

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

**Julie Di Martino**

Mount Sinai Icahn School of Medicine <https://orcid.org/0000-0002-2163-0736>

**Ana Rita Nobre**

Icahn School of Medicine at Mount Sinai

**Chandrani Mondal**

Mount Sinai School of Medicine

**Isrha Taha**

Medical School in St. Louis

**Eduardo Farias**

Icahn School of Medicine at Mount Sinai, New York, USA

**Elana Fertig**

Sidney Kimmel Comprehensive Cancer Center

**Alexandra Naba**

University of Illinois at Chicago <https://orcid.org/0000-0002-4796-5614>

**Julio Aguirre-Ghiso**

Icahn School of Medicine at Mount Sinai <https://orcid.org/0000-0002-6694-6507>

**Jose Javier Bravo-Cordero** (✉ [josejavier.bravo-cordero@mssm.edu](mailto:josejavier.bravo-cordero@mssm.edu))

Icahn School of Medicine at Mount Sinai, New York, USA

---

## Article

**Keywords:** Dormancy, tumor microenvironment, extracellular matrix, matrisome, type III collagen, DDR1, multiphoton imaging

**Posted Date:** June 22nd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-580847/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Nature Cancer on December 13th, 2021. See the published version at <https://doi.org/10.1038/s43018-021-00291-9>.



## **A type III collagen-rich ECM niche regulates tumor dormancy**

Julie S. Di Martino<sup>1</sup>, Ana Rita Nobre<sup>2,3</sup>, Chandrani Mondal<sup>1</sup>, Isra Taha<sup>4,5</sup>, Eduardo F. Farias<sup>1</sup>, Elana J. Fertig<sup>6</sup>, Alexandra Naba<sup>4,5</sup>, Julio A. Aguirre-Ghiso<sup>2</sup> and Jose Javier Bravo-Cordero<sup>1\*</sup>.

### **Affiliations:**

<sup>1</sup>Division of Hematology and Oncology, Department of Medicine, The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

<sup>2</sup>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

<sup>3</sup>Abel Salazar Biomedical Sciences Institute, University of Porto, Porto, Portugal

<sup>4</sup>Department of Physiology and Biophysics, University of Illinois at Chicago, Chicago, IL, USA

<sup>5</sup>University of Illinois Cancer Center, University of Illinois at Chicago, Chicago, IL, USA

<sup>6</sup>Departments of Oncology, Applied Mathematics and Statistics, and Biomedical Engineering, Johns Hopkins University, Baltimore, MA, USA

\*Corresponding author: Jose Javier Bravo-Cordero email: [josejavier.bravo-cordero@mssm.edu](mailto:josejavier.bravo-cordero@mssm.edu)

## ABSTRACT

Cancer cells disseminate from primary tumors and seed in distant organs, where they can remain dormant for many years before metastases become clinically detectable. Little is known about how extracellular matrix (ECM) sensing and remodeling can induce and sustain dormancy of disseminated tumor cells (DTCs). Further, whether dormant cells assemble ECM niches to sustain their phenotype is also an unanswered question. By using ECM proteomics, we found that dormant cancer cells assemble an ECM niche enriched in type III collagen. Assembly of a type III collagen-rich ECM is required to sustain tumor dormancy as its disruption restores proliferation of dormant cancer cells. Mechanistically, we show that type III collagen interacts with DDR1 to activate STAT1 signaling to induce and maintain dormancy. Second Harmonic Generation two-photon microscopy further reveals that the dormancy-to-reactivation transition is accompanied by changes in collagen three-dimensional architecture and type III collagen abundance. *In vivo*, exogenous type III collagen stops tumor growth by inducing dormancy. Type III collagen also prevents reawakening of residual dormant cells by prolonging their quiescence *in vivo*. Our data reveal a novel ECM-dependent mechanism by which dormant DTCs depend on the assembly of a pro-quiescence type III collagen-rich ECM niche. Manipulation of these mechanisms could serve as a self-sustained barrier to metastasis through DTC dormancy induction.

**Key words:** Dormancy, tumor microenvironment, extracellular matrix, matrisome, type III collagen, DDR1, multiphoton imaging

### One Sentence Summary:

Here we describe how cancer cells construct an extracellular matrix niche to induce and maintain dormancy, preventing their immediate growth.

## INTRODUCTION

The majority of cancer deaths are due to metastasis<sup>1</sup>. However, metastases can occur years or decades after primary tumor removal. This delay is a consequence of the activation of pathways that maintain disseminated tumor cells (DTCs) in distant organs in a non-proliferative state, also called tumor cell dormancy<sup>2-4</sup>. Dormancy is defined as a reversible growth arrest (G0/G1) that can be regulated by the interaction of DTCs with the microenvironment<sup>3</sup>. Several factors, including many extracellular matrix (ECM) bound (i.e. TGF $\beta$ II<sup>5</sup>, BMP4, BMP7<sup>2</sup>, INF- $\gamma$ <sup>6</sup>) as well as microenvironmental conditions, such as increased hypoxia<sup>7</sup>, have been involved in the activation of dormancy in solid tumors.

The ECM is a complex meshwork of proteins providing not only an anchor for tumor cells but also a source of signals that regulate tumor cell behavior during metastasis<sup>8</sup>, showing a dynamic reciprocity between tumor cells and the ECM<sup>9,10</sup>. Along these lines, the three-dimensional architecture of the ECM has been shown to correlate with the invasiveness and metastatic potential of tumors<sup>11-13</sup>. Recent proteomics studies have also profiled “the matrisome” (a collection of ECM and ECM-related proteins), of both primary tumors and metastasis<sup>14-17</sup>. These studies highlighted that tumor-cell-derived ECM molecules plays an important role in metastatic progression<sup>15,18-20</sup>. However, how tumor cells may tap into ECM-derived cues to induce and maintain dormancy is largely unknown.

Importantly, the ECM, a major component of the tumor microenvironment, also plays a role in determining the fate of DTCs. ECM proteins such as tenascin-C<sup>21</sup>, type I collagen<sup>22</sup> or fibronectin<sup>23</sup> 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, such as, thrombospondin 1 to maintain dormancy of cancer cells<sup>24</sup>. More recently, proteolysis of the ECM protein laminin-111 by the matrix metalloproteinase MMP9 secreted by neutrophils, was shown to interrupt dormancy and drive metastasis outgrowth through integrin  $\alpha$ 3 $\beta$ 1<sup>25</sup>, suggesting that remodeling of the ECM microenvironment can activate metastatic growth after dormancy. Although these studies point towards a role for the ECM and ECM remodeling as drivers of reactivation of growth from dormancy, little attention has been paid to the role of tumor cell-derived ECM and its role in dormancy. Moreover, the ability of DTCs to assemble a dormancy-permissive ECM niche to sustain their cellular state remains unknown.

Tumor cells are capable of sensing the ECM through transmembrane ECM receptors. While the integrin family of receptors is well-characterized and their role in the different stages of tumor progression has been described<sup>26</sup>, the role of other receptors, such as the 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, ablation of DDR1 induced

aggressiveness and metastasis<sup>27</sup>. Interestingly, DDR1 has been involved in reactivation after dormancy<sup>28</sup> and it has been also identified as part of a tumor cell quiescence signature<sup>29,30</sup> that predicted for dormancy and late relapse in ER+ breast tumors<sup>31</sup>, suggesting it may also have implications in tumor dormancy; however, its role in this context has not been explored.

The activation 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<sup>23,32</sup>. Activation of p38 lead to activation of quiescence and survival pathways and the induction of transcription factors such as NR2F1, that upregulate several quiescence genes<sup>30</sup>. Other pathways involving transcription factors such as SOX9, RAR $\beta$ , SOX2<sup>30,33</sup>, SMAD1/5<sup>5</sup> or STAT1<sup>6</sup> have been linked to the activation and initiation of dormancy suggesting that the molecular landscape of dormancy may include the coordination of multiple signaling cues and nodes.

In this study, we address the role of tumor cell-derived ECM on inducing and sustaining dormancy. We present evidence for a new mechanism of tumor dormancy, including the assembly of a pro-quiescence ECM by cancer cells. By using a combination of ECM proteomics, multiphoton microscopy imaging, and a cell cycle biosensor, we have defined the matrisome of dormant cells and identified type III collagen as one of the ECM proteins required to sustain dormancy. Moreover, we have identified a type III collagen/DDR1/STAT1 pathway as a signaling node required to establish a type III collagen-rich ECM niche. Intervention aimed at preventing the awakening of dormant cells has been suggested as a therapeutic strategy to prevent metastatic outgrowth<sup>30,34,35</sup> and it 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. They also illustrate a way by which therapeutic targeting of pro-metastatic ECM by correcting its composition to a pro-dormancy role can lead to prevention of local recurrence or metastatic outgrowth.

## RESULTS

### **The collagen ECM around dormant cancer cells is characterized by wavy fibers.**

First, we sought to determine the three-dimensional architectural organization of the ECM surrounding dormant and proliferative tumor cells. To do so, we performed multiphoton second harmonic generation (SHG) imaging and analyzed collagen fiber orientation in several cellular dormancy models and their proliferative counterparts: proliferative (T-HEp3) and dormant (D-HEp3) head and neck squamous carcinomas (HNSCC)<sup>5,29,36</sup>, and two models of proliferative and dormant mammary tumors (4T1 and D2.1A tumors and respective dormant counterparts 4T07<sup>37,38</sup> and D2.OR<sup>22,25</sup>). Our results show that

dormant tumors have an ECM with wavy collagen fibers, which have reduced collagen fiber alignment and orientation when compared to proliferative tumors (**Fig. 1A and B**).

Given the architectural differences observed between the ECM organization of dormant and proliferative tumors, we decided to further explore the ECM organization around dormant solitary cells *in vivo*. Residual dormant cells reside in the surgical margins after removal of the primary tumor and give rise to local recurrences after a dormancy period<sup>30</sup>. Multiphoton SHG imaging of residual cancer cells after primary tumor surgery showed that collagen fibers are poorly aligned and highly curly around individual tumor cells (**Fig. 1C**), similar to the phenotype described in Fig. 1, A and B. To confirm whether these residual cancer cells with poorly aligned and highly curly collagen fibers were dormant, we used a CDK2 biosensor (DHB-mVenus) that allows identification of each phase of the cell cycle within a cell using fluorescent microscopy (**Fig. 1D**). The fluorescent mVenus sensor shuttles between the cytoplasm and the nucleus upon CDK2 activation. CDK2 phosphorylates the sensor and promotes its nuclear/cytoplasmic shuttling<sup>39</sup>. As a result, G0-G1 cells present an exclusive 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 and solitary disseminated tumor cells (DTCs) in the lungs showed that solitary cells have a strong accumulation of the sensor in the nucleus, confirming that these cells are indeed dormant and arrested in G0/G1 (**Fig. 1E**). On the other hand, clusters of cells show a mixed population of cancer cells at different phases of the cell cycle confirming their proliferative status.

Our SHG imaging analysis has demonstrated a change in the ECM between dormant nodules and solitary cells vs. proliferative primary tumors and metastatic clusters. The increase of collagen fiber alignment around growing cells at the primary site (**Fig. 1E, fig. S1 A**) and in lung metastasis (**fig. S1 B**) suggests that proliferating cancer cells that escape dormancy are able to remodel the ECM favoring aligned collagen fibers; however, solitary dormant DTCs reside in an ECM organized in curly and unaligned fibers (**Fig. 1E, fig. S1.A, B**). Overall, these results show that collagen fibers around dormant cells are characterized by a non-linear alignment and a wavy architecture, and are able to switch to a high degree of linear orientation during cancer cell reactivation.

### **Dormant cancer cell-derived ECM is enriched in type III collagen**

As dormant and proliferative tumor ECMs are architecturally different, we next sought to determine if their composition, or "matrisome" differed. To do so, we performed mass spectrometry on decellularized T-HEp3 growing tumors and D-HEp3 dormant nodules (**fig S2 A to C and Supplementary Table 1**) (D-HEp3 dormant cells form small nodules *in vivo* while T-HEp3 forms growing tumor masses *in vivo*<sup>29,36</sup>). This differential analysis revealed that the matrisome of dormant cancer cells 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. 2A**). This enrichment in collagens was confirmed with Masson's trichrome staining of histological sections of these tumors and validated in the D2 mammary cancer model (**Fig. 2B**). Consistently, immunofluorescence staining of tumor sections shows a higher abundance of collagens (type I, III, IV, V and VI) in D-HEp3 as compared to proliferative T-HEp3 (**Fig. 2C**). The human-in-mouse xenograft system used further allowed us to determine the relative contribution of tumor cells (human protein sequence) and stromal cells (murine protein sequence)<sup>14</sup> to the production of the matrisome (**fig S2D, Supplementary Table1**). Importantly, examination of the collagen composition at the protein level in the tumor-cell compartment reveals that type III collagen is the most abundant collagen produced by dormant tumor cells and the only collagen enriched in dormant cells when compared to proliferative cells (**Fig. 2D, Supplementary Table 1**). Type III collagen was also enriched in the stromal compartment of dormant D-HEp3 nodules among other collagens (**Fig. S2E and Supplementary Table 1**).

Interestingly, we perform transcriptomic analysis of D-HEp3 and T-HEp3 tumors and found that the matrisome signature<sup>40</sup> is significantly dysregulated ( $p=4.97e-05$ ), with *COL3A1* among the most upregulated collagen genes in dormant cells as compared to proliferative cells (**Fig. 2E and Supplementary Table 2**).

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 produce by dormant cancer cells.

### **Type III collagen induces tumor dormancy**

Previous studies have shown that stromal type III collagen restricts metastasis formation from mammary tumors<sup>41</sup>, but this phenotype was not explored in the context of dormancy regulation. Given the high abundance of type III collagen in dormant tumors, we asked whether increasing type III collagen concentration in the microenvironment of highly proliferative tumors could drive cancer cells into dormancy and prevent their growth.

We first performed a series of *in vivo* experiments using a chicken chorioallantoic membrane (CAM) assay, a well characterized *in vivo* model to study tumor dormancy<sup>7,30,42</sup> (**fig. S3A**). T-HEp3 cells were inoculated on CAMs with or without type I, III, or IV collagen. We observed that only tumor cells co-injected with type III collagen grew significantly smaller tumors (**Fig. 3A**). Similar results were obtained in mice injected with either T-HEp3 (**Fig. 3B**), 4T1 or D2A1 (**fig. S3B and C**) with or without type III collagen. Importantly, the inhibition of growth *in vivo* is not a consequence of increased apoptosis as shown by TUNEL assays *in vivo* (**Fig. 3C**) and Annexin V *in vitro* (**fig. S3D**) but rather the induction of dormancy as shown by the increase in nuclear p27 dormancy marker<sup>7,30</sup> (**Fig. 3D**) and decrease in cell

division as shown by live cell imaging of T-HEp3 cells expressing the CDK2 biosensor (**Fig. 3E and Movies 1, 2 and 3**). Concentrations of type III collagen equal or greater than 1mg/ml inhibits cell proliferation *in vitro* and *in vivo* (**Fig. 3A, B, E, F and G and Fig. S3B**) and cells grown on type III collagen do not show increased cytotoxicity, apoptosis or decreased viability (**fig., S3E,F**).

Next, we sought to determine whether treating locally residual cancer cells (observed post-resection of a primary tumor) with a type III-collagen-containing bioengineered scaffold could prevent the awakening of residual dormant cancer cells. To do so, we used T-HEp3 cells expressing the CDK2 biosensor. Only 20% of the mice treated with type III collagen containing scaffolds present recurrences after tumor surgery, while 80% of tumors relapsed in the control group (**Fig. 3G**). Intravital imaging at the time of recurrences show that mice that have received the type III collagen scaffold present fewer cancer cells that are primarily in a dormant state, since the CDK2 sensor is nuclear (G0/G1), whereas mice having received the control scaffolds were colonized by growing masses with cells at different phases of the cell cycle (**Fig. 3H**).

### **Tumor cell-derived type III collagen defines ECM architecture and sustains dormancy**

The increased abundance of tumor cell-derived type III collagen observed in the ECM of dormant tumors (**Fig. 2**) correlates with a higher expression of *COL3A1* at the mRNA level and protein level in several dormancy models (**Fig. 4A and B**), suggesting that *COL3A1* is part of a gene network operating during dormancy in cancer cells.

To determine the role of *COL3A1* in sustaining dormancy of cancer cells we knocked down *COL3A1* in dormant D-HEp3 cells using siRNA (**Fig. 4,C and D**) and performed *in vivo* growth experiments. Depletion of *COL3A1* interrupts dormancy and restores the growth of D-HEp3. Interestingly, depletion of other collagen chains (*COL1A1*, *COL1A2*, *COL5A1*, *COL5A2*, *COL5A3*, *COL6A1*, *COL6A2*, *COL6A3*, *COL16A1*, *COL18A1*) that are also overexpressed in dormant cells have no discernable effect on tumor dormancy (**Fig. S4, A to D**). Moreover, *COL3A1* depletion in D-HEp3 induces a significant decrease in the expression of the dormancy-inducing gene *NR2F1* (**Fig.4 E**).

The changes in type III collagen content upon *COL3A1* depletion in tumor cells are accompanied by a significant increase in collagen fiber alignment (**Fig. 4F**), suggesting that type III collagen contributes to the wavy ECM organization observed in dormant cells. Finally, type III collagen was found accumulated around solitary T-HEp3 cells that spontaneously disseminated to the lungs, and significantly reduced in lung micro-metastasis (**Fig.4G**), suggesting that deposition of type III collagen by dormant DTCs may be a key step towards establishing a dormant niche. These results unveil for the first time a role of *COL3A1* in inducing and sustaining dormancy.

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

Given the dependence of dormant cells on type III collagen to sustain their dormant state (**Fig. 4**) and the ability of type III collagen to reprogram cancer cells into dormancy (**Fig. 3**), we next sought to determine which collagen receptors will be involved in binding to type III collagen. Interestingly, 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 (**Fig. 5A**).

We performed RNA sequencing and transcriptomics analysis to identify collagen receptors that were dysregulated in dormant cells compared with proliferative cells (**Supplementary Table 2**). We found that integrins alpha 1, alpha 2, alpha 10, beta 1 and the DDR1 receptors were significantly dysregulated in dormant D-HEp3 vs T-HEp3 (**Fig. 5B**). We next performed an *in vivo* siRNA screen in D-HEp3 cells to determine which receptors were necessary for cancer cells to remain dormant. DDR1 knockdown has the most significant and robust effect on restoring the growth of dormant cells (**Fig. 5C to E and fig. S5A and B**). No changes in DDR2 expression were observed upon DDR1 knockdown (**Fig. 5D and fig. SB**). We examined shRNA control and shRNA DDR1 depleted D-HEp3 cells that spontaneously disseminated from dormant nodules and found that DDR1 depleted cells formed metastasis whereas control cells remain as solitary cells (**Fig. 5F**). Similar results were observed in an experimental metastasis assay (**fig. S5C**).

Cell cycle analysis shows a significant decrease in the number of cells in G0 upon DDR1 knockdown (**fig. S5D**), further confirming the role of DDR1 in cell quiescence. *In vivo* rescue experiments in CAM assays using a full-length DDR1 construct confirm that DDR1 is necessary to sustain dormancy (**Fig. 5G**).

We observed that dormant cells have high DDR1 protein level when compared with proliferative counterparts in HNSCC and mammary tumor models (**Fig. S5E**). Thus, we hypothesized that increasing DDR1 levels in proliferative DDR1<sup>low</sup> cells may induce dormancy. To test this hypothesis, we overexpressed full-length DDR1 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. S5, F and G**).

Importantly, the binding of dormant cells to type III collagen but not type I or type IV is DDR1-dependent as shown by adhesion assay (**Fig 5H**). Together, these results suggest that the binding of DDR1 to type III collagen is responsible for the entry to dormancy.

Given the previous roles assigned to DDR1 in collagen remodeling<sup>43,44</sup> and early work by Vogel showing that DDR1 regulates gene expression of several ECM genes<sup>45</sup> we hypothesized that DDR1-driven dormancy could be mediated through regulation of the matrisome of dormant cells. DDR1-depleted

tumors have significant increase in alignment of the collagen ECM (**Fig. 5I**). Interestingly, cells depleted for DDR1 show a decreased *COL3A1* mRNA expression (**Fig 5J**). Furthermore, proteomics and immunostaining analyses confirmed that the expression of the type III collagen is affected upon DDR1 depletion (**Fig. 5K and L and Supplementary Table 3**). Moreover, DDR1 knockdown has a significant effect on the expression of the matrisome gene sets (**Fig. 5M and Supplementary Table 4**).

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

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

As DDR1 regulate *COL3A1* gene expression, we analyzed the RNA sequencing data to identify transcription factors that could regulate *COL3A1* levels and dormancy 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 highly expressed in dormant D-HEp3 compared with T-HEp3 and also downregulated upon DDR1 knockdown. Following this method, we identified STAT1, FOSB, IRF7, TRM25, TLE4 transcription factors as potential regulators of dormancy downstream of DDR1 (**Fig. 6A**). We performed *COL3A1* promoter analysis using CiiDER<sup>46</sup> to identify predicted binding sites for the five identified transcription factors.

Predicted binding sites were identified for STAT1, FOSB, IRF7 (**Fig. 6A**). Among them, only STAT1 regulated *COL3A1* gene expression (**Fig. 6B**). Importantly, STAT1 depletion restored growth of dormant cells and decreased DDR1 abundance at the protein level (**Fig 6C**). Similar results were obtained *in vivo* using a STAT1 inhibitor, Fludarabine (**fig. S6 A**).

Knowing the well characterized role of Janus kinase (JAKs) in the phosphorylation of STATs<sup>47</sup>, we tested the role of JAK 1, 2 and 3 on tumor cell dormancy. Inhibition of JAK2 using a selective inhibitor was able to restore the proliferation of dormant cells *in vivo* (**Fig. 6D**). We validated these results using a specific siRNA for JAK2 (**Fig. S6B**)

Interestingly, upon STAT1 downregulation we observed a shift in ECM organization from wavy to straight, and increased alignment (**Fig. 6E**), similar to the one previously observed upon *COL3A1* and DDR1 knockdown (**Fig. 4 and 5**).

Co-immunoprecipitations experiments showed that DDR1 and STAT1 are in complex in dormant cells when exposed to type III collagen (**Fig. 6F**). Additionally, immunofluorescence analysis of D-HEp3 plated on glass or on type III collagen shows an increase in nuclear localization of phosphoSTAT1 on type III collagen (**Fig. 6G**).

Downregulation of either DDR1 receptor or COL3A1 induces a downregulation of STAT1 expression at the RNA level (**fig. S6C and D**). This suggests a feedback loop where DDR1 / type collagen III in dormant cells may regulate STAT1 expression in order to sustain the dormant state of the cell.

Based on our observations, we propose a model where binding of type III collagen to DDR1 on the cell surface triggers STAT1 activation, possibly through the recruitment of JAK2 and STAT1 to the COL-III/DDR1 cell membrane complex to regulate COL3A1 expression. Increased expression of COL3A1 remodels the ECM by increasing its curliness and drives the cells into a dormant state (**Graphical abstract**).

## DISCUSSION

The study presented here sheds light into three underdeveloped areas of tumor biology: 1) The characterization of the matrisome of dormant and reactivated tumors that reveals potential targets in the ECM that mediate dormancy and relapse, 2) the role of tumor-cell derived ECMs, such as type III collagen, and the collagen receptor DDR1 on dormancy and 3) the potential role of the ECM as therapeutics for metastasis by preventing reawakening of cancer cells. Overall, our work describes a new mechanism of cellular dormancy, by which the assembly by cancer cells of a pro-quiescent ECM enriched in type III collagen could potentially be exploited to prevent tumor recurrences and metastasis. The concept of dormancy niches has been previously proposed and several reports have pointed towards stromal ECM components as constituents of these niches and drivers of dormancy (TGF $\beta$ 2<sup>5</sup>, TSP-1<sup>24</sup>, osteopontin<sup>48</sup>). In this study, we explored the concept of dormancy niches further and incorporate the idea that tumor cells themselves actively contribute to the assembly of pro-dormant ECM niches.

The idea of tumor-cell-derived ECM as a driver of metastasis has been previously proposed to explain differences in metastasis organ seeding. Work by the Hynes lab showed that cancer cell-derived ECM can determine tissue specificity for specific metastatic sites<sup>18</sup> and emphasized the importance of tumor-cell-derived ECM molecules in determining tropism of metastatic DTCs; however, its role on dormancy has not been explored before. Our data show that dormant cancer cells express a specific matrisome and interfering with matrisome components (as shown for type III collagen) can restore or ablate the proliferative capacity of cancer cells. We demonstrated that the ECM produced by dormant cancer cells contribute to establish a pro-dormancy niche and stimulate mechanisms that activate and reinforce dormancy. Our results also confirm the plasticity of dormancy and demonstrate the importance of the ECM microenvironment as a major driver of the dormancy-to-awakening transition. The identification of tumor cell-derived type III collagen as one of the main drivers of tumor dormancy opens the door to

explore other components of the matrisome identified in our proteomic analysis that could contribute to dormancy in conjunction with type III collagen.

Our data also adds information about how dormant cancer cell ECMs are formed. Type III collagen is a homotrimeric helix formed by three identical COL3A1 chains<sup>49</sup>. Studies showed that type III collagen fibrils form hybrid fibers with COL1A1 and the abundance of type III collagen can produce more flexible fibers and disorganized ECM<sup>50,51</sup>, similar to our results in dormant cells. Type III collagen is the most abundant collagen produced by dormant cells but our proteomics data also revealed that the stroma-derived ECM of dormant tumors is enriched in type III collagen (**fig. S2E and Supplementary Table 1**) suggesting that the dormant cells may instruct the microenvironment to increase type III collagen secretion by stromal cells, revealing more complex mechanisms involving the stroma compartment to orchestrate dormancy.

Previous work by the Volk lab showed that a decrease of type III collagen in the stromal microenvironment increases the aggressiveness of breast tumors, whereas upregulated expression is associated with increased survival in breast cancer patients, suggesting that type III collagen can limit metastasis<sup>41</sup>. Now, our data could mechanistically explain how type III collagen may mediate the suppression of metastasis observed previously<sup>41</sup> and add a molecular description on how type III collagen induces tumor cells into dormancy and prevents reawakening through a DDR1/STAT1 pathway.

Interestingly our data show a new function for the collagen receptor DDR1 in sustaining dormancy. Previous reports showed that DDR1 is required for restoration of growth after a dormancy period<sup>28</sup>. Our data complement these observations and highlight the complex regulation and function of DDR1 during metastasis progression. We have characterized the signaling pathway mediating dormancy and ECM remodeling through COLIII/DDR1/STAT1 and identify DDR1 as the receptor for Collagen Type III that drive activation of dormancy. While the mechanisms that regulate COL3A1 and DDR1 expression are not well understood, recent reports identified hypoxia as a regulator of collagen type III expression<sup>52,53</sup>. Also analysis of the hypoxic transcriptome of head and neck cancer cells showed that DDR1 expression increased upon hypoxia (GSE87456 from Brady et al.,<sup>54</sup>). We can speculate that hypoxia, previously shown to prime cells to become dormant<sup>7</sup> triggers the expression of *COL3A1* and *DDR1* to promote the formation of the pro-quiescence ECM niche.

Analysis of upstream regulators of STAT1 identified JAK2 as a regulator of dormancy. We can hypothesize that JAK2 activation of STAT1 downstream of DDR1 drives the activation of dormancy. STAT1 regulation of DDR1 and COL3A1 expression further suggests the activation of a signaling loop that reinforces dormancy.

Overall, while several signals have been shown to contribute to tumor dormancy such as TGF $\beta$ 2<sup>5</sup> and the retinoic acid pathway<sup>30</sup>, we believe that COLIII/DDR1/JAK2/STAT1 pathway may act in parallel with those signals to reinforce tumor dormancy.

Based on our data, we conclude that self-assembly of a collagen III rich ECM by dormant cells is a key step to prevent growth via the induction of a quiescence program. Which signals may affect Type III collagen abundance and organization to awaken dormant cells? We can hypothesize that changes in MMPs expression may drive collagen remodeling and contribute to awakening. Along those lines, D-HEp3 depleted for DDR1 and T-HEp3 have increased expression of MMPs (such as MMP9, that mediates Type III collagen degradation<sup>55</sup>) when compared with D-HEp3 (**Supplemental Table 2,4**), that may contribute to degrade the type III collagen ECM. Indeed, collagen type III levels are lower in lung metastasis when compared with solitary cells in the lung (**Fig. 4G**), suggesting that MMP activation may degrade the Type III collagen ECM niche contributing to the restoration of growth.

Finally, the treatment of wounds with collagen scaffolds has shown promising results and has been proposed as a therapeutic alternative for complex skin wounds<sup>56</sup>. Our data show that by enriching the microenvironment for collagen Type III, dormancy is activated and tumor growth is interrupted. Our *in vivo* data show that this dormancy mechanism through Collagen Type III enrichment works in immunodeficient and immunocompetent mice, arguing for a true tumor cell-ECM stroma crosstalk to regulate dormancy. Our studies open the door to think about possible treatments that could use type III collagen scaffolds to prevent local relapse of solid tumors by providing signals that will reprogram cancer cells into a permanent dormant state; this may be particularly relevant for head and neck squamous cell carcinoma (HNSCC) and breast tumors, two of the models we explored. As the biology of tumor dormancy gets uncovered and new specific drugs are developed, a combination of dormancy-inducing treatments with therapies that specifically kill dormant cells will ultimately prevent local recurrence and metastasis and pave the way to a cancer cure.

## References

1. Steeg, P. S. Targeting metastasis. *Nature Reviews Cancer* (2016). doi:10.1038/nrc.2016.25
2. Sosa, M. S., Bragado, P. & Aguirre-Ghiso, J. A. Mechanisms of disseminated cancer cell dormancy: an awakening field. *Nat. Rev. Cancer* **14**, 611–622 (2014).
3. Linde, N., Fluegen, G. & Aguirre-Ghiso, J. A. *The Relationship Between Dormant Cancer Cells and Their Microenvironment. Advances in Cancer Research* **132**, (Elsevier Inc., 2016).
4. Giancotti, F. G. Review Mechanisms Governing Metastatic Dormancy and Reactivation. *Cell* **155**, 750–764 (2013).
5. Bragado, P. *et al.* TGF- $\beta$ 2 dictates disseminated tumour cell fate in target organs through TGF- $\beta$ -RIII and p38 $\alpha$ / $\beta$  signalling. *Nat. Cell Biol.* **15**, 1351–1361 (2013).
6. Liu, Y. *et al.* Blockade of IDO-kynurenine-AhR metabolic circuitry abrogates IFN- $\gamma$ -induced immunologic dormancy of tumor-repopulating cells. *Nat. Commun.* **8**, 1–15 (2017).
7. Fluegen, G. *et al.* Phenotypic heterogeneity of disseminated tumour cells is preset by primary tumour hypoxic microenvironments. (2017). doi:10.1038/ncb3465
8. Kai, F. B., Drain, A. P. & Weaver, V. M. The Extracellular Matrix Modulates the Metastatic Journey. *Dev. Cell* **49**, 332–346 (2019).
9. Bissell, M. J., Hall, H. G. & Parry, G. How does the extracellular matrix direct gene expression? *J. Theor. Biol.* (1982). doi:10.1016/0022-5193(82)90388-5
10. Bissell, M. J. & Aggeler, J. Dynamic reciprocity: how do extracellular matrix and hormones direct gene expression? *Prog. Clin. Biol. Res.* (1987).
11. Eliceiri, K. *et al.* Automated quantification of aligned collagen for human breast carcinoma prognosis. *J. Pathol. Inform.* (2014). doi:10.4103/2153-3539.139707
12. Provenzano, P. P. *et al.* Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Med.* **4**, 38 (2006).
13. Conklin, M. W. *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
14. 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).
15. Naba, A., Clauser, K. R., Lamar, J. M., Carr, S. A. & Hynes, R. O. Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters. *Elife* **2014**, 1–23 (2014).
16. Socovich, A. M. & 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

17. Taha, I. N. & Naba, A. Exploring the extracellular matrix in health and disease using proteomics. *Essays Biochem.* **63**, 417–432 (2019).
18. Hebert, J. D. *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
19. 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
20. 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
21. 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).
22. Barkan, D. *et al.* Metastatic growth from dormant cells induced by a Col-I-enriched fibrotic environment. *Cancer Res.* **70**, 5706–5716 (2010).
23. Aguirre-Ghiso, J. A., 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).
24. Ghajar, C. M. *et al.* The perivascular niche regulates breast tumour dormancy. *Nat. Cell Biol.* **15**, 807–17 (2013).
25. Albregues, J. *et al.* Neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in mice. **4227**, (2018).
26. 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
27. 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).
28. Gao, H. *et al.* Multi-organ Site Metastatic Reactivation Mediated by Non-canonical Discoidin Domain Receptor 1 Signaling. *Cell* **166**, 47–62 (2016).
29. Adam, A. P. *et al.* Computational identification of a p38SAPK-regulated transcription factor network required for tumor cell quiescence. *Cancer Res.* **69**, 5664–5672 (2009).
30. Sosa, M. S. *et al.* NR2F1 controls tumour cell dormancy via SOX9- and RAR $\beta$ -driven quiescence programmes. *Nat. Commun.* **6**, 6170 (2015).
31. Kim, R. S. *et al.* Dormancy Signatures and Metastasis in Estrogen Receptor Positive and Negative Breast Cancer. *PLoS One* **7**, e35569 (2012).

32. Aguirre-Ghiso, J. A., 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).
33. Malladi, S. *et al.* Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. *Cell* **165**, 45–60 (2016).
34. Aguirre-Ghiso, J. a, Bragado, P. & Sosa, M. S. Metastasis awakening: targeting dormant cancer. *Nat Med* **19**, 276–277 (2013).
35. Ghajar, C. M. Metastasis prevention by targeting the dormant niche. *Nature reviews. Cancer* **15**, 238–247 (2015).
36. Aguirre-Ghiso, J. A., Ossowski, L. & Rosenbaum, S. K. 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).
37. Aslakson, C. J. & Miller, F. R. 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).
38. Gao, H. *et al.* The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. *Cell* **150**, 764–779 (2012).
39. Spencer, S. L. *et al.* XThe proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. *Cell* **155**, 369–383 (2013).
40. Naba, A. *et al.* The extracellular matrix: Tools and insights for the ‘omics’ era. *Matrix Biol.* **49**, 10–24 (2016).
41. Brisson, B. K. *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).
42. Ossowski, L. Plasminogen activator dependent pathways in the dissemination of human tumor cells in the chick embryo. *Cell* **52**, 321–328 (1988).
43. Coelho, N. M. *et al.* Discoidin Domain Receptor 1 Mediates Myosin-Dependent Collagen Contraction. *Cell Rep.* (2017). doi:10.1016/j.celrep.2017.01.061
44. Coelho, N. M. & McCulloch, C. A. Mechanical Signaling Through the Discoidin Domain Receptor 1 Plays a Central Role in Tissue Fibrosis. *Cell Adh. Migr.* **6918**, 00–00 (2018).
45. Faraci, E., Eck, M., Gerstmayer, B., Bosio, A. & Vogel, W. F. An extracellular matrix-specific microarray allowed the identification of target genes downstream of discoidin domain receptors. *Matrix Biol.* **22**, 373–381 (2003).
46. Gearing, L. J. *et al.* CiiiDER: A tool for predicting and analysing transcription factor binding sites. *PLoS One* **14**, e0215495 (2019).
47. Rawlings, J. S., Rosler, K. M. & Harrison, D. A. The JAK/STAT signaling pathway. *J. Cell Sci.*

- 117**, 1281–1283 (2004).
48. Boyerinas, B. *et al.* Adhesion to osteopontin in the bone marrow niche regulates lymphoblastic leukemia cell dormancy. *Blood* **121**, 4821–4831 (2013).
  49. Mouw, J. K., Ou, G. & Weaver, V. M. Extracellular matrix assembly: a multiscale deconstruction. *Nat. Rev. Mol. Cell Biol.* **15**, 771–785 (2014).
  50. Parkin, J. Des *et al.* The collagen III fibril has a ‘flexi-rod’ structure of flexible sequences interspersed with rigid bioactive domains including two with hemostatic roles. *PLoS One* **12**, e0175582 (2017).
  51. Fleischmajer, R., Douglas MacDonald, E., Perlish, J. S., Burgeson, R. E. & Fisher, L. W. Dermal collagen fibrils are hybrids of type I and type III collagen molecules. *J. Struct. Biol.* (1990). doi:10.1016/1047-8477(90)90110-X
  52. Hu, Y., Hu, D., Yu, H., Xu, W. & Fu, R. Hypoxia-inducible factor 1 $\alpha$  and ROCK1 regulate proliferation and collagen synthesis in hepatic stellate cells under hypoxia. *Mol. Med. Rep.* **18**, 3997–4003 (2018).
  53. Wang, X. *et al.* Hypoxic mast cells accelerate the proliferation, collagen accumulation and phenotypic alteration of human lung fibroblasts. *Int. J. Mol. Med.* **45**, 175–185 (2020).
  54. Brady, L. K. *et al.* Transcriptome analysis of hypoxic cancer cells uncovers intron retention in EIF2B5 as a mechanism to inhibit translation. *PLoS Biol.* **15**, e2002623 (2017).
  55. Veidal, S. S. *et al.* Matrix metalloproteinase-9-mediated type III collagen degradation as a novel serological biochemical marker for liver fibrogenesis. *Liver Int. Off. J. Int. Assoc. Study Liver* **30**, 1293–1304 (2010).
  56. Gould, L. J. Topical Collagen-Based Biomaterials for Chronic Wounds: Rationale and Clinical Application. *Adv. Wound Care* **5**, 19–31 (2016).
  57. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
  58. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
  59. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
  60. Naba, A., Clauser, K. R. & Hynes, R. O. Enrichment of Extracellular Matrix Proteins from Tissues and Digestion into Peptides for Mass Spectrometry Analysis. *J. Vis. Exp.* e53057 (2015). doi:10.3791/53057
  61. Di Martino, J. *et al.* 2D and 3D Matrices to Study Linear Invadosome Formation and Activity. *J. Vis. Exp.* (2017). doi:10.3791/54911

62. Franco-Barraza, J., Beacham, D. A., Amatangelo, M. D. & Cukierman, E. Preparation of Extracellular Matrices Produced by Cultured and Primary Fibroblasts. *Curr. Protoc. cell Biol.* **71**, 10.9.1-10.9.34 (2016).

### **Acknowledgments:**

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 Djedani 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 Microscopy 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 Bravo-Cordero, Aguirre-Ghiso, and Sosa laboratories for helpful discussions. 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.

### **Funding:**

This work was supported by a Susan G. Komen Career Catalyst Research (CCR18547848 to JJBC), an NCI Career Transition Award (K22CA196750 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).

**Author contributions:**

JDM designed and performed experiments, analyzed and interpreted the data and contributed to writing the manuscript. ARN performed cell sorting experiments. EF performed mouse tumor surgery experiments. CM performed lung metastasis experiments in MDA-MB-231 xenografts. 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. JAG 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 writing the manuscript. All authors contributed to editing of the manuscript.

**Competing interests:** Authors declare no competing interests.

## Figure legends:

**Graphical abstract.** ECM architecture dynamically changes between dormant and proliferative cells. 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 collagen ECM architecture around dormant cells. Binding of type III collagen to DDR1 activates JAK2/STAT1 signaling to activate dormancy and increases COL3A1 expression, establishing a pro-dormant ECM niche.

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

**(A)** Upper left panel: Representative SHG images from T-HEp3 primary tumor and D-HEp3 nodules grown in mice. Lower left panel: Representative Collagen OrientationJ output obtained from SHG images. More colors = more alignment. Color Scale bar represent fiber orientation. Scale bar 100 $\mu$ m. Upper right: normalized distribution of collagen fiber orientation. Lower right: Mode of fiber orientation (n= 5 tumors per condition)

**(B)** Left panel: OrientationJ visual output obtained from SHG images of D2.A1, D2.OR, 4T1 and 4T07 tumors grew in CAMs. Color Scale bar represent fiber orientation Scale bars, 100 $\mu$ m. Right: normalized distribution of collagen fiber orientation (n=4 tumors).

**(C)** Left panel: Representative two-photon images of T-HEp3-GFP primary tumors (SHG in gray, cells in green). Right panel: Remaining cancer cells around the primary tumor area after surgery. d refers to days. Scale bars, 100 $\mu$ m. Red square represents a magnified area. Scale bars, 50 $\mu$ m.

**(D)** 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 each phase of the cell cycle *in vivo* in T-HEp3 cells in nude mice. Scale bars, 5 $\mu$ m.

**(E)** Representative two-photon imaging of T-HEp3-DHB-mVenus grown in nude mice for two weeks and resected by surgery. T-HEp3 cells in green, H2B-RFP in red and SHG in gray. Scale bars, 50 $\mu$ m. Right: Percentage of cells in each phase of the cell cycle was assigned based on CDK2 sensor fluorescence intensity and location (nucleus versus cytoplasm) at surgery site and in lungs after

spontaneous dissemination from the primary site (n=5 mice). Bottom: A scheme showing the changes in ECM architecture as tumor cells grow is shown.

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

**(A)** Characterization of the matrisome of D-HEp3 and T-HEp3 tumors related to Supplementary Table 1C.

**(B)** Masson's trichrome staining of proliferative and dormant tumors grown in nude mice. Collagen fibers are stained in blue, tumor cells in red. Scale bars, 50 $\mu$ m.

**(C)** Immunofluorescence of D-HEp3 nodules and T-HEp3 tumors grown in mice for different collagens (red). Cancer cells were stained with vimentin (green), and DAPI (blue). Scale bars, 50 $\mu$ m.

**(D)** Relative abundance of tumor-cell-derived collagens identified in D-HEp3 and T-HEp3. Graph related to Supplementary Table 1H and I.

**(E)** Left: Enrichment plot for matrisome signature from RNA sequencing performed in D-HEp3 nodules and T-HEp3 tumors grown in nude mice. X-axis shows log<sub>2</sub>FC for T-HEp3 vs D-HEp3. Black bars represent matrisome genes. Right: List of significantly dysregulated collagen genes in D-HEp3 nodules and T-HEp3 tumors grown in nude mice. Data represented as log<sub>2</sub> fold change D-HEp3/T-HEp3 (related to Supplementary Table 2).

## **Fig. 3. Type III Collagen enriched microenvironments induces dormancy**

**(A)** Quantification of tumor growth *in vivo* in CAM assays. Each dot represent the number of tumor cells per tumor (one tumor per CAM). T-HEp3 were resuspended in DPBS or collagen I, III or IV before inoculation.

**(B)** Nude mice injected with T-HEp3 in DPBS (ctrl) in the right flank or in a 1mg/ml type III collagen III solution in the left flank. Representative picture of both tumors at day 19 after inoculation. Tumor growth was monitored for 19 days ( n=8 mice) and tumor weight was calculated after sacrifice (n=6 mice).

**(C)** TUNEL assay on tumors from (B). Scale bar: 100 pixels. Graph represents the quantification of the number of TUNEL positive cells per pixel<sup>2</sup> (n=4 mice).

**(D)** Left panel: Immunofluorescence of tumors from experiment in (B). vimentin (green), p27 (red) and DAPI (blue) is shown. Right: Quantification of at least 3 sections per tumors (n= 4 mice) of the number of p27(+) cells per field. Scale bar 10 $\mu$ m.

**(E)** Left panel: Time points from a 18 hr time lapse movie of T-HEp3 plated on Collagen I or III. (t = hours). CDK2-mVenus sensor (green) and 546-succinimidylester labelled collagen is shown (red). Scale bar 10 $\mu$ m. Right: Quantification of the number of mitosis per fields across 18 hrs (n= 3 independent movies with more than 10 cells per field). (Related to Movies 1, 2 and 3).

**(F)** MTT assay of T-HEp3 plated on different concentrations of type III collagen for 24 hrs.

**(G)** Scheme showing the experimental designed for treatment with Type III collagen bioengineer scaffolds. Representative SHG images of collagen type III and empty sponges are shown. Tumor growth was monitored before surgery and local tumor relapse was monitored post-surgery for 21 days. Tumor volume over time is shown.

**(H)** Representative images of the resected area after tumor surgery at the time mice were sacrifice. SHG signal from multiphoton imaging in gray and T-HEp3 CDK2 biosensor expressing cells in green. Yellow arrow point towards a G0 T-HEp3 cell (nuclear localization of CDK2 sensor). Scale bar 50 $\mu$ m.

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

**(A)** RT-qPCR for COL3A1 in different dormancy models.  $2^{-(dCT)}$  were calculated from technical duplicates from biological quadruplicates *in vivo* (Hep3 model) and triplicates *in vitro* (D2 model).

**(B)** Western blot for type III collagen in different dormancy models *in vivo* (CAMs).

**(C)** Quantification of tumors in CAM assays. Dots represent the number of tumor cells per tumor. D-HEp3 were either expressing a control siRNA, a single siRNA targeting COL3A1 or a pool of 4 siRNAs targeting COL3A1.

**(D)** Western blot for type III collagen and tubulin proteins extracted from CAM tumors. Representative image of CAM tumors 6 days after inoculation.

**(E)** RT-qPCR for NR2F1 (dormancy marker) in D-HEp3 siCTRL versus siCOL3A1.  $2^{-(dCT)}$  were calculated from technical duplicates from biological triplicates.

**(F)** Left panel: Representative SHG images and OrientationJ software's visual output obtained for each condition. Scale bars, 100 $\mu$ m. More colors = more alignment. Color Scale bar represent fiber orientation.

Upper right graph: normalized distribution of collagen fiber orientation. Lower right graph: Mode of fiber orientation.

**(G)** Top: Representative image of type III collagen staining of solitary DTCs and micrometastasis of spontaneously disseminated T-HEp3 cells in lungs of nude mice. DTCs were stained with vimentin (green), type III collagen (red) and Ki67 (grey) and DAPI (blue). Scale bars, 10 $\mu$ m. Bottom: quantification of type III collagen staining intensity per field between solitary cells and cluster of cells.

### **Fig. 5. DDR1 is required to sustain dormancy**

**(A)** Adhesion assay of D-HEp3 and T-HEp3 on type I, III and IV collagens.

**(B)** List of significantly changed ECM receptors between D-HEp3 and T-HEp3 cells. Data extracted from RNA sequencing of tumors and expressed as log<sub>2</sub> fold change D-HEp3/ T-HEp3. (\*) represent receptors known to interact with collagens. Related to Supplementary Table 2.

**(C)** Functional siRNA screen of collagen receptors for tumor growth in CAMs. D-HEp3 were transfected with either a control siRNA or siRNA targeting DDR1, ITGA1, ITGA2, ITGA10 and ITGB1. Graph represents the number of cells per CAM tumors where each dot is a tumor.

**(D)** Western blot for DDR1 and DDR2, on D-HEp3 expressing an siRNA control or an siRNA targeting the DDR1 receptor, extracted from CAMs.

**(E)** Nude mice were injected with T-HEp3, D-HEp3 shCTRL or D-HEp3 shDDR1. Tumor growth is monitored over time and tumor volume calculated. n= 5 tumors per condition. Western blot assessing DDR1 and DDR2 levels is shown on the right. DDR1 mRNA levels verified by RT-qPCR.

**(F)** Representative image of nude mice lungs 4 weeks after tumor removal surgery. SHG is displayed in gray and D-HEp3 cancer cells expressing GFP in green. Scale bar 50 $\mu$ m. Upper right table: percentage of mice presenting solitary cells or micrometastasis in the lung, 4 weeks after tumor cell orthotopic injections. Lower right table: percentage of mice presenting solitary cells or micrometastasis in the lungs, 4 weeks after tumor surgery (surgery performed 2 weeks after tumor cells orthotopic injection).

**(G)** D-HEp3 expressing a non-targeting sgRNA (NT) or an sgRNA against DDR1 were either transfected with an empty vector (EV) or DDR1b full length (+DDR1b) and incubated 6 days on CAM. Number of cells per tumor is quantified and represented on the graph on the left. Each dot represents a tumor. Western blot assessing the level of DDR1 expression is shown.

**(H)** Adhesion assay to collagen I, III or IV of D-HEp3 expressing a control sgRNA (NT) or an sgRNA against DDR1.

**(I)** Top panel shows representative SHG images of ECM of D-HEp3 cell inoculated in nude mice injected with either an shRNA control or targeting DDR1. SHG signal displayed in gray. Scale bar 100 $\mu$ m. Graph represents the mode of fiber orientation of each analyzed sample.

**(J)** RT-qPCR for COL3A1 from RNA extracted shRNA control or shRNA DDR1 tumors grown in nude mice.

**(K)** Relative collagen abundance of tumor-derived collagens identified in D-Hep3 shCTRL and D-HEp3 shDDR1 tumors. Related to Supplementary Table 3.

**(L)** Immunofluorescence on formalin-fixed paraffin embedded primary shRNA control or shRNA DDR1 tumors. Type III collagen is stained in red, Vimentin in green and DAPI in blue. Scale bar 50 $\mu$ m.

**(M)** Enrichment plot for matrisome signature from RNA sequencing performed in D-HEp3 shCTRL and D-HEp3 shDDR1 tumors grew in nude mice ( $p=7.68e-10$ ). X-axis shows log<sub>2</sub>FC for D-HEP3 shRNA DDR1 vs D-HEP3 shRNA Control. Black bars represent matrisome genes.

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

**(A)** Top: Transcription factors significantly dysregulated between T-HEp3 and D-HEp3 and significantly downregulated in D-HEp3 shDDR1 compared with D-HEp3 shCTRL. Values represent average read counts. Color scale range from dark blue (lowest value) to dark red (highest value) per line. Bottom: Map of predicted sites for transcription factors identified above using the CiiiDER tool.

**(B)** RT-qPCR for COL3A1 from RNA extracted from D-HEp3 cells *in vitro* expressing a control siRNA or an siRNA targeting STAT1, FOSB or IRF7.

**(C)** D-HEp3 expressing a control siRNA or a STAT1 siRNA were inoculated on CAMs. Number of cells per tumor is quantified and represented on the graph on the left. Each dot represents a tumor. Western blot assessing the level of STAT1 depletion, DDR1 and tubulin expression levels is shown on the right.

**(D)** D-HEp3 treated with selective JAK1, 2 or 3 inhibitors were inoculated on CAM. Number of cells per tumor is quantified and represented on the graph on the left. Each dot represents a tumor.

**(E)** Representative SHG images of control and siRNA STAT1 tumors. SHG signal is display in gray. Left panel: Representative images (n=5 tumors). Scale bar 100 $\mu$ m. Right: Mode of fiber orientation of each analyzed sample.

**(F)** Immunoprecipitation of endogenous DDR1 and co-immunoprecipitation of STAT1 in D-HEp3 incubated with type III collagen. Western blot for DDR1 and STAT1 is shown.

**(G)** Immunofluorescence of phospho-STAT1 in D-HEp3 plated either on glass or on Type III collagen. Scale bar 10 $\mu$ m. Quantification of the number of cells with nuclear pSTAT1 is shown.

**Supplementary Table 1. D-HEp3 and T-HEp3 proteomic data .** A) Samples. B) Complete MS output. C) Complete Matrisome. D) Normalization and enrichment. E) T-HEp3 matrisome. F) D-HEp3 matrisome. G) T-HEp3 vs D-HEp3 comparison. H) All collagens. I) Tumor-cell derived collagens.

**Supplementary Table 2. D-HEp3 and T-HEp3 RNA sequencing data**

**Supplementary Table 3. D-HEp3 shCTRL vs D-HEp3 shDDR1 proteomic data.** A) Samples. B) Complete MS output. C) Complete Matrisome. D) Normalization and enrichment. E) All collagens. F) Collagens\_norm to all.

**Supplementary Table 4. D-HEp3 shCTRL and shDDR1 RNA sequencing data**

**Supplementary Table 5. D-HEp3 and T-HEp3 transcription factors analysis**

**Supplementary Table 6. D-HEp3 shCTRL and shDDR1 transcription factors analysis**

**Supplementary Table 7. Primers, siRNA and shRNA sequences and antibodies lists**

**Movie S1.** D-HEp3 plated on type III collagen for 9hrs. Images were acquired every 30 minutes. Cells express a CDK2 sensor (green) and collagen was labeled in red. Scale bar 10 $\mu$ m.

**Movie S2.** T-HEp3 plated on type I collagen for 18hrs. Images were acquired every 30 minutes. Cells express a CDK2 sensor (green) and collagen was labeled in red. Scale bar 10 $\mu$ m.

**Movie S3.** T-HEp3 plated on type III collagen for 18hrs. Images were acquired every 30 minutes. Cells express a CDK2 sensor (green) and collagen was labeled in red. Scale bar 10 $\mu$ m.

**Fig. S1.**

**(A)** Images showing the imaging window design and implantation site in mice. Imaging windows were implanted on nude mice prior to tumor cell injection (n=5). T-HEp3-GFP were imaged with a multiphoton microscope 4 days after injection. Scale bar: 100 $\mu$ m. Zoom Scale bar: 50 $\mu$ m.

**(B)** Left panel: Multiphoton imaging of the lungs from nude mice with or without T-HEp3 GFP (orthotopic injection in the interscapular region of the neck) for 2 weeks. Scale bar 50 $\mu$ m. Right panel: Multiphoton imaging of NCG mice lungs. Mice were injected with MDA-MB-231 GFP in the mammary gland. SHG is represented in gray, cells in green. Scale bar 100 $\mu$ m.

**Fig. S2.**

**(A)** Schematic representation of the ECM enrichment pipeline.

**(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)** Pie charts represent, for each sample, the proportion of the mass-spectrometric signal intensity from matrisome (blue) and non-matrisome (grey) peptides, related to Supplementary Table 1B.

**(D)** Collagen species identified in D-HEp3 and T-HEp3 tumors in tumor and stroma compartment, related to Supplementary Table 1H and I.

**(E)** Relative collagen abundance of stroma derived collagen identified in D-HEp3 and T-HEp3, related to Supplementary Table 1H.

### **Fig. S3.**

**(A)** Cartoon representing the developmental stages of the chicken embryo, tumor inoculation and collection time points. Cells were inoculated on CAMs and grown for 6 days. Left graph: T-HEp3 and D-HEp3. Right graph: D2.A1 and D2.OR. Number of cells per tumor is quantified. Each dot represents a tumor.

**(B)** 8 week old nude mice were injected with 4T1 cells in DPBS (ctr) in their right mammary gland or in a 1mg/ml type III collagen solution in their left mammary gland. Tumor growth was monitored for 20 days.

**(C)** 8 week old nude mice were injected with D2.A1 cells in DPBS (ctr) in their right mammary gland or in a 1mg/ml type III collagen solution in their left mammary gland. Tumor growth was monitored for 20 days.

**(D)** T-HEp3 cells plated on plastic, type I collagen or type III collagen matrix were analyzed by FACS for apoptosis and necrosis using Annexin V and 7-AAD. Percentage of live cells is represented in green, apoptotic cells in yellow, necrotic cells in orange and dead cells in red.

**(E)** APOTOX assay of T-HEp3 plated on different concentrations of type III collagen for 24 hrs.

**(F)** APOTOX assay of T-HEp3 plated on different concentrations of type III collagen for 48 hrs.

### **Fig. S4.**

**(A)** D-HEp3 expressing a control siRNA or siRNA targeting COL1A1 or COL1A2 were inoculated on CAMs and grown for 6 days. Number of cells per tumor is quantified and represented on the graph. Each dot represents a tumor.

**(B)** D-HEp3 expressing a control siRNA or siRNA targeting COL5A1, COL5A2 or COL5A3 were inoculated on CAM and grown for 6 days. Number of cells per tumor is quantified and represented on the graph. Each dot represents a tumor.

**(C)** D-HEp3 expressing a control siRNA or siRNA targeting COL6A1, COL6A2 or COL6A3 were inoculated on CAM and grown for 6 days. Number of cells per tumor is quantified and represented on the graph. Each dot represents a tumor.

**(D)** D-HEp3 expressing a control siRNA or siRNA targeting COL16A1 or COL18A1 were inoculated on CAM and grown for 6 days. Number of cells per tumor is quantified and represented on the graph. Each dot represents a tumor.

**Fig. S5.**

**(A)** BM-HEp3<sup>5</sup> (dormant) were transfected with either a control siRNA or siRNA targeting DDR1 and grown in CAM for 6 days. Plot represents the number of cells per CAM tumors where each dot is a tumor. Western blot showing DDR1 and tubulin levels upon DDR1 depletion are shown.

**(B)** D2.OR stably expressing either a control shRNA or 3 different shRNAs targeting DDR1 were grown in the CAM for 6 days. Plot represents the number of cells per CAM tumors where each dot is a tumor. Western blot showing DDR1, DDR2 and tubulin levels upon DDR1 depletion are shown. Note that shRNA 1 and 2 only deplete DDR1.

**(C)** Nude mice were injected intracardially with D-HEp3 shCTRL or D-HEp3 shDDR1 expressing GFP. Lungs were analyzed 3 weeks after cell injection by using multiphoton imaging. Quantification of solitary or micro-metastasis per imaged field was performed (n=2 control mice and n=3 DDR1 knockdown mice).

**(D)** Percentage of G0 cells from D-HEp3 cells transfected with an siRNA ctrl or siRNA targeting DDR1. G0 cells were determined by FACS using Pyronin Y and Hoechst 33342 staining.

**(E)** Western blot for DDR1 and tubulin in T-HEp3, D-HEp3, BM-HEp3, D2.0R and D2.A1 cell lines

**(F)** T-HEp3 expressing a control empty plasmid or DDR1b full length were inoculated on the CAM and grown for 6 days. Number of cells per tumor is quantified and represented on the graph on the left. Each dot represents a tumor. Western blot assessing the level of DDR1, DDR1 and tubulin is shown on the right.

**(G)** T-HEp3 expressing a control empty plasmid (EV) or DDR1b full length were analyzed by FACS for apoptosis and necrosis using Annexin V and 7AAD. Parental T-HEp3 and T-HEp3 treated with transfection agent (jetPRIME) were used as a control of the transfection effect on cell death. Percentage of live cells is represented in green, apoptotic cells in yellow, necrotic cells in orange and dead cells in red.

**Fig. S6.**

**(A)** D-HEp3 treated with DMSO or STAT1 inhibitor (Fludarabine) were inoculated on CAM and grown for 6 days. Number of cells per tumor is quantified and represented on the graph on the left. Each dot represents a tumor. Western blot assessing the level of STAT1 phosphorylation is shown on the right.

**(B)** D-HEp3 expressing a control siRNA or siRNA targeting JAK1 or JAK2 were inoculated on CAM and grown for 6 days. Number of cells per tumor is quantified and represented on the plot. Each dot represents a tumor.

**(C)** RT-qPCR for *STAT1* from RNA extracted from D-HEp3 shCTRL or shDDR1 tumors *in vivo*.

**(D)** RT-qPCR for *STAT1*, *FOSB*, *IRF7* from RNA extracted from D-HEp3 cells *in vitro* expressing a control siRNA or targeting *COL3A1*.

## **MATERIALS AND METHODS**

All animal procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of the Icahn School of Medicine at Mount Sinai.

### **Cell culture**

All cell lines except the HEp3 cellular models were purchased from ATCC and Karmanos and IMPACT tests were performed for all cells injected in mice to check for mycoplasma and rodent pathogens. T-HEp3 (proliferative cells), D-HEp3 (dormant cells) and dormant BM-HEp3<sup>5</sup> (bone marrow clone) cells were kindly provided by Dr. Julio Aguirre-Ghiso and maintained in DMEM, 4.5g D-Glucose, 110mg/l sodium pyruvate, 10% PEAK serum heat inactivated (PS-FB1), 1% Penicillin/Streptomycin. 4T07 (dormant) were purchased from Karmanos Cancer Institute and maintained in DMEM, 4.5g/l D-Glucose Glutamax without sodium pyruvate, 10% Gemini Foundation serum (900-108), 1mM mixed nonessential amino acids, 1% Penicillin Streptomycin. 4T1 cells (proliferative) were purchased from ATCC and cultivated in RPMI Glutamax, 10% Gemini Foundation serum, 1% penicillin Streptomycin antibiotics. D2.A1 (proliferative) and D2.OR (dormant) were purchased from Karmanos Cancer Institute and cultivated in DMEM, 4.5g/l D-Glucose Glutamax without sodium pyruvate, 10% Gemini Foundation serum (900-108), 1mM mixed nonessential amino acids, 1% penicillin Streptomycin. All cells were maintained at 37°C and 5%CO<sub>2</sub>. MDA-MB-231 were purchased from ATCC and maintained in DMEM 4.5g/l D-Glucose Glutamax without sodium pyruvate, 10% Gemini Foundation serum (900-108), 1% Penicillin Streptomycin.

HEK293T cells were maintained in DMEM 4.5g/l D-Glucose Glutamax without sodium pyruvate, 10% Gemini Foundation serum (900-108), 1% penicillin Streptomycin and used to produce lentiviral particles.

### **Plasmids, transient transfections and cell line generation**

FG12-GFP (Addgene #14884) was kindly provided by Dr. Soengas, (CNIO, Spain) and used to make stable cell lines expressing GFP. pHIV-H2B-mRFP was purchased from Addgene (#18982) and used to make stable cell lines expressing H2B-mRFP. CSII-EF-DHB-mVenus (CDK2 sensor) was generously shared by Dr. Spencer, University of Colorado, USA and used to assess cell cycle phase of disseminated cells *in vivo*. Non targeting-CRISPR (NT) was provided by Dr. Louis Hodgson. DDR1-CRISPR was generated using a gRNA designed with GPP sgRNA designer from the Broad Institute cloned into the lentiCRISPR V2 purchased from Addgene (#52961) with BsmBI following the Zhang

laboratory GeCKO protocol and verified by sequencing using a U6 primer from Genewiz (5'-GACTATCATATGCTTACCGT-3').

Full length DDR1b-myc (pcDNA 3.1/Zeo DDR1b) was generously provided by Dr. Leitinger, Imperial College of London, UK. pcDNA 3.1 myc His A (-) was purchased from Thermofisher and used as empty vector control.

To perform transient transfection, 200,000 cells were seeded in 6 well plates a day prior transfection. On transfection day, 2ml of fresh appropriate complete media was added to the cells. A transfection mix was prepared with 200  $\mu$ l of jetPRIME® buffer, 2 $\mu$ g of plasmid of interest and 4 $\mu$ l of jetPRIME® reagent (Polyplus transfection ref 114-15) and added to the cells for 24 hrs. Cells were then seeded on collagen matrices or inoculated in chicken embryo CAMs.

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

### **siRNA and shRNA expression and CRISPR cell lines**

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

To perform the siRNA screen in CAMs, 30pmol of siRNA (purchased from Qiagen and IDT) were mixed with 6 $\mu$ l of lipofectamine RNAimax in 200 $\mu$ l 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 purchased from the AECOM plasmid bank and 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 5 $\mu$ g/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 1 $\mu$ g 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  $\mu$ l 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 passed the following three steps before library construction: (1) Nanodrop: for RNA purity (OD260/OD280). (2) Agarose gel electrophoresis: for RNA integrity and potential contamination. (3) Agilent 2100: check RNA integrity again.

### Library construction

Briefly, mRNA from eukaryotic organisms is purified from total RNA using poly-T oligo-attached magnetic beads. The mRNA is first fragmented randomly by addition of fragmentation buffer.

NEB library: Then first strand cDNA is synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H-). Second strand cDNA synthesis is subsequently performed using DNA Polymerase I and RNase H. Double-stranded cDNA is purified using AMPure XP beads. Remaining overhangs of the purified double-stranded cDNA are converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with hairpin loop structure is ligated to prepare for hybridization (Adaptor: P5/P7 is PCR primers and those primers are complementary to sequences on flow cell; Rd1/Rd2 SP are read1/read2 sequencing primers; Index is used for identifying different libraries). In order to select cDNA fragments of preferentially 150~200 bp in length, the library fragments are purified with AMPure XP system (Beckman Coulter, Beverly, USA). Finally, the final library is obtained by PCR amplification and purification of PCR products by AMPure XP beads.

After library construction, the library is diluted to 1.5ng/ml from the preliminary quantitative results by Qubit2.0 and detecting the insert size by Agilent 2100. qPCR is used to accurately quantify the library

effective concentration (> 2nM), in order to ensure the library quality. Libraries are fed into Illumina machines after pooling according to activity and expected data volume if library quality is up to standard.

Raw image data file from high-throughput sequencing (Illumina) was transformed to Sequenced Reads (called Raw Data or Raw Reads) by CASAVA base recognition (Base Calling). Raw Data is stored in FASTQ (fq) format files, which contain reads sequence and corresponding base quality. Differential expression analysis is performed on gene counts with DESeq<sup>57</sup> version 1.14.1 and enrichment statistics using the LIMMA<sup>58</sup> Wilcoxon gene set test on differential statistics for the NABA\_Matrisome pathways in MSigDB<sup>59</sup> version 6.0.

### **Mass spectrometry and proteomic analysis**

***Samples for ECM proteomic analysis:*** Samples used for the ECM proteomic study were as followed:

**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 <sup>14,60</sup> 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 2, 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 <sup>14,40</sup>). 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 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  $\mu$ m C18) and a self-pack 5  $\mu$ m tip analytical column (12 cm of 5  $\mu$ m 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 was performed using Matrisome Annotator (Naba et al., Matrix Biology, 2016<sup>40</sup>, 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  $\times$  75  $\mu$ m 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  $\geq 2$  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  $\geq 25$  and an isolation interference  $\leq 30$  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<sup>40</sup> 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, treated with a graded series of alcohol, 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.

### **Masson trichrome**

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, ref # 4112).

### **TUNEL assay**

To assess apoptosis in tissue sections we used a TUNEL assay. 5 μm 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, ref # 4112).

### **Immunoprecipitation**

Samples were lysed at room temperature with 50 mM HEPES (pH 7.4), 150 mM NaCl, 1% Brij98 (Sigma), 1.5 mM MgCl<sub>2</sub>, phosphatase and proteinase inhibitor cocktail (Halt™ Protease and Phosphatase Inhibitor reference 1861280) and supernatant collected after centrifugation for 15 minutes at 13,000g. Lysate containing 5 mg protein was incubated with 50 μl of Dynabeads protein A pre-bonded to 5 μg of control rabbit-IgG (Cell Signaling 2729s) or anti-DDR1 for 1.5 hrs at room temperature. For antibody binding protocol, follow reference protocol (ThermoFisher 10006D). Precipitation was performed with a Dynabead magnetic rack. The Dynabead-Ab-Ag complex was washed 2 times with washing buffer (supplied from the kit ThermoFisher 10006D) and resuspended in 20 μl of elution buffer and 10 μl of Biorad 4X Laemmli and incubated 10 minutes at 90°C. Fifty percent of total precipitates were loaded to western blotting in a 10% acrylamide gel and transfert for 1:30 hrs. at 120V on nitrocellulose membranes. For detecting STAT1, anti-STAT1 was used.

## **Western blot**

Cells were lysed 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). 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 MTT Cell Viability Assay Kit from Biotium reference #30006, following the supplier recommended protocol. Absorbance signal was measured with a plate reader SpectraMax M5e on a spectrophotometer 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 400<sub>Ex</sub>/505<sub>Em</sub>. Cytotoxicity was measured using 485<sub>Ex</sub>/520<sub>Em</sub>. 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 10 $\mu$ g/ml Hoechst 33342. The samples were incubated for 45 minutes in a water bath set at 37°C, and then incubated for an extra 15 minutes with Pyronin Y (Sigma 213519, 0.1 $\mu$ g/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 followed. 500,000 cells per condition were collected and washed twice in 100 $\mu$ l 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, 400 $\mu$ l 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/mg, and type IV collagen (nippi ASC 4-104-01) coated at 0.05mg/ml. 100 $\mu$ l 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 100 $\mu$ l 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 100 $\mu$ l 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 100 $\mu$ l 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. 50  $\mu$ l of cell suspension was inoculated per eggs using an insulin syringe. 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 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 of this article.

### **Inhibitor treatments**

Fludarabine (Stat 1 inhibitor) was purchased from Selleckchem (reference S1491), diluted in DMSO at 50mM and used for 24hrs at a final concentration of 20 $\mu$ M.

Filgotinib GLPG0634 (JAK 1 inhibitor) was purchased from Selleckchem (reference HY-18300), diluted in DMSO at 10 $\mu$ M and used for 24hrs at a final concentration of 10nM.

Fedratinib (JAK 2 inhibitor) was purchased from Selleckchem (reference HY-10409), diluted in DMSO at 10 $\mu$ M and used for 24hrs at a final concentration of 3nM.

Tofacitinib (JAK 3 inhibitor) was purchased from Selleckchem (reference HY-40354), diluted in DMSO at 10 $\mu$ M and used for 24hrs at a final concentration of 1nM.

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

## Mouse experiments

Human cell lines were injected in Nude female Athymic NCr-nu/nu from Charles Rivers (HEp3 model) or NCG (NOD-Prkdcem<sup>26Cd52</sup>Il2rgem<sup>26Cd22</sup>/NjuCrI) 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 lubricant Lubrifesh Major reference NDC 0904-6488-38 was applied on mice eye to prevent them from drying.

## Tumor growth studies

50,000 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 groups. 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  $V = \frac{4}{3}\pi r^3$  and plotted as single dot with SEM. Mice were sacrificed with CO<sub>2</sub> 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

750,000 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 ref # NDC 10019-360-40, provided by the Mount Sinai animal facility) and the tumor was cleaned with ethanol. Tumor was lifted from its center and an incision was made using small scissors. 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<sub>2</sub> and the area of resection was reopened to image residual cancer cells and lungs were collected and imaged to find spontaneous disseminated cells using multiphoton microscope. Tumors and lungs were fixed and paraffin-embedded for tissue staining.

## Imaging window intravital experiments

Ultem (polyetherimide) made windows were designed and made 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 (ref # 031477). Cancer cells were then injected under the glass in 50 $\mu$ l of DPBS using an insulin syringe. Cells were imaged four days after injection to visualize ECM organization at early stages of tumor formation.

### Intracardiac injection

Mice were anesthetized with isoflurane and intracardiac injections of 100,000 cells per animal were performed using echo guidance with a microultrasound machine FUJIFILM VisualSonics Inc. (FF-VSI) Transducer: MS-250, Frequency: 21MHz, Imaging Data Analysis Software: Vevo LAB 3.1.1 at Small Animal Imaging Core at BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai.

### Collagen Type III bioengineer scaffolds

500,000 cells (T-HEp3 CDK2 sensor and H2B-RFP) were resuspended in 100 $\mu$ l 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 CO<sub>2</sub> 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 were 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 50 $\mu$ l 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  $V = \frac{4}{3}\pi r^3$  and plotted as single dot with SEM. Mice were sacrificed with CO<sub>2</sub> 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 used 880nm for SHG and GFP imaging. Sequential 880nm and 1000nm used to acquire SHG, CDK2 sensor and H2B-RFP signals. Z-stacks were acquired with 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 (Di Martino et al.,<sup>61</sup> JOVE 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 CO<sub>2</sub>. Imaging was performed with Zeiss LSM880 and a 20x objective, taking one image of a 10- $\mu$ m z-stack, every 30 minutes for 18hrs. Number of mitosis per field were counted with the support of CDK2 sensor fluorescence.

### Image analysis

Images 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<sup>2</sup>).

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 vimentin was used to identify cancer cells.

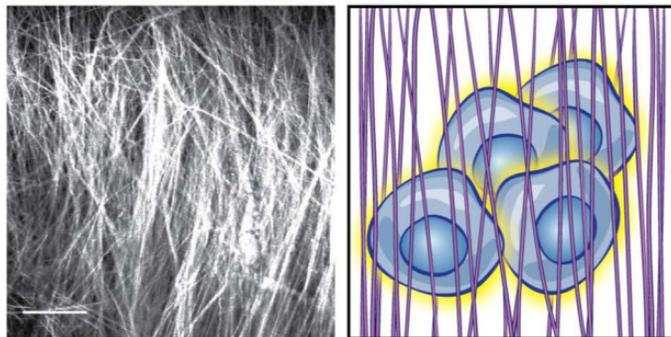
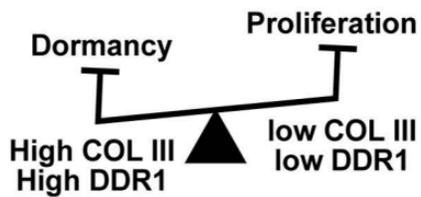
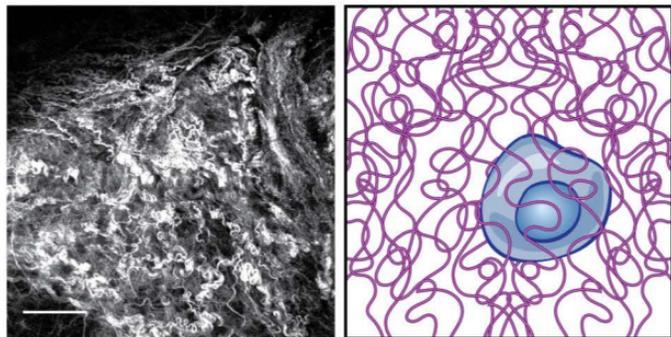
Collagen III intensity quantification: total intensity of the collagen channel was measured with ImageJ from images of solitary cells and clusters of T-HEp3 cells spontaneously disseminated in the lung of nude mice (n= 2 animals).

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  $\sigma$  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<sup>62</sup>.

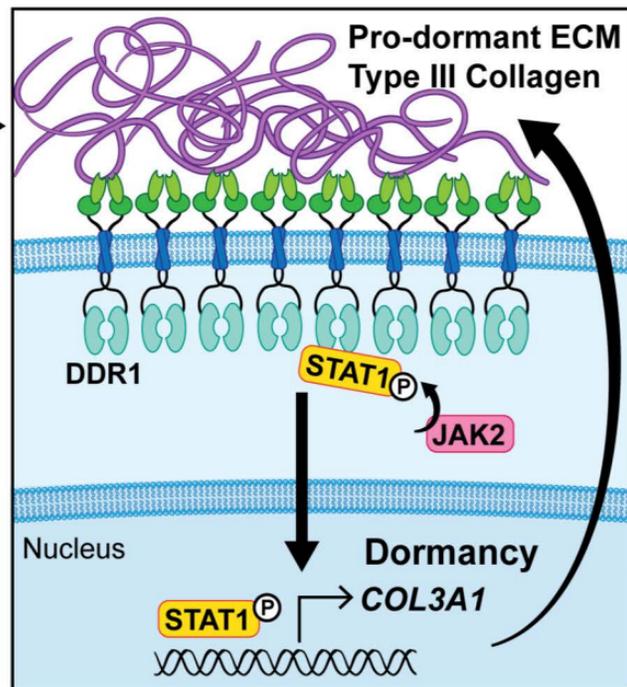
### **Statistical tests**

All experiments were performed at least 3 independent times. RT-qPCR data were reported as the mean  $\pm$  SEM of at least three experiments. All other data were reported as the median with interquartile range. Statistical significance ( $P < 0.05$  or less) was determined using a Mann Whitney test or ANOVA as appropriate and performed using GraphPad Prism version 5.00 for Windows, GraphPad Software, San Diego California USA, [www.graphpad.com](http://www.graphpad.com).

# Dormant



# Proliferative



*J Gregory ©2020 Mount Sinai Health System*

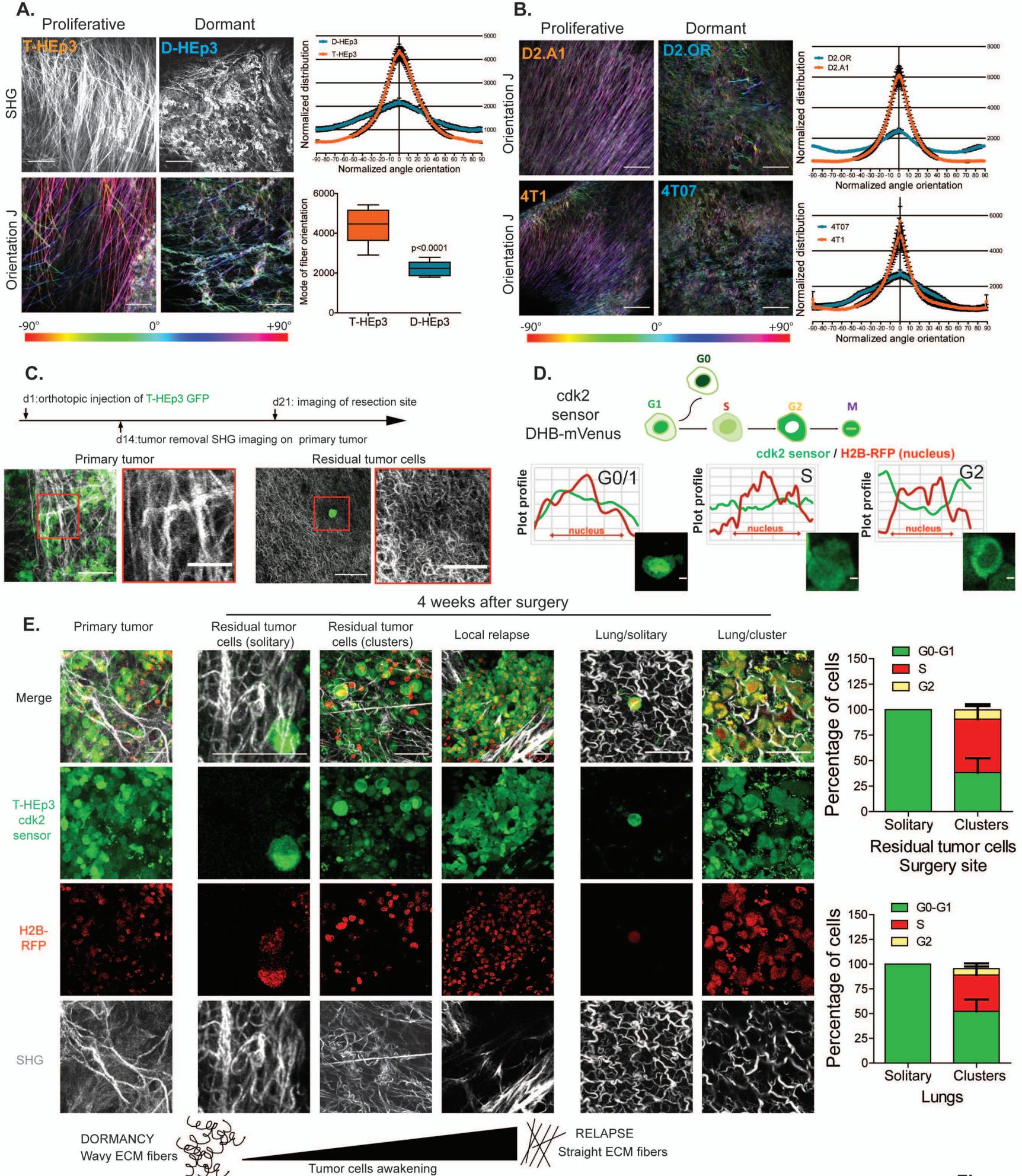
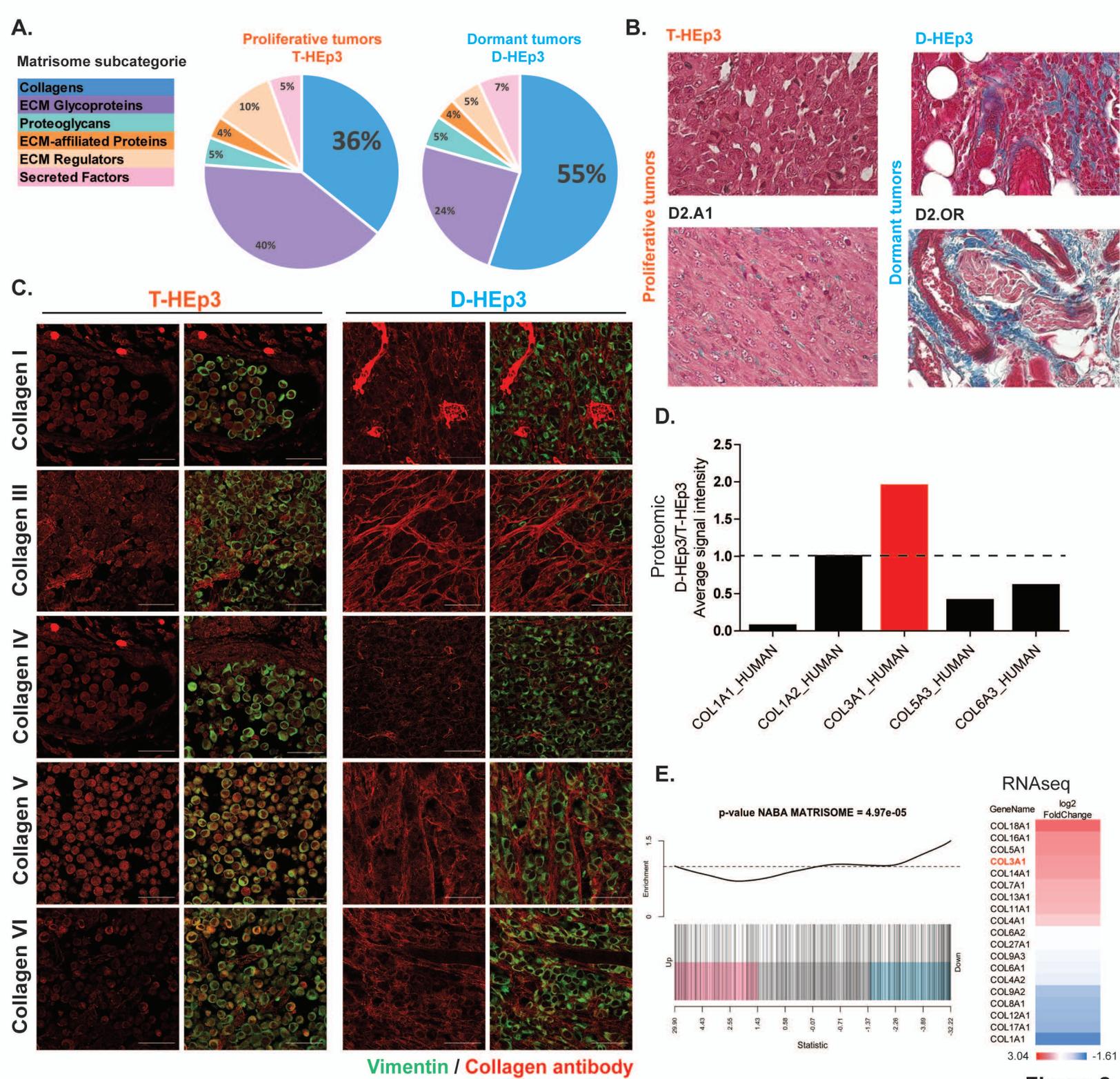


Figure 1.



**Figure 2.**

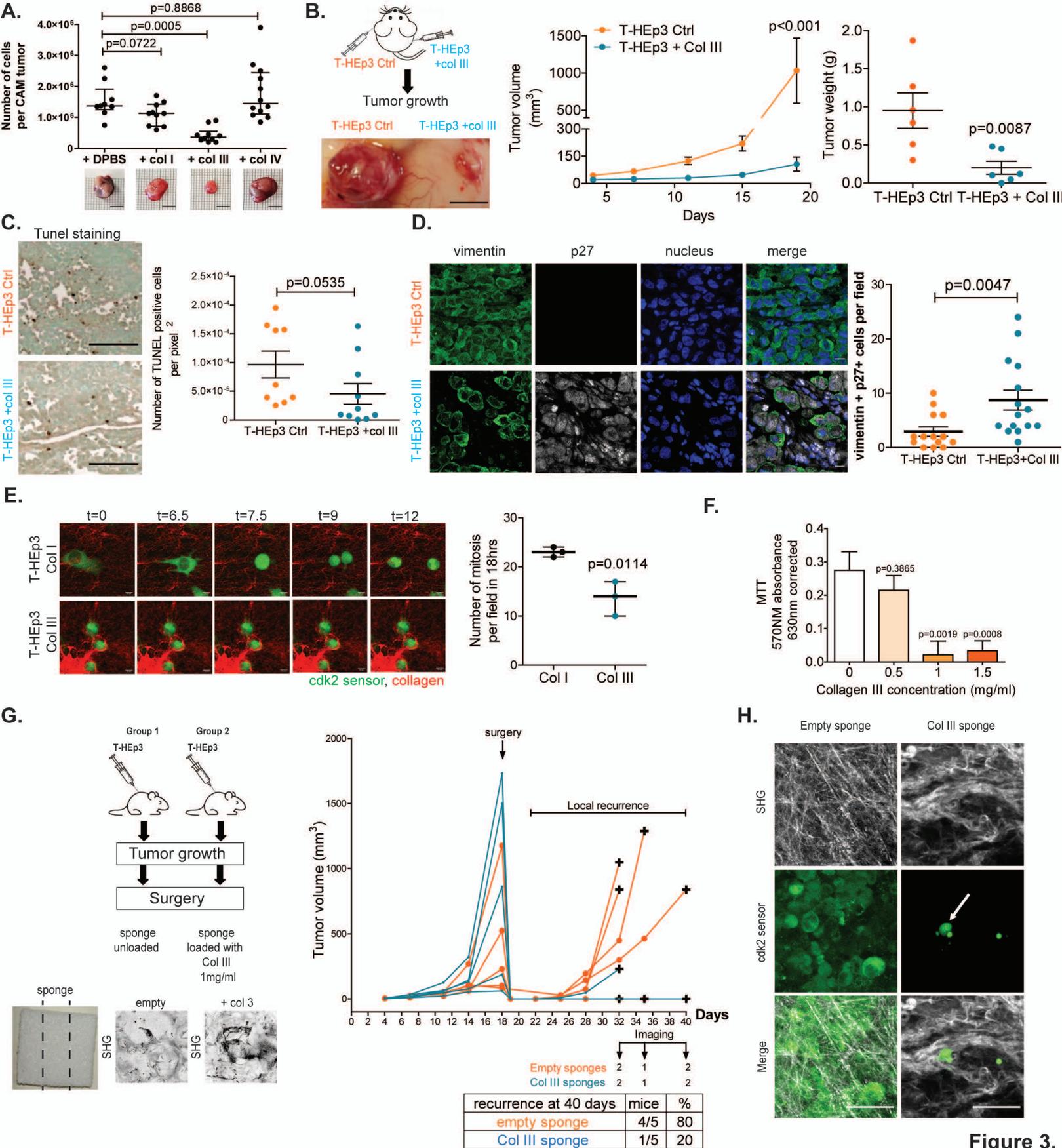


Figure 3.

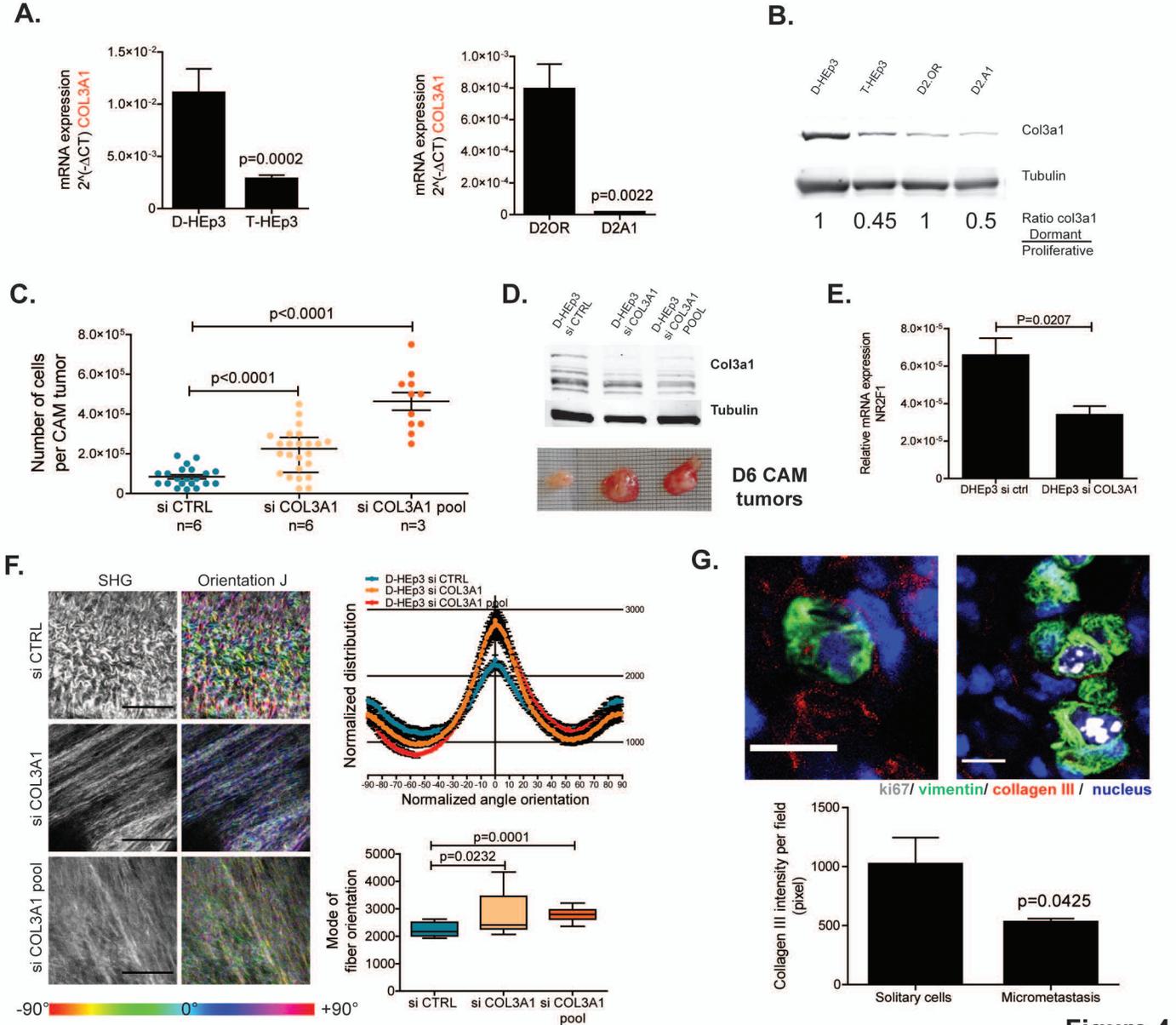
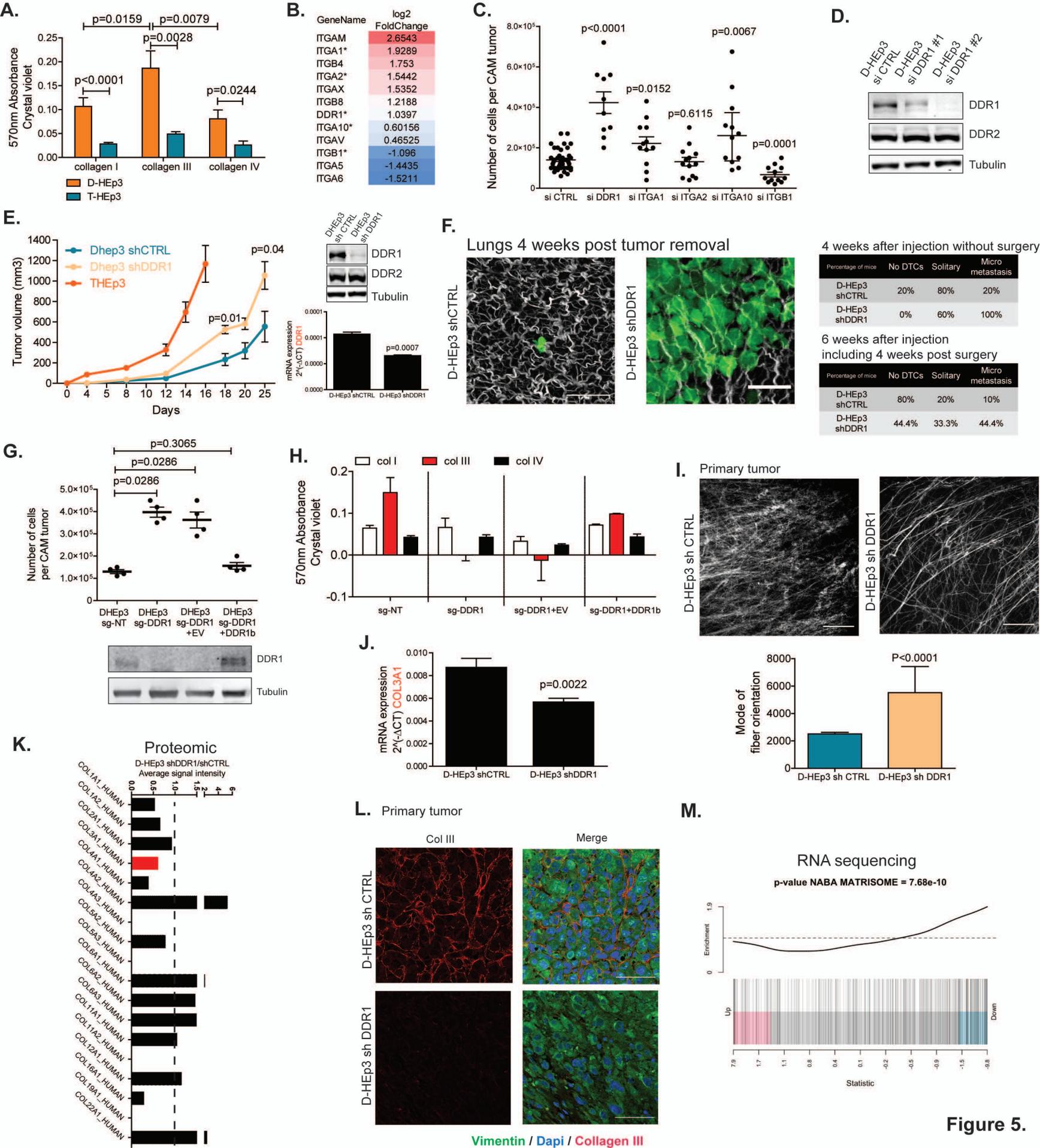
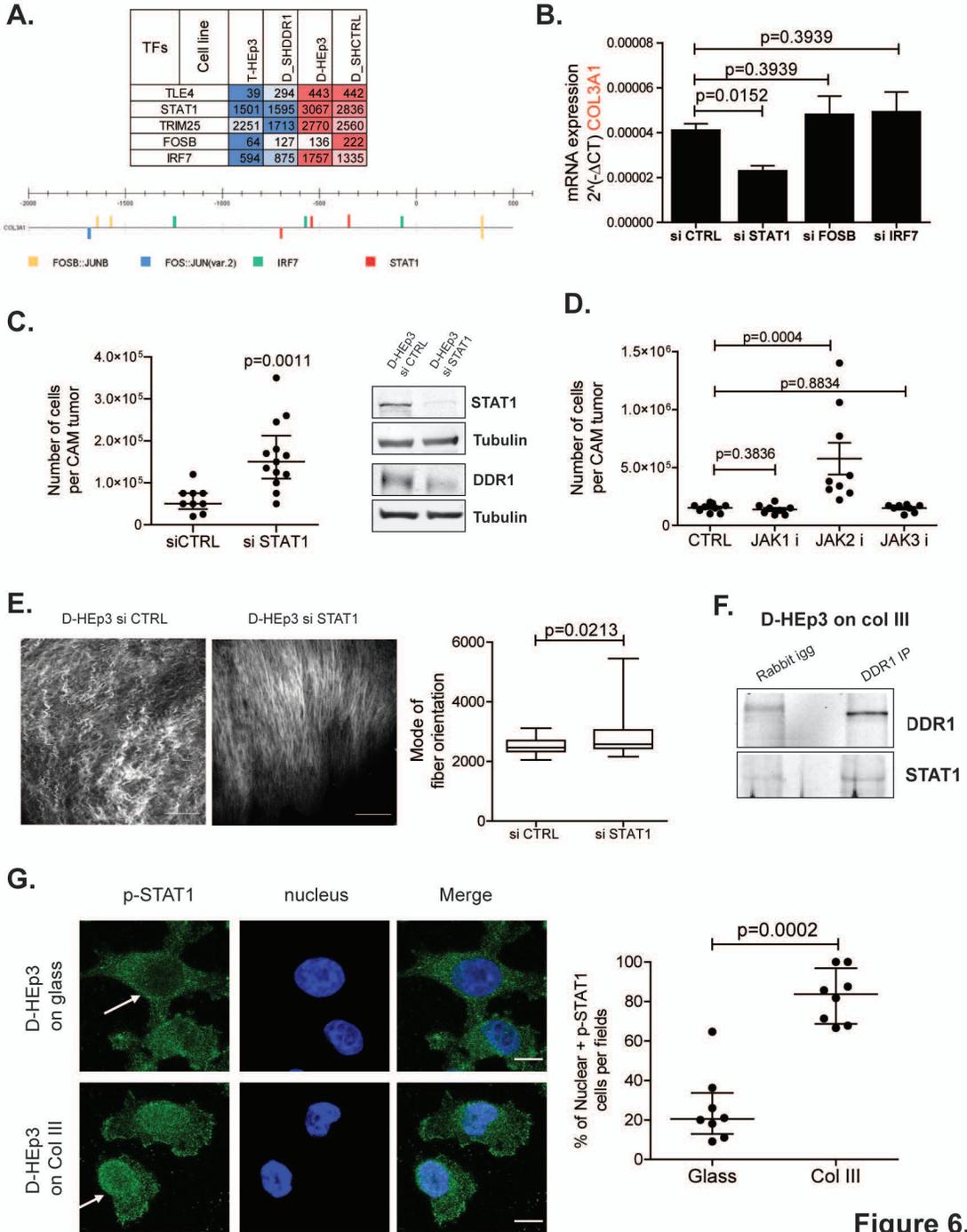
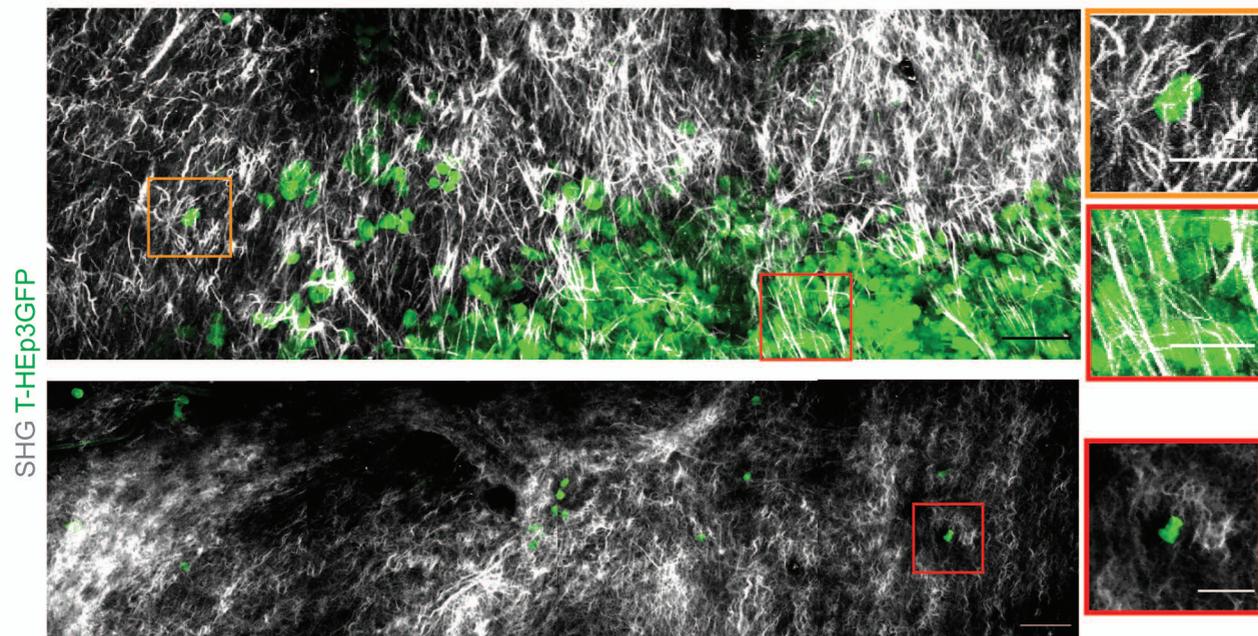
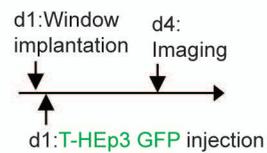
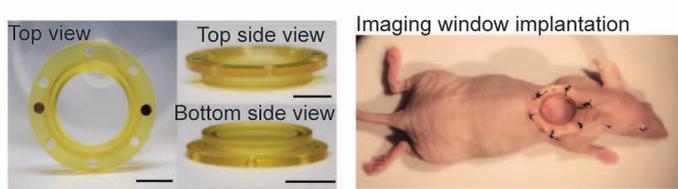
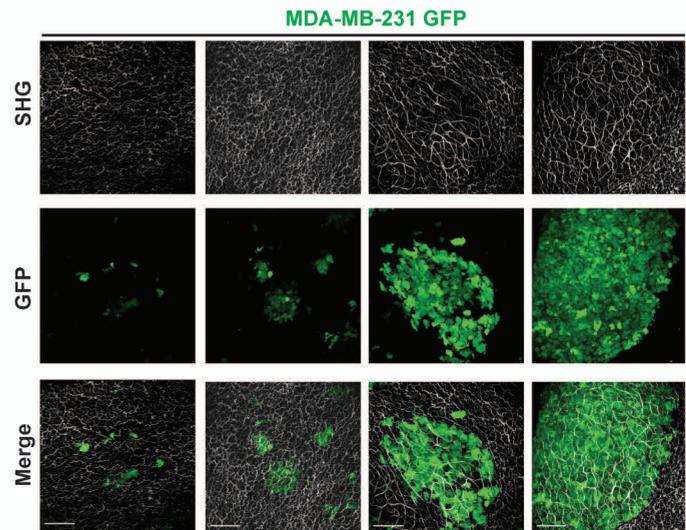
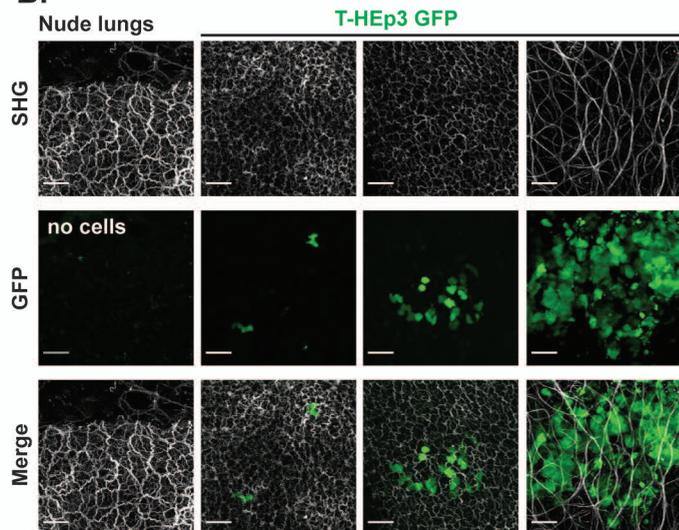


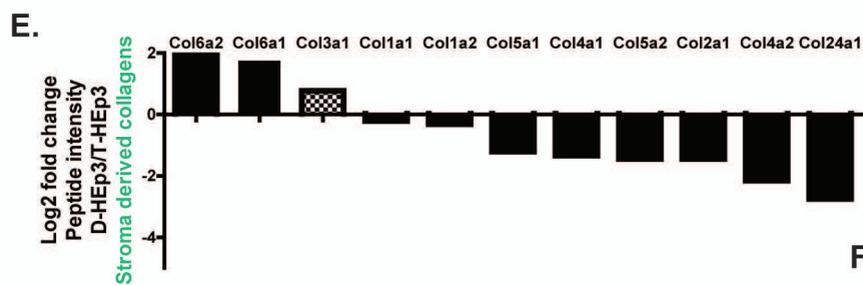
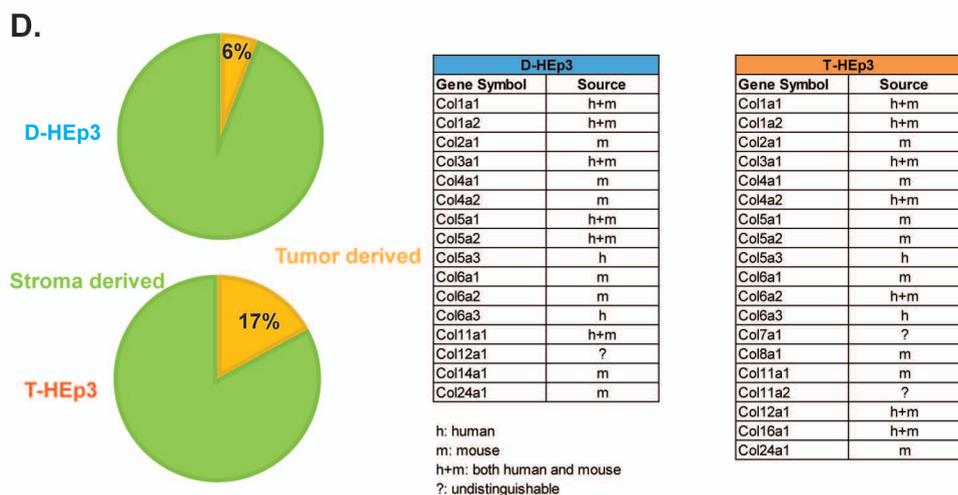
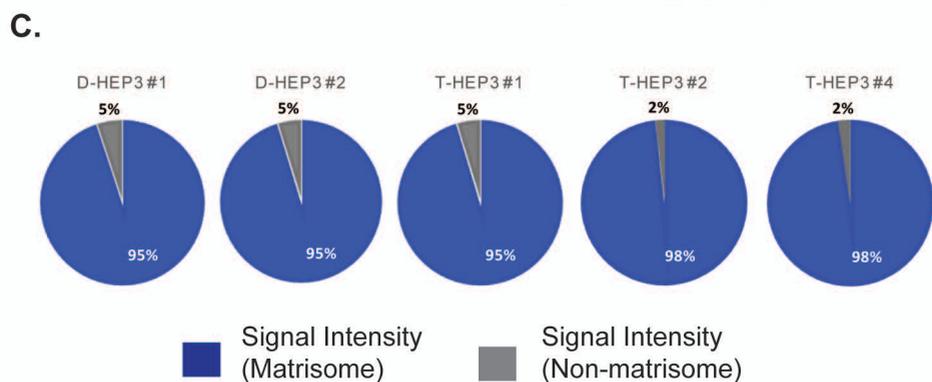
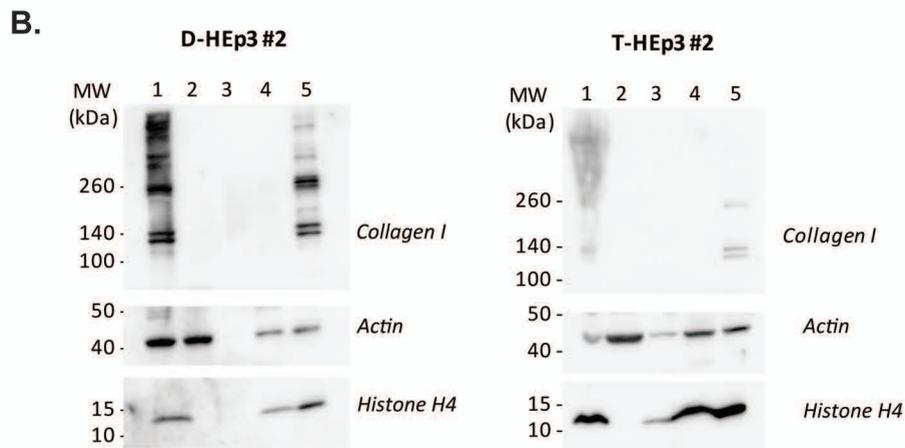
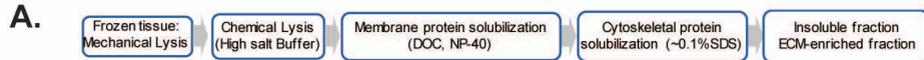
Figure 4.



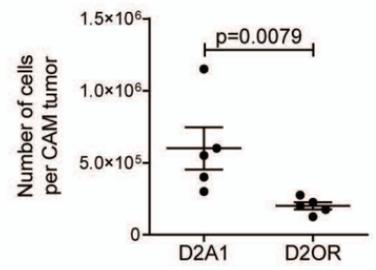
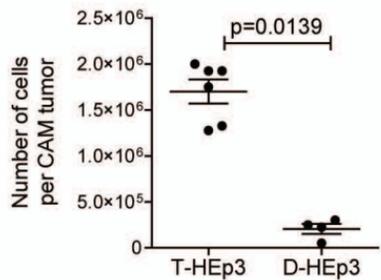
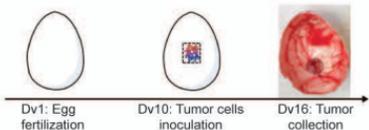
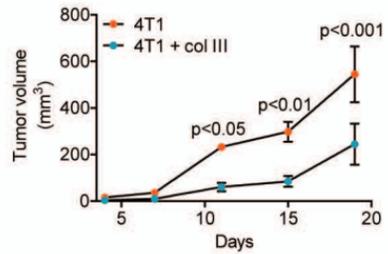
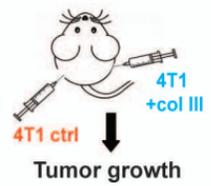
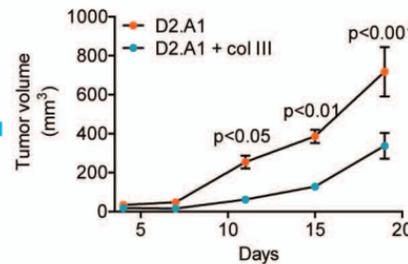
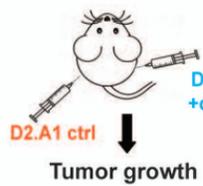
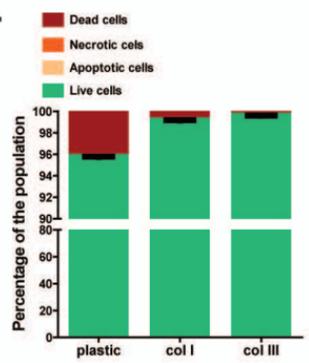
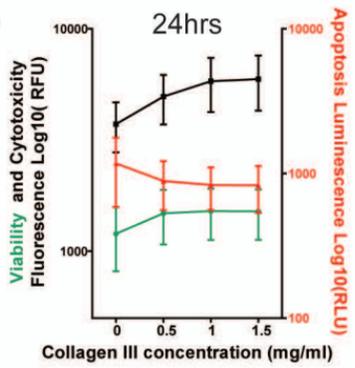
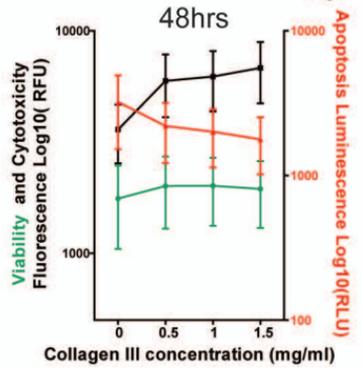


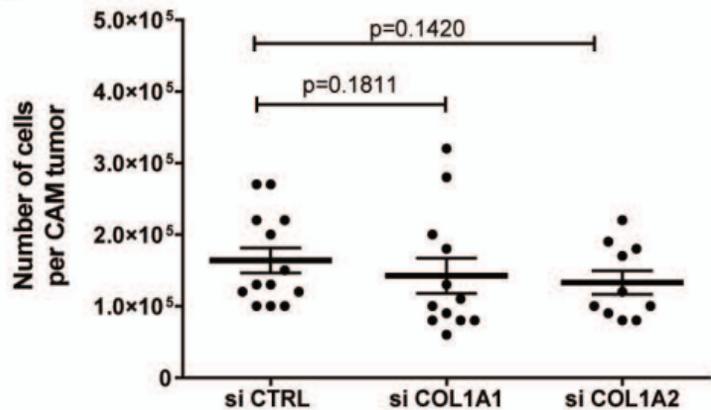
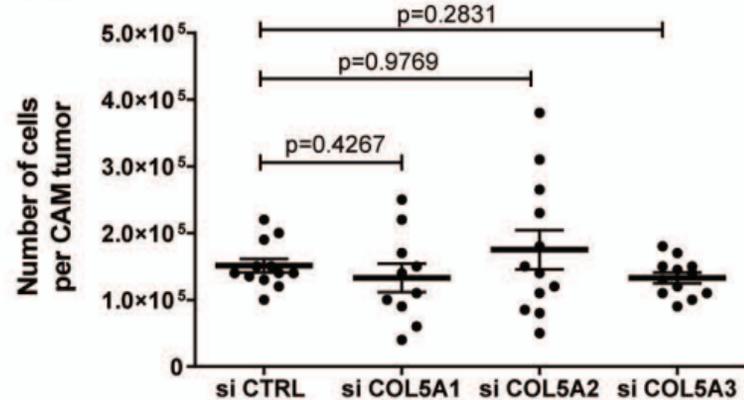
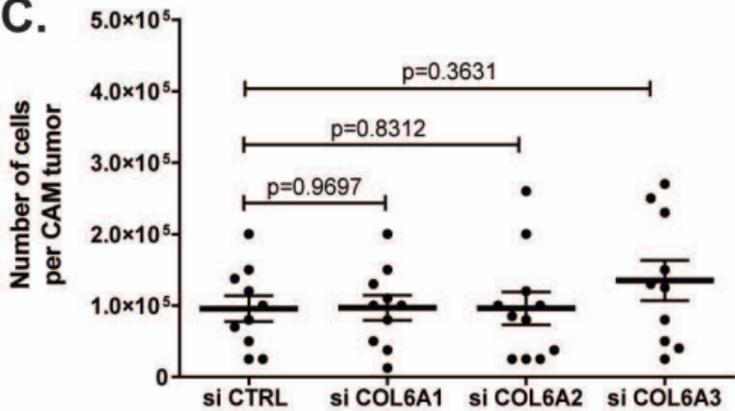
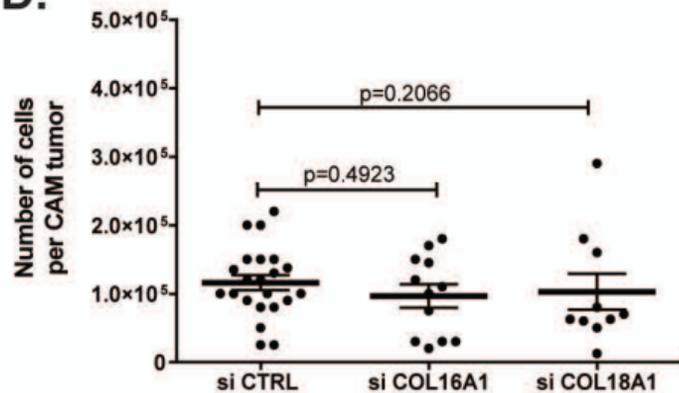
**Figure 6.**

**A.****B.****Fig. S1**



**Fig. S2**

**A.****B.****C.****D.****E.****F.****Fig. S3**

**A.****B.****C.****D.****Fig. S4**

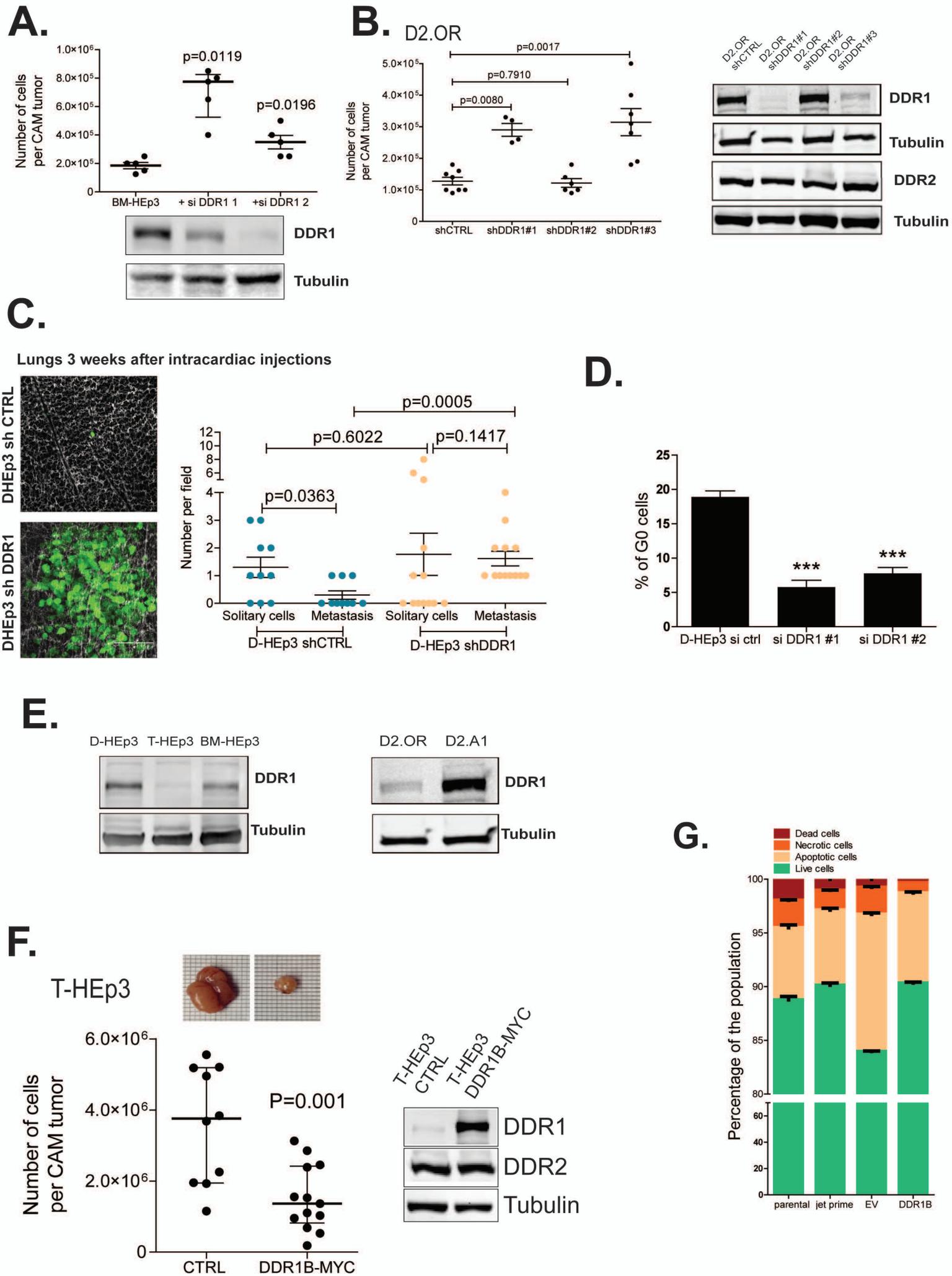
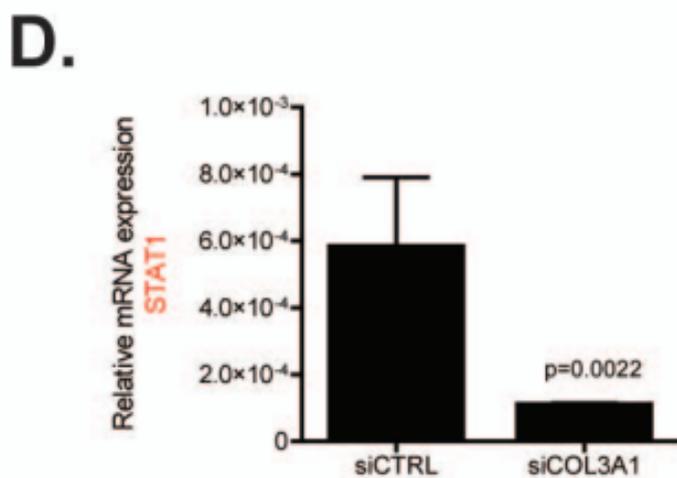
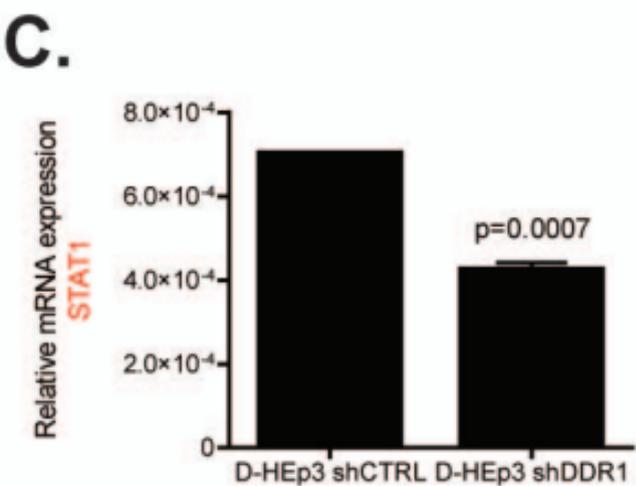
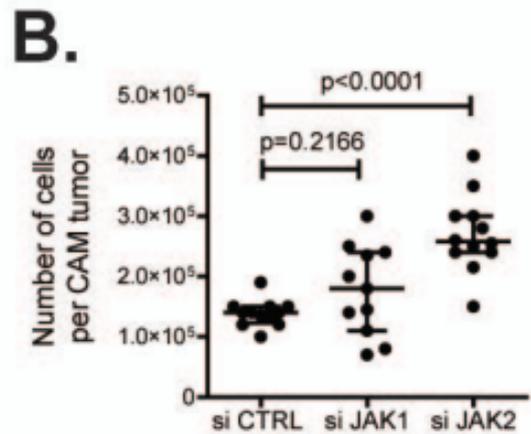
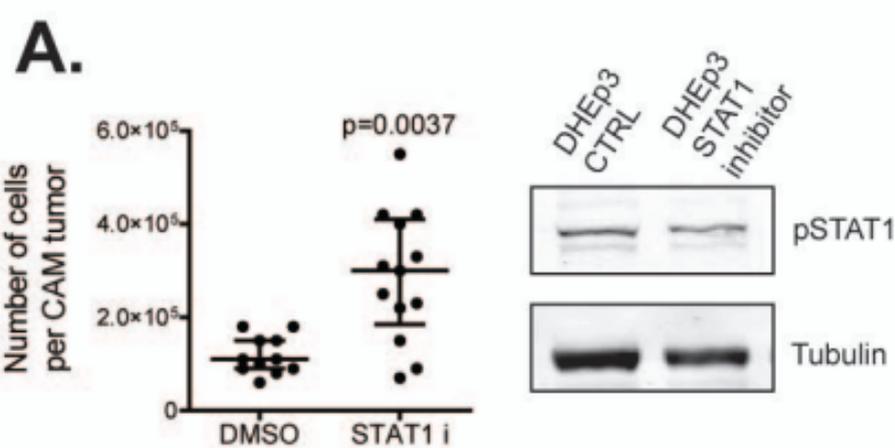


Fig. S5



**Fig. S6**

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable5.THEP3versusDHEp3TF.xlsx](#)
- [Supp.Table3.DHEp3shCTRLvsDHEp3shDDR1proteomics.xlsx](#)