sustains DTCs dormancy

As shown in our proteomics analysis, both stromal cells and cancer cells produce type III collagen (Fig. 2E, F). To understand the contribution of stroma-derived and tumor-derived ECM to the dormant state of these cancer cells, we established an *in vivo* co-culture system of fibroblasts expressing either control shRNA or an shRNA targeting COL3A1 and D-HEp3 cells expressing a CDK2 sensor. When inoculated alone in CAMs, fibroblasts depleted for *COL3A1* show ECM reorganization as fast as 24 hours after seeding. *COL3A1*-depleted fibroblasts are elongated and aligned with a straight collagen ECM as observed by SHG multiphoton imaging (Extended Data Fig. 4A). We quantified the percentage of D-HEp3 cells in each phase of the cell cycle when seeded with either control or type III collagen-depleted fibroblasts and found no difference in the distribution of cells in the different phases of the cell cycle (Fig. 4A). This suggests that once cancer cells have activated their dormancy program, they are no longer sensitive to external type III collagen produced by other cell types and can remain dormant even when the microenvironment is not supplying type III collagen.

Interestingly, there is a direct correlation between tumor cell-derived type III collagen abundance identified with ECM-enriched proteomics (Fig. 2E) and *COL3A1* expression in D-HEp3. *COL3A1* is among the most upregulated collagen genes in dormant D-HEp3 as compared to proliferative T-HEp3 (Fig. 4B–4D).

To determine if *COL3A1* expression by cancer cells is necessary to sustain cancer cell dormancy, we knocked down *COL3A1* in dormant D-HEp3 cells and performed *in vivo* growth experiments. Depletion of *COL3A1* interrupts dormancy and restores the growth of D-HEp3 in CAMs (Fig. 4E). *COL3A1* depletion is accompanied by an increase in phospho-histone H3 (Fig. 4F) in the HEp3 model and E-Cadherin expression as shown by increased staining in cell-cell junctions in the D2 model (Extended Data Fig. 4B).

COL3A1-depleted dormant cells injected into the tail vein of mice show a significantly higher proportion of cells proliferating as well as an increase in metastasis formation when compared with control (Fig. 4G). Interestingly, depletion of other collagen chains (*COL1A1, COL1A2, COL5A1, COL5A2, COL5A3, COL6A1, COL6A2, COL6A3, COL16A1, COL18A1*) that were also overexpressed in dormant cells have no discernible effect on disrupting tumor dormancy *in vivo* (Extended Data Fig. 4C).

To assess the role of tumor-derived COL3A1 in inducing dormancy, we designed a human *COL3A1* overexpression inducible vector that increases COL3A1 levels upon doxycycline stimulation (Fig. 4H). When T-HEp3 were treated with doxycycline *in vivo*, smaller tumors grow than the untreated controls. *COL3A1* overexpression was accompanied by a decrease in phospho-histone H3 phosphorylation. CDK2 sensor analysis also shows a higher proportion of G0 cells in the doxycycline treated condition overexpressing *COL3A1*. Collectively, these results show that overexpression of type III collagen in proliferative tumor cells drives cancer cells into dormancy.

The changes in type III collagen content upon *COL3A1* depletion in tumor cells are accompanied by a significant increase in collagen fiber alignment (Fig. 4I), suggesting that tumor-derived type III collagen (along with stromal derived type III collagen as shown in Extended Data Fig. 4A) contributes to the wavy ECM organization observed in dormant nodules. These results unveil a role for tumor-derived COL3A1 in inducing and sustaining dormancy of cancer cells and highlight the contribution of cancer cells to ECM remodeling and architecture.

DDR1 sustains dormancy through the regulation of type III collagen and the matrisome signature

Given the dependence of cancer cells on type III collagen to induce and sustain their quiescent state (Fig. 4) and the dormancy-inducing function of type III collagen (Fig. 3), we next determined which collagen receptors bind to and activate type III collagen-mediated signaling to trigger dormancy.

Dormant D-HEp3 cells are more adherent to collagens than T-HEp3 and bind more efficiently to type III collagen than type I or type IV collagens, while T-HEp3 adhere more to fibronectin (Fig. 5A, Extended Data Fig. 5A). Interestingly, when dormant cells lack expression of COL3A1 they lose their affinity to bind type III collagen ECM (Fig. 5B, Extended Data Fig. 5B). This is accompanied by decreased expression of the DDR1 collagen receptor (previously identified as part of a dormancy signature in breast cancer [37]) (Fig. 5C). In fact, DDR1 is overexpressed in HEp3 and D2 tumor dormancy cellular models compared with their proliferative counterparts (Fig. 5D).

To evaluate the role of DDR1 in dormancy of cancer cells we performed *DDR1* knockdown experiments (Fig. 5E, Extended Data Fig. 5C). In both dormancy models (D-HEp3, D2.0R) we found an outgrowth of *DDR1*-depleted cells when compared with control in CAM assays. Similar results were obtained using a dormant bone marrow-derived cancer cell clone *in vivo* (BM-HEp3) (Extended Data Fig. 5D). No changes in DDR2 expression were observed upon *DDR1* knockdown (Fig. 5E, Extended Data Fig. 5C).

We examined D-HEp3 expressing a control or a *DDR1* targeting shRNA for their ability to reinstate cancer cell growth in the lungs of mice. Tail vein injections of D-HEp3 cells depleted for *DDR1* form metastasis after three weeks, while control D-HEp3 cells remain as solitary or small clusters of cells (Fig. 5F). Interestingly, macroscopic metastases were identified in 60% of animals injected with *DDR1* depleted D-HEp3 cells (Fig. 5F, **right panel**). A concomitant increase in phospho-histone H3 levels (Fig. 5G) and a decrease in

G0 cells (Extended Data Fig. 5E) is observed upon *DDR1* depletion, confirming the role of DDR1 in sustaining quiescence.

In vivo, rescue experiments in CAM assays using a full-length DDR1 construct confirmed that DDR1 is necessary to sustain dormancy of D-HEp3 cells (Fig. 5H, Extended Data Fig. 5F).

Rescue experiments revealed that the dormancy phenotype cannot be rescued when D-HEp3 depleted for DDR1 overexpress a binding deficient mutant (W53A); however, the phenotype is restored with a kinase dead mutant (K655A) (Extended Data Fig. 5G). These results highlight the importance of the DDR1 collagen binding domain in sustaining dormancy. Inhibiting DDR1 kinase activity using Nilotinib does not allow dormancy escape, confirming the kinase domain is not required to sustain dormancy (Extended Data Fig. 5H). Importantly, the binding of D-HEp3 cells to type III collagen but not type I or type IV is DDR1-dependent as shown by an adhesion assay (Fig. 5I).

Given these results, we hypothesized that increasing *DDR1* levels in proliferative *DDR1*^{low} cells may induce dormancy. To test this hypothesis, we overexpressed full-length DDR1b in T-HEp3 cells and measured tumor growth *in vivo*. A decrease in tumor growth was observed upon *DDR1* overexpression with no induction of cell death (Fig. 5J, Extended Data Fig. 5I). On the other hand, depletion of *DDR1* in T-HEp3 cells prevented tumor growth (Extended Data Fig. 5J) and increased apoptosis (Extended Data Fig. 5K), suggesting that DDR1 may act as a pro-survival signal in proliferative cells. T-HEp3 treatment with Nilotinib showed similar results as DDR1 knockdown in T-HEp3 (Extended Data Fig. 5J), suggesting that the pro-survival role of DDR1 in proliferative cells is kinase-dependent (Extended Data Fig. 5L). Similar results are observed with reactivated dormant cells (R-HEp3) suggesting that dormant cells that restore growth also require DDR1 expression to survive (Extended Data Fig. 5M).

Together, these results suggest that DDR1 is necessary to enter and sustain dormancy in a type III collagen binding dependent and kinase-independent manner.

Based on previous roles assigned to DDR1 in collagen remodeling [52,53] and work showing DDR1 regulates gene expression of several ECM genes [54] we hypothesized that DDR1-driven dormancy could be mediated through regulation of the matrisome of dormant cells. *DDR1*-depleted tumors have a significant increase in collagen ECM alignment (Fig. 5K). Furthermore, proteomics and immunostaining analyses confirmed that type III collagen abundance is affected upon *DDR1* depletion *in vivo* (Fig. 5L). Interestingly, cells depleted for *DDR1* show a decrease in *COL3A1* mRNA expression (Fig 5M).

DDR1 knockdown has a significant effect on the expression of the matrisome gene sets (Extended Data Fig. 5N) as well as on the whole transcriptome of dormant cells, resembling the T-HEp3 transcriptome (Extended Data Fig. 5O).

These results reveal that DDR1 induces and sustains dormancy through type III collagen expression and ECM reorganization.

DDR1/STAT1 activation triggers type III collagen expression to sustain dormancy

As *DDR1* depletion induces a downregulation of *COL3A1* gene expression (Fig. 5M), we analyzed our RNA sequencing data to identify transcription factors that could regulate *COL3A1* expression downstream of DDR1. We analyzed transcription factors identified in RNA sequencing of T-HEp3, D-HEp3, D-HEp3 shCTRL and D-HEp3 shDDR1 (Supplementary Tables 5 and 6). We selected all transcription factors which were significantly overexpressed in dormant D-HEp3 compared with T-HEp3 as well as significantly downregulated upon DDR1 knockdown in reactivated cells compared with an shRNA control. Following this method, we identified the *STAT1, FOSB, IRF7, TRIM25, TLE4* transcription factors as potential regulators of dormancy downstream of DDR1 (Fig. 6A).

We performed *COL3A1* promoter analysis using CiiiDER [55] to identify predicted binding sites for the five identified transcription factors. Predicted binding sites were identified for STAT1, FOSB, IRF7 in the *COL3A1* promoter (Fig. 6B). Among them, only STAT1 seems to regulate *COL3A1* gene expression, as a significant downregulation of *COL3A1* is observed upon *STAT1* depletion in dormant cells but not upon *FOSB* or *IRF7* downregulation (Fig. 6C). Importantly, *STAT1* depletion restored growth of dormant D-HEp3 cells in CAMs and decreased DDR1 abundance at the protein level (Fig 6D). DDR1 promoter analysis using CiiiDER [55] revealed a STAT1 predicted binding site (Extended Data Fig. 6A). Similar results were obtained *in vivo* using a STAT1 inhibitor, Fludarabine (Fig. 6E).

Mice intravenously injected with *STAT1* depleted D-HEp3 cells expressing a CDK2 sensor show an increase in the proportion of small clusters of cells and a decrease in the number of single cells in the lung when compared to controls (Fig. 6F, Extended Data Fig. 6B).

Immunofluorescence analysis of D-HEp3 plated on glass or on type III collagen shows an increase in nuclear localization of phosphoSTAT1 when seeded on type III collagen (Fig. 6G). D-HEp3 expressing STAT1-GFP and H2B-RFP as a nuclear marker were seeded on CAMs in the presence of type III collagen or a DPBS control. After three days, most cells seeded with type III collagen in CAMs present a nuclear localization of STAT1 (Fig. 6H).

Interestingly, downregulation of either *DDR1* receptor or *COL3A1* induces a significant downregulation of *STAT1* expression at the RNA level (Extended Data Fig. 6C–D).

Finally, upon *STAT1* downregulation, we observed a shift in ECM organization from low-to-high degree of linear orientation (Fig. 6I), similar to previously observed upon *COL3A1* and *DDR1* knockdown (Fig. 4, 5).

Based on our observations, we propose a model where binding of DDR1 to type III collagen triggers STAT1 activation and nuclear translocation to regulate *COL3A1* expression. Increased expression of *COL3A1* remodels the ECM by increasing its curliness and drives the cells into a dormant state sustained by DDR1 binding

DISCUSSION

The study presented here investigates three underdeveloped areas of dormancy and metastasis biology: 1) the functional characterization of the matrisome produced by dormant and reactivated cancer cells that identifies type III collagen as a regulator of dormancy and relapse, 2) the identification of DDR1 as a mediator of type III collagen-derived dormancy signals and 3) the potential role of reprogramming metastatic niches by tuning the ECM to prevent reawakening of DTCs. Overall, our work describes a new mechanism of cellular dormancy, by which dormant cells assemble a pro-quiescence ECM enriched in type III collagen.

Our data show that dormant cancer cells secrete a unique matrisome that is enriched in type III collagen. By reducing the ability of dormant cells to produce type III collagen, their proliferative capacities are restored. Interestingly, enrichment for type III collagen in the tumor microenvironment drives tumor cells into dormancy, revealing a type III collagen-dependent mechanism controlling the fate of cancer cells.

Previous work showed that a decrease of type III collagen in the stromal microenvironment increases the aggressiveness of proliferative breast tumors [51], whereas upregulated *COL3A1* expression is associated with increased survival in breast cancer patients [56], suggesting that type III collagen can limit metastasis. ECM from *COL3A1*−/− fibroblasts and within tumors of *COL3A1*-deficient mice is denser and more aligned than in wild-type counterparts similar to our results obtained with dormant tumors and fibroblasts upon *COL3A1* depletion. Our data show that *COL3A1* depletion in the stromal or tumor compartment results in increased alignment of the collagen ECM, suggesting a synergistic effect between tumor-derived and stroma-derived ECM. Interestingly, *in vivo*, restoration of dormant cancer cells' growth is observed only when *COL3A1* is depleted from dormant cells, and not from fibroblasts. These data revealed that the stroma-derived type III collagen does not influence the dormant state, as cancer cells can remain dormant in a type III collagen low microenvironment as long as they express COL3A1. Our results argue thattype III collagen produced by dormant cells is sufficient to sustain their phenotype in a cell autonomous manner.

Early disseminated cancer cells (eDCC) in breast cancer models enter into prolonged dormancy before metastasis initiation. scRNAseq data of eDCC showed an upregulation of *COL3A1*[57], suggesting that *COL3A1* expression may contribute toprolonged dormancy of eDCC through mechanisms described in our study.

Previous reports showed that DDR1 is required to restore growth after a dormancy period [34]. Our data complement these observations and reveal a new function for DDR1 in sustaining dormancy as the receptor for type III collagen in dormant cells.

Altogether, our data show that cancer cells require DDR1-dependent kinase activity to proliferate and DDR1-kinase independent but collagen binding dependent activity to sustain dormancy. Additionally, STAT1 is activated upon type III collagen binding and regulates *DDR1* and *COL3A1* expression, further suggesting the activation of a signaling feedback loop that reinforces dormancy. Interestingly, recent work showed that DDR1 translocates to

the nucleus and binds to chromatin, regulating the transcription of COL IV [58], suggesting that additional DDR1-dependent mechanisms may regulate the ECM gene expression.

Overall, while several signals are demonstrated to contribute to tumor dormancy, including TGFβ2 [10] and the retinoic acid pathway [36], we believe that the COL3A1/DDR1/STAT1 pathway may act in parallel with these signals to reinforce tumor dormancy.

Our findings have potential clinical implications and may lead to a novel biomarker to predict tumor recurrences, as well as a therapeutic intervention to reduce local relapses. Previous studies using SHG imaging showed that breast tumors with poorly aligned collagen have a better prognosis and defined a series of signatures that can determine patient outcomes [18,19]. Our data now show that similar SHG measurements can be used to identify aggressive HNSCC based on collagen fiber orientation. Moreover, the reduced type III collagen expression in aggressive HNSCC tumors along with increased collagen fiber alignment could be used as potential measurements to predict tumor recurrence.

Finally, wound treatment with collagen scaffolds has shown promising results and has been proposed as a therapeutic alternative for complex skin wounds [59]. Our data show that by enriching the microenvironment in type III collagen, dormancy is activated, and tumor growth is suppressed. The intervention aimed at preventing the awakening of dormant cells has been suggested as a therapeutic strategy to prevent metastatic outgrowth [36,60,61] and has become part of a clinical trial (NCT03572387). Our studies demonstrate the potential therapeutic use of type III collagen to prevent the reawakening of cancer cells by inducing and maintaining cancer cell dormancy in the primary site; this may be particularly relevant for HNSCC and breast tumors, two of the models we explored. They also illustrate a way by which therapeutic targeting of pro-metastatic ECM by changing its composition to a pro-dormancy role can lead to the prevention of local recurrence or metastatic outgrowth. As the biology of tumor dormancy gets uncovered and new specific drugs are developed, a combination of dormancy-inducing treatments with therapies that specifically target dormant cells will ultimately prevent local recurrence and metastasis and pave the way to cancer remission.

MATERIALS AND METHODS

Further information on research design is available in the Nature Research Reporting Summary linked to this article. All mice procedures were ethically approved by the Institutional Animal Care and Use Committee (IACUC) of the Icahn School of Medicine at Mount Sinai.

Procedure involving chicken fertilized eggs were not subject to ethical approval as animals were sacrificed no later than day 16 post fertilization.

Cell culture

Al list of cell lines and appropriate culture media used is provided in Supplementary Table 7. All cell lines except the HEp3 cellular models were purchased for the study and IMPACT

tests were performed for all cells injected in mice to check for mycoplasma and rodent pathogens.

Plasmids, transient transfections, and cell line generation

Al list of plasmids used is provided in Supplementary Table 7. To perform transient transfection, 200,000 cells were seeded in 6 well plates a day prior to transfection. On transfection day, 2ml of fresh appropriate complete media was added to the cells. A transfection mix was prepared with 200 µl of jetPRIME® buffer, 2µg of plasmid of interest and 4µl of jetPRIME® reagent (Polyplus transfection reference 114–15) and added to the cells for 24 hrs. Cells were then seeded on collagen matrices or inoculated in chicken CAMs.

Fluorescent cell lines and CDK2 sensor cell lines were generated by lentiviral infection (48 hours) followed by cell sorting selection of mVENUS positive cells. Cell lines expressing shRNA were generated by lentiviral infection (48 hours) followed by two weeks of puromycin selection (final concentration puromycin: 5 µg/ml).

pLifeact-mRuby lentiviral plasmid [62] was used to stably label fibroblasts for CAM-coculture with mVenus labelled cancer cells. Cell lines expressing Lifeact-mRuby were generated by lentiviral infection (48 hours) followed by two weeks of puromycin selection (final concentration puromycin: 1 µg/ml).

siRNA and shRNA expression and CRISPR cell lines

siRNA, shRNA and sgRNA sequences are listed in Supplementary Table 7.

To perform the siRNA screen of collagen genes in CAMs, 30pmol of siRNA were mixed with 6µl of lipofectamine RNAimax in 200µl of OPTIMEM. Half of the mix was added to 1.5ml of DMEM 10% serum without antibiotics containing 400,000 cells in a 6 well plate incubated overnight at 37°C. Cells were used in CAMs the following day.

For all other siRNA experiments, we performed the transfection on non-adherent cells on day 1 of the experiment, then did a second transfection on day 2 on adherent cells and let the cells recover in fresh media, before using cells on day 4 to implant in CAMs or to extract RNA and proteins.

shRNAs in a pLKO.1 puromycin lentiviral plasmid were encapsulated in lentiviral particles using the 3rd generation system (pVSVG/GAG-POL/TAT/REV) in HEK 293T. Cells were infected for 48hrs and then selected using 5µg/ml puromycin for at least 2 weeks. Early passage stocks were made to start each experiment from similar passages.

pLentiCRISPR V2 DDR1 KD or non-targeting control (NT) were encapsulated in lentiviral particles and used to infect mammalian cells as described above.

RT-qPCR

Total RNA from cell culture or frozen tissues were extracted with a Trizol-chloroform method in a ratio 1:5 (chloroform: Trizol). Reverse transcription (RT) was performed with

1μg of total RNA with qScript cDNA SuperMix (QuantaBio) with random and oligo(dT) primers. Quantitative real time PCR was performed using the QuantStudio 3 Real-Time PCR System (Applied Biosystems) with PerfeCTa SYBR Green FastMix (QuantaBio) from 50ng of the RT in a final volume of 10 μl and a hybridization temperature of 60°C. Primer sequences are shown in Table 7. The threshold cycle (Ct) value for each gene was normalized to the Ct value for rRNA18S relative level of expression.

RNA sequencing

All samples were assessed with the NanoDrop™ 2000 (Thermo Scientific) for RNA purity (OD260/OD280), agarose gel electrophoresis for RNA integrity and potential contamination, and the Agilent 2100 Bioanalyzer (Agilent) for RNA integrity prior to library construction. Library construction was performed on mRNA purified from total RNA using poly-T oligo-attached magnetic beads and fragmented randomly by addition of fragmentation buffer. The library was generated with an NEBNext® Ultra™ RNA Library Prep Kit. After library construction, the library was diluted to 1.5ng/ml from quantitative results by Qubit2.0 (Invitrogen). Insert size was determined by the Agilent 2100 Bioanalyzer (Agilent), and qPCR was used to accurately quantify the library effective concentration (> 2nM) to ensure the library quality. An Illumina NovaSeq 6000 system (Illumina) was used for high-throughput sequencing. Raw image data file from high-throughput sequencing (Illumina) was transformed to fastq files of reads with CASAVA base recognition (Base Calling). Raw reads were aligned to hg19 with STAR (v2.5) and counts computed with HTSeq (v0.6.1). Differential expression analysis was performed on gene counts with DESeq57 version 1.14.1 and enrichment statistics using the LIMMA58 Wilcoxon gene set test on differential statistics for the NABA_Matrisome pathways in MSigDB59 version 6.0.

Mass spectrometry and proteomic analysis

Samples for ECM proteomic analysis: Samples used for the ECM proteomic study were as follows: 1) Proliferative (T-HEp3) vs dormant (D-HEp3) HNSCCs: T-HEp3 #1: 100 mg; T- HEp3 #2: 100 mg; T- HEp3 #4: 100 mg. Since D-HEp3 tumors were smaller in size, we combined several tumors to generate three pools: D-HEp3 #1: 43.3 mg; D-HEp3 #2: 49 mg; D-HEp3 #3: 55mg. Despite that, the amount of proteins obtained post-decellularization was still very low and we further combined D-HEp3 #1 and D-HEp3 #3 for the mass spectrometry analysis (see Supplementary Table 1A). 2) Control vs DDR1 knockdown D-HEp3 HNSCCs: 150mg – 200mg of D-HEp3 shCTRL tumors (n=4) and D-HEp3 shDDR1 tumors (n=4) were analyzed. (see Supplementary Table 3A).

Decellularization: Samples described above were mechanically disrupted using a Bullet Blender (Next Advance). Tumor lysates were then subjected to incubation in a series of buffer (CNCS Compartmental Protein Extraction Kit, Millipore) resulting in the extraction of soluble and mostly intracellular components and the enrichment for insoluble extracellular proteins (for more details, see [63,64]). The decellularization efficiency was monitored by western blot analysis monitoring the depletion of collagen I (Millipore, AB765P), actin (rabbit polyclonal antibody 14–4 generated in the Hynes lab at MIT), and histone H4 (Abcam, ab52178) across the 4 sequential extraction steps (see Supplementary Figure 2B, lanes 2 to 5).

Digestion of proteins into peptide: ECM-enriched protein samples were resuspended and reduced in a solution of 8M urea containing 10mM DTT (Pierce). Samples were alkylated with 25mM iodoacetamide (Pierce), deglycosylated with PNGaseF (New England BioLabs) and digested with mass-spectrometry-grade endopeptidase Lys-C (Pierce) and trypsin (Pierce) as previously described (for more details, see[46,64]). Samples were acidified and desalted according to standard procedure and stored at −80°C prior to proteomic analysis.

Mass spectrometry analysis: Mass spectrometry analysis of dormant (D-HEp3) vs proliferative (T-HEp3) tumors was performed at the Proteomics Core Facility at the Koch Institute for Integrative Cancer Research at MIT as follows: peptides were separated by reverse phase HPLC (Thermo Easy nLC1000) using a pre-column (made in house, 6 cm of 10 μm C18) and a self-pack 5 μm tip analytical column (12 cm of 5 μm C18, New Objective) over a 140-minute gradient before nanoelectrospray using a QExactive mass spectrometer (Thermo). Solvent A was 0.1% formic acid and solvent B was 80% MeCN/0.1% formic acid. The gradient conditions were 2–10% B (0–3 min), 10–30% B (3–107 min), 30–40% B (107–121 min), 40–60% B (121–126 min), 60–100% B (126–127 min), 100% B (127–137 min), 100–0% B (137–138 min), 0% B (138–140 min), and the mass spectrometer was operated in a data-dependent mode.

The parameters for the full scan MS were resolution of 60,000 across 350–2000 m/z, AGC 3e6, and maximum IT 50 ms. The full MS scan was followed by MS/MS for the top 15 precursor ions in each cycle with an NCE of 28 and dynamic exclusion of 30 s.

Raw mass spectral data files (.raw) were searched using Proteome Discoverer (Thermo) and Mascot version 2.4.1 (Matrix Science). Since the samples are human/mouse xenografts, the data was searched against both a Mus musculus (Mouse) database (17,032 sequences) and a Homo sapiens (Human) database (20,366 sequences). Mascot search parameters were: 10 ppm mass tolerance for precursor ions; 15 mmu for fragment ion mass tolerance; 2 missed cleavages of trypsin; fixed modification was carbamidomethylation of cysteine; variable modifications were methionine, lysine and proline oxidation, asparagine, and glutamine deamidation, N-term glutamine to pyroglutamate, carbamylation of N-term, tyrosine, serine, and threonine phosphorylation.

Only peptides with a Mascot score greater than or equal to 25 and an isolation interference less than or equal to 30 were included in the data analysis. Identification of ECM proteins in the mass spectrometry output were performed using Matrisome Annotator (Naba et al., Matrix Biology, 201640, available at http://matrisome.org).

Mass spectrometry analysis of control (D-HEp3 sh CTRL) vs DDR1 knockdown (D-HEp3 shDDR1) tumors was performed at Proteomics Core Facility at the University of Illinois at Chicago on a Thermo Fisher Orbitrap Velos Pro coupled with Agilent NanoLC system (Agilent, Santa Clara, CA). The LC columns (15 cm × 75 μm ID, Zorbax 300SB-C18) were purchased from Agilent. Samples were analyzed with a 120-min linear gradient (0–35% acetonitrile with 0.1% formic acid) and data were acquired in a data-dependent manner in which MS/MS fragmentation was performed on the top 10 intense peaks of

every full MS scan. Full MS scans were acquired in the Orbitrap mass analyzer over m/z 350–1800 range with resolution 30,000 (m/z 400). The target value was 1.00E+06. The ten most intense peaks with charge state ≥ 2 were fragmented in the HCD collision cell with normalized collision energy of 35%, these peaks were then excluded for 30 s after 2 counts within a mass window of 10 ppm. Tandem mass spectrum was acquired in the Orbitrap mass analyzer with a resolution of 7,500. The target value was 5.00E+04. The ion selection threshold was 5,000 counts, and the maximum allowed ion accumulation times were 500 ms for full scans and 250 ms for HCD. RAW files were converted into. mgf files using MSConvert (ProteoWizard). Database search was carried out using Mascot server version 2.6.2 (from Matrix Science). Mascot search parameters were: 10 ppm mass tolerance for precursor ions; 100 mmu for fragment-ion mass tolerance; two missed cleavages of trypsin; fixed modification was carbamidomethylation of cysteine; and variable modifications were oxidized methionine, deamidation of asparagine, pyro-glutamic acid modification at N-terminal glutamine, and hydroxylation of lysine and proline. Only peptides with a Mascot score ≥ 25 and an isolation interference ≤ 30 were included in the data analysis.

Mass spectrometry output were further annotated to identify ECM and non-ECM components using the Matrisome Annotator R-script we previously developed [46] and freely available at http://matrisome.org.

The raw mass spectrometry datasets are available upon request and will be made publicly available via deposition to the ProteomeXchange Consortium upon acceptance of the manuscript.

Immunofluorescence

Cells were plated on coverslips overnight, fixed with 4% paraformaldehyde for 10 min, permeabilized with 0.2% Triton X-100 for 10 min, and incubated for one hour at room temperature with primary antibodies listed in Table 7, in 1X PBS 4% BSA. After primary antibody incubation, cells were washed in 1X PBS and incubated with secondary antibodies in PBS 4% BSA for 30 min at room temperature, and then washed in 1X PBS three times with a final wash in distilled water. Coverslips were mounted on slides using Invitrogen™ Fluoromount-G™ polymerizing medium.

Immunofluorescence in paraffin embedded tissue

For tissue staining, 5mm sections were deparaffinized in xylene 10min, treated with a graded series of alcohol (100%, twice 5min, 95% twice 4min, 70% twice 3min), rehydrated in PBS, and subjected to heat-induced antigen retrieval in 10 mM citrate buffer (pH 6.0) (20 minutes on program HIGH of pressure cooker reference Cuisinart® 6 qt. Electric Pressure Cooker). Sections were preincubated with 10% normal donkey serum in 0.5% Tween-20 PBS for 2 hours at room temperature, incubated with primary antibodies in 1% donkey serum and 0.5% Tween-20 PBS buffer overnight at 4°C, washed 3 times in PBS and incubated in fluorescently labeled secondary antibodies (AlexaFluor, Molecular Probes) for 2 hours at room temperature, and in DAPI (Biotium 40011) to label the DNA. Tissue was washed in

PBS 1X three times with a final wash in distilled water. Coverslip were mounted on slides using Invitrogen™ Fluoromount-G™ polymerizing medium.

Immunohistochemistry on Tissue microarray

TMA reference T343a, HN242b, HNT1021, HN802c, HN811b and HN483at were purchased from Biomax. All TMA were first imaged using multiphoton microscopy using the tile function (4 columns by 4 rows to collect data from the entire core tissue) to collect a second harmonic signal. TMA reference T343a, HN242b, HNT1021 were used to stain for type III collagen. Staining was performed as described below.

5mm sections were deparaffinized in xylene 10min, treated with a graded series of alcohol (100%, twice 5min, 95% twice 4min, 70% twice 3min), rehydrated in PBS, and subjected to heat-induced antigen retrieval in 10 mM citrate buffer (pH 6.0) (20 minutes on program HIGH of pressure cooker reference Cuisinart® 6 qt. Electric Pressure Cooker). Endogenous peroxidases were inhibited by 5min incubation with 3%H_2O_2, followed by a PBS 1X 0.5% tween wash. Tissues were then blocked for 30min at room temperature with 2.5% Horse serum and solution A from an Avidin/Biotin blocking kit (reference SP-2001). Primary antibody against COL3A1 was incubated overnight at 4°C in a humidified chamber (dilution 1/100) in 2.5% Horse serum and solution B from an Avidin/Biotin blocking kit (reference SP-2001). Tissues were washed twice in PBS 1X 0.5% tween prior to secondary antibody incubation. Secondary antibodies from VECTASTAIN elite ABC universal kit (reference PK-7200) were incubated for 30min at room temperature following manufacturer recommendations. Tissues were then washed twice in PBS 1X 0.5% tween and incubated 30min at room temperature with the ABC reagent followed by a wash in PBS 1X 0.5% tween. DAB (reference SK-4105) was prepared following the manufacturer recommendations and incubated for 10min at room temperature in the dark. Tissues were then washed in distilled water and counter-colored using Harris hematoxylin solution (reference sigma HHS32–1L) following the manufacturer protocol. Then, tissues were sealed with coverslips using Richard Allan™ Scientific Mounting medium (ThermoFisher, reference # 4112).

Masson trichrome staining

To stain collagen content in tumors we performed Masson trichrome staining. 5 μm sections were used. Thermo Scientific™ Richard-Allan Scientific™ Masson Trichrome Kit (reference 87019) was used following the supplier recommended protocol. Coverslips were mounted on slides using Richard Allan™ Scientific Mounting medium (ThermoFisher, reference # 4112).

Picrosirius red staining

Picrosirius red staining was performed using Kiernan protocol.

5mm sections were deparaffinized in xylene 10min, treated with a graded series of alcohol (100%, twice 5min, 95% twice 4min, 70% twice 3min), rehydrated in PBS. Nuclei were stained with a Wiegert's hematoxylin kit (reference sigma 1.15973.0002) following the manufacturer's protocol. Tissues were then stained for one hour at room temperature with

picrosirius red solution (Sirius red, reference sigma 365548 (0.5g diluted in 500ml of picric acid (reference sigma P6744–1GA)). Tissues were then washed twice in acidified water (0.5% glacial acetic acid diluted in distilled water). Tissues were dehydrated with 3 changes of 100% ethanol, washed in xylene, and mounted on slides using Richard Allan™ Scientific Mounting medium (ThermoFisher, reference # 4112).

TUNEL assay

To assess apoptosis in tissue sections we used a TUNEL assay. 5μm sections were used. Abcam TUNEL Assay Kit - HRP-DAB kit (ab206386) was used following the supplier recommended protocol. Coverslips were mounted on slides using Richard Allan™ Scientific Mounting medium (ThermoFisher, reference # 4112).

Western blot

All western blots were performed on cells prior to their use for *in vivo* experiments to verify the level of protein depletion following the different treatments. Cells were lysed in radio-immunoprecipitation assay (RIPA) buffer (25 mM Tris HCl, pH 7.5, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, and 0.1% SDS), sonicated, and quantified using DC™ Protein Assay (Biorad). Sixty micrograms of total protein were incubated at 95°C for 5 min in Laemmli 4X (Bio-Rad), and loaded onto a 6% or 10% SDS-PAGE gel. Proteins were transferred onto a nitrocellulose membrane (Bio-Rad) using wet transfer at 120V for 1.5 hours. Ponceau S solution (sigma P7170–1L) was performed for each blot to validate even transfer across the membrane. Membranes were blocked with 1% bovine serum albumin (Fisher BP1600–100) in 1X TBS-T and probed with primary antibody overnight. Membranes were then washed in 1X TBS-T and incubated with the corresponding LICOR secondary antibody (1:15,000 dilution) for 1 hour at room temperature, and signals were acquired and quantified with the Odyssey system (LI-COR Biosciences).

MTT assay

First, 50μl of type III collagen (nippi PSC3–00-20) was polymerized in 96 well plates for 4hrs at 37°C at different concentrations: 0.5, 1 and 1.5 mg/ml resuspended in 1X D-PBS with calcium and magnesium. 2500 cells were added on top of collagen coating and incubated for 24hrs. Then, we used an MTT Cell Viability Assay Kit from Biotium (reference #30006) and followed the supplier recommended protocol. Absorbance signal was measured on a spectrophotometer SpectraMax M5e at 570 nm. Background absorbance was measured at 630 nm. Background absorbance was subtracted from signal absorbance to obtain normalized absorbance values. Each data point was performed in triplicate.

ApoTox-Glo™ Triplex Assay

First, 50 μl of type III collagen (nippi PSC3–00-20) was polymerized in 96 well plates for 4hrs at 37°C at different concentrations: 0.5, 1 and 1.5 mg/ml resuspended in 1X D-PBS with calcium and magnesium. 5000 cells were added on top of collagen coating and incubated for 24hrs or 48hrs. Absorbance signal was measured with a plate reader SpectraMax M5e on a spectrophotometer. Viability was measured using 400Ex/505Em.

Cytotoxicity was measured using 485Ex/520Em. Apoptosis was measured by luminescence. Each data point was performed in triplicate.

FACS

G0 analysis—1,000,000 cells per condition were resuspended in 1 ml of DMEM media containing 10μg/ml Hoechst 33342. The samples were incubated for 45 minutes in a 37°C water bath, and then incubated for an extra 15 minutes with Pyronin Y (Sigma 213519, 0.1μg/ml final concentration). Samples were then transferred on ice and analyzed by flow cytometry (BD LSRFortessa). At least 100,000 cells were recorded per sample in triplicate. A dot plot with Hoechst in the X axis and pyronin Y in the Y axis shows a distribution of DNA and RNA content with G0 cells having lower Pyronin Y signal. Data were analyzed using FlowJo_V10.

Apoptosis/necrosis—BioLegend APC Annexin V detection kit with 7-AAD (reference 640930) was used to determine apoptosis and necrosis by FACS as follows. 500,000 cells per condition were collected and washed twice in 100ul BioLegend's cell staining buffer. Cells were then resuspended in Annexin V binding buffer with 5ml of APC Annexin V and 5ml of 7-AAD for 15 minutes in the dark at room temperature. Finally, 400ml of binding buffer was added to each sample and each sample was analyzed by flow cytometry (BD LSRFortessa). At least 100,000 cells were recorded per sample in triplicates. Data were analyzed using FlowJo_V10.

Adhesion assay

Adhesion assays were performed in 96 well plates. Each condition was performed in triplicate with empty well and no cell controls for each matrix. Type I collagen I (Corning 354236) and type III collagen (nippi PSC3–00-20) were coated at 0.5mg/ml, type IV collagen (nippi ASC 4–104-01) coated at 0.05mg/ml and fibronectin (Invitrogen 33010–018) coated at 100μg/ml. 100μl of each matrix was added per well and polymerized for 4hrs at 37°C. Excess of matrix was then removed. Cells were detached with EDTA 2mM in 1X PBS no calcium no magnesium, diluted in DMEM 1% serum and counted with trypan blue. 50,000 cells were seeded per wells in 100μl total volume. Plates were incubated for 30 minutes at 37°C. After incubation, the wells were washed three times with 1X PBS and remaining cells fixed with 5% glutaraldehyde (Millipore 354400) for 20min at room temperature. Glutaraldehyde was rinsed with one 1X PBS wash. Cells were stained with 100μl of Crystal violet (0.1%w/v in MES 200mM pH6.0) for 1hr at room temperature under constant rocking. Excess of crystal violet was removed by three washes with deionized water and remaining staining was solubilized in 100μl of 10%v/v of glacial acetic acid. Absorbance was read at 570nm with plate reader SpectraMax M5e. An average of no cells control is subtracted from cells absorbance for each matrix type.

CAM assay

Fertilized eggs (from Charles Rivers) were inoculated with 150,000 cells at day 10 of chicken embryo development. Beforehand, cells were trypsinized and their concentration was determined with Countess II (Invitrogen) using Trypan blue stain 0.4%. Viable cells were used to prepare cell suspension in 1X D-PBS without calcium and magnesium.

Chicken eggs (day 10 of development) were lying on their side and CAMs dropped by puncturing the top and side (air bag side) shell and sucking up air from the air bag creating an air pocket on the top of the CAM. 50 µl of cell suspension was inoculated per CAM using an insulin syringe (reference BD328468). Eggs were incubated at 37°C in humid atmosphere for 6 days before tumor collection. At day 16 of chicken embryo development, eggs were opened, and tumors harvested. To evaluate tumor growth, tumors were minced and digested for 20min at 37°C with Collagenase/ BSA solution (2.5%BSA, 0.15% collagenase in PBS +Mg +Ca filtered) from Clostridium histolyticum (reference Sigma Aldrich C0130). Number of cells per tumors were estimated by counting tumor cells recognized by their larger diameter (compared with chicken host cells). To evaluate protein expression, tumors were collected and grounded on ice in radio-immunoprecipitation assay (RIPA) buffer (25 mM Tris HCl, pH 7.5, 150 mM NaCl, 1% IGEPAL, 1% sodium deoxycholate, and 0.1% SDS), sonicated, and quantified using DC™ Protein Assay (Biorad) and blotted following western blot protocol section (see above).

Inhibitor treatments in CAMs—Fludarabine (Stat 1 inhibitor) was purchased from Selleckchem (reference S1491), diluted in DMSO at 50mM and used for 24hrs at a final concentration of 20µM.

Cells were pretreated overnight before inoculation in chicken fertilized eggs with Fludarabine 20µM. Control cells were treated with similar DMSO concentration. Tumors were then processed as described in CAM assay section of this article.

Nilotinib (DDR1 inhibitor) was purchased from Selleckchem (reference AMN-107) diluted in DMSO at 2.5mM and used 24hrs at a final concentration of 10µM. Cells were pretreated overnight before inoculation in chicken fertilized eggs with Nilotinib 10µM. Control cells were treated with similar DMSO concentration. Tumors were then processed as described in CAM assay section of this article.

hCOL3 overexpression in CAMs—T-HEp3 cells tetON-hCOL3A1 expressing CDK2 sensor were pretreated *in vitro* for 24hrs with doxycycline (1µg/ml final concentration) prior eggs inoculation. Chicken eggs (day 10 of development) were lying on their side and CAMs dropped by puncturing the top and side (air bag side) shell and sucking up air from the air bag creating an air pocket on the top of the CAM. CAMs were inoculated with 150,000 cells per egg using an insulin syringe. Doxycycline was added to the treated condition at the time of inoculation. Remaining cells were used to extract proteins and validated *COL3A1* overexpression. CAMs were then treated with 100µl of doxycycline 1µg/ml at day 2 and 5 post inoculation. Control CAMs were treated with vehicle sterile distilled water. Eggs were opened and tumors collected at day 7 post inoculation. Tumors were imaged with a multiphoton microscope and then digested for 20min at 37°C with Collagenase/BSA solution (2.5%BSA, 0.15% collagenase in PBS +Mg +Ca filtered) from Clostridium histolyticum (reference Sigma Aldrich C0130). Number of cells per tumors were estimated by counting tumor cells recognized by their larger diameter (compared with chicken host cells).

CAMs co-culture of fibroblasts and cancer cells—To co-culture fibroblasts and cancer cells. Fibroblasts stably express a Lifeact-mRuby vector and cancer cells express the CDK2 sensor mVenus allowing us to distinguish both cell types using a multiphoton microscope. 200,000 fibroblasts were mixed with 150,000 cancer cells and inoculated per egg after CAMs were dropped as previously described. Eggs were then incubated for 24hrs prior imaging.

Control eggs inoculated with fibroblasts only, were inoculated with 200,000 cells per egg following the previously described protocol and analyzed 24hrs after seeding.

Mouse experiments

Human cell lines were injected in Nude female Athymic NCr-nu/nu from Charles Rivers (HEp3 model) or NCG (NOD-Prkdcem26Cd52Il2rgem26Cd22/NjuCrl) strain from Charles River Labs (MDA-MB-231) between 8 and 12 weeks old. Murine cell lines were injected in BALB/cAnNCR female from Charles Rivers between 8 and 12 weeks old. Each time mice were anesthetized and lubricant Lubrifesh Major (reference NDC 0904–6488-38) was applied to their eyes to prevent drying.

Tumor growth studies—The maximal tumor size/burden of 1500mm3 was not exceeded.

500,000 cells (T-HEp3, D-HEp3, D-HEp3 shCTRL and shDDR1) were injected in 1X sterile DPBS orthotopically in the interscapular region of the neck of nude mice. 5 mice were used per group. Tumor sizes were monitored every 4 days measuring the largest side of the tumor with a Fisherbrand™ Traceable™ Digital Carbon Fiber Caliper in millimeters. Tumor volume was calculated following the formula $V=4/3\pi r^3$ and plotted as single dot with SEM. Mice were sacrificed with CO_2 when tumor diameter reached 1 cm. At day 25, all remaining animals were sacrificed. Tumors and lungs were imaged with multiphoton microscope and paraffin-embedded for tissue staining.

Tumor removal surgery procedure—500,000 cells (T-HEp3 GFP or expressing CDK2 sensor and H2B-RFP) were injected in 100μl of 1X sterile DPBS orthotopically in the interscapular region of the neck of nude mice. 5 mice were used per group. Tumor sizes were monitored every 4 days for 2 weeks. At day 14 after tumor cell injections, tumors were removed under procedure sterile hood. Mice were anesthetized with isoflurane (Baxter reference # NDC 10019–360-40, provided by the Mount Sinai animal facility) and the skin around the tumor was cleaned with ethanol. Tumor was lifted from its center and an incision was made using small scissors. All macroscopic tumor tissue was removed. The wound was closed using 9mm AutoClips® (reference 205016) from Braintree Scientific. Clips were removed 10 days after surgery. Two to four weeks after surgery, mice were sacrificed with CO_2 and the area of resection was reopened to image residual cancer cells (far from scar tissue) and lungs were collected and imaged to find spontaneous DTCs using a multiphoton microscope. Tumors and lungs were fixed and paraffin-embedded for tissue staining.

Imaging window intravital experiments—Ultem (polyetherimide)-made windows were designed and created in house at Mount Sinai. These windows were designed by the addition of a ring of holes to suture the window to the animal skin. We implanted the

window in the back of the neck of the animal prior to injecting tumor cells. At day 1, mice were anesthetized with isoflurane and an incision in the back of the neck was made with a scalpel (blade #10) and a window containing a no 1.5 coverslip was placed in the wound. The window was attached to the skin with Ethicon silk black braided attached to a FS-2 19MM 3/8C reverse cutting needle and each suture was glued with a drop of skin glue VETCLOSE (reference # 031477). Cancer cells were then injected under the glass in 50μl of DPBS using an insulin syringe (reference BD328468). Cells were imaged four days after injection to visualize ECM organization at early stages of tumor formation.

Tail vein injections—8-week-old nude mice were injected in the tail vein using an insulin syringe (reference EXEL 26028). 1,000,000 cells were injected per animal in a total volume of 100μl of DPBS 1X. Mice were then monitored for 5 days (siRNA experiments) or 3 weeks (shRNA experiments) prior to sacrifice and lungs were imaged using a multiphoton microscope.

Collagen Type III bioengineer scaffolds—500,000 cells (T-HEp3 CDK2 sensor and H2B-RFP) were resuspended in 100μl of 1X sterile DPBS and injected in the interscapular region of the neck of nude mice. 5 mice were used per group. Tumor sizes were monitored every 4 days for 19 days. At day 19 after tumor cell injections, tumors were removed following the surgery procedure described in the previous section. One group of mice received a dental sponge rehydrated in DPBS and the other group received a dental sponge rehydrated in DPBS and incubated with type III collagen 1mg/ml polymerized for 4hrs at 37°C. Dental sponge reference Gelfoam® size 4 was purchased from Pfizer. They were cut in three pieces, each animal received one piece of sponge. The wound was closed using 9mm AutoClips® (reference 205016) from Braintree Scientific. Clips were removed 10 days after surgery. Local tumor relapse was monitored up to 40 days (20 days after surgery). Mice were sacrificed with CO2 when tumors reached 1cm in diameter. At day 40, all remaining mice with no local relapses were sacrificed. For each animal, the area of resection was reopened to image residual cancer cells where the sponge was places by using a multiphoton microscope.

Collagen Type III co-injection assays

In CAMs, 150,000 T-HEp3 cells were mixed with 1.5 mg/ml of type I or III or 0.25 mg/ml IV collagen in DPBS or DPBS only and 50μl were inoculated per CAM using an insulin syringe. Six days after inoculation, tumors were collected and processed as described in the section CAM assay of this materials and methods.

In mice, 750,000 T-HEp3 were injected in nude mice with 1.5mg/ml type III collagen or in DPBS. Mice were injected in each side of their flank, one side with DPBS only and the other side with type III collagen.

100,000 4T1 were injected in the mammary gland of BALB/cAnNCR mice with 1.5mg/ml type III collagen or in DPBS. Mice were injected in each side of their inguinal mammary gland, one side with DPBS only and the other side with type III collagen.

2,000,000 D2.A1 were injected in the mammary gland of BALB/cAnNCR mice with 1.5mg/ml type III collagen or in DPBS. Mice were injected in each side of their inguinal mammary gland, one side with DPBS only and the other side with type III collagen.

Tumor sizes were monitored every 4 days measuring the largest side of the tumor with a Fisherbrand™ Traceable™ Digital Carbon Fiber Caliper in millimeters. Tumor volume was calculated following the formula $V=4/3\pi r^3$ and plotted as single dot with SEM. Mice were sacrificed with CO2 when tumor diameter reached 1 centimeter.

Microscopy

Leica SP5 DMI—Leica TCS SP5 confocal microscope was used to image immunofluorescence of cells and of tissue sections. It has four lasers: a UV Diode (405nm), an Argon laser (458nm, 476nm, 488nm, 514nm), a DPSS laser (561nm), and a HeNe laser (647nm). Its spectrophotometer scan head uses up to five PMT detectors to detect five different tunable emission wavelengths, allows simultaneous imaging with an additional transmitted light detector. 40x and 63x objectives were used.

Olympus FV1000MPE multiphoton—Multiphoton has a Coherent Chameleon Vision II laser, tunable from 680nm to 1080nm, which was used for SHG imaging of *in vivo* and *ex vivo* samples using a water immersion 25x objective (NA=1.05). Excitation line 880nm was used for SHG and GFP imaging. Sequential 880nm and 1000nm was used to acquire SHG, CDK2 sensor and H2B-RFP signals. Z-stacks were acquired with a 5μm z-step from top to bottom of the sample with a zoom of 1.

Zeiss Axio Imager: Zeiss Axio was used to image Masson trichrome using 40x and brightfield microscopy.

Zeiss Axio Scan.Z1 slide scanner: Slide scanner was used to acquire large images of tissue sections for TUNEL assay. Objective Plan-Apochromat 40x/0.95 Korr M27 was used with brightfield contrast method and a light source TL LED lamp 300% intensity. Exposure time 200ms. Depth of focus 1.2mm. Binning mode 1,1.

Live-cell microscopy—T-HEp3 cells stably expressing the CDK2 sensor were plated on 14-mm glass-bottom dishes, No. 1.5 thickness (MatTek) coated with 546 fluorescent type I or III collagen ([65] for detailed protocol to label collagen). Cells were imaged 1 hr after seeding in DMEM 4.5g D-Glucose 110mg/l sodium pyruvate, 10% PEAK serum heat inactivated, 1%PenStrep at 37°C with CO2. Imaging was performed with Zeiss LSM880 and a 20x objective, taking one image of a 10-μm z-stack, every 30 minutes for 18hrs. Number of mitosis per field were counted with the support of CDK2 sensor fluorescence.

Leica DMI 8—Leica DMi8 with polarized light filter and HC PL Fluotar 20x/0.40 Ph1 “C” objective and DFC450C camera was used to image picrosirius staining.

Image analysis

All representative images showed in figures were chosen randomly from sets of images including at least 10 images per conditions.—Image analysis was performed with FIJI version ImageJ 1.52p NIH, USA, http://imageJ.nih.gov/ij

CDK2 sensor quantification: From multiphoton images we used H2B-RFP to visualize the nucleus of cancer cells and determined the phase of the cell cycle following CDK2 sensor fluorescence. During G0/G1, the sensor is in the nucleus of the cells. Cells in S phase present an equal distribution of the sensor across the nucleus and cytoplasm, and fluorescence is excluded from the nucleus in the G2 phase of the cell cycle. We quantified number of cells in each phase of the cell cycle per field from 5 animals for the surgery site and 4 animals for the lung quantification.

TUNEL quantification: number of TUNEL positive cells were counted per entire tissue section (2 sections per animal from 5 animals per condition) and reported to the total tissue area quantified (in pixel^2).

p27+ quantification: Number of cells positive for p27 staining in their nucleus was quantified per field from 3 images per tissue section from 4 independent animals where T-HEp3 tumors and T-HEp3+type III collagen paired tumors were imaged. Only cells positive for vimentin staining were quantified as human vimentin was used to identify cancer cells.

Collagen III intensity quantification: total intensity of the collagen channel was measured with ImageJ from images of solitary cells and clusters of T-HEp3 cells spontaneously disseminated in the lung of nude mice (n= 2 animals).

Collagen alignment quantification: OrientationJ plugin written by Daniel Sage at the Biomedical Image Group (BIG), EPFL, Switzerland (http://bigwww.epfl.ch/demo/orientation/) was used to determine collagen alignment. OrientationJ Analysis was used with a local window σ of 1 pixel with a cubic spline gradient to color code the images and OrientationJ Distribution was used to generate graphs and statistics normalized to the mode following protocol described in [66].

Picrosirius red image analysis: Polarized light images were transformed to CIELAB coordinates using the FIJI software. The output image L (Lstar) was used to calculate the total collagen content. The a axis output image was used to extract the green component (negative values indicating green pixels) of the images corresponding to type III collagen thin fibers used to calculate the percentage of images occupied by green pixels.

Statistics and Reproducibility

The Investigators were not blinded to allocation during experiments and outcome assessment. No statistical method was used to predetermine sample size.

No data were excluded from the analyses.

All experiments were performed at least 3 independent times with at least 3 biological replicates. All data are reported as the mean ± SEM of at least three independent

experiments. Statistical significance ($P < 0.05$ or less) was determined using appropriate statistical test and performed using GraphPad Prism version 9.2.0 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com. Tests and n are described in the legends of each figure.

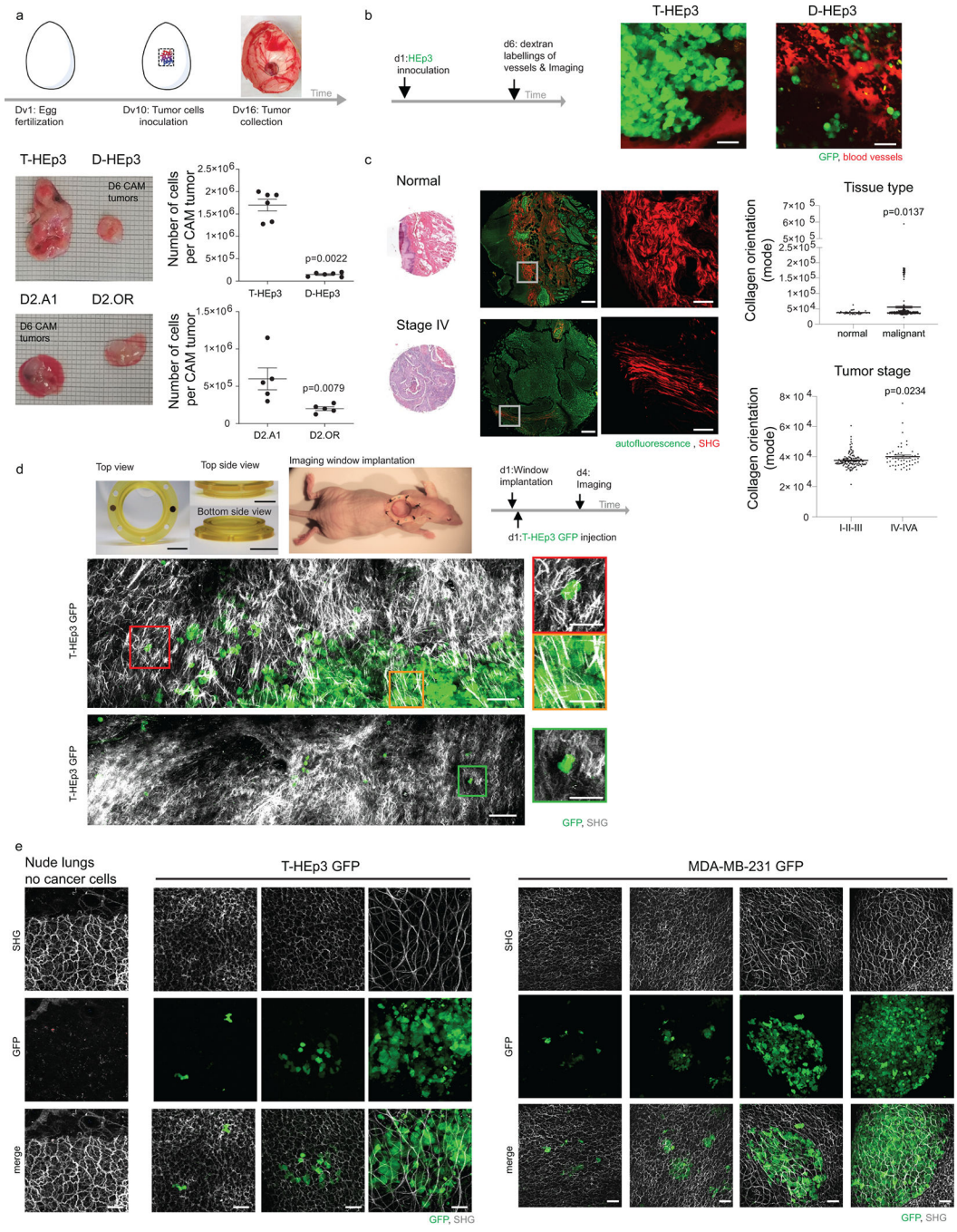
Adjusted p-values of differential expression of genes in Supplementary Table 2 and 4 (3 biological replicates per group) were calculated using the DESeq2 R package (v2_1.6.3). P-values were adjusted using the Benjamini and Hochberg's approach for controlling the False Discovery Rate (FDR), and the threshold of differential gene expression was padj < 0.05.

Data availability

The raw mass spectrometry proteomic data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository [67] with the dataset identifiers PXD019185 (T-HEp3 and D-HEp3 tumors) and PXD018883 (shControl and shDDR1 D-HEp3 tumors).

RNA–seq data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE182890.

Extended Data



Extended Data Fig. 1. Supportive Data to Main Figure 1

All numerical data are presented as mean +/−SEM.

(a) Experimental design of CAMs experiments. d refers to days on timeline scheme. Left panel: Representative images of day 6 collected tumors. Right panel: Top graph: T-HEp3 and D-HEp3 (n= 6 independent CAMs). Bottom graph: D2.A1 and D2.OR (n=

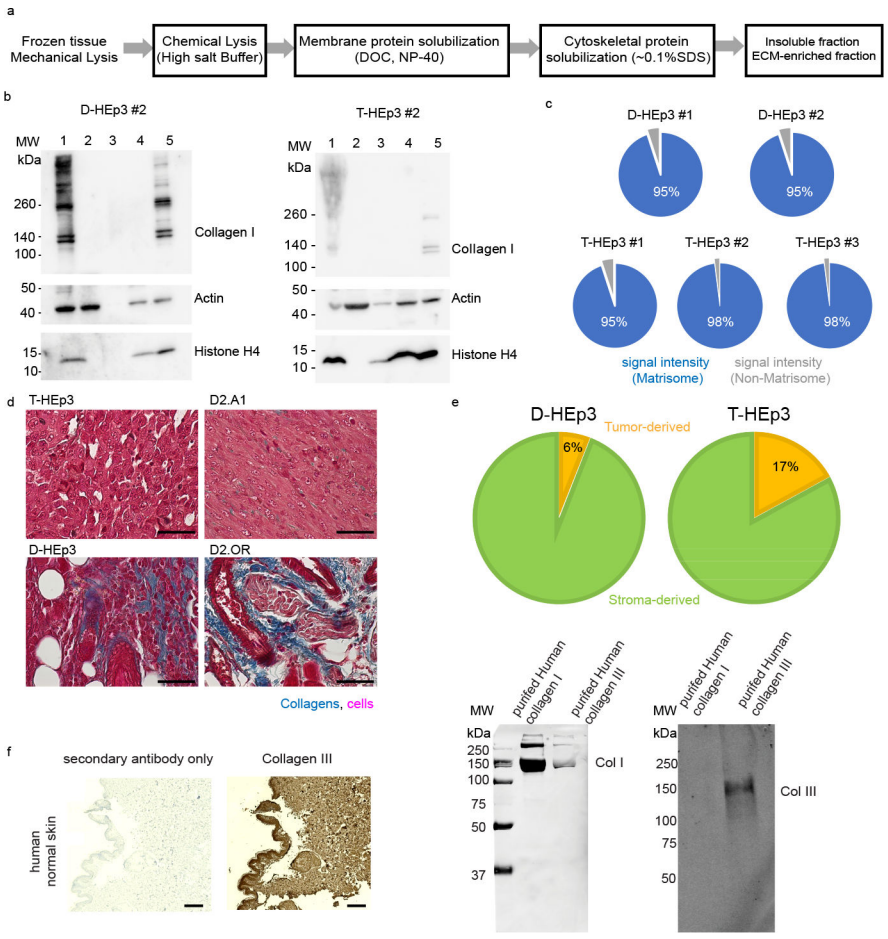
5 independent CAMs). Number of cells per tumor compared with an unpaired two-tailed Mann-Whitney test with 95% confidence level.

(b) Representative multiphoton images of T-HEp3 and D-HEp3 CAM tumors. Scale bar, 50***μ***m. d refers to days on timeline scheme.

(c) Tissue microarray SHG analysis. Left panel: representative images of normal tissue versus stage IV HNSCC ECM architecture. Right images are a zoom of white squares on left image. Scale bars, 200***μ***m. Scale bar zoom, 50 ***μ***m. Right panel: Collagen orientation between normal tissues (n=43 samples) and malignant HNSCC (n=289 samples) and between stage I to III (n=130 samples) and stage IV and IVa (n=53 samples). Data were compared using an unpaired two-tailed Mann-Whitney test with 95% confidence level.

(d) Imaging window design and implantation site in mice (n=5). Representative images of T-HEp3-GFP in primary site. Scale bar, 100***μ***m. Zoom Scale bar, 50***μ***m. d refers to days on timeline scheme.

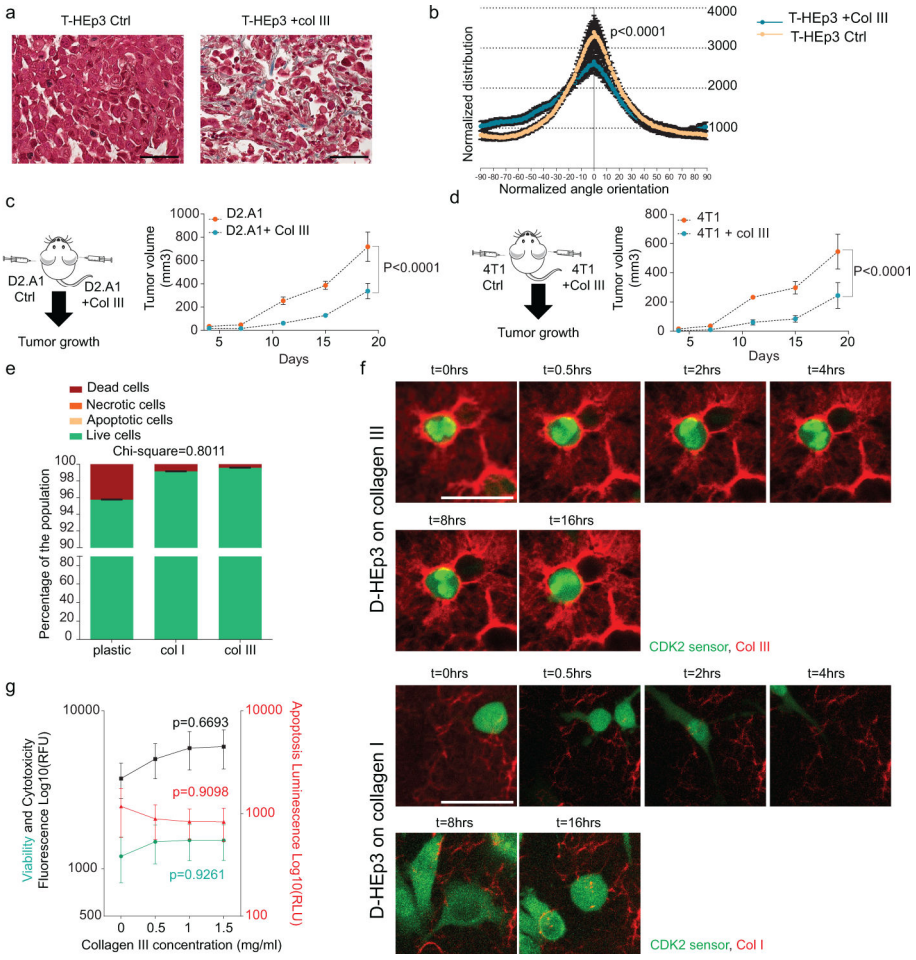
(e) Left panel: Nude mice lung representative images with or without T-HEp3 GFP spontaneously disseminated cells. Scale bar, 50***μ***m. Right panel: NCG mice lungs representative images with MDA-MB-231 GFP spontaneously disseminated cells. Scale bar, 50***μ***m.



Extended Data Fig. 2. Supportive Data to Main Figure 2

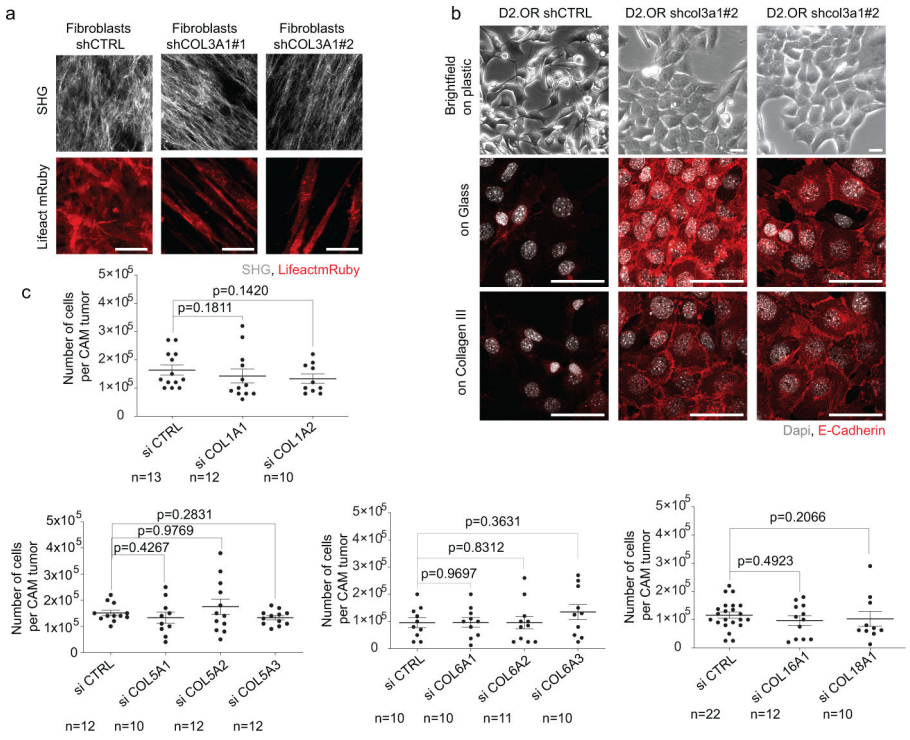
(a) ECM enrichment pipeline for mass spectrometry.

(b) ECM-enrichment validation by western blot before mass spectrometry analysis. Removal of intracellular components and ECM enrichment via sequential decellularization (lanes 2–4) from the total tissue lysate (1) was monitored by immunoblotting for actin (cytoskeleton protein) and histones (nuclear proteins). The remaining insoluble fraction (5) was highly enriched for ECM proteins (collagen I) and largely depleted for intracellular components.
(c) Proportion of the mass-spectrometric signal intensity from matrisome (blue) and non-matrisome (grey) peptides for each sample, related to Supplementary Table 1B.
(d) Masson's trichrome staining of proliferative and dormant mice tumors. Scale bars, 50μm.
(e) Percentage of tumor-derived and stroma-derived ECM d in D-HEp3 and T-HEp3 mice tumors, related to Supplementary Tables 1H and I.
(f) Collagen III staining specificity tested in immunohistochemistry staining on human skin tissues (Scale bar, 100μm) and by western blot using purified native human collagen I and III.



Extended Data Fig. 3. Supportive Data to Main Figure 3
All numerical data are presented as mean +/−SEM.
(a) Representative images of Masson's Trichrome from T-HEp3 mice tumors with or without type III collagen co-injection. Scale bar, 50μm.

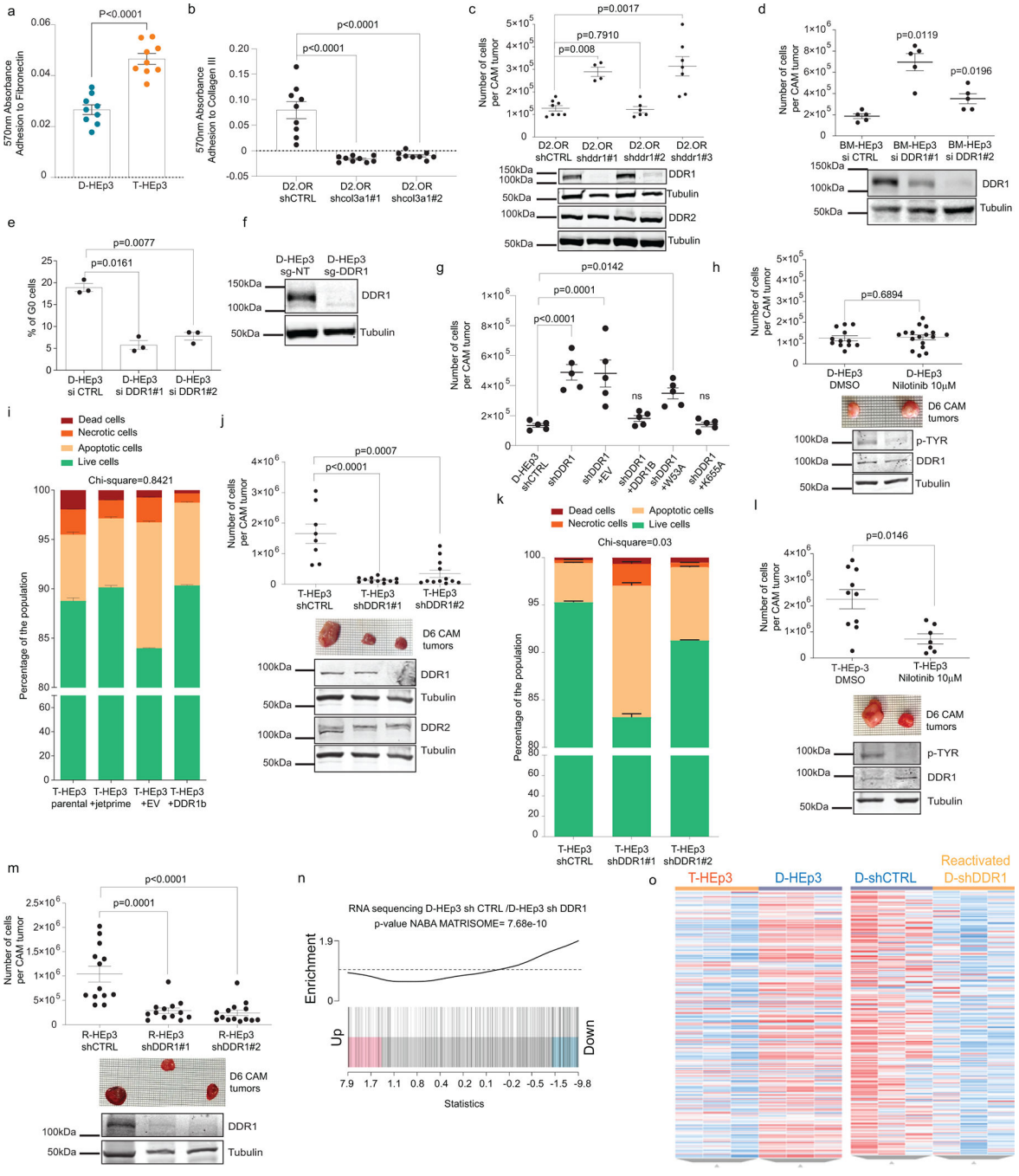
(b) Normalized distribution of collagen fiber orientation from tumors presented in A (n= 5 independent tumors per group, 2 images analyzed per tumors). Cumulative distributions were compared using an unpaired two-tailed Kolmogorov Smirnov test with 95% confidence level.
(c) Tumor growth of D2.A1 +/− type III collagen co-injection (n= 5 mice per group). Curves were compared using a two-way ANOVA with mixed model effects analysis and a Bonferroni correction and a 95% confidence interval.
(d) Tumor growth of 4T1 +/− type III collagen co-injection (n= 5 mice per group). Curves were compared using a two-way ANOVA with mixed model effects analysis and a Bonferroni correction and a 95% confidence interval.
(e) FACS analysis for percentage of T-HEp3 live cells (green), dead cells (red), apoptotic cells(yellow), and necrotic cells (orange) plated on plastic, type I collagen, or type III collagen matrix (n= 3 independent experiments). Distributions were compared using a Chi-squared test with 95% confidence interval.
(f) Time points from an 18hrs time lapse movie of D-HEp3 plated on type I or III collagen. (t=hours). Scale bar, 10 ***μ***m. Related to Supplemental Movies 3 and 4.
(g) APOTOX assay of T-HEp3 plated on different concentrations of type III collagen for 24hrs (n= 3 independent experiments).



Extended Data Fig. 4. Supportive Data to Main Figure 4
(a) Representative multiphoton images of MRC5 fibroblasts shRNA CTRL or expressing 2 independent shRNA targeting COL3A1 seed in CAMs for 24hrs. Scale bar, 50***μ***m.
(b) Top panel: representative brightfield images of D2.OR shRNA CTRL or expressing 2 independent shRNAs targeting col3a1 in vitro. Bottom panel: immunofluorescence of

D2.OR shRNA CTRL or expressing 2 independent shRNAs targeting COL3A1 in vitro for E-cadherin. Scale bar, 50μm.

(c) Number of cells per tumor for D-HEp3 expressing a control siRNA or siRNA targeting COL1A1, COL1A2, COL5A1, COL5A2, COL5A3, COL6A1, COL6A2, COL6A3, COL16A1 or COL18A1 in CAMs. (n= number of CAMs per group are described in the graphs). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. All numerical data are presented as mean +/−SEM.



Extended Data Fig. 5. Supportive Data to Main Figure 5

All numerical data are presented as mean +/−SEM.

(a) Adhesion assay for T-HEp3 and D-HEp3 to fibronectin. (n= 3 independent experiment with triplicates). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.

(b) Adhesion assay for D2.OR expressing an shRNA control or targeting COL3A1 to type III collagen. (n= 3 independent experiment with triplicates). Data were compared using unpaired A tumor-derived type III collagen-rich ECM niche regulates tumor cell dormancytwo-tailed Mann-Whitney test with 95% confidence level.

(c) Upper panel: Number of cells per CAM tumors of D2.OR shCTRL or sh DDR1 (n=8 independent CAMs shCTRL, n=4 shDDR1#1, n=5 sh DDR1#2, n=7 shDDR1#3). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Note that shRNA 1 and 3 only deplete DDR1. Lower panel: Western blot showing DDR1, DDR2 and tubulin levels upon DDR1 depletion.

(d)Upper panel: Number of cells per CAMs tumors of BM-HEp3 (dormant) expressing a control siRNA or siRNA targeting DDR1. (n=5 independent CAMs per condition). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Lower panel: Western blot showing DDR1 and tubulin levels upon DDR1 depletion.

(e) Percentage of G0 cells from D-HEp3 cells expressing a control siRNA or siRNA targeting DDR1. (n=3 independent experiments). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.

(f) Western blot for DDR1 and tubulin in D-HEp3 NT sgRNA control and expressing an sgRNA against DDR1.

(g) Number of cells per CAM tumors in D-HEp3 shCTRL or shDDR1, or shDDR1 rescued with overexpression of either an empty vector (EV), a DDR1b full length, a binding deficient mutant (W53A) or a kinase dead mutant (K655A) (n= 5 independent CAMs). Data were compared using an ordinary one-way ANOVA test with multiple comparison to shCTRL condition with 95% confidence level.

(h) Number of cells per CAM tumors of D-HEp3 +/− Nilotinib treatment (n=12 control CAMs and n=17 treated CAMs) Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Representative tumors and Western blot showing phospho-Tyrosin (pTYR), total DDR1 and tubulin levels upon nilotinib treatment are displayed below.

(i) FACS analysis for percentage of T-HEp3 live cells (green), dead cells (red), apoptotic cells(yellow), and necrotic cells (orange), treated with jetPRIME only or expressing a control empty plasmid (EV) or DDR1b full length (n= 3 independent experiments). Distributions were compared using a one-tailed Chi-squared test with 95% confidence interval.

(j) Number of T-HEp3 cells per CAM tumors expressing a control shRNA or an shRNA targeting DDR1 (n= 8 control CAMs and n=12 DDR1 depleted CAMs). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Representative tumors and Western blot showing DDR1, DDR2 and tubulin levels upon DDR1 depletion are displayed below.

(k) FACS analysis for percentage of T-HEp3 live cells (green), dead cells (red), apoptotic cells(yellow), and necrotic cells (orange), expressing either a control shRNA or an shRNA

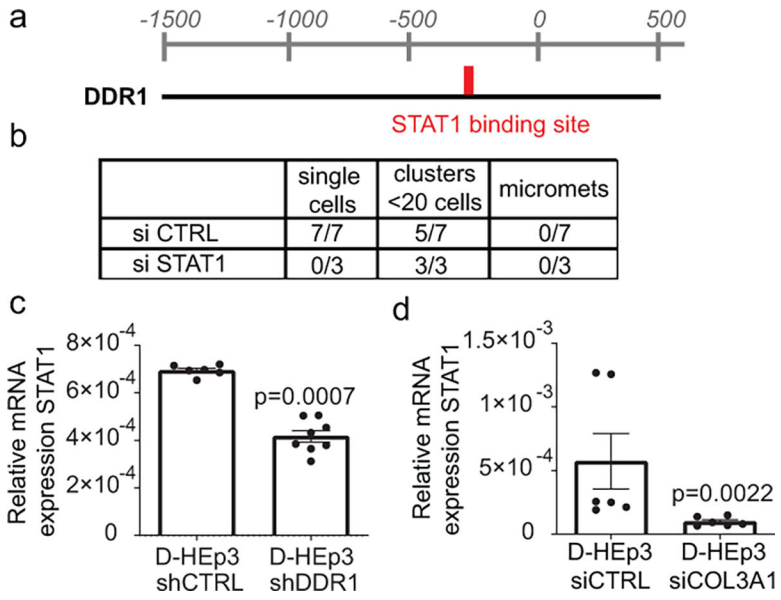
targeting DDR1 (n= 3 independent experiments). Distributions were compared using a one-tailed Chi-squared test with 95% confidence interval.

(l) Number of cells per CAM tumors of T-HEp3 +/− Nilotinib treatment (n=10 control CAMs and n=7 treated CAMs). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Representative tumors and Western blot showing phospho-Tyrosin (pTYR), total DDR1 and tubulin levels upon nilotinib treatment are displayed below.

(m) Number of R-HEp3 cells per tumors in CAM expressing a control shRNA or an shRNA targeting DDR1. (n= 13 control CAMs, n=13 shDDR1#1, n=14 shDDR1#2). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Representative tumors and Western blot showing DDR1and tubulin levels upon DDR1 depletion are displayed below.

(n) Enrichment plot for matrisome signature from RNA sequencing performed in D-HEp3 shRNA CTRL and D-HEp3 shDDR1 mice tumors (p=7.68e-10). X-axis shows log2FC for D-HEP3 shRNA DDR1 vs D-HEp3 shRNA CTRL. Black bars represent matrisome genes. Related to supplemental table 4.

(o) Heat map related to tables 2 and 4 where the entire transcriptome is displayed and organized by alphabetical order of genes. T-HEp3 and Reactivated D-shDDR1 cells show similar profiles compared with D-HEp3 and D-shCTRL conditions. Heat maps were generated using the Biojupie tool (https://maayanlab.cloud/biojupies/) [67].



Extended Data Fig. 6. Supportive Data to Main Figure 6

All numerical data are presented as mean +/−SEM.

(a) Map of predicted sites for STAT1 in DDR1 promoter region using the CiiiDER tool.

(b) Number of mice presenting single cells, clusters of less than 20 cells or micromets in their lungs after tail vein injection of D-HEp3 +/− si STAT1.

(c)RT-qPCR for STAT1 from RNA extracted from D-HEp3 shRNA CTRL or shDDR1 tumors in vivo (n= 3 independent RNA extraction from 3 different tumors, in duplicate).

Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.

(d) RT-qPCR for STAT1 from RNA extracted from D-HEp3 cells in vitro expressing a control siRNA or siRNA targeting COL3A1 (n= 3 independent RNA extractions in duplicate). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

JJBC, EJF, and AN would like to thank the NCI and Sage Bionetwork's Interdisciplinary Approaches to Cancer Metastasis workshop for inspiring this project.

The authors would like to thank Sabrina Spencer for providing the DHB-Venus plasmid, Birgit Leitinger for sharing DDR1 plasmids, Marisol Soengas for the FG12-GFP plasmid, Louis Hodgson for HEK cells, non-targeting CRISPR controls and lentiviral packaging plasmids, Eduardo Farias for teaching mice surgery, Mansour Djedaini and Selma Bekri for teaching FACS, Brian Wu for building the plastic imaging window, and Julie Cheung for teaching the CAM model. We acknowledge the Advanced Bioimaging Core and the Flow Cytometry Core from Mount Sinai. We want to thank Jill Gregory for her illustration of the graphical abstract. We thank the Aguirre-Ghiso, and Sosa laboratories for helpful discussions. We also would like to thank Tiphaine Martin for revising the statistical analysis through the paper.

The authors would like to thank Dr. Hui Chen from the Mass Spectrometry Core facility at the University of Illinois at Chicago and the Dr. George Chlipala from the Research Informatics Core facility at the University of Illinois at Chicago for their technical assistance with the analysis of D-HEp3 shCTRL vs D-HEp3 shDDR1 tumors and Richard Schiavoni from the Proteomics Core Facility at the Koch Institute for Integrative Cancer Research at MIT and Karl Clauser from the Broad Institute for their assistance with the analysis of T-HEp3 and D-HEp3 tumors.

This work was supported by a Susan G. Komen Career Catalyst Research (CCR18547848 to JJBC), an NCI Career Transition Award (K22CA196750 to JJBC), an NCI R01 (R01CA244780 to JJBC), the Tisch Cancer Institute NIH Cancer Center grant (P30-CA196521), the Schneider-Lesser Foundation Award (to JJBC), a Stony Brook-Mount Sinai pilot award (to JJBC). CM received support from NIH T32 CA078207 Training Program in Cancer Biology.

This work was partially supported by a start-up fund from the Department of Physiology and Biophysics at the University of Illinois at Chicago to AN. IT is the recipient of a Research Grant from the Honors College at the University of Illinois at Chicago and a LASURI award from the College of Liberal Arts and Sciences at the University of Illinois at Chicago

Proteomics services were provided by the UIC Research Resources Center Mass spectrometry Core which was established in part by a grant from The Searle Funds at the Chicago Community Trust to the Chicago Biomedical Consortium and by the Proteomics Core Facility of the Koch Institute for Integrative Cancer Research at MIT supported in part by a Cancer Center Support Grant from the NCI. Bioinformatic analyses of the proteomics data were performed by the UIC Research Informatics Core, supported in part by the National Center for Advancing Translational Sciences (NCATS, Grant UL1TR002003).

JAAG and ARN were supported by grants from National Institutes of Health (NIH)/National Cancer Institute (NCI) (CA109182, CA196521). JAAG is Samuel Waxman Cancer Research Foundation Investigator. EJF was supported by NCI (P30 CA006973).

References

1. Steeg PS Targeting metastasis. Nature Reviews Cancer (2016) doi:10.1038/nrc.2016.25.
2. Sosa MS, Bragado P & Aguirre-Ghiso JA Mechanisms of disseminated cancer cell dormancy: an awakening field. Nat. Rev. Cancer 14, 611–622 (2014). [PubMed: 25118602]
3. Linde N, Fluegen G & Aguirre-Ghiso JA The Relationship Between Dormant Cancer Cells and Their Microenvironment. Advances in Cancer Research vol. 132 (Elsevier Inc., 2016).

4. Giancotti FG Review Mechanisms Governing Metastatic Dormancy and Reactivation. Cell 155, 750–764 (2013). [PubMed: 24209616]
5. Holmgren L, O'Reilly MS & Folkman J Dormancy of micrometastases: balanced proliferation and apoptosis in the presence of angiogenesis suppression. Nat. Med 1, 149–153 (1995). [PubMed: 7585012]
6. Townson JL & Chambers AF Dormancy of solitary metastatic cells. Cell Cycle 5, 1744–1750 (2006). [PubMed: 16861927]
7. Yeh AC & Ramaswamy S Mechanisms of Cancer Cell Dormancy--Another Hallmark of Cancer? Cancer Res. 75, 5014–5022 (2015). [PubMed: 26354021]
8. Phan TG & Croucher PI The dormant cancer cell life cycle. Nat. Rev. Cancer 20, 398–411 (2020). [PubMed: 32488200]
9. Fluegen G et al. Phenotypic heterogeneity of disseminated tumour cells is preset by primary tumour hypoxic microenvironments. (2017) doi:10.1038/ncb3465.
10. Bragado P et al. TGF-β2 dictates disseminated tumour cell fate in target organs through TGF-β-RIII and p38α/β signalling. Nat. Cell Biol 15, 1351–1361 (2013). [PubMed: 24161934]
11. Gao H et al. The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. Cell 150, 764–779 (2012). [PubMed: 22901808]
12. Kobayashi A et al. Bone morphogenetic protein 7 in dormancy and metastasis of prostate cancer stem-like cells in bone. J. Exp. Med 208, 2641–2655 (2011). [PubMed: 22124112]
13. Liu Y et al. Blockade of IDO-kynurenine-AhR metabolic circuitry abrogates IFN-γ-induced immunologic dormancy of tumor-repopulating cells. Nat. Commun 8, 1–15 (2017). [PubMed: 28232747]
14. Kai FB, Drain AP & Weaver VM The Extracellular Matrix Modulates the Metastatic Journey. Dev. Cell 49, 332–346 (2019). [PubMed: 31063753]
15. Bissell MJ, Hall HG & Parry G How does the extracellular matrix direct gene expression? J. Theor. Biol (1982) doi:10.1016/0022-5193(82)90388-5.
16. Bissell MJ & Aggeler J Dynamic reciprocity: how do extracellular matrix and hormones direct gene expression? Prog. Clin. Biol. Res (1987).
17. Eliceiri K et al. Automated quantification of aligned collagen for human breast carcinoma prognosis. J. Pathol. Inform (2014) doi:10.4103/2153-3539.139707.
18. Provenzano PP et al. Collagen reorganization at the tumor-stromal interface facilitates local invasion. BMC Med. 4, 38 (2006). [PubMed: 17190588]
19. Conklin MW et al. Aligned collagen is a prognostic signature for survival in human breast carcinoma. Am. J. Pathol (2011) doi:10.1016/j.ajpath.2010.11.076.
20. Naba A et al. The Matrisome: In Silico Definition and In Vivo Characterization by Proteomics of Normal and Tumor Extracellular Matrices. Mol. Cell. Proteomics 11, M111.014647–M111.014647 (2012).
21. Naba A, Clauser KR, Lamar JM, Carr SA & Hynes RO Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters. Elife 2014, 1–23 (2014).
22. Socovich AM & Naba A The cancer matrisome: From comprehensive characterization to biomarker discovery. Seminars in Cell and Developmental Biology (2019) doi:10.1016/j.semcdb.2018.06.005.
23. Taha IN & Naba A Exploring the extracellular matrix in health and disease using proteomics. Essays Biochem. 63, 417–432 (2019). [PubMed: 31462529]
24. Hebert JD et al. Proteomic Profiling of the ECM of Xenograft Breast Cancer Metastases in Different Organs Reveals Distinct Metastatic Niches. Cancer Res. 1–12 (2020) doi:10.1158/0008-5472.can-19-2961.
25. Tian C et al. Proteomic analyses of ECM during pancreatic ductal adenocarcinoma progression reveal different contributions by tumor and stromal cells. Proc. Natl. Acad. Sci. U. S. A (2019) doi:10.1073/pnas.1908626116.
26. Tian C et al. Cancer-cell-derived matrisome proteins promote metastasis in pancreatic ductal adenocarcinoma. Cancer Res. canres.2578.2019 (2020) doi:10.1158/0008-5472.can-19-2578.

27. Oskarsson T et al. Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. Nat. Med 17, 867–874 (2011). [PubMed: 21706029]
28. Barkan D et al. Metastatic growth from dormant cells induced by a Col-I-enriched fibrotic environment. Cancer Res. 70, 5706–5716 (2010). [PubMed: 20570886]
29. Aguirre-Ghiso JA, Liu D, Mignatti A, Kovalski K & Ossowski L Urokinase receptor and fibronectin regulate the ERK(MAPK) to p38(MAPK) activity ratios that determine carcinoma cell proliferation or dormancy in vivo. Mol. Biol. Cell 12, 863–879 (2001). [PubMed: 11294892]
30. Ghajar CM et al. The perivascular niche regulates breast tumour dormancy. Nat. Cell Biol 15, 807–17 (2013). [PubMed: 23728425]
31. Albrengues J et al. Neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in mice. 4227, (2018).
32. Hamidi H & Ivaska J Every step of the way: integrins in cancer progression and metastasis. Nat. Rev. Cancer (2018) doi:10.1038/s41568-018-0038-z.
33. Takai K et al. Discoidin domain receptor 1 (DDR1) ablation promotes tissue fibrosis and hypoxia to induce aggressive basal-like breast cancers. Genes Dev. 32, 244–257 (2018). [PubMed: 29483153]
34. Gao H et al. Multi-organ Site Metastatic Reactivation Mediated by Non-canonical Discoidin Domain Receptor 1 Signaling. Cell 166, 47–62 (2016). [PubMed: 27368100]
35. Adam AP et al. Computational identification of a p38SAPK-regulated transcription factor network required for tumor cell quiescence. Cancer Res. (2009) doi:10.1158/0008-5472.CAN-08-3820.
36. Sosa MS et al. NR2F1 controls tumour cell dormancy via SOX9- and RARβ-driven quiescence programmes. Nat. Commun 6, 6170 (2015). [PubMed: 25636082]
37. Kim RS et al. Dormancy signatures and metastasis in estrogen receptor positive and negative breast cancer. PLoS One 7, 1–8 (2012).
38. Aguirre-Ghiso JA, Estrada Y, Liu D & Ossowski L ERKMAPK activity as a determinant of tumor growth and dormancy; regulation by p38SAPK. Cancer Res. 63, 1684–1695 (2003). [PubMed: 12670923]
39. Malladi S et al. Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. Cell 165, 45–60 (2016). [PubMed: 27015306]
40. Aguirre-Ghiso JA, Ossowski L & Rosenbaum SK Green fluorescent protein tagging of extracellular signal-regulated kinase and p38 pathways reveals novel dynamics of pathway activation during primary and metastatic growth. Cancer Res. 64, 7336–7345 (2004). [PubMed: 15492254]
41. Aslakson CJ & Miller FR Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. Cancer Res. 52, 1399–1405 (1992). [PubMed: 1540948]
42. Ossowski L Plasminogen activator dependent pathways in the dissemination of human tumor cells in the chick embryo. Cell 52, 321–328 (1988). [PubMed: 3125981]
43. Aguirre Ghiso JA, Kovalski K & Ossowski L Tumor dormancy induced by downregulation of urokinase receptor in human carcinoma involves integrin and MAPK signaling. J. Cell Biol 147, 89–104 (1999). [PubMed: 10508858]
44. Montagner M et al. Crosstalk with lung epithelial cells regulates Sfrp2-mediated latency in breast cancer dissemination. Nat. Cell Biol 22, 289–296 (2020). [PubMed: 32094692]
45. Spencer SL et al. XThe proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. Cell 155, 369–383 (2013). [PubMed: 24075009]
46. Naba A et al. The extracellular matrix: Tools and insights for the ‘omics’ era. Matrix Biology (2016) doi:10.1016/j.matbio.2015.06.003.
47. Vogel B, Siebert H, Hofmann U & Frantz S Determination of collagen content within picrosirius red stained paraffin-embedded tissue sections using fluorescence microscopy. MethodsX 2, 124–134 (2015). [PubMed: 26150980]
48. Rittié L Method for Picrosirius Red-Polarization Detection of Collagen Fibers in Tissue Sections. Methods Mol. Biol 1627, 395–407 (2017). [PubMed: 28836216]

49. Wegner KA, Keikhosravi A, Eliceiri KW & Vezina CM Fluorescence of Picrosirius Red Multiplexed With Immunohistochemistry for the Quantitative Assessment of Collagen in Tissue Sections. J. Histochem. Cytochem. Off. J. Histochem. Soc 65, 479–490 (2017).
50. Coelho PGB, Souza M. V. de, Conceição LG, Viloria MIV & Bedoya SAO Evaluation of dermal collagen stained with picrosirius red and examined under polarized light microscopy. An. Bras. Dermatol 93, 415–418 (2018). [PubMed: 29924246]
51. Brisson BK et al. Type III Collagen Directs Stromal Organization and Limits Metastasis in a Murine Model of Breast Cancer. Am. J. Pathol 185, 1471–1486 (2015). [PubMed: 25795282]
52. Coelho NM et al. Discoidin Domain Receptor 1 Mediates Myosin-Dependent Collagen Contraction. Cell Rep. (2017) doi:10.1016/j.celrep.2017.01.061.
53. Coelho NM & McCulloch CA Mechanical Signaling Through the Discoidin Domain Receptor 1 Plays a Central Role in Tissue Fibrosis. Cell Adh. Migr 6918, 00–00 (2018).
54. Faraci E, Eck M, Gerstmayer B, Bosio A & Vogel WF An extracellular matrix-specific microarray allowed the identification of target genes downstream of discoidin domain receptors. Matrix Biol. 22, 373–381 (2003). [PubMed: 12935821]
55. Gearing LJ et al. CiiiDER: A tool for predicting and analysing transcription factor binding sites. PLoS One 14, e0215495 (2019). [PubMed: 31483836]
56. Beck AH, Espinosa I, Gilks CB, van de Rijn M & West RB The fibromatosis signature defines a robust stromal response in breast carcinoma. Lab. Invest 88, 591–601 (2008). [PubMed: 18414401]
57. Aguirre-Ghiso J et al. A Mesenchymal-like Program of Dormancy controlled by ZFP281 Serves as a Barrier To Metastatic Progression of Early Disseminated Cancer Cells. Res. Sq (2021) doi:10.21203/rs.3.rs-145308/v1.
58. Chiusa M et al. The Extracellular Matrix Receptor Discoidin Domain Receptor 1 Regulates Collagen Transcription by Translocating to the Nucleus. J. Am. Soc. Nephrol 30, 1605–1624 (2019). [PubMed: 31383731]
59. Gould LJ Topical Collagen-Based Biomaterials for Chronic Wounds: Rationale and Clinical Application. Adv. Wound Care 5, 19–31 (2016).
60. Aguirre-Ghiso J. a, Bragado P & Sosa MS Metastasis awakening: targeting dormant cancer. Nat Med 19, 276–277 (2013). [PubMed: 23467238]
61. Ghajar CM Metastasis prevention by targeting the dormant niche. Nature reviews. Cancer vol. 15 238–247 (2015). [PubMed: 25801619]
62. Juin A et al. Discoidin domain receptor 1 controls linear invadosome formation via a Cdc42-Tuba pathway. J. Cell Biol 207, 517–533 (2014). [PubMed: 25422375]
63. Naba A, Clauser KR & Hynes RO Enrichment of Extracellular Matrix Proteins from Tissues and Digestion into Peptides for Mass Spectrometry Analysis. J. Vis. Exp (2015) doi:10.3791/53057.
64. Naba A et al. The Matrisome: *In Silico* Definition and *In Vivo* Characterization by Proteomics of Normal and Tumor Extracellular Matrices. Mol. Cell. Proteomics (2012) doi:10.1074/mcp.M111.014647.
65. Di Martino J et al. 2D and 3D Matrices to Study Linear Invadosome Formation and Activity. J. Vis. Exp 350, 78–89 (2017).
66. Franco-Barraza J, Beacham DA, Amatangelo MD & Cukierman E Preparation of extracellular matrices produced by cultured and primary fibroblasts. Curr. Protoc. Cell Biol (2016) doi:10.1002/cpcb.2.
67. Perez-Riverol Y et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 47, D442–D450 (2019). [PubMed: 30395289]

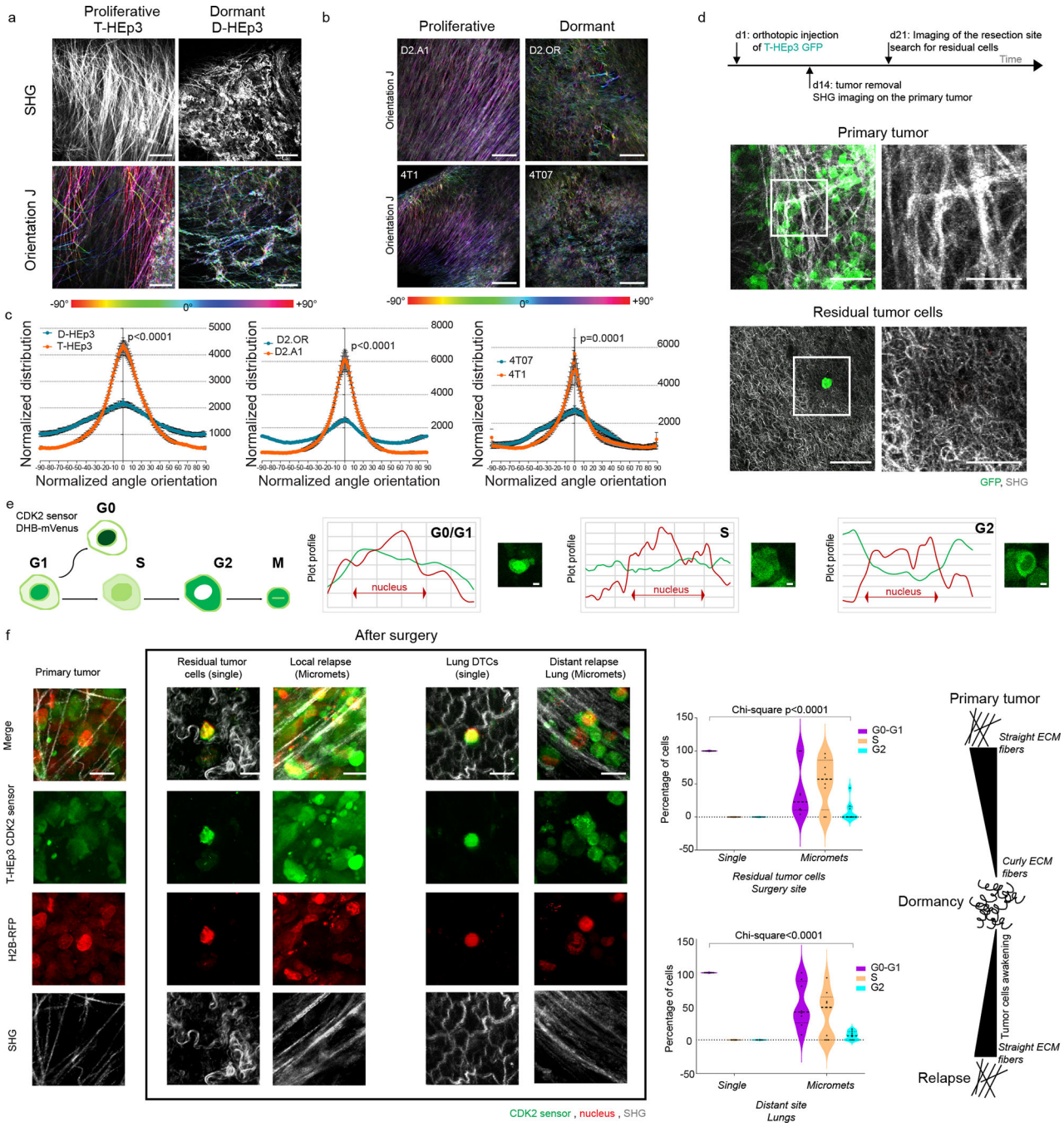


Fig. 1. Characterization of the ECM architecture around dormant cells

All numerical data are presented as mean +/−SEM.

(a) Upper panel: Representative SHG images from T-HEp3 and D-HEp3mice tumors. Lower panel: Representative Collagen OrientationJ output obtained from SHG images. More colors = less alignment. Color Scale bar represent fiber orientation. Scale bar, 100***μ***m. (n=5 tumors per group)

(b) OrientationJ visual output obtained from SHG images of D2.A1, D2.OR, 4T1 and 4T07 CAMs tumors. Scale bars, 100***μ***m. (n=5 tumors per group)

(c) Normalized distribution of collagen fiber orientation from tumors presented in A and B. (n= 5 mice for D-HEp3/T-HEp3, 2 images per animal; n= 5 CAMs for D2.OR/D2.A1 and 4T07/4T1, 2 images per animal). Cumulative distributions were compared using an unpaired two-tailed Kolmogorov-Smirnov test with 95% confidence level. Data are presented as mean values +/− SEM.

(d) Top panel: Representative multiphoton images of T-HEp3-GFP primary tumors. Bottom panel: Residual local tumor cells after surgery. Scale bars, 100***μ***m. Red square represents a magnified area. Scale bars, 50***μ***m. (n=5 mice). d refers to days on timeline scheme.

(e) CDK2 sensor (DHB-mVenus) fluorescence dynamics correlated with phase of the cell cycle. Example of plot profile for CDK2 sensor and H2B-RFP fluorescence intensity during each phase of the cell cycle based on imaging data. Inserts are representative images of the sensor for each phase of the cell cycle *in vivo* in T-HEp3 mice tumor. Scale bars, 5***μ***m.

(f) Left panel: Representative two-photon imaging of T-HEp3-DHB-mVenus in mice. T-HEp3 cells in green, H2B-RFP in red and SHG in gray. Scale bars, 20***μ***m. Middle panel: Violin plots of percentage of cells in each phase of the cell cycle at surgery site and in lungs (n=5 mice, 38 single cells were counted per condition). Distributions were compared using a One-tailed Chi-squared test with 95% confidence interval. Right panel: A scheme showing the ECM architecture changes as cancer progresses.

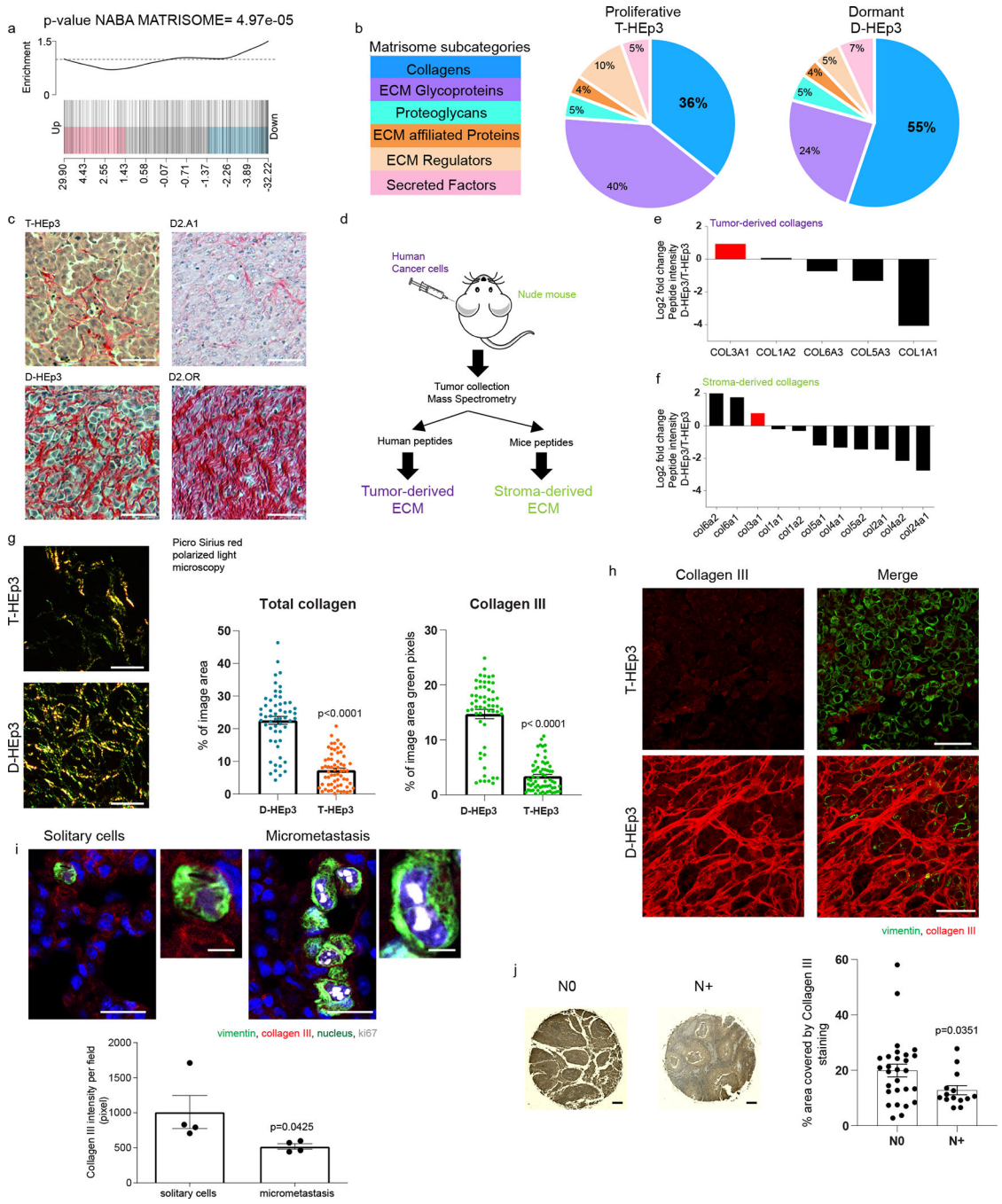


Fig. 2. Proteomic analysis of the ECM of dormant and proliferative tumors

All numerical data are presented as mean +/−SEM.

(a) Enrichment plot for matrisome signature from RNA sequencing performed in D-HEp3 and T-HEp3 mice tumors. X-axis shows log2FC for T-HEp3 versus D-HEp3. Black bars represent matrisome genes. (n= 3 tumors per group) (related to Supplementary Table 2).

(b) Proteomic characterization of the matrisome of D-HEp3 and T-HEp3 tumors related to Supplementary Table 1C. Each color represents a subcategory of the matrisome signature.

(c) Picro Sirius red staining of proliferative and dormant mice tumors. Collagen fibers are stained in pink, nucleus in dark grey.) Scale bars, 50***μ***m.
(d) Scheme of the mass spectrometry experiment allowing the separation between the tumor-derived and stroma-derived ECM.
(e) Relative abundance of tumor-derived collagens identified in D-HEp3 and T-HEp3 tumors (n= 3 tumors per group) (related to Supplementary Table 1).
(f) Relative abundance of stroma-derived collagens identified in D-HEp3 and T-HEp3 tumors (n= 3 tumors per group) (related to Supplementary Table 1).
(g) Polarized-light microscopy images of Picrosirius red staining from proliferative and dormant mice tumors. Thick collagen fibers are colored in red to yellow (equivalent to type I collagen) and thin fibers are colored in green (equivalent to type III collagen) Scale bars, 50***μ***m. Graphs from left to right: (1) Percentage of total collagen covered area per field comparing T-HEp3 and D-HEp3 tumors. (n=56 images for D-HEp3: n=63 images for T-HEp3 over 4 mice) (2) Analysis of the green component (type III collagen) of images using CIELAB module in ImageJ comparing T-HEp3 and D-HEp3 tumors. (n=56 images for D-HEp3: n=63 images for T-HEp3 over 4 mice). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.
(h) Immunofluorescence of D-HEp3 and T-HEp3 mice tumors for type III collagen (red). Cancer cells were stained with vimentin (green). Scale bars, 50***μ***m.
(i) Top: Representative image of type III collagen staining of solitary DTCs and micrometastasis of spontaneously disseminated T-HEp3 cells in mice lungs. DTCs were stained with vimentin (green), type III collagen (red) and Ki67 (grey) and DAPI (blue). Scale bars, 20***μ***m. Red insert shows a zoom of the single cell. Scale bars, 5***μ***m Bottom: quantification of type III collagen staining intensity per same size field between solitary cells and micrometastasis (n= 4 images from 4 mice lungs). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.
(j) Tissue microarray immunohistochemistry staining of type III collagen. Left panel: representative images of type III collagen staining in N0 cohort (no positive lymph node) versus N+ (1 or more positive lymph nodes) in HNSCC patient samples. Type III collagen in brown and nucleus in blue. Scale bars, 100 ***μ***m. Right panel: Percentage of area covered by type III collagen staining between N0 (n=28 samples) and N+ HNSCC (n=14 samples). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.

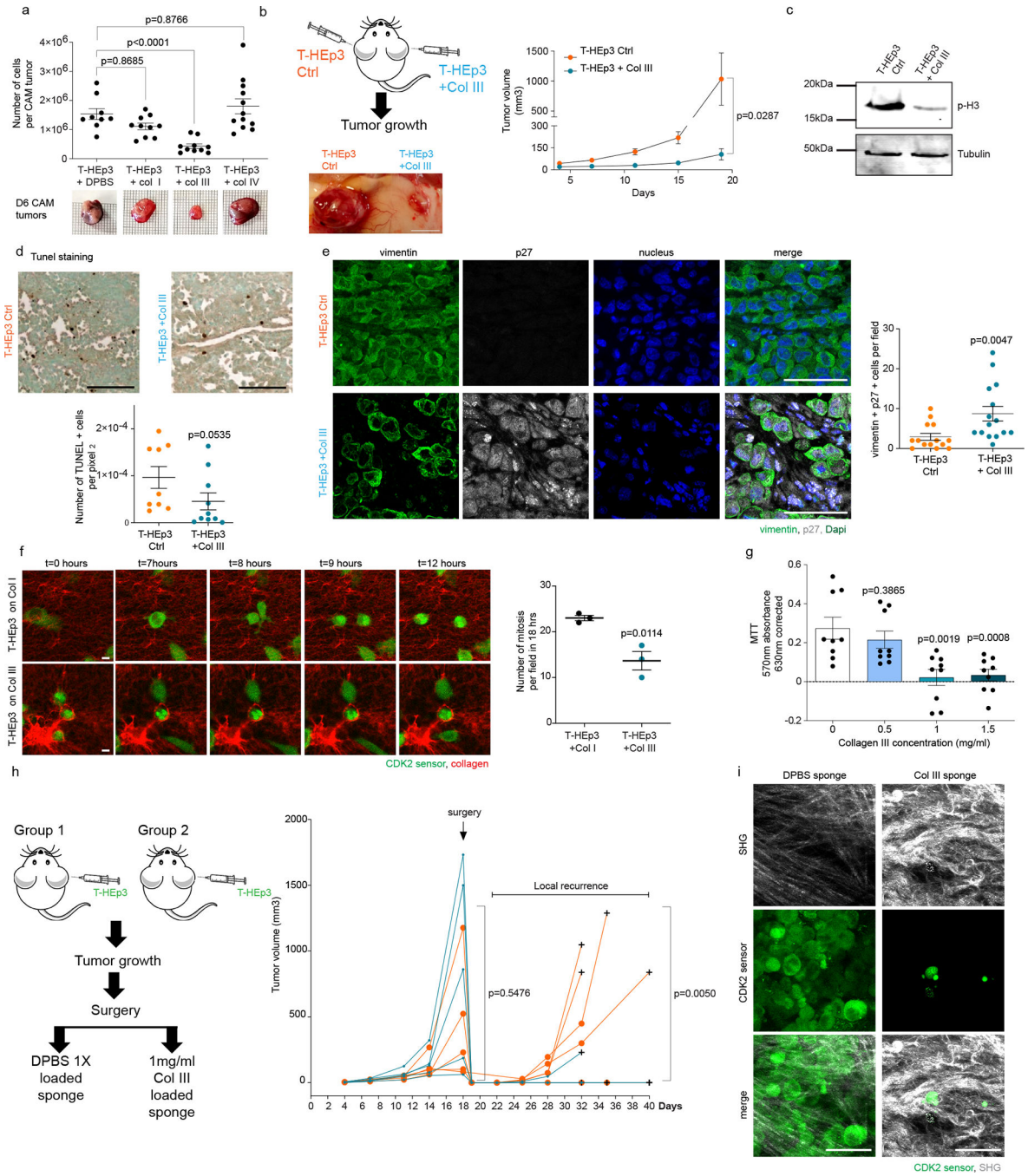


Fig. 3. Type III collagen enriched microenvironments induces dormancy

All numerical data are presented as mean +/−SEM.

(a) Number of cells per CAM tumor for T-HEp3 co-injected with DPBS (n= 9 CAMs), collagen I (n=10 CAMs), collagen III (n=10) or collagen IV(n=12). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.

(b) Representative picture of both tumors at day 19 after inoculation in a same animal. Tumor growth over time (n=8 mice control n= 6 mice co-injected with type III collagen).

Curves were compared using a one-tailed two-way ANOVA with mixed model effects analysis and a Bonferroni correction and a 95% confidence interval.
(c) Western blot for phospho-histone H3 and tubulin loading control to compare T-HEp3 plated on plastic versus type III collagen for 24hrs *in vitro*.
(d) TUNEL assay on tumors from (b). Scale bar, 200 microns. Graph represents the quantification of the number of TUNEL positive cells per pixel^2 (n=9 sections from 4 mice tumors control, n=10 sections from 4 mice co-injected with type III collagen). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.
(e) Left panel: Immunofluorescence of tumors from experiment in (b). vimentin (green), p27 (gray) and DAPI (blue). Right panel: Number of p27(+) cells per field (n= 14 images from 4 mice control, n= 15 images from 4 mice co-injected with type III collagen). Scale bar, 50μm. Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.
(f) Left panel: Time points from an 18hrs time lapse movie of T-HEp3 plated on type I or III collagen. (t=hours). CDK2-mVenus sensor (green) and 546-succinimidylester labelled collagen is shown (red). Scale bar, 10μm. Right: Quantification of the number of mitosis per field across 18hrs (n= 3 independent movies with more than 10 cells per field per condition). Data were compared using an unpaired two-tailed T-test. (Related to Supplemental Movies 1 and 2).
(g) MTT assay of T-HEp3 plated on different concentrations of type III collagen for 24 hrs. (n= 3 independent experiments in triplicate per condition). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.
(h) Experimental design for treatment with type III collagen bioengineered scaffolds. Tumor growth over time before surgery and local tumor relapse over time, post-surgery for 21 days. (n=5 mice per group).Groups were compared prior to surgery using unpaired two-tailed Mann-Whitney test with 95% confidence level and relapses compared using a mixed model effects analysis with a Bonferroni correction and a 95% confidence interval.
(i) Representative images of the resected area after tumor surgery at the time mice were sacrificed. SHG signal in gray and T-HEp3 CDK2 biosensor expressing cells in green. White arrow points towards a G0 T-HEp3 cell (nuclear localization of CDK2 sensor). Scale bar, 50μm.

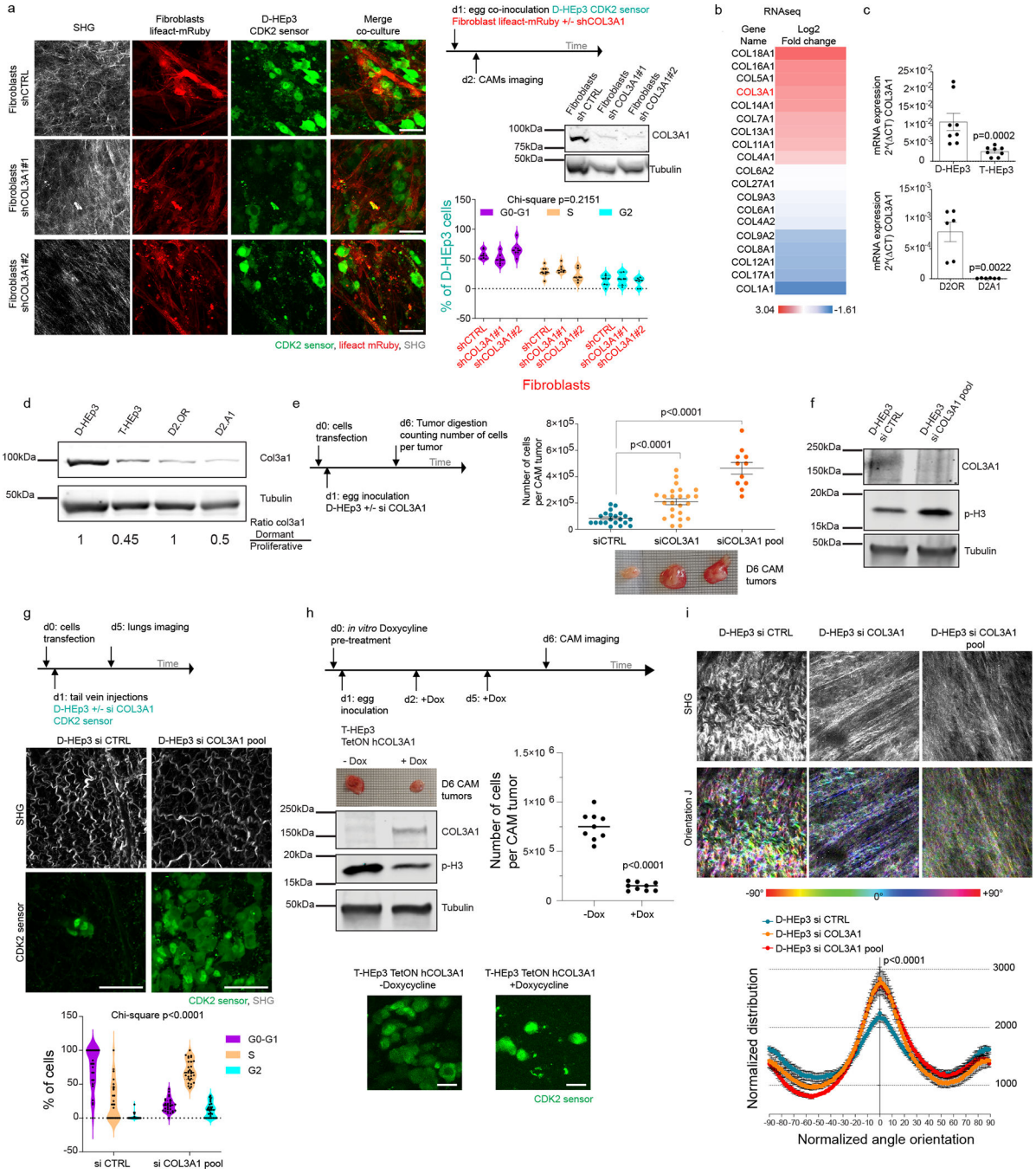


Fig. 4. Tumor-cell-derived type III collagen regulates dormancy

All numerical data are presented as mean +/−SEM.

(a) Left panel: Representative multiphoton images of fibroblasts expressing an shRNA CTRL or targeting *COL3A1* co-cultured with D-HEp3 expressing CDK2 sensor in CAMs for 24hrs. SHG signal displayed in grey, fibroblasts inred, cancer cells in green. Scale bar, 50***μ***m. Top right panel: Western blot for type III collagen and tubulin comparing fibroblasts +/−sh*COL3A1*. Bottom right panel: Percentage of D-HEp3 cells in each phase of the cell

cycle. (n= 4 CAMs per group, at least 5 images per CAM). Distributions were compared using a One-tailed Chi-squared test with 95% confidence interval..

(b) List of significantly dysregulated collagen genes in D-HEp3 nodules and T-HEp3 tumors grown in nude mice. Data represented as log2 fold change D-HEp3/T-HEp3 (related to Supplementary Table 2).

(c) COL3A1 mRNA expression in different dormancy models. (HEp3 model n= 4 mice tumor in duplicates) and (D2 model n=3 independent RNA extraction for cultured cells in duplicates). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.

(d) Western blot for type III collagen and tubulin loading control in different dormancy models *in vivo* (CAMs).

(e) Number of cells per tumors in CAM assays(siCTRL n= 21 CAMs, siCOL3A1 n= 24 CAMs, siCOL3A1 pool n= 11CAMs). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Representative image of CAM tumors 6 days after inoculation are displayed below. d refers to days on timeline scheme.

(f) Western blot for type III collagen, phospho-histone H3 and tubulin comparing D-HEp3 +/− si*COL3A1*.

(g) Top panel: Representative multiphoton images of D-HEp3 siCTRL and siCOL3A1 disseminated in the lungs. Scale bar, 50μm, a 0.5 pixel radius minimum filter was applied to the green channel. Bottom panel: Percentage of cells in each phase of the cell cycle.(n=7 mice per group, at least 5 images per lung analyzed). Distributions were compared using a one-tailed Chi-square test with 95% confidence interval. d refers to days on timeline scheme.

(h) Upper left panel: Representative CAMs tumors for each condition and Western blot for COL3A1, phospho-Histone H3 and tubulin. Upper right panel: number of T-HEp3 expressing CDK2 sensor and dox-inducible human COL3A1 cells per CAM tumor +/− doxycycline treatment. (n=9 CAMs per group) Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Lower panel: Representative images of the different conditions are displayed, Scale bar, 20 μm. d refers to days on timeline scheme.(i) Top panel: Representative SHG images and OrientationJ software's visual output obtained for each condition. Images extracted from tumors grown on CAMs. Scale bars, 100μm. More colors = less alignment. Color Scale bars represent fiber orientation. Lower panel: Normalized distribution of collagen fiber orientation. (n= 3 tumor per group, siCTRL n= 14 images, si COL3A1 n=12 images, siCOL3A1 pool n=15 images). Cumulative distributions were compared using an unpaired two-tailed Kolmogorov-Smirnov test with 95% confidence level.

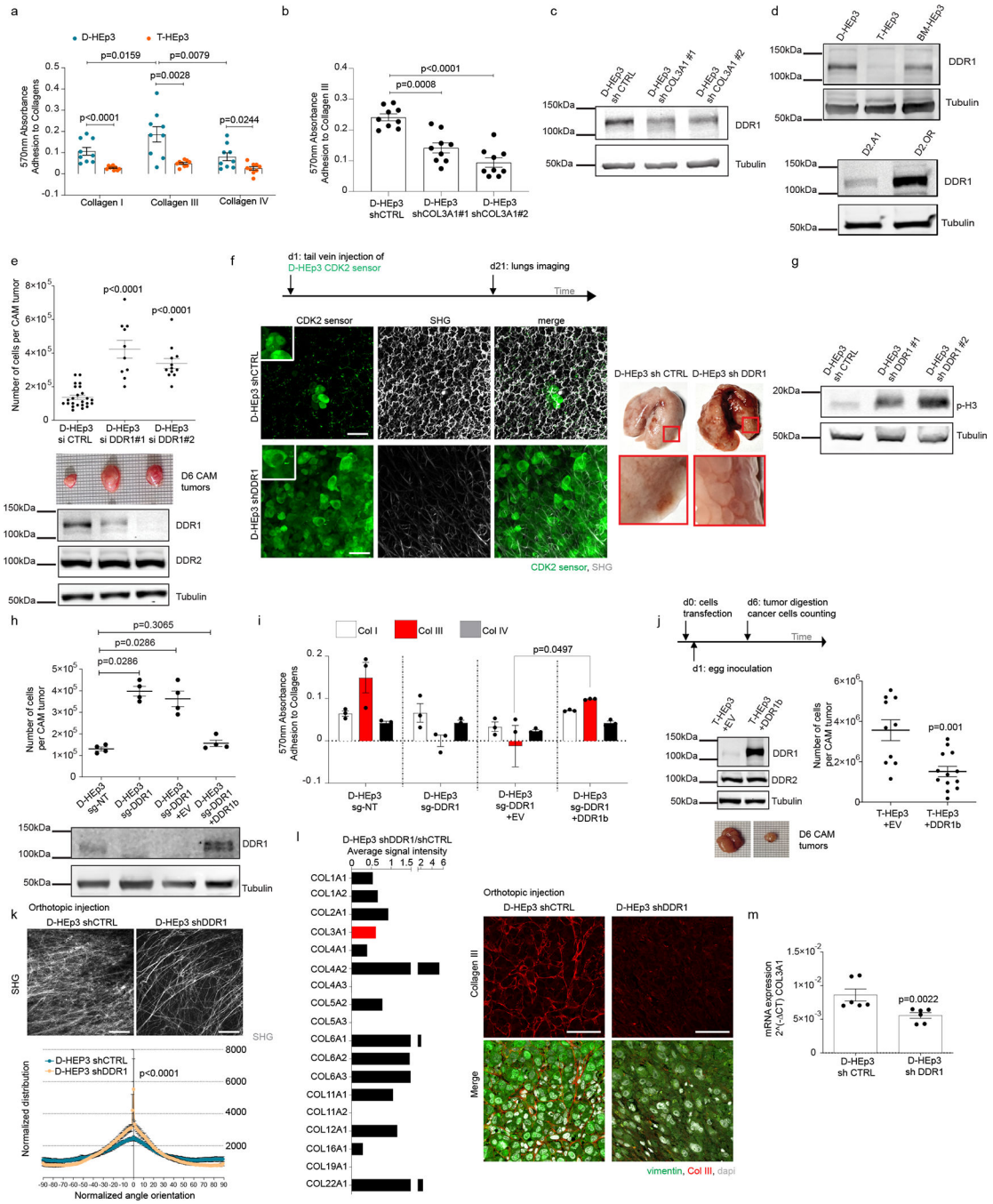


Fig. 5. DDR1 is required to sustain dormancy

All numerical data are presented as mean +/−SEM.

(a) Adhesion assay of D-HEp3 and T-HEp3 on type I, III and IV collagens. (n= 3 independent experiment with triplicates). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.

(b) Adhesion assay of D-HEp3 expressing an shRNA control or targeting COL3A1 on type III collagen. (n= 3 independent experiment with triplicates). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.

(c) Western blot for DDR1 and tubulin, on D-HEp3 expressing an shRNA control or targeting *COL3A1.*
(d) Western blot for DDR1 and tubulin from different dormancy models.
(e) Upper panel: Number of D-HEp3 cells per CAM tumors(si CTRL n= 24 CAMs, siDDR1#1 n=10 CAMs, si DDR1#2 n=12 CAMs)Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.
Lower panel: Representative images of tumors for each condition and Western blot showing DDR1, DDR2 and tubulin levels upon DDR1 depletion.
(f) Left panel: Representative multiphoton images of D-HEp3 shRNA CTRL and shDDR1 disseminated in the lungs. Scale bar, 50***μ***m. A 0.5 pixel radius minimum filter was applied to the green channel. Right panel: Representative images of lungs and inserts show macrometastases in 60% of the shDDR1 group. d refers to days on timeline scheme.
(g) Western blot for phospho-histone H3 and tubulin, on D-HEp3 expressing an shRNA control or targeting DDR1.
(h) Upper panel: Number of D-HEp3 cells per CAM tumor (n= 4 CAM tumors per group). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Lower panel : Western blot for DDR1 and tubulin.
(i) Adhesion assay to collagen I, III or IV of D-HEp3 expressing a control sgRNA (NT) or an sgRNA against DDR1. (n= 3 independent experiment with triplicates).Data were compared using a one-tailed two-way ANOVA with mixed model effects analysis and a Bonferroni correction and a 95% confidence interval
(j) Right panel: Number of T-HEp3 cells per CAM tumor (+EV n= 10 CAMs, +DDR1b n=13 CAMs). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Left panel: Western blot for DDR1, DDR2 and tubulin and representative CAM tumors.d refers to days on timeline scheme.
(k) Top panel:Representative SHG images of ECM of D-HEp3 mice tumors +/−shDDR1. Scale bar, 100***μ***m. Distribution of fiber orientation of each analyzed sample. (n= 5 tumors per group2 images analyzed per tumor). Cumulative distributions were compared using an unpaired two-tailed Kolmogorov-Smirnov test with 95% confidence level.
(l) Left panel: Relative collagen abundance of tumor-derived collagens identified in D-HEp3 +/− shDDR1 tumors. Related to Supplementary Table 3. Right panel: Representative immunofluorescence onD-HEp3 +/− shRNA DDR1 mice tumors. (n=3 tumors stained) Scale bar, 50***μ***m.
(m) COL3A1 mRNA expression from D-HEp3 tumors (n=3 independent RNA extractions from 3 different tumors in duplicates). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.

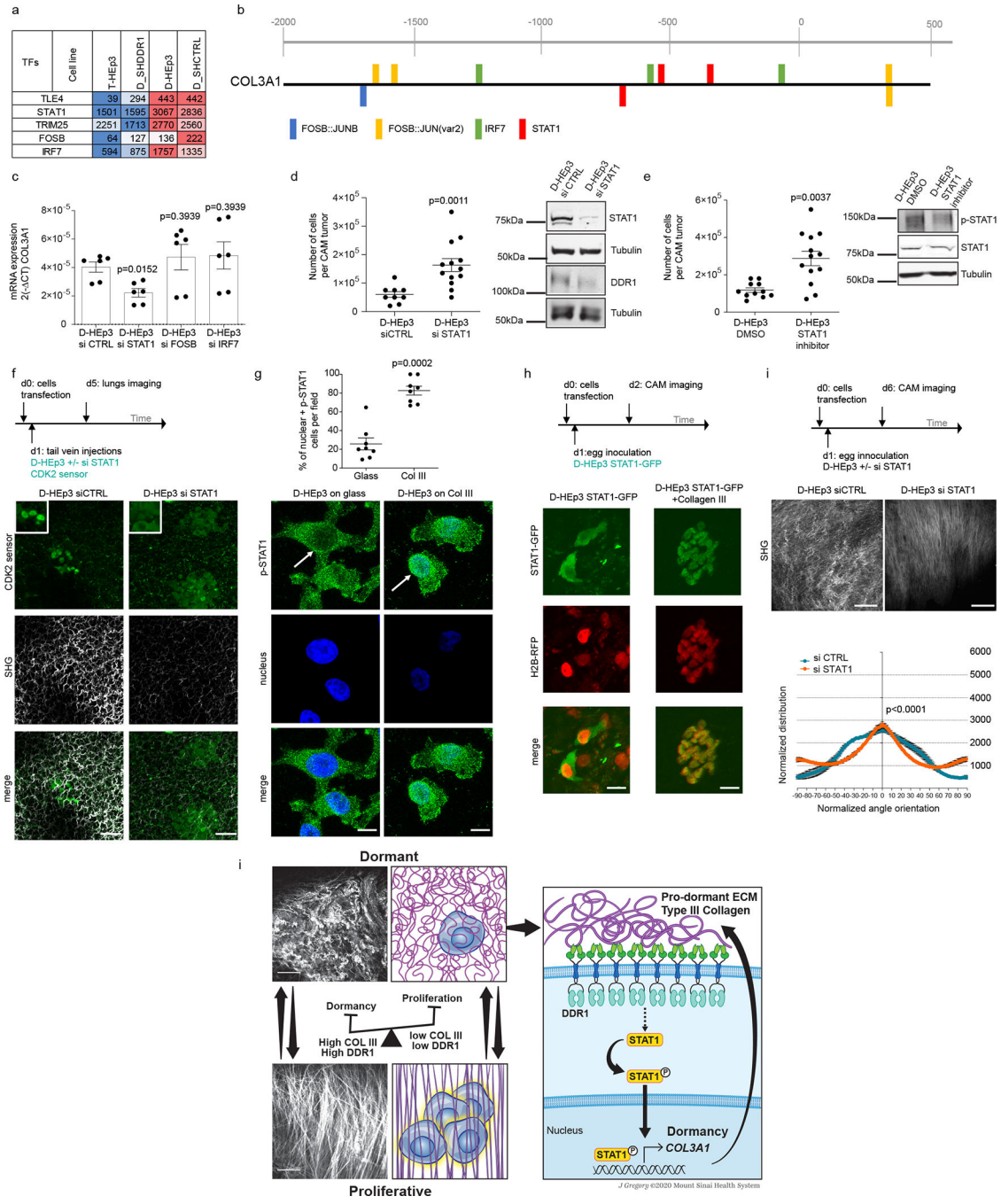


Fig. 6. A DDR1/STAT1 pathway regulates dormancy and COL3A1 expression

All numerical data are presented as mean +/−SEM.

(a) Transcription factors significantly dysregulated between T-HEp3 and D-HEp3 and significantly downregulated in D-HEp3 shDDR1 compared with D-HEp3 shRNA CTRL. Values represent average read counts. Color scales range from dark blue (lowest value) to dark red (highest value) per line (n= 3 tumors per group). Related to Supplemental tables 5 and 6.

(b) Map of predicted sites for transcription factors identified above in COL3A1 promoter using the CiiiDER tool.
(c) COL3A1 mRNA expression from D-HEp3 +/− si against the different transcription factors (n=3 independent RNA extractions from 3 different tumors in duplicates). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.
(d) Left panel: Number of D-HEp3 cells per CAM tumor +/− siSTAT1 (si CTRL n= 9 CAMs, siSTAT1 n= 13 CAMs) Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level. Right panel: Western blot for STAT1, DDR1 and tubulin expression.
(e) Left panel: Number of D-HEp3 cells per CAM tumor +/− Fludarabine (DMSO n= 11 CAMs, Fludarabine n= 13 CAMs). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence
(f): Representative multiphoton images of D-HEp3 siCTRL and siSTAT1 disseminated in the lungs. Scale bar 50μm. A 0.5 pixel radius minimum filter was applied to the green channel.d refers to days on timeline scheme.
(g)Upper panel: number of cells with nuclear pSTAT1 per field. (n= 8 fields of view, 155 cells counted on glass and 87 cells counted on type III collagen). Data were compared using unpaired two-tailed Mann-Whitney test with 95% confidence level.
Lower panel: Immunofluorescence of phospho-STAT1 in D-HEp3 plated either on glass or on Type III collagen. Scale bar 10μm.
(h) Representative images of D-HEp3 stably expressing H2B-RFP and transfected with a STAT1-GFP plasmid and seeded on CAMs with DPBS or type III collagen (1.5mg/ml) for 3 days. (n= 3 CAMs, at least 5 images per egg). Scale bar, 20 μm. d refers to days on timeline scheme.
(i) Top panel: Representative SHG images of control and siRNA STAT1 CAM tumors. Scale bar, 100μm. Bottom panel: Normalized distribution of collagen fiber orientation from CAM tumors. (sictrl n=9 images from 4 independent CAMs, siSTAT1 n= 12 images from 4 independent CAMs) Cumulative distributions were compared using an unpaired two-tailed Kolmogorov-Smirnov test with 95% confidence level. **(j) Working model:** ECM architecture dynamically changes between dormant and proliferative tumors. The ECM around dormant cells is characterized by a wavy collagen matrix that reorganizes into a highly aligned matrix upon dormant cell awakening. Dormant cell ECM is enriched in type III collagen that contributes towards establishing the non-linear collagen ECM architecture around dormant cells. DDR1 binding to type III collagen activates STAT1 signaling to activate dormancy and increases COL3A1 expression in DTCs, establishing a pro-dormant ECM niche.

Table 1:

Number of mice with local recurrence after sponge treatment

Recurrence at 40 days	Mice	Percentage
Empty sponge	4/5	80
Col III sponge	1/5	20

Table 2:

Number of mice with lungs DTCs following tail vein injections.

Related to results in Figure 4g			
Data presented as number of animals	Single cells	Clusters <20 cells	Micromets
Si CTRL	7/7	5/7	0/7
Si COL3A1 pool	0/7	3/7	6/7
Related to results in Figure 5f			
Data presented as number of animals	Single cells	Clusters <20 cells	Micromets
Sh CTRL	3/4	1/4	1/4
Sh DDR1	0/4	3/4	4/4