Endocytic recycling is central to circadian collagen fibrillogenesis and disrupted in fibrosis

Authors

[1,2]Joan Chang*, [1]Adam Pickard, [1]Jeremy A. Herrera, [2]Sarah O'Keefe, [1]Richa Garva, [2]John Knox, [1]Thomas A. Jowitt, [1]Matthew Hartshorn, [1]Anna Hoyle, [3]Lewis Dingle, [2]Madeleine Coy, [4]Cédric Zeltz, [3]Jason Wong, [3]Adam Reid, [5]Rajamiyer V. Venkateswaran, [1]Yinhui Lu, [1]Patrick Caswell, [2]Stephen High, [4]Donald Gullberg, and [1]Karl E. Kadler*

Affiliations

[1]Wellcome Centre for Cell-Matrix Research, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK.

[2]Division of Molecular and Cellular Function, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK.

[3]Blond McIndoe Laboratories, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK.

[4]Department of Biomedicine and Centre for Cancer Biomarkers, Norwegian Center of Excellence, University of Bergen, Norway.

[5]Manchester University National Health Service Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK.

* Co-corresponding authors:

joan.chang@manchester.ac.uk (JC)

karl.kadler@manchester.ac.uk (KEK)

Abstract

Collagen-I fibrillogenesis is crucial to health and development, where dysregulation is a hallmark of fibroproliferative diseases. Here, we show that collagen-I fibril assembly required a functional endocytic system that recycles collagen-I to assemble new fibrils. Endogenous collagen production was not required for fibrillogenesis if exogenous collagen was available, but the circadian-regulated vacuolar protein sorting (VPS) 33b and collagen-binding integrin α11 subunit were crucial to fibrillogenesis. Cells lacking VPS33B secrete soluble collagen-I protomers but were deficient in fibril formation, thus secretion and assembly are separately controlled. Overexpression of VPS33B led to loss of fibril rhythmicity and over-abundance of fibrils, which was mediated through integrin α11β1. Endocytic recycling of collagen-I was enhanced in human fibroblasts isolated from idiopathic pulmonary fibrosis, where VPS33B and integrin α11 subunit were overexpressed at the fibrogenic front; this correlation between VPS33B, integrin α11 subunit, and abnormal collagen deposition was also observed in samples from patients with chronic skin wounds. In conclusion, our study showed that circadian-regulated endocytic recycling is central to homeostatic assembly of collagen fibrils and is disrupted in diseases.

Introduction

Collagen fibrils account for ~25% of total body protein mass [1] and are the largest protein polymers in vertebrates [2]. The fibrils can exceed centimeters in length and are organized into elaborate networks to provide structural support for cells. It is unclear how the fibrils are formed and how this process goes awry in collagen pathologies, such as fibrosis. A key mechanism, cell surface mediated fibrillogenesis, has been suggested to occur via indirect binding to fibronectin fibrils, or via direct assembly by collagen-binding integrins [3-5]. Interestingly, fibrils can be reconstituted *in vitro* from purified collagen (reviewed in [6]) but the assembly process is not controlled to the extent seen *in vivo* in terms of number, size and organization; this suggests a much tighter cellular control over the process. Additional support for tight cellular control of collagen fibril formation comes from electron microscope observations of collagen fibrils at the plasma membrane of embryonic avian and rodent tendon fibroblasts [7, 8], where the end (tip) of a fibril is enclosed within a plasma membrane invagination termed a fibripositor [9]. Previously we identified vacuolar protein sorting (VPS) 33b (a regulator of SNARE-dependent membrane fusion in the endocytic pathway) as a circadian clock regulated protein involved in collagen homeostasis [10]. VPS33B forms a protein complex with VIPAS39 (VIPAR) [11], and mutations in the *VPS33B* gene cause arthrogryposis-renal dysfunction-cholestasis (ARC) syndrome [12]. Here, death usually occurs within the first year of birth, accompanied with renal insufficiency, jaundice, multiple congenital anomalies, and predisposition to infection [13]. One proposed disease-causing mechanism is abnormal post-Golgi trafficking of lysyl hydroxylase 3 (LH3, PLOD3), which catalyzes the hydroxylation of lysyl residues in collagen to form hydroxylysine residues. It also has hydroxylysyl galactosyltransferase and galactosylhydroxylysyl glucosyltransferase activities, which creates attachment sites for carbohydrates that is crucial in stabilizing intramolecular and intermolecular crosslinks within the collagen structure, and thus is essential for collagen homeostasis during development [14, 15]. All these observations points to a central role for the endosome, situated in proximity of the plasma membrane and acting as a hub for a complex assortment of vesicles, in sorting collagen molecules to different fates.

Previous studies of collagen endocytosis have focused on collagen degradation and signaling, identifying additional collagen-binding proteins such as Endo180 and MRC1 [16-20], as well as a role for non-integrin collagen-receptors involved in mediating signaling activities such as DDR1/DDR2

(reviewed in [21]). Here, we showed that collagen uptake by fibroblasts is circadian rhythmic. Importantly, instead of degrading endocytosed collagen, cells utilize endocytic recycling of exogenous collagen-I to assemble new fibrils, even in the absence of endogenous collagen production. Further we showed that the secretion of soluble collagen protomers is separate from, and can occur independently to, collagen fibril assembly. We identified VPS33B and integrin α11 subunit (part of the collagen-binding integrin α11β1 heterodimer [4, 22, 23]) as central molecules specific to fibril formation. Finally, we showed that in idiopathic pulmonary fibrosis (IPF), a life threatening disease with unknown trigger/mechanism where lung tissue is replaced with collagen fibrils, integrin α11 subunit and VPS33B are located at the invasive fibroblastic focus where collagen fibril rapidly accumulates [24]; IPF fibroblasts also have elevated endocytic recycling of exogenous collagen-I, despite having similar collagen-I expression level to normal fibroblasts. Together, these results provide novel insights into the importance of the circadian clock-regulated endosomal system in normal fibrous tissue homeostasis, where fibrillogenesis occurs with both endogenous collagen and/or recycled scavenged exogenous collagen. This work also highlights that collagen utilization, rather than production (i.e. assembled into fibrils vs. protomeric secretion), is central to the maintenance of homeostasis, and dysregulation leads to disease.

Keywords: collagen-I, endocytic recycling, fibrillogenesis, integrin α11, circadian, trafficking, VPS33B, fibrosis, chronic wound

Results

Collagen-I is taken up into punctate structures within the cell, and reassembled into fibrils

To study collagen-I endocytosis and the fate of endocytosed collagen, we made Cy3 or Cy5 labeled collagen-I (Cy3-colI, Cy5-colI) from commercial rat-tail collagen [25]. We confirmed the helicity of these collagen with an expected molecular weight corresponding to a heterotrimeric mature collagen-I without the propeptides, and showed that the process of fluorescence-labelling did not alter the collagen trimer secondary structure or stability (**Supplementary Figure 1A-C**). We incubated Cy3-colI with explanted murine tail tendons for 3 days, then imaged the core of the tendon. Cells in tendon (nuclei marked with Hoechst stain) showed a clear uptake of Cy3-colI (**Figure 1A**, yellow box, indicated by yellow arrowheads). Interestingly, fibrillar Cy3 fluorescence signals that extend across cells were also observed, suggestive of these being extracellular fibrils (**Figure 1A**, grey box, indicated by white arrows). Pulse-chase experiments were performed where Cy3-colI was added to the tendons for 3 days, and then removed from the media. This was followed by 5FAM-colI for a further 2 days (**Figure S1D**). The results showed distinct areas where only Cy3-colI was observed (**Figure S1D**, yellow box); this persistence indicated that not all collagen endocytosed by cells is directed for degradation. 5FAM-positive and 5FAM/Cy3-positive fibril-like structures were also observed, confirming that collagen-I taken up by tissues is reincorporated into the matrix (**Figure S1D**, grey boxes). To confirm these findings in 2D cultures, we incubated immortalized murine tendon fibroblast cultures (iTTF) with labelled Cy3-colI. Flow cytometry analysis revealed that most cells had taken up Cy3-colI after overnight incubation (**Figure S1E**). Further time course analyses revealed a time-dependent and concentration-dependent increase of collagen-I uptake (**Figure 1B**). Cells incubated with Cy3-ColI for one hour followed by trypsinization were then analyzed using flow cytometry imaging (flow imaging), revealing that collagen-I is endocytosed by the cells into distinct puncta (**Figure 1C**). Similar results were obtained using Cy5-labelled collagen-I, indicating that the fluorescence label itself did not alter cellular response (**Figure S1F, G, H**). To confirm that the labelled collagen-I is in fact endocytosed into the fibroblasts and not only associated with the cell surface when cells are still attached, we performed live-imaging on cells incubated with Cy3-colI. Time lapse images showed Cy3-colI congregate in

113 structures within the cell, as indicated with white arrows (**Figure S1I**). Additionally, we transduced iTTF
114 with GFP-tagged Rab5 and confirmed co-localisation of Cy5-colI with Rab5-positive intracellular
115 structures (**Figure 1D, yellow arrows**). We then incubated iTTFs with Cy3-colI for one hour, before
116 trypsinizing and replating the cells to ensure any Cy3-colI signal detected hereon originated from the
117 endocytosed pool. Immunofluorescence (IF) staining indicated co-localization of endogenous
118 collagen-I with Cy3-labeled collagen-I in extracellular fibrillar structures after 72hrs, compared to a
119 striking lack of extracellular collagen-I (endogenous or labeled) at 18hrs (**Figure 1E, Figure S2A**). These
120 results demonstrated that exogenous collagen-I can be taken up by cells both *in vitro* and in tissues
121 *ex vivo*, and recycled into fibrils.

122 We considered the possibility that the fluorescent fibril-like structures were the result of spontaneous
123 cell-free fibrillogenesis of the added Cy3-colI. Thus, we incubated Cy3-colI at concentrations spanning
124 the 0.4 μg/mL critical concentration for cell-free *in vitro* fibril formation [26], with or without cells
125 (**Figure 1F**). In cell-free cultures, Cy3-positive fibrils were observed at 5 μg/mL, 1 μg/mL and 0.5 μg/mL
126 in a dose-dependent manner, with no discernible Cy3-positive fibrils at 0.1 μg/mL Cy3-colI (**Figure 1F**,
127 right panel; **Figure 1G**). However, in the presence of cells, Cy3-positive fibrillar structures were
128 observed at the cell surfaces at all concentrations of Cy3-colI examined, including 0.1 μg/mL (**Figure**
129 **1F** zoom ins of red box expanded to the left, highlighted by white arrows; **Figure 1H, I**). These results
130 indicated that the cells actively took up and recycled Cy3-colI into extracellular fibrils.

131 **The endocytic pathway controls collagen-I secretion and fibril assembly**

132 We then investigated the route that collagen-I is taken up. Prior research identified macropinocytosis
133 [27], which can be modelled with high-molecular weight dextran. We hypothesized that there will be
134 co-localisation of labelled collagen-I and labelled 70kDa dextran within the cell if they are entering
135 through the same route, however surprisingly we see little co-localization (**Figure S2B**). We then
136 performed a receptor saturation experiment, where unlabelled collagen-I was added in excess to
137 labelled collagen-I (10:1) during the 1-hour incubation, before flow imaging analysis of the fibroblasts.
138 Flow imaging showed that labelled collagen was at the cell-periphery and not intracellular when
139 saturation occurs, suggestive of a receptor-mediated process (**Figure S2C**). Thus, we used Dyngo4a
140 [28] to inhibit clathrin-mediated endocytosis, where Dyngo4a treatment (20 μM) leads to ~ 60%
141 reduction in Cy3-colI uptake relative to control (**Figure S2D**), without affecting cell viability (**Figure**
142 **S2E**). We then investigated the effects of Dyngo4a treatment on the ability of wild-type fibroblasts to
143 assemble collagen fibrils. Cells were treated with Dyngo4a for 48 h before fixation and IF against
144 collagen-I antibody. A significant reduction (average 60%) in the number of collagen-I fibrils assembled
145 was observed (**Figure 2A**), indicative of endocytosis playing a key role in collagen-I fibrillogenesis.
146 Conditioned media (CM) from these treated cells were also collected and assessed. To our surprise,
147 the amount of soluble collagen-I secreted was also greatly reduced (**Figure 2B**). qPCR analyses
148 revealed that *Col1a1* mRNA was significantly reduced in Dyngo4a-treated cells (**Figure S2F**),
149 suggesting a potential feedback mechanism between endocytosis and collagen-I synthesis, secretion,
150 and fibrillogenesis. Interestingly, whilst fibronectin (FN1) mRNA was significantly lower in Dyngo4a-
151 treated cells (**Figure S2F)**, and the intensity of FN1 signal appears lower, the amount of FN1 fibrils
152 deposited (as determined by the area occupied by fibrils) was not significantly impacted (**Figure 2C**).
153 Taken together, these results showed that inhibiting endocytosis in fibroblasts does not lead to
154 accumulation of soluble collagen-I protomers in the extracellular space. Rather, endocytosis impacts
155 collagen fibrillogenesis and the transcriptional control on collagen-I and fibronectin.

156 **Collagen-I endocytosis is circadian clock regulated, and recycling alone can generate fibrils**

157 Previously, we have shown that clock-synchronized fibroblasts synthesize collagen fibrils in a circadian
158 rhythmic manner [10]; the results here thus far indicate an involvement of endocytosis in collagen
159 fibrillogenesis. Therefore, we hypothesized that collagen-I endocytosis may also be circadian clock
160 regulated. Time-series flow cytometry analyses of fibroblasts incubated with Cy3-colI revealed that the
161 level of Cy3-colI endocytosed by the cells is rhythmic, with a periodicity of 23.8hrs as determined by
162 MetaCycle analysis (**Figure 2D**). When corrected to the running average, the rhythmic nature of Cy3-
163 colI uptake is accentuated (**Figure S2F)**. We noted that, when compared to the number of fibrils
164 produced over time, the peak time of uptake happens before peak fibril numbers (**Figure 2E**). These
165 data suggest that the cells may be endocytosing exogenous collagen under circadian control and
166 holding it in the endosomal compartment, then trafficking the collagen to the plasma membrane for
167 fibril formation.

168 To eliminate the possibility that the fluorescent fibril-like structures were due to attachment of
169 fluorescently-labeled collagen protomers to pre-existing fibrils already deposited by cells, fibrils
170 already deposited by cells, we performed an siRNA-mediated knockdown against *Col1a1* (siCol1a1)
171 to target endogenous collagen production. Fibroblasts were then analyzed by IF using anti-collagen-I
172 or anti-fibronectin (FN1) antibodies. Control (scrambled siRNAs, scr) cells synthesized collagen-I and
173 fibronectin (**Figure 3A**, top row) with defined fibrillar structures. In contrast, cells treated with siCol1a1
174 synthesized lower-levels of collagen-I, with significantly few collagen-I fibrillar structures as well as
175 lower intracellular collagen-I signal (**Figure 3A**, bottom row; quantification plots). A ~ 90% reduction
176 in *Col1a1* mRNA was confirmed using qPCR (**Figure S3A**).

177 Flow cytometry analysis of cells incubated with Cy3-colI revealed that siCol1a1 fibroblasts retained the
178 ability to endocytose exogenous collagen (**Figure S3B**). To assess if siCol1a1 fibroblasts can assemble
179 a collagen fibril with exogenous collagen, Cy3-colI was incubated with scr or siCol1a1 cells for 1 hour,
180 before trypsinization and replating to ensure no cell surface associated Cy3-colI remains. Within 3 days,
181 Cy3-positive collagen fibrils were observed in both scr (**Figure 3B**, left column) and siCol1a1 cells
182 (**Figure 3B**, right column). The observation of fibrillar Cy3 signals in siCol1a1 cells showed that the
183 cells can repurpose collagen into fibrils without the requirement for intrinsic collagen-I production (red
184 arrow, **Figure 3B**). The cells were also probed for collagen-I using an anti-collagen-I antibody with a
185 secondary antibody conjugated to Alexa-488 (**Figure 3B**, second row). In scr cells we detected multiple
186 fibrillar structures that did not contain Cy3-colI (**Figure 3B,** white arrows), indicating fibrils derived from
187 endogenous collagen-I. In the siCol1a1 cells we detected a more diffuse signal and puncta within the
188 cell, but no discernible Alexa-488-only collagen-I fibrils (**Figure 3B**). These results indicate that
189 fibroblasts can assemble collagen-I fibrils from recycled exogenous collagen-I. Importantly, siCol1a1
190 cells treated with Dyngo4a could not effectively form Cy3-collagen-containing fibrils under these
191 conditions, indicative of a requirement for endocytic-recycling of exogenous collagen-I for
192 fibrillogenesis, when endogenous levels are insufficient (**Figure 3C**). To confirm these findings, we
193 isolated primary tendon fibroblasts from the Col1a2-CreERT2::Col1a1-fl/fl (termed CKO) mouse that
194 had either been treated with tamoxifen (CKO+) to produce fibroblasts that cannot synthesise collagen-
195 I or from untreated mice to yield matched controls (CKO-). Fibroblasts were then analysed by
196 immunofluorescence microscopy using anti-collagen-I antibodies. As expected CKO- cells synthesised
197 collagen-I, and CKO+ cells showed no collagen-I expression (**Figure 4A**). Quantitative PCR analysis of
198 the cells confirmed the absence of Col1a1 mRNAs containing exon 6 to exon 8 sequences, as expected
199 from the location of the LoxP sites in the Col1a1 gene (**Figure S3C**).

200 Flow analysis of cells incubated with Cy3-colI revealed collagen knockout fibroblasts retained the
201 ability to endocytose labelled collagen (**Figure S3D**). To assess whether collagen knockout fibroblasts

202 were able to assemble a collagen fibril, Cy3-colI was incubated with cells before the cells were released
203 by trypsin and replated, to ensure that Cy3 signals arose only from Cy3-colI endocytosed by cells.
204 Within 3 days, Cy3-positive collagen fibres could be observed in both CKO- (**Figure 4B, left column**)
205 and CKO+ cells (**Figure 4B, right column**). The observation of fibrillar Cy3 signals in CKO+ cells
206 showed that the cells can repurpose collagen into fibrils without the requirement for intrinsic collagen-
207 I (**red arrows, Figure 4B**). The cells were also probed for collagen-I using an anti-collagen-I antibody
208 with a secondary antibody conjugated to Alexa-488 (**Figure 4B, second row**). In CKO- cells we
209 detected a diffuse signal over the entire cell and an indication of fibrillar structures. In the CKO+ cells
210 we detected fibrillar structures at the periphery of the cell, and the same diffuse signal seen in CKO-
211 cells (**Figure 4B, white arrows**). We suspect that the diffuse signal observed in CKO+ cells is due to
212 incomplete labelling of collagen-I in our preparation of Cy3-colI, as complete labelling of lysine
213 residues would cause collagen-I to be unable to form fibrils [25]. Nonetheless, Dyngo4a treatment led
214 to a lack of Cy3-colI fibrils (**Figure 4C, right panel and quantification plot)**; thus, these results
215 demonstrated that fibroblasts can effectively form fibrils from exogenous collagen alone.

216 **VPS33B controls collagen fibril formation but not protomeric secretion**

217 VPS33B is situated in a post-Golgi compartment where it is involved in endosomal trafficking with
218 multiple functions including extracellular vesicle formation [29], modulation of p53 signaling [30], and
219 maintenance of cell polarity [31], dependent on cell type. We previously identified VPS33B to be a
220 circadian-controlled component of collagen-I homeostasis in fibroblasts [10]. As a result, we decided
221 to further investigate its role in endosomal recycling of collagen-I.

222 We confirmed our previous finding that CRISPR-knockout of VPS33B (VPSko, as verified by western
223 blot analysis and qPCR analysis, **Figure S4A and B**) in tendon fibroblasts led to fewer collagen fibrils
224 without impacting proliferation (**Figure S4C**), as evidenced by both electron microscopy (**Figure 5A**)
225 and IF imaging (**Figure 5B**). Decellularized matrices showed that VPSko fibroblasts produced less
226 matrix by mass than control, which was mirrored by a reduction in hydroxyproline content (**Figure 5C**).
227 We then stably overexpressed VPS33B in fibroblasts (VPSoe), as confirmed by western blot, qPCR
228 analyses, and flow cytometry of transfected cells (**Figure S4D-F**). IF staining indicated a greater number
229 of collagen-I fibrils in VPSoe cells (**Figure 5D**), although the mean total matrix and mean total
230 hydroxyproline in VPSoe cultures was not significantly higher than control (**Figure 5E**). As VPSoe cells
231 showed equivalent proliferation to controls (**Figure S4G**), this suggests VPSoe specifically enhanced
232 the assembly of collagen fibrils. We then performed time-series IF on synchronized cell cultures to
233 quantify the number of collagen fibrils formed. Tendon fibroblasts exhibited a ~ 24 h rhythmic
234 fluctuation in collagen-I fibril numbers, whereas VPSoe cells continuously deposited collagen-I fibrils
235 over the 55-h period (**Figure 5F**). MetaCycle analyses indicated that fluctuation of fibril numbers in
236 control and VPSoe cultures occurred at a periodicity of 22.7 h and 28.0 h respectively (**Figure S4H**),
237 indicating that continuous expression of VPS33B leads to loss of collagen-I fibril circadian homeostasis.
238 Interestingly, when assessing the levels of secreted soluble collagen-I protomers from control and
239 VPSoe cells, VPSoe CM have lower levels of soluble collagen-I (**Figure 5G**). In addition, when VPS33B
240 is knocked-down using siRNA, there is an elevation of soluble collagen-I secreted; siVPS33B on VPSoe
241 cells also increased the levels of collagen-I in CM (**Figure S4I, J**), whilst VPS33B levels are not
242 correlated with *Col1a1* expression levels (**Figure S4K**). This finding indicates that VPS33B is specifically
243 involved in collagen-I fibril assembly, and not secretion or translation. The reduction of secreted
244 soluble collagen-I in VPSoe cells further supports that VPS33B directs collagen-I towards fibril
245 assembly.

246 **VPS33B-positive intracellular structures contain collagen-I**
247 We then utilized the split-GFP system [32, 33] to investigate how VPS33B specifically directs collagen-
248 I to fibril assembly. Here, a GFP signal will be present if VPS33B co-traffics with collagen-I. We, and
249 others, have previously demonstrated that insertion of tags at the N-terminus of proα2(I) or proα1(I)
250 chain does not interfere with collagen-I folding or secretion [34, 35]. To determine which terminus of
251 the VPS33B protein GFP1-10 should be added, we performed computational ΔG analysis [36], which
252 predicts that VPS33B contains two regions of extended hydrophobicity towards its C-terminus (**Figure**
253 **S5A**). The first region (residues 565-587, denoted TMD1) has a ΔG of -0.62 kcal/mol, consistent with a
254 single pass type IV transmembrane domain (TMD), known as a tail-anchor. This arrangement locates
255 the short C-terminus of VPS33B inside the lumen of the endoplasmic reticulum (ER), and subsequently
256 within the endosomal compartment [37]. The second region (residues 591-609, denoted HR1) is
257 significantly less hydrophobic, with a ΔG of +2.61 kcal/mol (**Figure S5B**). Its presence raises the
258 possibility that VPS33B may contain two TMDs that, depending on their relative membrane topologies,
259 result in either a luminal or cytosolic C-terminus (**Figure S5C**).

260 To define the topology of VPS33B, we used a well-established *in vitro* system where newly synthesized
261 and radiolabeled proteins of interest are inserted into the membrane of ER-derived canine pancreatic
262 microsomes, and created constructs with ER luminal modification of either endogenous N-
263 glycosylation sites (N54, N545 in VPS33B), or artificial sites in an appended OPG2 tag (N2, N15 in
264 residues 1-18 of bovine rhodopsin, Uniprot: P02699) as a robust reporter for membrane protein
265 topology in the ER (**Figure S5D, E**) [38]. Due to the N-terminal region of VPS33B being highly
266 aggregation-prone in our *in vitro* system (**Figure S5F, G**), we created three additional chimeric proteins
267 comprised of different regions of VPS33B and Sec61β to investigate the topology of VPS33B (**Figure**
268 **S5H**). Whilst a small proportion of each chimera continued to pellet when synthesized in the absence
269 of ER-derived microsomes (**Figure S5I,** lanes 3, 6, 9), N-glycosylated species were now clearly
270 identifiable for each of the chimeras tested.

271 In chimera 1, given the respective efficiency of membrane insertion of the TA region and N-
272 glycosylation of the C-terminal OPG tag of Sec61βOPG2 (**Figure S5I,** lanes 1-2), we attribute the N-
273 glycosylated species to the C-terminal translocation of OPG2 tag, although it is evident that the
274 remaining short stretch of VPS33B (residues 414-564) still impedes efficient membrane insertion, likely
275 due to aggregation (**Figure S5I**, lanes 3-5) [39, 40]. For chimera 2, the efficient modification of its
276 distal N-glycosylation site (N15 of the OPG2 tag) but inefficient use of the proximal site (N2 of the
277 OPG2 tag) reflects the latter residues' close proximity to the ER membrane [41], and supports the *bona*
278 *fide* membrane insertion of the VPS33B TMD1 and thus translocation of the C-terminal OPG2 tag into
279 the ER lumen (**Figure S5I**, lanes 6-8). Interestingly, the inclusion of HR1 to the C-terminus of TMD1 in
280 chimera 3 results in a substantial qualitative reduction in the amount of protein that is N-glycosylated
281 (**Figure S5I**, lanes 7 and 10). Further, for the fraction of chimera 3 that is modified, the majority is
282 doubly N-glycosylated; most likely due to the extra length provided by HR1, resulting in both N-
283 glycosylation sites on the OPG2 tag (N2 and N15) now being efficiently modified [41].

284 The clear reduction in the proportion of N-glycosylated species obtained with chimera 3 compared to
285 chimera 2 (**Figure S5I**) indicates that only a small proportion of HR1 and the appended OPG2 tag are
286 successfully translocated into the ER lumen. This is likely due to a combination of the low proportion
287 of hydrophobic residues and the presence of several charged and polar amino acids in HR1[42]. We
288 thus propose that, for a minority of VPS33B, TMD1 is inserted into the ER membrane as a tail-anchored
289 region with a luminal HR1 that most likely remains associated with the inner leaflet of the bilayer (N-
290 cytosolic, C-luminal; **Figure 6A**, right). In contrast, the majority of VPS33B likely assumes a 'hairpin' like

291 conformation in the ER membrane where both its N- and C-termini remain in the cytosol (N-cytosolic,
292 C-cytosolic) with either: i) a partially membrane inserted TMD1, with HR1 associated with the outer
293 leaflet of the ER membrane (**Figure 6A**, left); or ii) a fully membrane inserted TMD1 followed by the
294 marginally hydrophobic HR1 which may span the membrane through the formation of stabilizing
295 hydrogen bonds with TMD1 (**Figure 6A**, middle) [43].

296 To test our topology findings, we inserted BFP at either the N- (VPSnBFP) or C-terminus (VPScBFP) of
297 VPS33B. In cells expressing VPSnBFP, the fluorescence signal was diffuse throughout the cell body,
298 and in some cases appeared to be completely excluded from circular structures (**Figure 6B**, left),
299 suggestive of protein mistargeting. In contrast, VPScBFP-expressing cells have punctate blue
300 fluorescence signals and peripherally blue structures (**Figure 6B**, right), supportive of the topology
301 findings. Previous studies have also tagged VPS33B at the C-terminus [11]. Thus, the 214-residue N-
302 terminal fragment (GFP1-10) was cloned onto VPS33B, and the 17-residue C-terminal peptide (GFP11)
303 was cloned onto the proα1(I) chain of collagen-I as previously described [34] (GFP11-proα1(I), **Figure**
304 **6C**).

305 Fibroblasts stably expressing VPS-barrel and GFP11-proα1(I) were imaged to identify any GFP
306 fluorescence. The results showed puncta throughout the cell body (**Figure 6D**). Intriguingly, GFP
307 fluorescence was also observed at the cell periphery (**Figure 6D**, green box zoom). IF staining of
308 endogenous VIPAS39 (a known VPS33B-interacting partner [11]) also revealed co-localization of
309 VIPAS39 with collagen-I in intracellular punctate structures, where in some of the co-localized puncta,
310 the signal of VIPAS39 is strongest surrounding collagen-I (**Figure 6E**, zoom), suggesting that VIPAS39
311 is encasing collagen-I, not within the lumen but present in proximity with the external membrane of
312 these structures. VPS33B has been demonstrated to interact with VIPAS in regions before the TMD1
313 site [44]; thus, in all suggested VPS33B topologies herein, VIPAS39 will still be able to interact with
314 VPS33B, and encase collagen-I-containing intracellular structures.

315 The relationship between VPS33B levels and collagen fibril numbers was confirmed using
316 endogenously-tagged Dendra2-collagen-I expressing 3T3 cells [45], where notably the number of
317 Dendra2-positive fibrils significantly increased when VPS33BcBFP was expressed; however, the
318 average length of the fibril was not significantly different, suggesting that VPS33B is important in fibril
319 initiation but not elongation (**Figure 6F**). Incubation of iTTFs expressing only the VPS-barrel with CM
320 collected from iTTFs expressing only GFP11-proα1(I) revealed that the GFP signal was detected only
321 in intracellular structures and not along the periphery of the cells, where endocytosis takes place
322 (**Figure 6G**). Flow cytometry analyses of fluorescently labelled-colI endocytosed by control, VPSko,
323 and VPSoe fibroblasts also revealed no consistent change in uptake by VPSko or VPSoe cells (**Figure**
324 **6H**). Taken together, these results suggest that VPS33B interacts with endocytosed collagen-I within
325 the cell and and trafficks with collagen-I to the extracellular space, and is not involved with collagen-I
326 endocytosis. VPSko cells replated after Cy5-colI uptake have conspicuously fewer fibrils when
327 compared to control. In contrast, VPSoe cells have shorter but more Cy5-colI fibrils (**Figure 5I**),
328 highlighting the role of VPS33B in recycling endocytosed collagen-I to initiate collagen fibrillogenesis.

329 **Integrin chain α11 mediates VPS33B-dependent fibrillogenesis**

330 Having identified VPS33B as a driver for collagen-I fibril formation but not protomeric secretion, we
331 used biotin cell surface labeling coupled with mass spectrometry protein identification to identify other
332 proteins that may be involved in this process at the cell surface. VPSko and VPSoe fibroblasts were
333 analyzed using a "shotgun" approach. In total, 4121 proteins were identified in total lysates (**Table**
334 **S1**), and 1691 proteins in the enriched-for-surface-protein samples (**Table S2**). Gene Ontology (GO)

335 Functional Annotation analysis [46, 47] identified the top 25 enriched terms (based on p-values) with
336 the top 5 terms all associated with "extracellular" or "cell surface" (**Figure 7A**), indicative that the
337 biotin-labeling procedure had successfully enriched proteins at the cell surface interacting or
338 associated with the extracellular matrix (ECM).

339 We then interrogated the differences between VPSko and control cells, visualizing the results in a semi-
340 quantitative manner using spectral counting (**Table S3**). Conspicuously, collagen α1(I) and α2(I) chains
341 were detected at a reduced level at the surface of VPSko cells (**Figure 7B**). The reduction of α1(V) and
342 α2(V) chains (which constitute type V collagen) from the cell surface supports the long-standing view
343 that collagen-V nucleates collagen-I containing fibrils. Gene ontology pointed to integral components
344 of the plasma membrane, which included several integrins. Whilst many integrins were detected in all
345 samples there was an absence of integrin α1 and integrin α11 subunit in VPSko cultures (**Figure 7C**).
346 It is well established that integrins are cell-surface molecules that interact extensively with the ECM
347 [48], with integrin α11β1 functioning as a major collagen binding integrin on fibroblasts [23], which is
348 expressed during development, and is upregulated in subsets of myofibroblasts in tissue and tumor
349 fibrosis [5, 49, 50]. Importantly, whilst the level of integrin β1 chain detected was also reduced in VPSko
350 cultures, it was not as drastic as the reduction of integrin α11 chain. This is likely due to the
351 promiscuous nature of integrin β1 subunit, that is being able to partner with other α subunits for
352 functions other than collagen-I interaction. The absence of integrin α11 subunit from VPSko cells
353 suggested a link between VPS33B-mediated collagen fibrillogenesis and integrin α11**β**1.

354 VPS33B was conspicuous by its absence from cell-surface labeling studies of control samples (**Figure**
355 **7D**). We have previously struggled to detect VPS33B protein using mass spectrometry [10] and
356 postulate that its absence could be due to low abundance. Regardless, VPS33B was detected at the
357 cell surface in VPSoe samples along with PLOD3 (**Figure 7D**), a lysyl hydroxylase involved in stabilizing
358 collagen and previously identified to be delivered by VPS33B [14]. Complementary western blot
359 analysis of total cell lysates showed that expression of integrin α11 subunit is significantly reduced in
360 VPSko cells and is elevated in VPSoe cells (**Figure 7E**). This correlation between VPS33B and integrin
361 α11 expression levels was confirmed at the mRNA level (**Figure 7F**), inferring a link between VPS33B,
362 collagen fibril, and integrin α11 abundance. We verified this by siRNA knockdown of *Itga11* in control
363 and VPSoe fibroblasts, where knockdown efficiency is confirmed by qPCR (**Figure 7G**). siItga11 in both
364 control and VPSoe cells reduced the number of collagen-I fibrils in culture (**Figure 7H**). Thus, even with
365 elevated levels of VPS33B, integrin α11 subunit is required for collagen-I fibrillogenesis. Interestingly,
366 knocking-down integrin α11 subunit increased exogenous collagen-I uptake as demonstrated by flow
367 cytometry (**Figure 7I**). Thus, we propose that both VPS33B and integrin α11 are involved in directing
368 collagen-I protomers to the formation of collagen-I fibrils, and not mediators of collagen-I endocytosis.

369 **Integrin α11 subunit is localized to the fibroblastic focus of idiopathic pulmonary fibrosis**

370 We next investigated if VPS33B and integrin α11 are involved in human fibrotic diseases, where
371 accumulation of collagen fibrils is a disease hallmark. Lung fibroblasts isolated from control individuals
372 or individuals suffering from IPF were cultured and the mRNA levels of *COL1A1*, *ITGA11*, and *VPS33B*
373 determined (**Figure 8A**). Although *COL1A1* was similarly expressed, the levels of *VPS33B* and *ITGA11*
374 were significantly increased in IPF fibroblasts compared to control. We next assessed the endocytic
375 capacities of collagen-I in human lung fibroblasts, and confirmed co-localisation of exogenous Cy5-
376 colI with Rab5-GFP (**Figure S6A**). We also found that IPF fibroblasts endocytosed significantly more
377 Cy5-labelled exogenous collagen-I when compared to control fibroblasts (**Figure 8B**, left); uptake of
378 Cy3-labelled exogenous collagen-I was also elevated, albeit not significantly (**Figure 8B**, right).
379 Subsequent culture of cells that have taken up exogenous collagen-I revealed that IPF fibroblasts made

380 significantly more fluorescently-labeled collagen fibrils, indicating enhanced recycling of endocytosed
381 collagen-I to generate new fibrils (**Figure 8C**). The relationship between VPS33B, ITGA11 and
382 endocytic recycling of collagen-I for fibrillogenesis was further confirmed using siRNA (**Figure S6B**),
383 where knock-down of either proteins led to a significant decrease in recycling of exogenous Cy5-ColI
384 (**Figure 8D**). This *in vitro* observation was also represented in IPF pathology, where patient-derived
385 lung samples showed enrichment of collagen-I which overlaps with both integrin α11 subunit and
386 VPS33B within the IPF hallmark lesion (termed the fibroblastic focus [24], encircled by red dotted line,
387 **Figure 9A**). In sites of emerging fibrotic remodeling, integrin α11 subunit and VPS33B are also
388 detected (red asterisks, **Figure 9B**; additional N = 3 IPF specimens, **Figure S7A**), indicating a role for
389 VPS33B/integrin α11 chain in collagen-I deposition in the context of IPF. In control lung samples,
390 collagen-I and VPS33B were present whereas integrin α11 subunit was detected at negligible levels
391 (**Figure 8C, Figure S7B**). These results indicate that the proteins required for assembly of collagen into
392 fibrils (i.e. integrin α11, VPS33B) are present at the fibrotic fronts of IPF, and that enhanced endocytic
393 recycling of collagen-I by fibroblasts may be a disease-potentiating mechanism in IPF.

394 **VPS33B and integrin α11 subunit levels are elevat3ed in chronic skin wounds**

395 Fibrosis has been described as a dysregulation of the normal wound-healing process [51]. Thus, we
396 postulated that chronic skin wounds, a similar chronic inflammation unresolved wound-healing
397 condition, may also share a similar pathological molecular signature as IPF. Immunohistochemical (IHC)
398 staining revealed that in human chronic skin wounds, the expressions of integrin α11 subunit and
399 VPS33B are elevated when compared to normal skin areas taken from the same patient (**Figure 9**),
400 where expression was evident at both the fibrotic wound margins and perivascular tissues, indicative
401 of areas under constant collagen remodeling, suggestive of a similar dysregulation in collagen
402 fibrillogenesis pathway utilizing VPS33B and integrin α11.

403 Discussion

404 In this study we have identified an endocytic recycling mechanism for type I collagen fibrillogenesis
405 that is under circadian regulation. Integral to this process is VPS33B (a circadian clock-regulated
406 endosomal tethering molecule) and integrin α11 subunit (a collagen-binding transmembrane receptor
407 when partnered with integrin β1 subunit). Collagen-I co-traffics with VPS33B to the plasma membrane
408 for fibrillogenesis, which requires integrin α11. These proteins are all enhanced at active sites of
409 collagen pathologies such as fibrosis and chronic skin wounds, suggestive of a common disease
410 mechanism.

411 Previous research [16, 52-55] has focused on degradation or signaling as the endpoint of collagen-I
412 endocytosis. However, our results show that collagen-I endocytosis contributes towards fibril
413 formation, which is reduced when endocytosis is inhibited. The decision between degrading or
414 recycling endocytosed collagen may depend on cell type, microenvironment, or collagen type itself.
415 In complex environments, for example a wound, degradation of damaged collagen molecules and
416 recycling of structurally sound collagen molecules could be beneficial. A caveat in this present study
417 is the use of Dyngo4a to inhibit endocytosis. It was demonstrated that there may be non-dynamin
418 targeting effects with Dyngo4a when comparing the effects of dynamin triple knock-out and drug
419 treatment (e.g. fluid-phase endocytosis and membrane ruffling) [56]; here we have shown that collagen
420 uptake is likely through receptor-mediated and not fluid-phase endocytosis, although whether
421 periphery membrane ruffling alters collagen deposition remains to be investigated. As total secretion
422 did not appear to be affected by Dyngo4a treatment, as indicated by Ponceau stain on western blots
423 of CM, we take this to suggest that Dyngo4a in this study is not affecting overall secretion and is mostly
424 acting on endocytosis. Further studies targeting multiple endocytosis routes, either through genetic

425 manipulation or drug treatments, will help shed light on how collagen uptake is controlled, and if
426 different forms of collagen (i.e. protomers, fibril) require different mechanisms. Nonetheless, the
427 requirement of VPS33B for fibril assembly confirms the previously-observed rhythmicity of collagen
428 fibril formation by fibroblasts, *in vivo* and *in vitro* [10]. Our discovery that the endosomal system is
429 involved in fibril assembly explains how collagen fibrils can be assembled in the absence of
430 endogenous collagen synthesis, namely, if collagen can be retrieved from the extracellular space. We
431 also observed a delay between maximum uptake and maximum fibril numbers. A possible function of
432 this delay may be to increase the concentration of collagen-I in readiness for more efficient fibril
433 initiation. This is supported by our findings that fibroblasts can utilize a solution of Cy3-colI lower than
434 the 0.4 μg/mL critical concentration for fibril formation to assemble fibrils. Of note, the comparison
435 between uptake and fibril numbers are not directly from the same experiments due to experimental
436 limitations; the Cy3/Cy5-colI uptake we observed did not reflect total collagen-I levels, as endogenous
437 collage-I was present as well. Nonetheless, we postulate that the fates of the collagen-I that was taken
438 up by fibroblasts are likely: 1) recycled for fibril formation, 2) re-secreted as protomers, or if they are
439 damaged, 3) degradation. Future research will be required to determine the molecular mechanisms
440 that determines these fates. Surface biotin-labelling in fibroblasts identified two integrin α subunits
441 (α1 and α11) to be absent in VPSko cells which could not make collagen fibrils. There are four known
442 collagen-binding integrins (α1β1, α2β1, α10β1, α11β1), where α2β1 and α11β1 have affinity for fibrillar
443 collagens and are expressed in fibroblasts *in vivo* [23]. α1β1, in contrast, has been demonstrated *in*
444 *vitro* to display higher affinity for collagen-IV than collagen-I [57]; it has a wide expression pattern *in*
445 *vivo*, including basement membrane-associated smooth muscle cells, strongly suggesting that
446 collagen-IV is the major ligand for α1β1 integrin instead. Thus, we focused on the effects of integrin
447 α11 subunit. Interestingly, both α1 and α11 subunits were also reduced in VPSoe cells as detected by
448 mass spectrometry; however, validation by western blots indicated that protein levels of integrin α11
449 follows VPS33B and collagen fibril numbers. We postulate that this discrepancy detected through MS
450 might be due to integrin dynamics at the plasma membrane. Regardless, VPS33B and integrin α11β1
451 are essential for targeting collagen-I to fibril assembly and not for the secretion of soluble triple helical
452 collagen-I molecules (i.e. protomers); there is also no evidence to support an active role for VPS33B or
453 α11β1 in collagen endocytosis. We have also demonstrated involvement of VPS33B and integrin α11
454 subunit in IPF, particularly at sites of disease progression (fibroblastic foci). The presence of integrin
455 α11 subunit confirms a recent study utilizing spatial proteomics to decipher regions of IPF tissues,
456 where in addition to integrin α11 the authors also identified proteins involved in collagen biogenesis
457 specific to the fibroblastic foci [58]. Crucially, in our study, levels of *Col1a1* transcript were unchanged
458 between control and IPF fibroblasts, and IHC demonstrated that collagen-I and VPS33B were prevalent
459 in control lungs. This highlights that the production of collagen-I does not equate to fibril formation;
460 indeed, here VPS33B is required for normal collagen homeostasis in the lung, and it is the combination
461 of VPS33B and integrin α11 subunit, coupled with elevated endocytic recycling of exogenous collagen
462 at the fibroblastic focus, that promotes excess collagen-I fibril assemblies, which is a hallmark of
463 fibrosis. Furthermore, this relationship between excessive collagen-I, VPS33B, and integrin α11 is also
464 observed in chronic skin wounds, where, similar to lung fibrosis, there is a chronic unresolved wound-
465 healing response; this is suggestive of a common pathological molecular pathway between the two
466 organs, based on enhanced endocytic recycling of collagen-I protomers directed to fibrillogenesis.

467 The discovery that VPS33B is required for collagen fibril formation but not collagen protomer secretion
468 highlights an important distinction between 'collagen secretion' and 'collagen fibrillogenesis',
469 especially in the context of collagen pathologies (e.g. fibrosis, cancer metastasis). In this context, it is
470 crucial to keep in mind that elevated collagen levels (i.e. collagen transcripts or protomer levels) does

471 not equate to the formation of an insoluble fibrillar network. This distinction was also suggested in a
472 recent study on *Pten*-knockout mammary fibroblasts, where SPARC protein acts via fibronectin to
473 affect collagen fibrillogenesis but not secretion [59]. Of note, whilst SPARC and fibronectin were
474 detected in our previously-reported time-series mass spectrometry analyses [10], they were not
475 circadian clock rhythmic in tendon.

476 The nucleation of a collagen-I fibril is expected to involve collagen-V, a minor fibril-forming collagen
477 that is necessary for the appearance of collagen-I fibrils *in vivo* [60]. Another long-standing view is that
478 fibronectin may tether collagen to a fibronectin-binding integrin and thereby function as a
479 proteolytically cleavable anchor ([61], reviewed by [62]). Whilst the present study did not focus on the
480 role of fibronectin, there is a slight indication that fibronectin is decreased when collagen-I is knocked
481 down, which IF staining suggested that collagen-I and fibronectin do not completely overlap, and that
482 inhibition of endocytosis did not affect fibronectin fibril numbers. Further, biotin-surface labeling mass
483 spectrometry analysis in the present study indicated that the levels of the major fibronectin-binding
484 integrins (integrins α5β1, αVβ3) were not drastically altered between control, VPSko, and VPSoe
485 fibroblasts, indicating that VPS33B controls collagen fibrillogenesis via a fibronectin-independent
486 route. We have recently demonstrated that integrin α11β1 localizes to fibrillar adhesions and
487 contributes to collagen assembly in a mechano-regulated and tensin-1 dependent manner,
488 independent of α5β1 and fibronectin [63]. In these studies, the α11β1 mediated collagen assembly
489 did not appear to depend on endocytosis. In the future it will be important to establish which factors
490 governs endocytosis-dependent and endocytosis-independent α11β1-mediated collagen
491 fibrillogenesis. Interestingly, in fibronectin knockout liver cells, exogenous collagen-V could be added
492 to initiate collagen-I fibrillogenesis [64], which suggests that other matrix proteins could also be
493 recycled in a similar manner as collagen-I. Whilst collagen α1(V) could only be detected in surface
494 biotin-labeled control but not VPSko or VPSoe cells, collagen α2(V) was only detected in VPSoe cells.
495 Of note, we have previously shown that collagen α2(V) protein was rhythmic and in phase with collagen
496 α1(I) and collagen α2(I). Whether collagen-V co-traffics with collagen-I in VPS33B-compartments to
497 fibrillogenic sites at the cell periphery, or collagen-V is directed to the nucleation site by another route,
498 is an important question to address in future studies. Indeed, endocytic recycling of other factors
499 required for collagen fibrillogenesis (e.g. other nucleation factors such as collagen-V, collagen-XI,
500 integrins), in addition to collagen-I protomers, may be required for fibril assembly by fibroblasts.

501 Previous research has demonstrated that collagen-I can be secreted within minutes [9], whilst other
502 studies have demonstrated a relatively slow emergence of collagen-I fibrils over days in culture [45].
503 The delay in assembly of fibrils could be due to the requirement to reach the threshold concentration
504 of collagen needed for fibrillogenesis [26]. Procollagen-I can be converted to collagen-I within the
505 secretory pathway [9, 65] in a process that requires giantin for intracellular N-terminal processing of
506 procollagen-I [66]. Thus, it is possible that triple helical collagen-I protomers destined for fibril
507 formation are: 1) directed straight to a VPS33B-positive compartment from the Golgi apparatus, or 2)
508 quickly secreted and then recaptured, via endocytosis, prior to recycling for fibrillogenesis. Such a
509 mechanism would provide the cell with additional opportunities to determine the number of fibrils,
510 the rate at which they are to be initiated, and their required growth in the extracellular matrix (see
511 **Figure 10**). We have not identified the specific proteins that mediate collagen-I uptake into the cell,
512 however as inhibition of clathrin-mediated endocytosis did not completely inhibit collagen
513 endocytosis, it is likely that collagen-I has multiple routes to enter the cells [18-21]. With regards to
514 the fibrillogenesis route, although a recent model of VPS33B/VIPAS39 protein complexes does not
515 consider the possibility that VPS33B can associate directly with vesicular membranes [67], our *in vitro*

516 studies suggest its hydrophobic C-terminal region can act as a tail-anchor domain. Importantly, the
517 proposed model of the VPS33B/VIPAS39 complex suggests the C-terminus of VPS33B would be
518 available for membrane association, consistent with our *in vitro* findings. Further studies will be
519 required to establish whether differences between our models reflect alternative experimental
520 systems, yeast vs. mammalian, and/or the behavior of VPS33B alone vs. in complex with VIPAS39. We
521 found that endogenous VIPAS39 encases collagen-I in intracellular puncta, consistent with a model
522 where VPS33B and VIPAS39 co-traffic with collagen-I at the same endosomal compartment (see also
523 [67]). Our data suggest two membrane bound populations of VPS33B, the majority with a hairpin-like
524 structure and cytosolic C-terminus and a minority that spans the membrane with the C-terminus inside
525 the lumen. Whether or not alternative topologies of VSP33B are associated with different biological
526 functions remains to be determined. Nonetheless, here we have identified a previously unknown role
527 for endocytosis in collagen fibrillogenesis, which is exploited in collagen pathologies.

528 Author contributions

529 JC, AP and KEK conceived the project. KEK supervised the experiments. AP, JC, RG, JK, TAJ
530 performed experiments and interpreted data. JAH consented and biopsied patients. JAH, LD, MC
531 performed experiments, MH, PC performed super-resolution live imaging, analysis, and data
532 interpretation, YL performed electron microscopy and analysis, CZ produced integrin α11 antibodies
533 and performed antibody screening, DG, RVV, AH provided reagents and data interpretation, and SH
534 and SO'K performed the VPS33B ΔG analysis, topology analyses, and data interpretation. All authors
535 contributed to writing the manuscript.

536 Acknowledgements

537 The proteomics was performed at the Biological Mass Spectrometry Facility with the assistance of
538 Stacey Warwood and Ronan O'Cualain, imaging was performed at the Bioimaging Facility, flow
539 cytometry analyses were performed at the Flow Cytometry Facility with the assistance of Gareth
540 Howell, all in the Faculty of Biology, Medicine and Health (University of Manchester). This report is
541 independent research supported by the North West Lung Centre Charity and National Institute for
542 Health Research Clinical Research Facility at Manchester University NHS Foundation Trust. The views
543 expressed in this publication are those of the author(s) and not necessarily those of the NHS, the North
544 West Lung Centre Charity, National Institute for Health Research or the Department of Health. The
545 authors would like to acknowledge the Manchester Allergy, Respiratory and Thoracic Surgery Biobank
546 and the North West Lung Centre Charity for supporting this project. We would also like to thank the
547 study participants for their contribution. The authors would like to acknowledge and thank Prof David
548 Stephens for his feedback on this manuscript. Some illustrations were created with Biorender.com.

549 **Funding**

550 Wellcome Senior Investigator Award 110126/Z/15/Z (KEK)

551 MRC Career Development Award MR/W016796/1 (JC)

552 Wellcome Centre Award 203128/Z/16/Z (KEK)

553 Wellcome Investigator Award in Science 204957/Z/16/Z (SH)

554 Norwegian Centre of Excellence grant 223250 (DG)

555 Norwegian Cancer Society grant 223052 (DG)

556 Nasjonalföreningen for folkhelsen grant, project 16216 (DG)

557 Wellcome Institutional Strategic Support fund 204796/Z/16/Z (AR, JW)

558 Wellcome Translational Partnership Award 209741/Z/17/Z (AR, JW)

559 Declaration of interests

560 The authors declare no competing interests.

561 Materials and Methods

562 Procurement of human lung tissue and human lung-derived fibroblasts

563 The use of human lung tissue was approved by the National Health Research Authority with patient
564 consent (NRES14/NW/0260). The specimens used for this study met the criteria for Idiopathic
565 Pulmonary Fibrosis (IPF) diagnosis [68], as previously described [69]. 4 IPF patient samples and 2
566 control lung samples were used in this study. The patient-derived fibroblasts, isolated from 5 IPF
567 patient samples and 5 control samples used here were a kind gift from Professor Peter Bitterman and
568 Professor Craig Henke at the University of Minnesota. Briefly, control and IPF-derived fibroblasts
569 (isolated from were established by allowing fibroblasts to propagate from lung tissue explants as
570 previously described [70]). The fibroblasts used in the described experiments were between passages
571 4 and 6.

572 Procurement of human skin tissues

573 Human chronic wound samples (as defined by [71]) were collected alongside healthy skin samples from
574 4 consenting patients undergoing surgical debridement and reconstruction at Manchester University
575 NHS Foundation Trust. All procedures were approved by the National Health Research Authority
576 through the ComplexWounds@Manchester Biobank (NRES 18/NW/0847).

577 Cell culture

578 Unless otherwise stated, all cell culture reagents were obtained from Gibco, and all cells were
579 maintained at 37 °C, 5% CO_2 in a humidified incubator. Immortalized mouse tail tendon fibroblasts
580 (iTTF, isolated from 8-10 weeks old C57BL/6 male mice [10]), and NIH3T3 cells with Dendra2-tagged
581 collagen-I expression (Dendra2-3T3 [45]) were cultured in DMEM/F12 with sodium bicarbonate and L-
582 glutamine (supplemented with 10% fetal calf serum (FCS), 200 μM L-ascorbate-2-phosphate (Sigma),
583 and 10,000 U/mL penicillin/streptomycin), and high-glucose DMEM with sodium bicarbonate and L-
584 glutamine (supplemented with 10% new born calf serum, 200 μM L-ascorbate-2-phosphate, and
585 10,000 U/mL penicillin/streptomycin) media respectively. Human patient-derived fibroblasts were
586 cultured in low glucose DMEM with sodium bicarbonate and L-glutamine (supplemented with 10%
587 FCS, 200 μM L-ascorbate-2-phosphate, and 10,000 U/mL penicillin/streptomycin). HEK293T cells were
588 cultured in DMEM with sodium bicarbonate and L-glutamine, supplemented with 10% FCS.

589 To synchronize cells in culture, 100 μM Dexamethasone was added to each sample, then incubated
590 for an hour before the media was removed and replaced with fresh culture media. This marks t=0 in
591 the experiments (i.e. post-synchronisation time 0, or PS0).

592 Murine tail tendons were isolated by first removing the skin from the mouse tails by cutting it from tail
593 root and degloving the whole tail. Once tendons become exposed, tendon fibers were pulled out
594 using sterile forceps, breaking every other vertebra. The vertebra was then trimmed with scalpels
595 leaving just the tail tendons. These were then placed on a cell culture insert and equilibrated in full
596 DMEM/F12 culture media (see above section) overnight, before the fluorescently-labelled collagen
597 were added at 5μg/ml for 3 days, washed extensively with PBS, and then removed before embedding
598 in low-melting-point agarose for confocal imaging (Leica SP8 upright confocal microscopy). For the

599 pulse-chase experiment, Cy3-labelled collagen was first added and incubated, before tendons were
600 washed and 5FAM-labelled collagen was added for a further 48 hours prior to imaging.

601 **CRISPR-Cas9-mediated knockout**

602 iTTFs were treated with CRISPR-Cas9 to delete *VPS33B* gene as previously described [10]. Gene
603 knockout was confirmed by western blotting and qPCR.

604 **Constructs**

605 The cDNA for mouse VPS33B (Uniprot: P59016) was purchased from Genscript (OMu07060D). Sec61β
606 modified with a C-terminal OPG2 tag (residues 1-18 of bovine rhodopsin (Uniprot: P02699) was as
607 previously described [72]. An artificial N-glycosylation site (VPS33B-A606T) and OPG2-tag (VPS33B-
608 OPG2) were incorporated into parent VPS33B by site-directed mutagenesis (Stratagene QuikChange,
609 Agilent Technologies) using the relevant forward and reverse primers (Integrated DNA Technologies;
610 see reagents table) and confirmed by DNA sequencing (GATC, Eurofins Genomics). Linear DNA
611 templates were generated by PCR and mRNA transcribed using T7 polymerase or SP6 RNA
612 polymerase as appropriate (see reagents table). In order to improve the signal intensity of all
613 radiolabeled proteins, an additional five methionine residues (5M) were appended to the C-terminus
614 of all VPS33B linear DNA templates by PCR. All primer combinations used for mutagenesis and PCR
615 are listed in the tables below.

616 **Transfection and stable infection of over-expression vectors in cells**

617 Three different vectors were used to stably over-express VPS33B protein in iTTF and 3T3 cells –
618 untagged VPS33B overexpression (with red fluorescent protein expression as selection), N-terminus
619 BFP-tagged VPS33B overexpression, and C-terminus BFP-tagged VPS33B overexpression. All proteins
620 were cloned into pLV V5-Luciferase expression vector, where luciferase was excised from the backbone
621 prior to gel purification, or into pCMV expression vector, followed by Gibson Assembly® (NEB) [73].
622 Lentiviral particles were generated In HEK293T cells, and cells were infected in the presence of 8
623 µg/mL polybrene (Millipore). Cells were then sorted using Flow Cytometry to isolate RFP-positive or
624 BFP-positive cells. Alternatively, cells were treated with puromycin (Sigma) or G418 (Sigma)
625 depending on selection marker on the vectors.

626 **Fluorescence labeling of collagen-I**

627 Cy3 or Cy5 NHS-ester dyes (Sigma) were used to fluorescently label rat tail collagen-I (Corning), using
628 a previously described method [25]. Briefly, 3 mg/mL collagen-I gels were made, incubated with 50
629 mM borate buffer (pH 9) for 15 min, followed with incubation with either Cy3-ester or Cy5-ester (Sigma)
630 dissolved in borate buffer in dark overnight at 4 °C, gently rocking. Dyes were then aspirated and 50
631 mM Tris (pH 7.5) were added to quench the dye reaction, and incubated in the dark rocking for 10
632 min. Gels were washed with PBS (with calcium and magnesium ions) 6x, incubating for at least 30 min
633 each wash. The gels were then resolubilized using 500 mM acetic acid and dialyzed in 20 mM acetic
634 acid.

635 **Circular dichroism of rat tail collagen-I**

636 Circular dichroism spectra of rat tail collagen in 10 mM acetic acid was recorded on a Jasco J810
637 spectrometer using 0.2mm quartz coverslip sample holders. Spectra were recorded using
638 approximately 0.5 mg/ml rat tail collagen for WT, Cy3 labelled and Cy5 labelled collagens. The melting
639 curves were performed by monitoring at 223 nm only which is the positive triple helical peak and
640 recording between 30 and 70 degrees Celsius, and capturing data points every 30 seconds with a
641 temperature increment of 1 degree every 60 seconds in a 1mm quartz cuvette. All outputs are

642 recorded in machine units rather than being converted to molar units because of the difficulty in
643 ascertaining accurate concentrations of labelled collagen.

644 **Mass Photometry of rat tail collagen-I**

645 The Refeyn mass photometer was used to assess the mass of the collagen molecules. The instrument
646 was calibrated with BSA and PTX3 with a mass range of 67, 135 and 340 kDa. The WT, Cy3 labelled
647 and Cy5 labelled collagens were diluted to ~20 nM in 10 mM acetic acid and counts read over the
648 course of 60-seconds. The ratiometric contrast was converted to mass using the calibration standards.

649 **RNA isolation and quantitative real-time PCR**
650 RNA was isolated using TRIzol Reagent (Thermo Fisher Scientific) following manufacturer's protocol,
651 and concentration was measured using a NanoDrop OneC (Thermo Fisher Scientific). Complementary
652 DNA was synthesized from 1 µg RNA using TaqMan™ Reverse Transcription kit (Applied Biosystems)
653 according to manufacturer's instructions.

654 SensiFAST SYBR kit reagents were used in qPCR reactions. Primer sequences can be found in reagents
655 section.

656 **Protein extraction and western blotting**
657 For lysate experiments, proteins were extracted using urea buffer (8 M urea, 50 mM Tris-HCl pH7.5,
658 supplemented with protease inhibitors and 0.1% β-mercaptoethanol). For conditioned media (CM)
659 media, cells were plated out at 200,000 cells per 6-well plate, and left for 48-72 h before 250 µl was
660 sampled. Samples were mixed with 4xSDS loading buffer with 0.1% β-mercaptoethanol and boiled at
661 95°C for 5 min. The proteins were separated on either NuPAGE Novex 10% polyacrylamide Bis–Tris
662 gels with 1XNuPAGE MOPS SDS buffer or 6% Tris-glycine gels with 1XTris-glycine running buffer (all
663 Thermo Fisher Scientific), and transferred onto polyvinylidene difluoride (PVDF) membranes (GE
664 Healthcare). The membranes were blocked in 5% skimmed milk powder in PBS containing 0.01%
665 Tween 20. Antibodies were diluted in 2.5% skimmed milk powder in PBS containing 0.01% Tween 20.
666 The primary antibodies used were: rabbit polyclonal antibody (pAb) to collagen-I (1:1,000; Gentaur),
667 mouse mAb to vinculin (1:1000; Millipore), rabbit pAb to integrin α11 subunit (1:1000; see [74]) and
668 mouse mAb to VPS33B (1:500; Proteintech). Horseradish-peroxidase-conjugated antibodies and
669 Pierce ECL western blotting substrate (both from Thermo Fisher Scientific) were used and reactivity
670 was detected on GelDoc imager (Biorad). Alternatively, Licor goat-anti-mouse 680, mouse-anti-rabbit
671 800 were used and reactivity detected on an Odyssey Clx imager.

672 ***In vitro* ER import assays**
673 Translation and ER import assays were performed in nuclease-treated rabbit reticulocyte lysate
674 (Promega) as previously described [38, 39, 75]: briefly, in the presence of EasyTag EXPRESS ^{35}S Protein
675 Labelling Mix containing [^{35}S] methionine (Perkin Elmer) (0.533 MBq; 30.15 TBq/mmol), 25 µM amino
676 acids minus methionine (Promega), 6.5% (v/v) nuclease-treated ER-derived canine pancreatic
677 microsomes (from stock with OD_{280}=44/mL) or an equivalent volume of water, and 10% (v/v) of *in vitro*
678 transcribed mRNA (~1 µg/µL) encoding relevant precursor proteins. Translation reactions for
679 Sec61βOPG2 (20 µL; 1X) were performed at 30°C for 30 min whereas VPS33B and its variants (20 µL;
680 2X) were performed at 30°C for 1 h. Irrespective of the precursor protein, all translation reactions were
681 finished by incubating with 0.1 mM puromycin for 10 min at 30°C to ensure translation termination
682 and the ribosomal release of newly synthesized proteins prior to analysis. Microsomal membrane-
683 associated fractions were recovered by centrifugation through an 80 µL high-salt cushion (0.75 M
684 sucrose, 0.5 M KOAc, 5 mM $Mg(OAc)_2$ and 50 mM HEPES-KOH, pH 7.9) at 100,000 g for 10 min at
685 4°C, the pellet suspended directly in SDS sample buffer and, where indicated, treated with 1000 U of

686 endoglycosidase Hf (New England Biolabs, P0703S). All samples were denatured for 10 min at 70°C
687 and resolved by SDS–PAGE (10% or 16% PAGE, 120 V, 120-150 min). Gels were fixed for 5 min (20%
688 MeOH, 10% AcOH), dried for 2 h at 65°C, and radiolabeled products were visualized using a Typhoon
689 FLA-700 (GE Healthcare) following exposure to a phosphorimaging plate for 24–72 h. Images were
690 opened using Adobe Photoshop and annotated using Adobe Illustrator.

691 **Flow cytometry and imaging**

692 Cy3- or Cy5-tagged collagen was added to cells and incubated at 37°C, 5% CO_2 in a humidified
693 incubator for predetermined lengths of times before washing in PBS, trypsinized, spun down at 2500
694 rpm for 3 min at 4 °C, and resuspended in PBS on ice. For labelled dextran experiment, 200 µg/ml
695 Oregon green™ 488-labelled 70kDa dextran and 10 µg/ml Cy5-coll were added to the cells together
696 and incubated for 1 hr; for unlabelled collagen-I saturation experiments, 100 µg/ml rat-tail collagen-I
697 were added together with 10 µg/ml labelled collagen-I and incubated for 1 hr. Cells were then
698 processed as described above. Cells were analyzed for Cy3-/Cy5-collagen/488-dextran uptake using
699 LSRFortessa (BD Biosciences). For imaging, cells were prepared as described, and analyzed on
700 Amnis® ImageStream®^{X}Mk II (Luminex).

701 **Decellularization of cells to obtain extracellular matrix**

702 Cells were seeded out at 50,000 in a 6-well plate and cultured for 7 days before decellularization.
703 Extraction buffer (20mM NH_4OH, 0.5% Triton X-100 in PBS) was gently added to cells and incubated
704 for 2 minutes. Lysates were aspirated and the matrix remaining in the dish were washed gently twice
705 with PBS, before being scraped off into ddH_2O for further processing.

706 **Hydroxyproline assay**

707 Samples were incubated overnight in 6 M HCl (diluted in ddH_2O (Fluka); approximately 1 mL per 20
708 mg of sample) in screw-top tubes (StarLab) in a sand-filled heating block at 100 °C covered with
709 aluminum foil. The tubes were then allowed to cool down and centrifuged at 12,000 g for 3 min.
710 Hydroxyproline standards were prepared (starting at 0.25 mg/mL; diluted in ddH2O) and serially
711 diluted with 6 M HCl. Each sample and standard (50 µL) were transferred into fresh Eppendorf tubes,
712 and 450 µL chloramine T reagent [0.127 g chloramine T in 50% N-propanol diluted with ddH2O; made
713 up to 10 mL with acetate citrate buffer (120 g sodium acetate trihydrate, 46 g citric acid, 12 mL glacial
714 acetic acid, 34 g sodium hydroxide) adjusted to pH 6.5 and then made to 1 litre with dH_2O; all reagents
715 from Sigma] was added to each tube and incubated at room temperature for 25 min. Ehrlich's reagent
716 (500 µL; 1.5 g 4-dimethylaminobenzaldehyde diluted in 10 mL N-propanol:perchloric acid (2:1)) was
717 added to each reaction tube and incubated at 65 °C for 10 min and then the absorbance at 558 nm
718 was measured for 100 µL of each sample in a 96-well plate format.

719 **Electron microscopy**

720 Unless otherwise stated, incubation and washes after EM fixation were done at room temperature.
721 Cells were plated on top of ACLAR films and allows to deposit matrix for 7 days. The ACLAR was then
722 fixed in 2% glutaraldehyde/100mM phosphate buffer (pH 7.2) for at least 2 h and washed in ddH_2O 3
723 x 5 min. The samples were then transferred to 2% osmium (v/v)/1.5% potassium ferrocyanide (w/v) in
724 cacodylate buffer (100 mM, pH 7.2) and further fixed for 1 h, followed by extensive washing in ddH_2O.
725 This was followed by 40 minutes of incubation in 1% tannic acid (w/v) in 100 mM cacodylate buffer,
726 and then extensive washes in ddH_2O. and placed in 2% osmium tetroxide for 40 min. This was followed
727 by extensive washes in ddH_2O. Samples were incubated with 1% uranyl acetate (aqueous) at 4 °C for
728 at least 16 h, and then washed again in ddH_2O. Samples were then dehydrated in graded ethanol in
729 the following regime: 30%, 50%, 70%, 90% (all v/v in ddH_2O) for 8 min at each step. Samples were

730 then washed 4 x 8 min each in 100% ethanol, and transferred to pure acetone for 10 min. The samples
731 were the infiltrated in graded series of Agar100Hard in acetone (all v/v) in the following regime: 30%
732 for 1 h, 50% for 1 h, 75% for overnight (16 h), 100% for 5 h. Samples were then transferred to fresh
733 100% Agar100Hard in labeled moulds and allowed to cure at 60 °C for 72 h. Sections (80 nm) were
734 cut and examined using a Tecnai 12 BioTwin electron microscope.

735 **Surface biotinylation-pulldown mass spectrometry**

736 Cells were grown in 6 well plates for 72 h prior to biotinylating. Briefly, cells around 90% confluence
737 were kept on ice and washed in ice-cold PBS, following with incubation with ice-cold biotinylating
738 reagent (prepared fresh, 200 µg/mL in PBS, pH 7.8) for 30 min at 4 °C gently shaking. Cells were then
739 washed twice in ice-cold TBS (50 mM Tris, 100 mM NaCl, pH 7.5), and incubated in TBS for 10 min at
740 4 °C. Cells were then lysed in ice-cold 1% TritonX (in PBS, with protease inhibitors) and lysates cleared
741 by centrifugation at 13,000 g for 10 min at 4 °C. Supernatant was then transferred to a fresh tube and
742 1/5 of the lysates were kept as a reference sample. Streptavidin-Sepharose beads was aliquoted into
743 each sample and incubated for 30 min at 4 °C rotating. Beads were then washed 3x in ice cold PBS
744 supplemented with 1% TritonX, followed by one final wash in ice cold PBS and boiled at 95 °C for 10
745 min in 2x sample loading buffer, followed by centrifugation at 13,000 g for 5 min follow by sample
746 preparation for Mass Spectrometry analysis. The protocol used for sample preparation was as
747 described previously [69]. Mass spectrometry results files were exported into Proteome Discoverer
748 (PD) for identification and spectral counting. All searches included the fixed modification for
749 carbamidomethylation on cysteine residues resulting from IAA treatment to prevent cysteine bonding.
750 The variable modifications included in the search were oxidized methionine (monoisotopic mass
751 change, +15.955 Da); hydroxylation of asparagine, aspartic acid, proline, or lysine (monoisotopic mass
752 change, +15.955 Da); and phosphorylation of threonine, serine, and tyrosine (79.966 Da). A maximum
753 of 2 missed cleavages per peptide was allowed. The minimum precursor mass was set to 350Da with
754 a maximum of 5000. Precursor mass tolerance was set to 10ppm, fragment mass tolerance was 0.02Da
755 and minimum peptide length was 6. Peptides were searched against the Swissprot database using
756 Sequest HT with a maximum false discovery rate of 1%. Proteins required a minimum FDR of 1% and
757 were filtered to remove known contaminants and to have at least 2 unique peptides. Missing values
758 were assumed to be due to low abundance. For spectral counting, enriched samples were normalized
759 and spectral matches were compared. Peptide spectral matches were only included if they were
760 calculated to have a q-value of less than 0.01.

761 **Imaging and immunofluorescence**

762 For fluorescence live imaging of Cy3-colI internalisation, iTTF cells were seeded for 24 h onto Ibidi µ-
763 plates and stained using CellMask™ Green Plasma Membrane Stain (Invitrogen C37608) according to
764 the manufacturers protocol. Cy3-colI was added to cells and images were collected in a 37°C chamber
765 using a Zeiss 3i spinning disk (CSU-X1, Yokagowa) confocal microscope with a 63x/1.40 Plan-
766 Apochromat oil objective. 7.5 µm z-stacks with a step size of 0.5 µm were captured at 2 minute intervals
767 using 488 nm (100 power, 150 ms exposure) and 561 nm (100 power, 100 ms exposure) lasers with
768 SlideBook 6.0 software (3i) and a front illuminated Prime sCMOS camera. Imaris 9.9.1 software was
769 then used to generate 3D reconstructions.

770 For fixed immunofluorescence imaging, cells plated on coverslips or Ibidi µ-plates were fixed with
771 100% methanol at -20°C and then permeabilized with 0.2% Triton-X in PBS. Primary antibodies used
772 were as follows: rabbit pAb collagen-I (1:400, Gentaur OARA02579), rabbit pAb VIPAS (1:50,
773 Proteintech 20771-1-AP), mouse mAb FN1 (1:400, Sigma F6140 or Abcam ab6328). Secondary
774 antibodies conjugated to Alexa-488, Alexa-647, Cy3, and Cy5 were used (ThermoScientific), and nuclei

775 were counterstained with DAPI (Sigma). Coverslips were mounted using Fluoromount G (Southern
776 Biotech). Images were collected using a Leica SP8 inverted confocal microscope (Leica) using an
777 x63/0.50 Plan Fluotar objective. The confocal settings were as follows: pinhole, 1 Airy unit; scan speed,
778 400 Hz bidirectional; format 1024 x 1024 or 512 x 512. Images were collected using photomultiplier
779 tube detectors with the following detection mirror settings: DAPI, 410-483; Alexa-488, 498-545nm;
780 Cy3, 565-623 nm; Cy5, 638-750 nm using the 405 nm, 488 nm, 540 nm, 640 nm laser lines.
781 Alternatively, images were collected using EVOS™ M7000 (Thermo Fisher) using 40x/1.3na oil
782 objective. Images were collected in a sequential manner to minimize bleed-through between channels.

783 For fluorescence live-imaging, cells were plated onto Ibidi μ-plates and imaged using Zeiss LSM880
784 NLO (Zeiss). For split-GFP experiment, cells were seeded for 24 h onto Ibidi μ-plates before imaging
785 with an Olympus IXplore SpinSR (Olympus) with 100x oil magnification. Prior to imaging, media was
786 changed to FluoroBrite media with the appropriate supplements.

787 **Immunohistochemistry**

788 Human lung samples were formalin-fixed and paraffin-embedded (FFPE). 5-μm FFPE sections were
789 obtained and mounted onto Superfrost® Plus (Thermo Scientific™) slides and subjected to antigen
790 heat retrieval using citrate buffer (Abcam, ab208572), in a pre-heated steam bath for 20 minutes,
791 before cooling to room temperature in a water bath for 20 min. Slides were then treated with 3-4%
792 hydrogen peroxide (Leica Biosystems RE7101) for 10 min, blocked in SuperBlock™ (TBS) blocking
793 buffer for a minimum of 1 h (Thermo Scientific™; 37581), and probed with primary antibodies
794 overnight at 4°C in 10% SuperBlock™ solution in Tris Buffered Saline Tween 20 solution (TBS-T, pH
795 7.6). Primary antibodies used were as follows: Collagen-I A1/A2 (Rockland Immunochemicals Inc 600-
796 401-103.0.5), integrin α11 (integrin α11 mAb 210F4 [76]), VPS33B (Atlas antibodies).

797 After overnight incubation, specimens were subjected to Novolink Polymer Detection Systems (Leica
798 Biosystems RE7270-RE, as per the manufacturer's recommendations), with multiple TBS-T washes.
799 Sections were developed for 5 min with DAB Chromagen (Cell Signaling Technology®, 11724/5)
800 before being counterstained with haematoxylin for 1 minute, followed by acid alcohol and blueing
801 solution application. Slides were dehydrated through sequential ethanol and xylene before being
802 cover-slipped with Permount mounting medium (Thermo Scientific™, SP15).

803 **Histological imaging**

804 Stained slides were imaged using a DMC2900 Leica camera along with Lecia Application Suite X
805 software (Leica).

806 ## QUANTIFICATION AND STATISTICAL ANALYSIS

807 Data are presented as the mean ± s.e.m. unless otherwise indicated in the figure legends. The sample
808 number 'N' indicates the number of independent biological samples in each experiment, and 'n'
809 indicates the number of technical repeats, and are indicated in the figure legends. Data were analyzed
810 as described in the legends. The data analysis was not blinded, apart from quantification of fibril
811 numbers over time, and differences were considered statistically significant at $P<0.05$, using Student's
812 t-test or One-way Anova, unless otherwise stated in the figure legends. Analyses were performed using
813 Graphpad Prism 8 or 10.2.0 software. Significance levels are: * $P<0.05$; ** $P<0.01$; *** $P<0.005$;
814 ****$P<0.0001$. Where applicable, normality test was performed using the Shapiro-Wilk method. For
815 periodicity, analysis was performed using the MetaCycle package[77] in the R computing environment
816 [78] with the default parameters.

817 **Reagents**

REAGENT or RESOURCE	SOURCE	IDENTIFIER

Antibodies		
rabbit polyclonal antibody (pAb) to collagen-I	Aviva Systems Biology	Cat# OARA02579, RRID:AB_10873334
rabbit polycloncal antibody (pAb) to collagen-I	Kerafast	Cat# ENH018
mouse mAb to mouse vinculin	Sigma-Aldrich	Cat# V9131, RRID:AB_477629
rabbit pAb to mouse integrin α11 subunit	This paper	N/A
mouse mAb to mouse VPS33B	Proteintech	Cat# 12195-1-AP, RRID:AB_2215198
Rabbit pAb to mouse VIPAS	Proteintech	Cat# 20771-1-AP, RRID:AB_10695764
Mouse mAb to FN1	Sigma-Aldrich	Cat# F6140, RRID:AB_476981
Mouse mAb to FN1	Abcam	Cat# ab6328, RRID:AB_305428
Rabbit pAb to human VPS33B	Atlas antibodies	Cat# HPA040415, RRID:AB_10795419
Mouse mAb integrin α11 210F4	DOI: 10.1002/cjp2.148	N/A
Rabbit pAb to Collagen-I	Rockland Immunochemicals	600-401-103.0.5, RRID:AB_217595
IRDye® 680RD_Goat anti-Mouse_IgG (H + L)	Li-Cor	Cat# 925-68070, RRID:AB_2651128
IRDye® 800RD_Goat anti-Rabbit_IgG (H + L)	Li-Cor	Cat# 925-32211, RRID:AB_2651127
Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP	Thermo Fisher Scientific	Cat# 32430, RRID:AB_1185566
Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP	Thermo Fisher Scientific	Cat# 32460, RRID:AB_1185567
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Cyanine3	Thermo Fisher Scientific	Cat# A10520, RRID:AB_2534029
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Cyanine5	Thermo Fisher Scientific	Cat# A10523, RRID:AB_2534032
Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Cyanine3	Thermo Fisher Scientific	Cat# A10521, RRID:AB_2534030
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-21236, RRID:AB_2535805
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	Cat# A-11029, RRID:AB_2534088
Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	Cat# A-11034, RRID:AB_2576217
Bacterial and virus strains		
pLV-V5-VPS33B-GFP1-10	This paper	N/A
pLV-V5-VPS33B-BFP	This paper	N/A
pLV-V5-BFP-VPS33B	This paper	N/A
pCMV-SBP-GFP11-COL1A1	This paper	N/A
pcDNA3.1+/C-(K)-DYK	This paper	N/A
Biological samples		
Human control lung tissues	Manchester University NHS Foundation Trust	NRES14/NW/0260

Human IPF lung tissues	Manchester University NHS Foundation Trust	NRES14/NW/0260
Human chronic skin wound tissue	Manchester University NHS Foundation Trust	NRES 18/NW/0847
Human healthy skin wound tissue	Manchester University NHS Foundation Trust	NRES 18/NW/0847
Chemicals, peptides, and recombinant proteins		
Recombinant S. pyogenes Cas9 nuclease protein	IDT	1081059
Rat tail collagen-I	Corning	354259
Dyngo®4a	Abcam	Ab120689
TRIzol reagent	Invitrogen	15596-018
Oregon green™ 488-Dextran, 70kDa	Life technologies	D7173
Cy®3 NHS Ester	Sigma	GEPA13101
Cy®5 NHS Ester	Sigma	GEPA15101
Critical commercial assays		
Gibson Assembly Cloning kit	NEB	E5510S
Site-Directed Mutagenesis QuikChange kit	Agilent	200513
Novolink Polymer Detection Systems	Leica Biosystems	RE7270-RE
Rabbit reticulocyte lysate, nuclease treated	Promega	L4960
TaqMan™ Reverse Transcription kit	Applied Biosystems	N8080234
SensiFAST™SYBR® No-ROX kit	Bioline	BIO-98005
Experimental models: Cell lines		
NIH3T3-Dendra2-Col1	This paper	N/A
Immortalized mouse tail tendon fibroblasts	This paper	N/A
Primary human IPF lung fibroblasts	University of Minnesota	N/A
Primary human control lung fibroblasts	University of Minnesota	N/A
HEK293T	This paper	N/A
Oligonucleotides		
siCol1a1, Mission® esiRNA	Merck	EMU069551
siItga11, Mission® esiRNA	Merck	EMU042761
Primers for site-directed mutagenesis	See primer list	N/A
Primers for transcription templates	See primer list	N/A
Primers for qPCR reactions	See primer list	N/A
Recombinant DNA		
VPS33B mouse cDNA	Genscript	Omu07060D
Sec61β modified with a C-terminal OPG2 tag	This paper	N/A
VPS33B A-606-T	This paper	N/A
VPS33B-OPG2	This paper	N/A
VPS33B-Sec61βTMD	This paper	N/A
VPS33B-Sec61βTMD-OPG2	This paper	N/A
Software and algorithms		
Fiji ImageJ	doi:10.1038/nmeth.2019	https://imagej.net/software/fiji/
MetaCycle	doi:10.1093/bioinformatics/btw405	https://github.com/gangwug/MetaCycle

Scaffold Proteome Software	doi:10.1002/pmic.200900437	https://www.proteomesoftware.com/products
Graphpad Prism 8	https://www.graphpad.com/scientific-software/prism/	https://www.graphpad.com/scientific-software/prism/
Biorender	https://biorender.com/	https://biorender.com/

818

819 Primer list for cDNAs generated by site-directed mutagenesis.

New cDNA	Cloned cDNA	Vector	Forward primer (5′-3′)	Reverse primer (5′-3′)
VPS33B A-606-T	VPS33B	pcDNA3.1+/C-(K)-DYK	ACTGCTGTTACAAACAGTACCCGCCTCATGGAAGCC	GGCTTCCATGAGGCGGGTACTGTTTGTAACAGCAGT
VPS33B-OPG2	VPS33B	pcDNA3.1+/C-(K)-DYK	GCCAACGGAACAGAAGGACCAAACTTCTACGTACCATTCAGCAACAAAACAGGCTAATCCGATTACAAGGATGACGAC	CTATTAGCCTGTTTTGTTGCTGAATGGTACGTAGAAGTTTGGTCCTTCTGTTCCGTTGGATTTCACCTCACTCATGGCTTC
VPS33B-Sec61βTMD	VPS33B	pcDNA3.1+/C-(K)-DYK	GTATTGGTTATGTGTCTTCTGTTCATCGCTTCTGTATTTATGTTGCACATTTGGGGCAAGTACACTCGTTCGTAGCTGCGCCTCATCTTGGTGGTGTTCC	CTACGAACGAGTGTACTTGCCCCAAATGTGCAACATAAATACAGAAGCGATGAACAGAAGACACATAACCAATACTGACTCACTGGAAGCCTTGTCTTCC
Chimera 3 (VPS33B-Sec61βTMD-OPG2)	VPS33B-Sec61βTMD	pcDNA3.1+/C-(K)-DYK	AACGGAACAGAAGGACCAAACTTCTACGTACCATTCAGCAACAAAACAGGCTAATAGCTGCGCCTCATCTTGGTGGTGTTCCTG	CTATTAGCCTGTTTTGTTGCTGAATGGTACGTAGAAGTTTGGTCCTTCTGTTCCGTTCGAACGAGTGTACTTGCCCCAAATGTGCAAC

820

821 Primer list for PCR reactions used to create transcription templates.

Substrate	cDNA	Vector	Forward primer (5′-3′)	Reverse primer (5′-3′)	Promoter
414-564-5M	VPS33B	pcDNA3.1+/C-(K)-DYK	GCCAACGGAACAGAAGGACCAAACTTCTACGTACCATTCAGCAACAAAACAGGCTAATCCGATTACAAGGATGACGAC	CTACATCATCATCATCATTGACTCACTGGAAGCCTTGTC	SP6
414-587-5M	VPS33B	pcDNA3.1+/C-(K)-DYK	GCCAACGGAACAGAAGGACCAAACTTCTACGTAC	CTACATCATCATCATCATCAGGA	SP6

			CATTCAGCAACA AAACAGGCTAAT CCGATTACAAGG ATGACGAC	AGCGCAGGGCT GATAT	
414-FL-5M	VPS33B	pcDNA3.1+/ C-(K)-DYK	GCCAACGGAAC AGAAGGACCAA ACTTCTACGTAC CATTCAGCAACA AAACAGGCTAAT CCGATTACAAGG ATGACGAC	CTACATCATCAT CATCATGGATTT CACCTCACTCAT GGCTTC	SP6
414-N603-FL-5M	VPS33B-N603	pcDNA3.1+/ C-(K)-DYK	GCCAACGGAAC AGAAGGACCAA ACTTCTACGTAC CATTCAGCAACA AAACAGGCTAAT CCGATTACAAGG ATGACGAC	CTACATCATCAT CATCATGGATTT CACCTCACTCAT GGCTTC	SP6
414-FL-OPG2-5M	VPS33B-OPG2	pcDNA3.1+/ C-(K)-DYK	GCCAACGGAAC AGAAGGACCAA ACTTCTACGTAC CATTCAGCAACA AAACAGGCTAAT CCGATTACAAGG ATGACGAC	CTACATCATCAT CATCATGCCTGT TTTGTTGCTGAA TGGTACGTAGA AGTTTGGTCCTT CTGTTCCGTT	SP6
1-564-5M	VPS33B	pcDNA3.1+/ C-(K)-DYK	CGCAAATGGGC GGTAGGCGTG	CTACATCATCAT CATCATTGACTC ACTGGAAGCCT TGTC	T7
1-587-5M	VPS33B	pcDNA3.1+/ C-(K)-DYK	CGCAAATGGGC GGTAGGCGTG	CTACATCATCAT CATCATCAGGA AGCGCAGGGCT GATAT	T7
1-FL-5M	VPS33B	pcDNA3.1+/ C-(K)-DYK	CGCAAATGGGC GGTAGGCGTG	CTACATCATCAT CATCATGGATTT CACCTCACTCAT GGCTTC	T7
1-N603-5M	VPS33B-N603	pcDNA3.1+/ C-(K)-DYK	CGCAAATGGGC GGTAGGCGTG	CTACATCATCAT CATCATGGATTT CACCTCACTCAT GGCTTC	T7
1-FL-OPG2-5M	VPS33B-OPG2	pcDNA3.1+/ C-(K)-DYK	CGCAAATGGGC GGTAGGCGTG	CTACATCATCAT CATCATGCCTGT TTTGTTGCTGAA TGGTACGTAGA AGTTTGGTCCTT CTGTTCCGTT	T7

822 Primer list for qPCR reactions.

Gene (ms – mouse, hu – human)	Forward primer (5′-3′)	Reverse primer (5′-3′)

msCol1a1	AGAGCATGACCGATGGATTC	AGGCCTCGGTGGACA
msItga11	AGATGTCGCAGACTGGCTTT	CCCTAGGTATGCTGCATGGT
msRplp0	ACTGGTCTAGGACCCGAGAAG	CTCCCACCTTGTCTCCAGTC
msGapdh	CAGCCTCGTCCCGTAGACAA	CAATCTCCACTTTGCCACTGC
msVPS33B	GCATTCACAGACACGGCTAAG	ACACCACCAAGATGAGGCG
huCOL1A1	GGGATTCCCTGGACCTAAAG	GGAACACCTCGCTCTCCA
huGAPDH	GAGTCAACGGATTTGGTCGT	GACAAGCTTCCCGTTCTCAG
huACTB	GATCATTGCTCCTCCTGAGC	AAAGCCATGCCAATCTCATC
huITGA11	CACGACATCAGTGGCAATAAG	GACCCTTCCCAGGTTGAGTT
huVPS33B	GAGCTGCCTGACTTCTCCAT	GCTTGTCTACTTCGTGTTGCTG

References

1. Smejkal GB, Fitzgerald C. Revised estimate of total collagen in the human body. International Journal of Proteomics & Bioinformatics. 2017;2:001-2.
2. Craig AS, Birtles MJ, Conway JF, Parry DA. An estimate of the mean length of collagen fibrils in rat tail-tendon as a function of age. Connect Tissue Res. 1989;19(1):51-62. PubMed PMID: 2477190.
3. Kadler KE, Hill A, Canty-Laird EG. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. Curr Opin Cell Biol. 2008;20(5):495-501. Epub 2008/07/22. doi: 10.1016/j.ceb.2008.06.008. PubMed PMID: 18640274; PubMed Central PMCID: PMCPMC2577133.
4. Musiime M, Chang J, Hansen U, Kadler KE, Zeltz C, Gullberg D. Collagen Assembly at the Cell Surface: Dogmas Revisited. Cells. 2021;10(3). Epub 2021/04/04. doi: 10.3390/cells10030662. PubMed PMID: 33809734; PubMed Central PMCID: PMCPMC8002325.
5. Zeltz C, Kusche-Gullberg M, Heljasvaara R, Gullberg D. Novel roles for cooperating collagen receptor families in fibrotic niches. Curr Opin Cell Biol. 2023;85:102273. Epub 2023/11/03. doi: 10.1016/j.ceb.2023.102273. PubMed PMID: 37918273.
6. Kadler KE, Holmes DF, Trotter JA, Chapman JA. Collagen fibril formation. Biochem J. 1996;316 (Pt 1):1-11. Epub 1996/05/15. doi: 10.1042/bj3160001. PubMed PMID: 8645190; PubMed Central PMCID: PMCPMC1217307.
7. Trelstad RL, Hayashi K. Tendon collagen fibrillogenesis: intracellular subassemblies and cell surface changes associated with fibril growth. Dev Biol. 1979;71(2):228-42. Epub 1979/08/01. doi: 10.1016/0012-1606(79)90166-0. PubMed PMID: 499658.
8. Birk DE, Trelstad RL. Extracellular compartments in tendon morphogenesis: collagen fibril, bundle, and macroaggregate formation. J Cell Biol. 1986;103(1):231-40. Epub 1986/07/01. doi: 10.1083/jcb.103.1.231. PubMed PMID: 3722266; PubMed Central PMCID: PMCPMC2113791.
9. Canty EG, Lu Y, Meadows RS, Shaw MK, Holmes DF, Kadler KE. Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon. J Cell Biol. 2004;165(4):553-63. Epub 2004/05/26. doi: 10.1083/jcb.200312071. PubMed PMID: 15159420; PubMed Central PMCID: PMCPMC2172348.
10. Chang J, Garva R, Pickard A, Yeung CC, Mallikarjun V, Swift J, et al. Circadian control of the secretory pathway maintains collagen homeostasis. Nat Cell Biol. 2020;22(1):74-86. Epub 2020/01/08. doi: 10.1038/s41556-019-0441-z. PubMed PMID: 31907414.
11. Hunter MR, Hesketh GG, Benedyk TH, Gingras AC, Graham SC. Proteomic and Biochemical Comparison of the Cellular Interaction Partners of Human VPS33A and VPS33B. J Mol Biol.

857 2018;430(14):2153-63. Epub 2018/05/21. doi: 10.1016/j.jmb.2018.05.019. PubMed PMID: 29778605;
858 PubMed Central PMCID: PMCPMC6005816.
859 12. Gissen P, Johnson CA, Morgan NV, Stapelbroek JM, Forshew T, Cooper WN, et al. Mutations
860 in VPS33B, encoding a regulator of SNARE-dependent membrane fusion, cause arthrogryposis-renal
861 dysfunction-cholestasis (ARC) syndrome. Nat Genet. 2004;36(4):400-4. Epub 2004/03/31. doi:
862 10.1038/ng1325. PubMed PMID: 15052268.
863 13. Nezelof C, Dupart MC, Jaubert F, Eliachar E. A lethal familial syndrome associating
864 arthrogryposis multiplex congenita, renal dysfunction, and a cholestatic and pigmentary liver disease.
865 J Pediatr. 1979;94(2):258-60. Epub 1979/02/01. doi: 10.1016/s0022-3476(79)80839-2. PubMed PMID:
866 762621.
867 14. Banushi B, Forneris F, Straatman-Iwanowska A, Strange A, Lyne AM, Rogerson C, et al.
868 Regulation of post-Golgi LH3 trafficking is essential for collagen homeostasis. Nat Commun.
869 2016;7:12111. Epub 2016/07/21. doi: 10.1038/ncomms12111. PubMed PMID: 27435297; PubMed
870 Central PMCID: PMCPMC4961739.
871 15. Salo AM, Wang C, Sipila L, Sormunen R, Vapola M, Kervinen P, et al. Lysyl hydroxylase 3 (LH3)
872 modifies proteins in the extracellular space, a novel mechanism for matrix remodeling. J Cell Physiol.
873 2006;207(3):644-53. Epub 2006/02/01. doi: 10.1002/jcp.20596. PubMed PMID: 16447251.
874 16. Madsen DH, Engelholm LH, Ingvarsen S, Hillig T, Wagenaar-Miller RA, Kjoller L, et al.
875 Extracellular collagenases and the endocytic receptor, urokinase plasminogen activator receptor-
876 associated protein/Endo180, cooperate in fibroblast-mediated collagen degradation. J Biol Chem.
877 2007;282(37):27037-45. Epub 2007/07/12. doi: 10.1074/jbc.M701088200. PubMed PMID: 17623673.
878 17. Madsen DH, Leonard D, Masedunskas A, Moyer A, Jurgensen HJ, Peters DE, et al. M2-like
879 macrophages are responsible for collagen degradation through a mannose receptor-mediated
880 pathway. J Cell Biol. 2013;202(6):951-66. Epub 2013/09/11. doi: 10.1083/jcb.201301081. PubMed
881 PMID: 24019537; PubMed Central PMCID: PMCPMC3776354.
882 18. Madsen DH, Jurgensen HJ, Siersbaek MS, Kuczek DE, Grey Cloud L, Liu S, et al. Tumor-
883 Associated Macrophages Derived from Circulating Inflammatory Monocytes Degrade Collagen
884 through Cellular Uptake. Cell Rep. 2017;21(13):3662-71. Epub 2017/12/28. doi:
885 10.1016/j.celrep.2017.12.011. PubMed PMID: 29281816; PubMed Central PMCID:
886 PMCPMC5753792.
887 19. Rainero E. Extracellular matrix endocytosis in controlling matrix turnover and beyond:
888 emerging roles in cancer. Biochem Soc Trans. 2016;44(5):1347-54. Epub 2016/12/03. doi:
889 10.1042/BST20160159. PubMed PMID: 27911717.
890 20. Lee TH, McKleroy W, Khalifeh-Soltani A, Sakuma S, Lazarev S, Riento K, et al. Functional
891 genomic screen identifies novel mediators of collagen uptake. Mol Biol Cell. 2014;25(5):583-93. Epub
892 2014/01/10. doi: 10.1091/mbc.E13-07-0382. PubMed PMID: 24403604; PubMed Central PMCID:
893 PMCPMC3937085.
894 21. Leitinger B. Discoidin domain receptor functions in physiological and pathological conditions.
895 Int Rev Cell Mol Biol. 2014;310:39-87. Epub 2014/04/15. doi: 10.1016/B978-0-12-800180-6.00002-5.
896 PubMed PMID: 24725424; PubMed Central PMCID: PMCPMC4021107.
897 22. Tiger CF, Fougerousse F, Grundstrom G, Velling T, Gullberg D. alpha11beta1 integrin is a
898 receptor for interstitial collagens involved in cell migration and collagen reorganization on
899 mesenchymal nonmuscle cells. Dev Biol. 2001;237(1):116-29. Epub 2001/08/24. doi:
900 10.1006/dbio.2001.0363. PubMed PMID: 11518510.
901 23. Zeltz C, Gullberg D. The integrin-collagen connection - a glue for tissue repair? J Cell Sci.
902 2016;129(6):1284. Epub 2016/07/22. doi: 10.1242/jcs.188672. PubMed PMID: 27442113.
903 24. Herrera J, Forster C, Pengo T, Montero A, Swift J, Schwartz MA, et al. Registration of the
904 extracellular matrix components constituting the fibroblastic focus in idiopathic pulmonary fibrosis. JCI
905 Insight. 2019;4(1). Epub 2019/01/11. doi: 10.1172/jci.insight.125185. PubMed PMID: 30626754;
906 PubMed Central PMCID: PMCPMC6485370.
907 25. Doyle AD. Fluorescent Labeling of Rat-tail Collagen for 3D Fluorescence Imaging. Bio Protoc.
908 2018;8(13). Epub 2018/08/17. doi: 10.21769/BioProtoc.2919. PubMed PMID: 30112424; PubMed
909 Central PMCID: PMCPMC6089535.

26. Kadler KE, Hojima Y, Prockop DJ. Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. Assay of critical concentration demonstrates that collagen self-assembly is a classical example of an entropy-driven process. J Biol Chem. 1987;262(32):15696-701. Epub 1987/11/15. PubMed PMID: 3316206.
27. Nazemi M, Yanes B, Martinez ML, Walker HJ, Pham K, Collins MO, et al. The extracellular matrix supports breast cancer cell growth under amino acid starvation by promoting tyrosine catabolism. PLoS Biol. 2024;22(1):e3002406. Epub 2024/01/16. doi: 10.1371/journal.pbio.3002406. PubMed PMID: 38227562; PubMed Central PMCID: PMCPMC10791009.
28. McCluskey A, Daniel JA, Hadzic G, Chau N, Clayton EL, Mariana A, et al. Building a better dynasore: the dyngo compounds potently inhibit dynamin and endocytosis. Traffic. 2013;14(12):1272-89. Epub 2013/09/13. doi: 10.1111/tra.12119. PubMed PMID: 24025110; PubMed Central PMCID: PMCPMC4138991.
29. Huang D, Sun G, Hao X, He X, Zheng Z, Chen C, et al. ANGPTL2-containing small extracellular vesicles from vascular endothelial cells accelerate leukemia progression. J Clin Invest. 2021;131(1). Epub 2020/10/28. doi: 10.1172/JCI138986. PubMed PMID: 33108353; PubMed Central PMCID: PMCPMC7773400.
30. Liang Z, Liu Z, Cheng C, Wang H, Deng X, Liu J, et al. VPS33B interacts with NESG1 to modulate EGFR/PI3K/AKT/c-Myc/P53/miR-133a-3p signaling and induce 5-fluorouracil sensitivity in nasopharyngeal carcinoma. Cell Death Dis. 2019;10(4):305. Epub 2019/04/05. doi: 10.1038/s41419-019-1457-9. PubMed PMID: 30944308; PubMed Central PMCID: PMCPMC6447525.
31. Wang C, Cheng Y, Zhang X, Li N, Zhang L, Wang S, et al. Vacuolar Protein Sorting 33B Is a Tumor Suppressor in Hepatocarcinogenesis. Hepatology. 2018;68(6):2239-53. Epub 2018/05/08. doi: 10.1002/hep.30077. PubMed PMID: 29729199.
32. Magliery TJ, Wilson CG, Pan W, Mishler D, Ghosh I, Hamilton AD, et al. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. J Am Chem Soc. 2005;127(1):146-57. Epub 2005/01/06. doi: 10.1021/ja046699g. PubMed PMID: 15631464.
33. Cabantous S, Terwilliger TC, Waldo GS. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. Nat Biotechnol. 2005;23(1):102-7. Epub 2004/12/08. doi: 10.1038/nbt1044. PubMed PMID: 15580262.
34. McCaughey J, Stevenson NL, Cross S, Stephens DJ. ER-to-Golgi trafficking of procollagen in the absence of large carriers. J Cell Biol. 2019;218(3):929-48. Epub 2018/12/28. doi: 10.1083/jcb.201806035. PubMed PMID: 30587510; PubMed Central PMCID: PMCPMC6400576.
35. Calverley BC, Kadler KE, Pickard A. Dynamic High-Sensitivity Quantitation of Procollagen-I by Endogenous CRISPR-Cas9 NanoLuciferase Tagging. Cells. 2020;9(9). Epub 2020/09/16. doi: 10.3390/cells9092070. PubMed PMID: 32927811; PubMed Central PMCID: PMCPMC7564849.
36. Hessa T, Meindl-Beinker NM, Bernsel A, Kim H, Sato Y, Lerch-Bader M, et al. Molecular code for transmembrane-helix recognition by the Sec61 translocon. Nature. 2007;450(7172):1026-30. Epub 2007/12/14. doi: 10.1038/nature06387. PubMed PMID: 18075582.
37. Rabu C, Schmid V, Schwappach B, High S. Biogenesis of tail-anchored proteins: the beginning for the end? J Cell Sci. 2009;122(Pt 20):3605-12. doi: 10.1242/jcs.041210. PubMed PMID: 19812306; PubMed Central PMCID: PMCPMC2758799.
38. O'Keefe S, High S, Demangel C. Biochemical and Biological Assays of Mycolactone-Mediated Inhibition of Sec61. Methods Mol Biol. 2022;2387:163-81. doi: 10.1007/978-1-0716-1779-3_16. PubMed PMID: 34643911.
39. O'Keefe S, Zong G, Duah KB, Andrews LE, Shi WQ, High S. An alternative pathway for membrane protein biogenesis at the endoplasmic reticulum. Commun Biol. 2021;4(1):828. Epub 20210701. doi: 10.1038/s42003-021-02363-z. PubMed PMID: 34211117; PubMed Central PMCID: PMCPMC8249459.
40. Zong G, Hu Z, O'Keefe S, Tranter D, Iannotti MJ, Baron L, et al. Ipomoeassin F Binds Sec61alpha to Inhibit Protein Translocation. J Am Chem Soc. 2019;141(21):8450-61. Epub 20190515. doi: 10.1021/jacs.8b13506. PubMed PMID: 31059257; PubMed Central PMCID: PMCPMC6627486.

962 41. Nilsson IM, von Heijne G. Determination of the distance between the oligosaccharyltransferase
963 active site and the endoplasmic reticulum membrane. J Biol Chem. 1993;268(8):5798-801. PubMed
964 PMID: 8449946.
965 42. Whitley P, Grau B, Gumbart JC, Martinez-Gil L, Mingarro I. Folding and Insertion of
966 Transmembrane Helices at the ER. Int J Mol Sci. 2021;22(23). Epub 20211126. doi:
967 10.3390/ijms222312778. PubMed PMID: 34884581; PubMed Central PMCID: PMCPMC8657811.
968 43. Meindl-Beinker NM, Lundin C, Nilsson I, White SH, von Heijne G. Asn- and Asp-mediated
969 interactions between transmembrane helices during translocon-mediated membrane protein
970 assembly. EMBO Rep. 2006;7(11):1111-6. Epub 20060929. doi: 10.1038/sj.embor.7400818. PubMed
971 PMID: 17008929; PubMed Central PMCID: PMCPMC1679787.
972 44. Tornieri K, Zlatic SA, Mullin AP, Werner E, Harrison R, L'Hernault S W, et al. Vps33b pathogenic
973 mutations preferentially affect VIPAS39/SPE-39-positive endosomes. Hum Mol Genet.
974 2013;22(25):5215-28. Epub 2013/08/07. doi: 10.1093/hmg/ddt378. PubMed PMID: 23918659;
975 PubMed Central PMCID: PMCPMC3842178.
976 45. Pickard A, Adamson A, Lu Y, Chang J, Garva R, Hodson N, et al. Collagen assembly and
977 turnover imaged with a CRISPR-Cas9 engineered Dendra2 tag. bioRxiv. 2018. doi:
978 https://doi.org/10.1101/331496
979 46. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists
980 using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57. Epub 2009/01/10. doi:
981 10.1038/nprot.2008.211. PubMed PMID: 19131956.
982 47. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the
983 comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13. Epub
984 2008/11/27. doi: 10.1093/nar/gkn923. PubMed PMID: 19033363; PubMed Central PMCID:
985 PMCPMC2615629.
986 48. Humphries JD, Byron A, Humphries MJ. Integrin ligands at a glance. J Cell Sci. 2006;119(Pt
987 19):3901-3. Epub 2006/09/22. doi: 10.1242/jcs.03098. PubMed PMID: 16988024; PubMed Central
988 PMCID: PMCPMC3380273.
989 49. Zeltz C, Gullberg D. The integrin-collagen connection - a glue for tissue repair? J Cell Sci.
990 2016;129(4):653-64. doi: 10.1242/jcs.188672. PubMed PMID: 27442113.
991 50. Zeltz C, Gullberg D. alpha11beta1: a mesenchymal collagen-binding integrins with a central
992 role in tissue and tumor fibrosis. In: Gullberg D, Eble JA, editors. Integrins in Health and Disease.
993 Biology of Extracellular Matrix. Biology of Extracellular Matrix: Springer; 2023. p. 235-54.
994 51. Wynn TA. Integrating mechanisms of pulmonary fibrosis. J Exp Med. 2011;208(7):1339-50.
995 doi: 10.1084/jem.20110551. PubMed PMID: 21727191; PubMed Central PMCID: PMCPMC3136685.
996 52. Knowles GC, McKeown M, Sodek J, McCulloch CA. Mechanism of collagen phagocytosis by
997 human gingival fibroblasts: importance of collagen structure in cell recognition and internalization. J
998 Cell Sci. 1991;98 (Pt 4):551-8. Epub 1991/04/01. PubMed PMID: 1650378.
999 53. Everts V, Aronson DC, Beertsen W. Phagocytosis of bone collagen by osteoclasts in two cases
1000 of pycnodysostosis. Calcif Tissue Int. 1985;37(1):25-31. Epub 1985/01/01. doi: 10.1007/BF02557674.
1001 PubMed PMID: 3922593.
1002 54. Arora PD, Wang Y, Bresnick A, Dawson J, Janmey PA, McCulloch CA. Collagen remodeling
1003 by phagocytosis is determined by collagen substrate topology and calcium-dependent interactions of
1004 gelsolin with nonmuscle myosin IIA in cell adhesions. Mol Biol Cell. 2013;24(6):734-47. Epub
1005 2013/01/18. doi: 10.1091/mbc.E12-10-0754. PubMed PMID: 23325791; PubMed Central PMCID:
1006 PMCPMC3596245.
1007 55. Vijayan RS, Arnold E, Das K. Molecular dynamics study of HIV-1 RT-DNA-nevirapine complexes
1008 explains NNRTI inhibition and resistance by connection mutations. Proteins. 2014;82(5):815-29. Epub
1009 2013/11/01. doi: 10.1002/prot.24460. PubMed PMID: 24174331; PubMed Central PMCID:
1010 PMCPMC4502926.
1011 56. Park RJ, Shen H, Liu L, Liu X, Ferguson SM, De Camilli P. Dynamin triple knockout cells reveal
1012 off target effects of commonly used dynamin inhibitors. J Cell Sci. 2013;126(Pt 22):5305-12. Epub
1013 2013/09/21. doi: 10.1242/jcs.138578. PubMed PMID: 24046449; PubMed Central PMCID:
1014 PMCPMC3828596.

1015 57. Eble JA, Golbik R, Mann K, Kuhn K. The alpha 1 beta 1 integrin recognition site of the
1016 basement membrane collagen molecule [alpha 1(IV)]2 alpha 2(IV). EMBO J. 1993;12(12):4795-802.
1017 Epub 1993/12/01. doi: 10.1002/j.1460-2075.1993.tb06168.x. PubMed PMID: 8223488; PubMed
1018 Central PMCID: PMCPMC413926.
1019 58. Herrera JA, Dingle L, Montero MA, Venkateswaran RV, Blaikley JF, Lawless C, et al. The UIP/IPF
1020 fibroblastic focus is a collagen biosynthesis factory embedded in a distinct extracellular matrix. JCI
1021 Insight. 2022;7(16). Epub 2022/07/20. doi: 10.1172/jci.insight.156115. PubMed PMID: 35852874;
1022 PubMed Central PMCID: PMCPMC9462507.
1023 59. Jones CE, Sharick JT, Colbert SE, Shukla VC, Zent JM, Ostrowski MC, et al. Pten regulates
1024 collagen fibrillogenesis by fibroblasts through SPARC. PLoS One. 2021;16(2):e0245653. Epub
1025 2021/02/04. doi: 10.1371/journal.pone.0245653. PubMed PMID: 33534863; PubMed Central PMCID:
1026 PMCPMC7857610.
1027 60. Wenstrup RJ, Smith SM, Florer JB, Zhang G, Beason DP, Seegmiller RE, et al. Regulation of
1028 collagen fibril nucleation and initial fibril assembly involves coordinate interactions with collagens V
1029 and XI in developing tendon. J Biol Chem. 2011;286(23):20455-65. Epub 2011/04/07. doi:
1030 10.1074/jbc.M111.223693. PubMed PMID: 21467034; PubMed Central PMCID: PMCPMC3121453.
1031 61. Taylor SH, Yeung CY, Kalson NS, Lu Y, Zigrino P, Starborg T, et al. Matrix metalloproteinase
1032 14 is required for fibrous tissue expansion. Elife. 2015;4:e09345. Epub 2015/09/22. doi:
1033 10.7554/eLife.09345. PubMed PMID: 26390284; PubMed Central PMCID: PMCPMC4684142.
1034 62. Moses M, Chang J, Hansen U, Kadler KE, Zeltz C, Gullberg D. Collagen assemble at the cell
1035 surface: Dogmas revisited. Cells. 2021;in press.
1036 63. Musiime M, Erusappan PM, Cukierman E, Chang J, Molven A, Hansen U, et al. Fibroblast
1037 integrin alpha11beta1 is a collagen assembly receptor in mechanoregulated fibrillar adhesions. Matrix
1038 Biol. 2024;under revision.
1039 64. Moriya K, Bae E, Honda K, Sakai K, Sakaguchi T, Tsujimoto I, et al. A fibronectin-independent
1040 mechanism of collagen fibrillogenesis in adult liver remodeling. Gastroenterology. 2011;140(5):1653-
1041 63. Epub 2011/02/16. doi: 10.1053/j.gastro.2011.02.005. PubMed PMID: 21320502; PubMed Central
1042 PMCID: PMCPMC3081910.
1043 65. Canty-Laird EG, Lu Y, Kadler KE. Stepwise proteolytic activation of type I procollagen to
1044 collagen within the secretory pathway of tendon fibroblasts in situ. Biochem J. 2012;441(2):707-17.
1045 Epub 2011/10/05. doi: 10.1042/BJ20111379. PubMed PMID: 21967573; PubMed Central PMCID:
1046 PMCPMC3430002.
1047 66. Stevenson NL, Bergen DJM, Lu Y, Prada-Sanchez ME, Kadler KE, Hammond CL, et al. Giantin
1048 is required for intracellular N-terminal processing of type I procollagen. J Cell Biol. 2021;220(6). Epub
1049 2021/05/05. doi: 10.1083/jcb.202005166. PubMed PMID: 33944912; PubMed Central PMCID:
1050 PMCPMC8103548.
1051 67. Liu RJY, Al-Molieh Y, Chen SZ, Drobac M, Urban D, Chen CH, et al. The Sec1-Munc18 protein
1052 VPS33B forms a uniquely bidirectional complex with VPS16B. J Biol Chem. 2023;299(6):104718. Epub
1053 2023/04/17. doi: 10.1016/j.jbc.2023.104718. PubMed PMID: 37062417; PubMed Central PMCID:
1054 PMCPMC10208892.
1055 68. Lynch DA, Sverzellati N, Travis WD, Brown KK, Colby TV, Galvin JR, et al. Diagnostic criteria
1056 for idiopathic pulmonary fibrosis: a Fleischner Society White Paper. Lancet Respir Med. 2018;6(2):138-
1057 53. Epub 2017/11/21. doi: 10.1016/S2213-2600(17)30433-2. PubMed PMID: 29154106.
1058 69. Herrera JA, Mallikarjun V, Rosini S, Montero MA, Lawless C, Warwood S, et al. Laser capture
1059 microdissection coupled mass spectrometry (LCM-MS) for spatially resolved analysis of formalin-fixed
1060 and stained human lung tissues. Clin Proteomics. 2020;17:24. Epub 2020/06/23. doi: 10.1186/s12014-
1061 020-09287-6. PubMed PMID: 32565759; PubMed Central PMCID: PMCPMC7302139.
1062 70. Herrera J, Beisang DJ, Peterson M, Forster C, Gilbertsen A, Benyumov A, et al. Dicer1
1063 Deficiency in the Idiopathic Pulmonary Fibrosis Fibroblastic Focus Promotes Fibrosis by Suppressing
1064 MicroRNA Biogenesis. Am J Respir Crit Care Med. 2018;198(4):486-96. Epub 2018/03/27. doi:
1065 10.1164/rccm.201709-1823OC. PubMed PMID: 29579397; PubMed Central PMCID:
1066 PMCPMC6118029.

1067 71. Kyaw BM, Jarbrink K, Martinengo L, Car J, Harding K, Schmidtchen A. Need for Improved
1068 Definition of "Chronic Wounds" in Clinical Studies. Acta Derm Venereol. 2018;98(1):157-8. doi:
1069 10.2340/00015555-2786. PubMed PMID: 28880344.
1070 72. McKenna M, Simmonds RE, High S. Mechanistic insights into the inhibition of Sec61-
1071 dependent co- and post-translational translocation by mycolactone. J Cell Sci. 2016;129(7):1404-15.
1072 Epub 20160211. doi: 10.1242/jcs.182352. PubMed PMID: 26869228; PubMed Central PMCID:
1073 PMCPMC4852723.
1074 73. Campeau E, Ruhl VE, Rodier F, Smith CL, Rahmberg BL, Fuss JO, et al. A versatile viral system
1075 for expression and depletion of proteins in mammalian cells. PLoS One. 2009;4(8):e6529. Epub
1076 2009/08/07. doi: 10.1371/journal.pone.0006529. PubMed PMID: 19657394; PubMed Central PMCID:
1077 PMCPMC2717805.
1078 74. Popova SN, Rodriguez-Sanchez B, Liden A, Betsholtz C, Van Den Bos T, Gullberg D. The
1079 mesenchymal alpha11beta1 integrin attenuates PDGF-BB-stimulated chemotaxis of embryonic
1080 fibroblasts on collagens. Dev Biol. 2004;270(2):427-42. Epub 2004/06/09. doi:
1081 10.1016/j.ydbio.2004.03.006. PubMed PMID: 15183724.
1082 75. O'Keefe S, Roboti P, Duah KB, Zong G, Schneider H, Shi WQ, et al. Ipomoeassin-F inhibits the
1083 in vitro biogenesis of the SARS-CoV-2 spike protein and its host cell membrane receptor. J Cell Sci.
1084 2021;134(4). Epub 20210219. doi: 10.1242/jcs.257758. PubMed PMID: 33468620; PubMed Central
1085 PMCID: PMCPMC7904091.
1086 76. Smeland HY, Askeland C, Wik E, Knutsvik G, Molven A, Edelmann RJ, et al. Integrin
1087 alpha11beta1 is expressed in breast cancer stroma and associates with aggressive tumor phenotypes.
1088 J Pathol Clin Res. 2020;6(1):69-82. Epub 2019/10/13. doi: 10.1002/cjp2.148. PubMed PMID:
1089 31605508; PubMed Central PMCID: PMCPMC6966706.
1090 77. Wu G, Anafi RC, Hughes ME, Kornacker K, Hogenesch JB. MetaCycle: an integrated R
1091 package to evaluate periodicity in large scale data. Bioinformatics. 2016;32(21):3351-3. Epub
1092 2016/10/30. doi: 10.1093/bioinformatics/btw405. PubMed PMID: 27378304; PubMed Central PMCID:
1093 PMCPMC5079475.
1094 78. Ihaka R, Gentleman R. R: A Language for Data Analysis and Graphics. Journal of
1095 Computational and Graphical Statistics. 1996;5(3):299-314. doi: 10.1080/10618600.1996.10474713.
1096

Figure 1

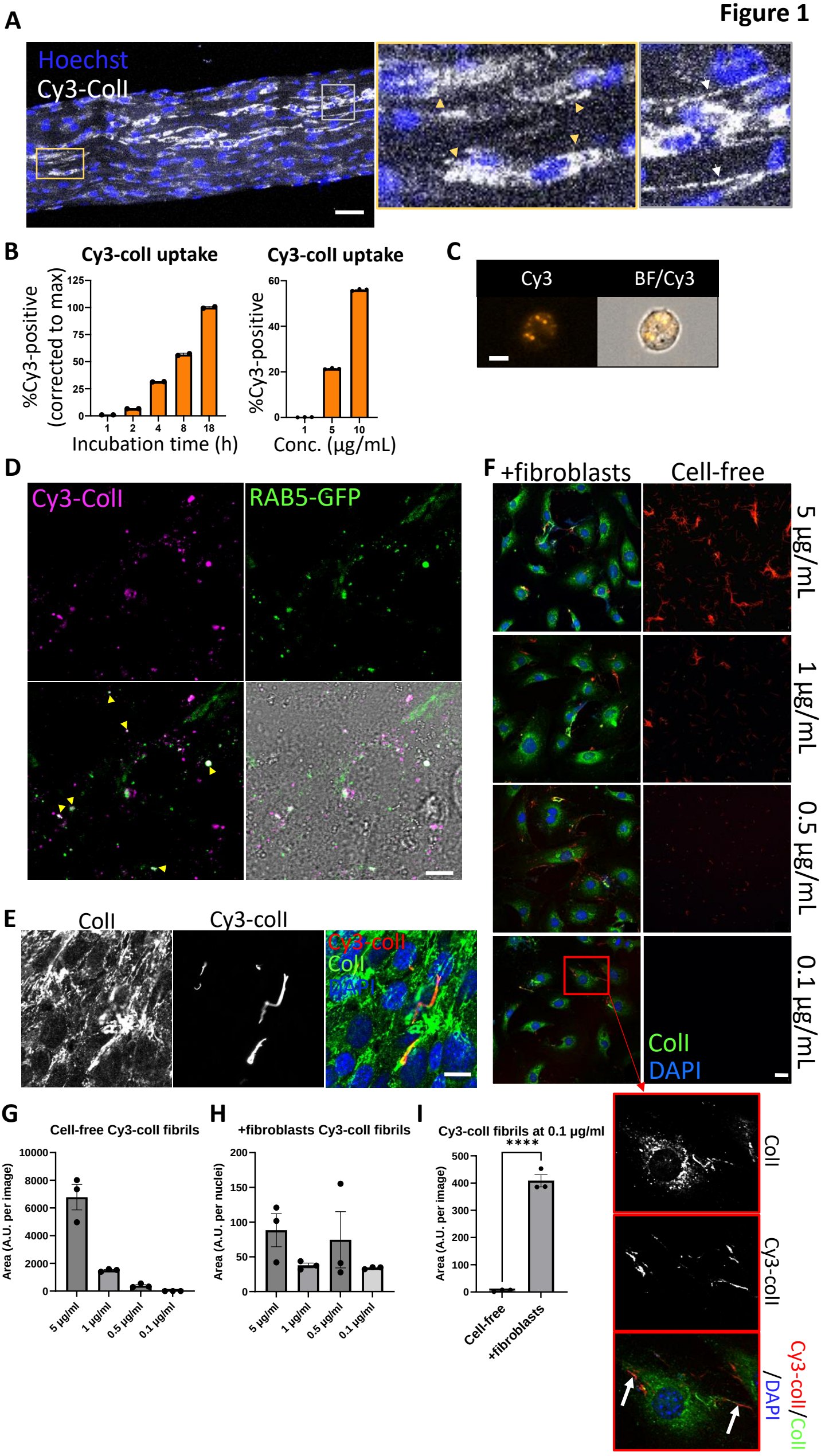

A
Hoechst
Cy3-Coll
B
Cy3-coll uptake
%Cy3-positive (corrected to max)
Incubation time (h)
Cy3-coll uptake
%Cy3-positive
Conc. (µg/mL)
C
Cy3
BF/Cy3
D
Cy3-Coll
RAB5-GFP
E
Coll
Cy3-coll
Cy3-coll
Coll
DAPI
F
+fibroblasts
Cell-free
5 µg/mL
1 µg/mL
0.5 µg/mL
0.1 µg/mL
Coll
DAPI
G
Cell-free Cy3-coll fibrils
Area (A.U. per image)
5 µg/ml
1 µg/ml
0.5 µg/ml
0.1 µg/ml
H
+fibroblasts Cy3-coll fibrils
Area (A.U. per nuclei)
5 µg/ml
1 µg/ml
0.5 µg/ml
0.1 µg/ml
I
Cy3-coll fibrils at 0.1 µg/ml
****
Area (A.U. per image)
Cell-free
+fibroblasts
Coll
Cy3-coll
Cy3-coll/Coll /DAPI

Figure 1: Collagen-I is endocytosed and reassembled into fibrils

A. Fluorescent images of tail tendon incubated with Cy3-colI for 5 days, showing presence of collagen-I within the cells, and fibril-like fluorescence signals outside of cells. Hoechst stain was used to locate cells within the tendon. Area surrounded by yellow box expanded on the right, and cells with Cy3-colI present intracellularly pointed out by yellow triangles. Area surrounded by grey box expanded on the right, and fibril-like fluorescence signals indicated with white arrows. Scale bar = 50 µm. Representative of N = 3.

B. Bar chart showing an increase of percentage of fluorescent iTTFs incubated with 1.5 µg/mL Cy3-colI over time (left), and an increase of percentage of fluorescent iTTFs incubated with increasing concentration of Cy3-colI for one hour (right), suggesting a non-linear time-dependent and dose-dependent uptake pattern. N = 3.

C. Flow cytometry imaging of iTTFs incubated with 5 µg/mL Cy3-labeled collagen-I for one hour, showing that collagen-I is taken up by cells and held in vesicular-like structures. Images acquired using ImageStream at 40x magnification. Scale bar = 10 μm. Cy3 – Cy3 channel, BF/Cy3 – merged image of BF and Cy3. Representative of >500 cells images collected per condition.

D. Fluorescent images of iTTFs transduced with Rab5-GFP and incubated with Cy3-labeled collagen-I. Yellow arrows point to labelled collagen co-localizing with Rab5 in intracellular structures. Representative of N = 3. Scale bar = 10 µm.

E. Fluorescent images of iTTFs incubated with 5 µg/mL Cy3-colI for one hour, trypsinized and replated in fresh media, and further cultured for 72 h. Top labels denotes the fluorescence channel corresponding to proteins detected. Merged image colour channels as denoted on top left. Representative of N>3. Scale bar = 20 µm.

F. Fluorescent image series of Cy3-colI incubated at different concentrations for 72 h, either cell-free (right panel), or with iTTFs (+fibroblasts, left panel). Representative of N = 3. Scale bar = 20 µm. Red box - zoomed out to the bottom left and separated according to fluorescence channel. White arrows highlighting Cy3-positive fibrils assembled by fibroblasts when incubated with 0.1 µg/mL Cy3-colI.

G. Quantification of the area of Cy3-positive fibrils in cell-free cultures, quantified per image area. N=3.

H. Quantification of the area of Cy3-positive fibrils in +fibroblasts cultures, corrected to number of nuclei per image area. N=3.

I. Comparison of total area of Cy3-positive fibrils in cell-free and +fibroblast cultures at 0.1 µg/mL concentration, as quantified per image area. N=3. *****p <0.0001.*

Figure 2

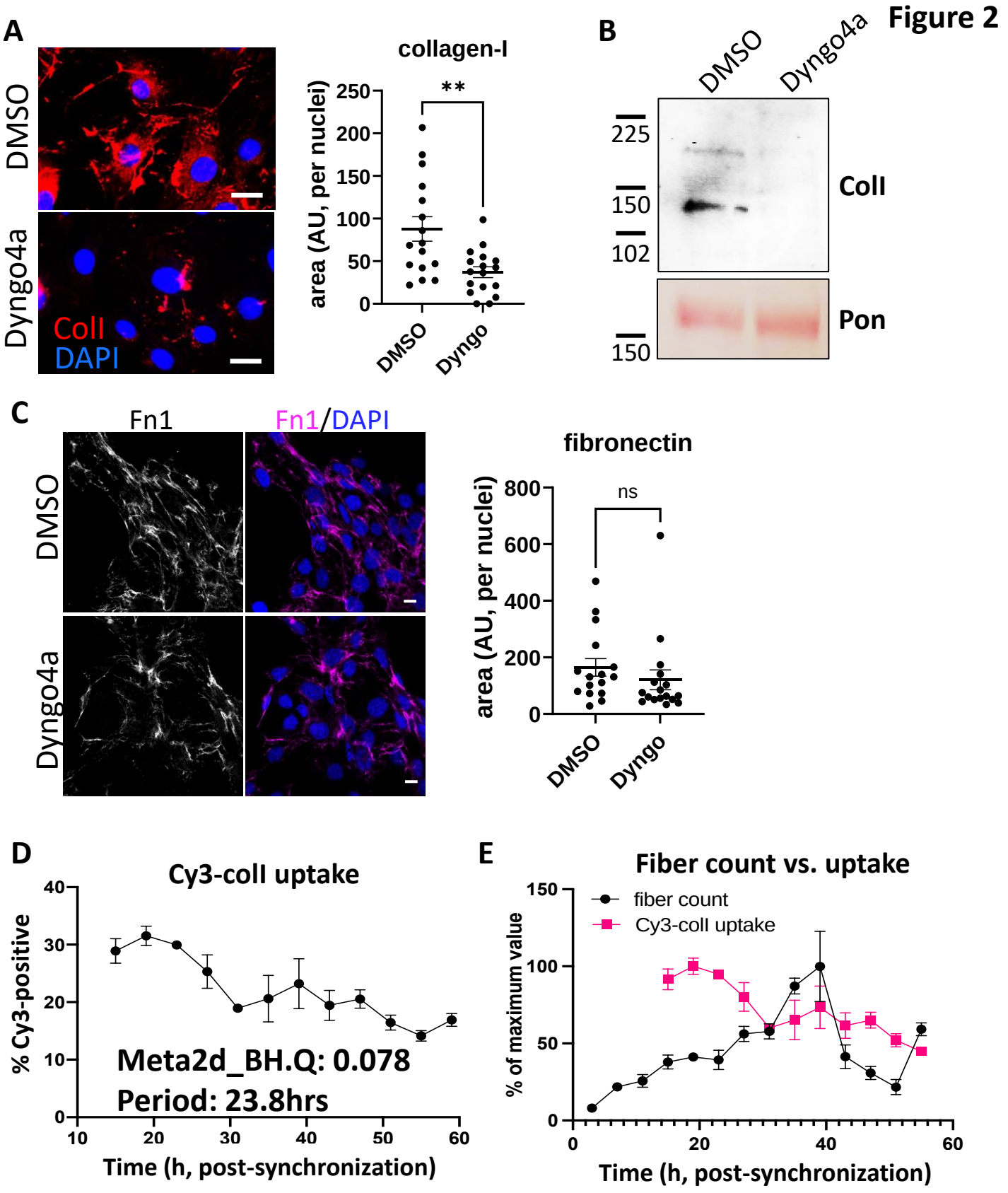

A
DMSO
Dyngo4a
ColI
DAPI
collagen-I
area (AU, per nuclei)
250
200
150
100
50
0
**
DMSO
Dyngo
B
DMSO
Dyngo4a
225
150
102
ColI
150
Pon
C
Fn1
Fn1/DAPI
DMSO
Dyngo4a
fibronectin
area (AU, per nuclei)
800
600
400
200
0
ns
DMSO
Dyngo
D
Cy3-ColI uptake
% Cy3-positive
40
30
20
10
0
Meta2d_BH.Q: 0.078
Period: 23.8hrs
10
20
30
40
50
60
Time (h, post-synchronization)
E
Fiber count vs. uptake
fiber count
Cy3-ColI uptake
% of maximum value
150
100
50
0
0
20
40
60
Time (h, post-synchronization)

Figure 2: Inhibition of endocytosis leads to changes in collagen-I homeostasis, and endocytosis is a rhythmic event

A. Left: fluorescent images of collagen-I (red) counterstained with DAPI (blue) in iTTFs treated with DMSO (top) or Dyng4a (bottom) for 72 h. Scale bar = 20 µm. Right: quantification of area occupied by collagen-I fibrils, corrected to number of nuclei. N = 3 with 5 images from each experiment ** p = 0.0025.

B. Western blot analysis of conditioned media taken from iTTFs treated with DMSO or Dyng4a for 72 h, showing a decrease in collagen-I secretion. Top: probed with collagen-I antibody (Col-I), bottom: counterstained with Ponceau (Pon) as control. Protein molecular weight ladders to the left (in kDa). Representative of N = 3.

C. Left: fluorescent images of fibronectin (magenta) counterstained with DAPI (blue) in iTTFs treated with DMSO (top) or Dyng4a (bottom) for 72 h. Scale bar = 20 µm. Right: quantification of area occupied by fibronectin fibrils, corrected to number of nuclei. N = 3 with 5 images from each experiment.

D. Percentage Cy3-colI taken up by synchronized iTTFs over 48 h. Meta2d analysis indicates a circadian rhythm of periodicity of 23.8 h. Bars show mean ± s.e.m. of N = 3 per time point.

E. Percentage of Cy3-colI taken up by synchronized iTTFs, corrected to the maximum percentage uptake of the time course (pink, bars show mean ± s.e.m. of N=3 per time point), compared to the percentage collagen fibril count over time, corrected to the maximum percentage fibril count of the time course (black, fibrils scored by two independent investigators. Bars show mean± s.e.m. of N=2 with n = 6 repeats at each time point).

Figure 3

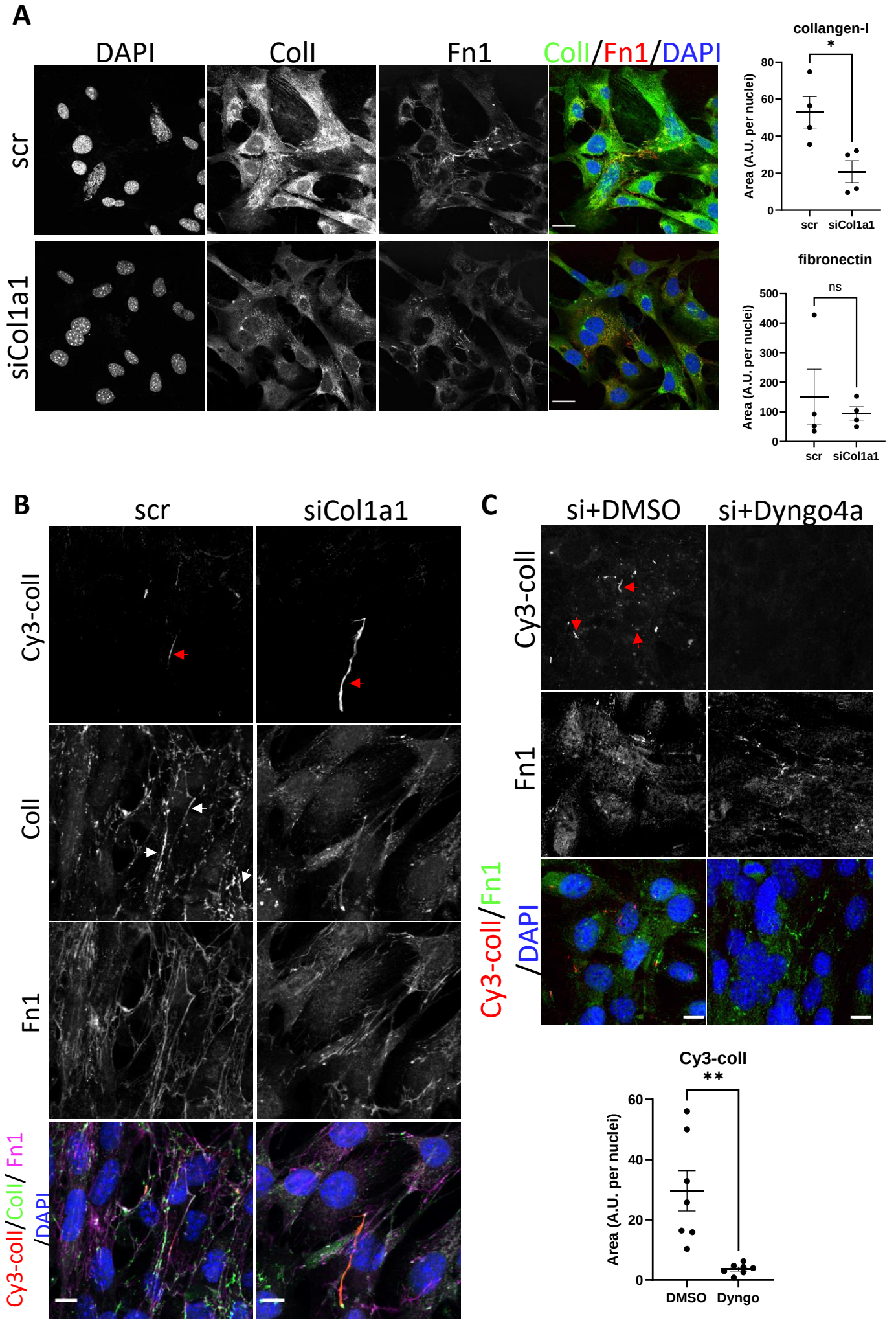

A
DAPI
Coll
Fn1
Coll/Fn1/DAPI
scr
siCol1a1
collangen-I
*
Area (A.U. per nuclei)
80
60
40
20
0
scr
siCol1a1
fibronectin
ns
500
400
300
200
100
0
B
scr
siCol1a1
Cy3-coll
Coll
Fn1
Cy3-coll/Coll/ Fn1
/DAPI
C
si+DMSO
si+Dyngo4a
Cy3-coll
Fn1
Cy3-coll/Fn1
/DAPI
Cy3-coll
**
60
40
20
0
DMSO
Dyngo

Figure 3: Collagen-I recycling can generate fibrils

A. Fluorescent image series of iTTFs treated with scrambled control (top panel, scr), and siRNA against col1a1 (bottom panel, siCol1a1). Labels on top denotes the fluorescence channel corresponding to proteins detected (ColI – collagen-I Fn1 – fibronectin); . Quantification of collagen-I and fibronectin signal to the right. Representative of N = 4. Scale bar = 25 µm. * $p = 0.021$.

B. Fluorescent image series of scr (left column) and siCol1a1 (right column) iTTFs incubated with Cy3-colI. Labels on left denotes the fluorescence channel(s) corresponding to proteins detected (ColI – collagen-I Fn1 – fibronectin). Cy3-colI fibrils highlighted by red arrows, and collagen-I fibrils highlighted by white arrows. Both scr cells and siCol1a1 cells can take up exogenous collagen-I and recycle to form collagen-I fibril. Representative of N>3. Scale bar = 10 μm.

C. Fluorescent image series of siCol1a1 iTTFs treated with DMSO control (left) or Dyngo4a (right) during Cy3-colI uptake, followed by further culture for 72 h. Labels on left denotes the fluorescence channel corresponding to proteins detected (ColI – collagen-I Fn1 – fibronectin). Quantification of Cy3-colI signal to the bottom. Dyngo4a treatment led to a reduction of Cy3-colI fibrils. Representative of N>3. Scale bar = 20 µm. ** $p = 0.0022$

Figure 4

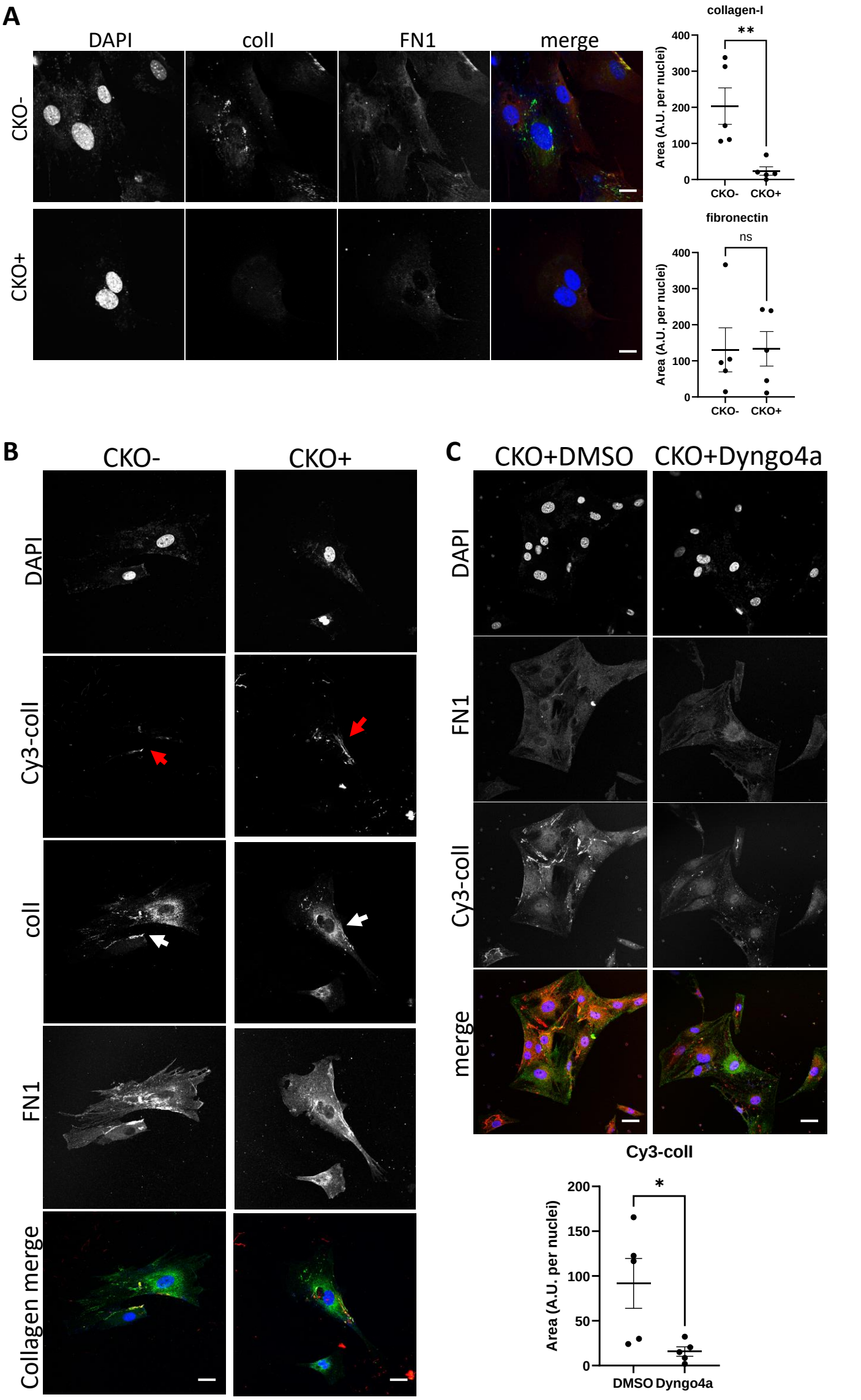

A
DAPI
coll
FN1
merge
CKO-
CKO+
collagen-I
**
Area (A.U. per nuclei)
400
300
200
100
0
CKO-
CKO+
fibronectin
ns
Area (A.U. per nuclei)
400
300
200
100
0
CKO-
CKO+
B
CKO-
CKO+
DAPI
Cy3-coll
coll
FN1
Collagen merge
C
CKO+DMSO
CKO+Dyngo4a
DAPI
FN1
Cy3-coll
merge
Cy3-coll
*
Area (A.U. per nuclei)
200
150
100
50
0
DMSO
Dyngo4a

Figure 4: Fibroblasts without endogenous collagen-I can effectively make fibrils by endocytic-recycling of exogenous collagen.

A. Fluorescent images of primary tail tendon fibroblasts isolated from control mice (top panel, CKO-), and tamoxifen-treated collagen-knockout mice (bottom panel, CKO+). Labels on top denotes the fluorescence channel corresponding to proteins detected. Quantification of collagen-I and fibronectin fluorescence signal to the right. Representative of N=3. Scale bar = 10μm. ** p = 0.0084.
B. Fluorescent images of CKO-/CKO+ tail tendon fibroblasts incubated with Cy3-colI. Labels on top denotes the fluorescence channel corresponding to proteins detected. Cy3-colI fibril highlighted by red arrows, and collagen-I fibril highlighted by white arrows. Both CKO- and CKO+ cells can take up exogenous collagen-I and recycle to form collagen-I fibrils. Representative of N>3. Scale bar = 25μm.
C. Fluorescent image series of CKO+ tail tendon fibroblasts treated with DMSO control (left) or Dyngo4a (right) during Cy3-colI uptake, followed by further culture for 72 h. Labels on left denotes the fluorescence channel corresponding to proteins detected. Quantification of Cy3-colI signal to the bottom. Dyngo4a treatment led to a significant reduction of Cy3-colI fibrils. Representative of N>3. Scale bar = 25 µm. * p = 0.00273.

Figure 5

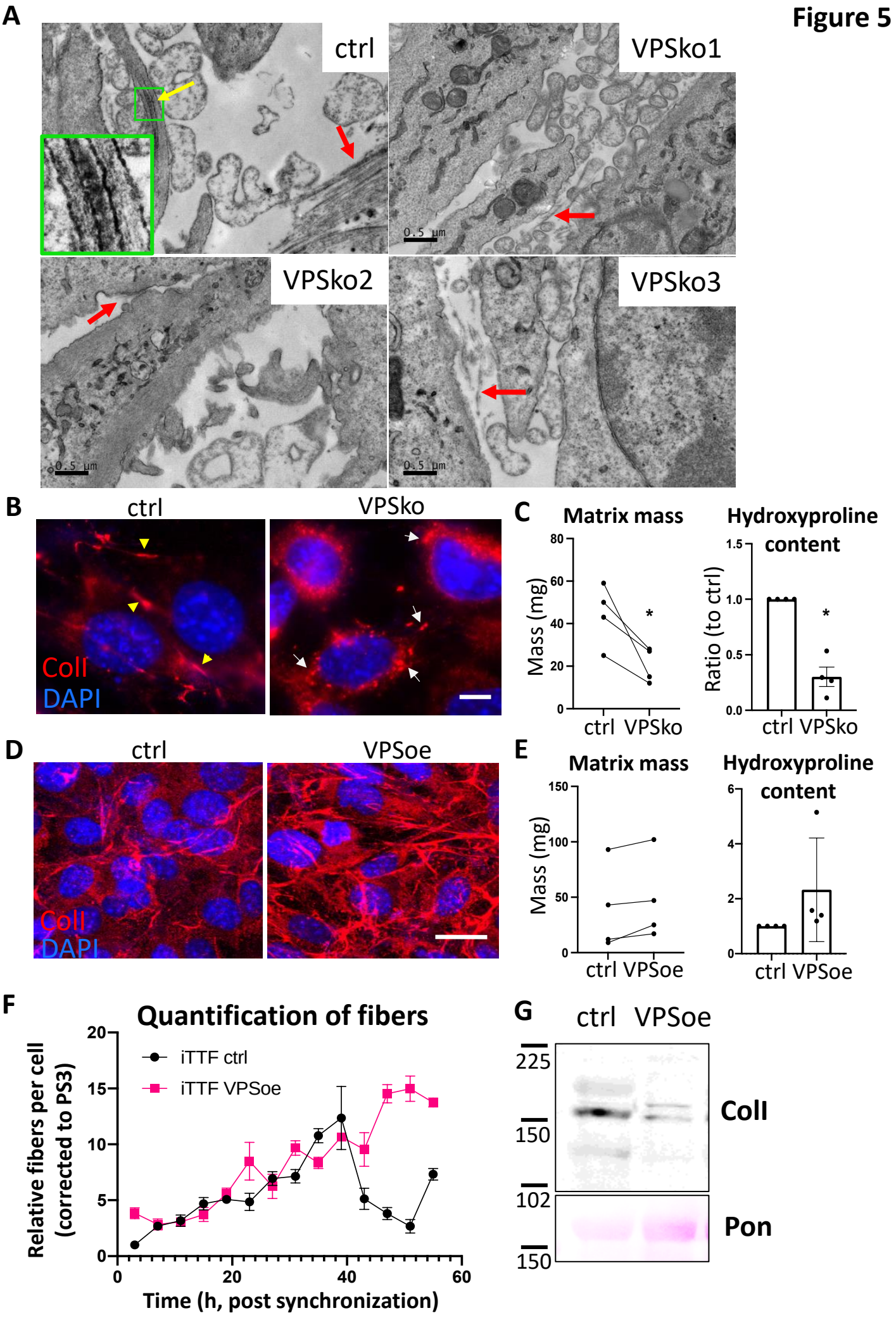

A
ctrl
VPSko1
VPSko2
VPSko3
0.5 µm
B
ctrl
VPSko
Coll
DAPI
C
Matrix mass
Mass (mg)
ctrl VPSko
Hydroxyproline content
Ratio (to ctrl)
ctrl VPSko
D
ctrl
VPSoe
Coll
DAPI
E
Matrix mass
Mass (mg)
ctrl VPSoe
Hydroxyproline content
ctrl VPSoe
F
Quantification of fibers
iTTF ctrl
iTTF VPSoe
Relative fibers per cell (corrected to PS3)
Time (h, post synchronization)
G
ctrl VPSoe
225
150
102
150
Coll
Pon

Figure 5: VPS33B controls collagen fibril formation at the plasma membrane in a rhythmic manner

A. Electron microcopy images of fibroblasts plated on Aclar and grown for a week before fixation and imaging. Ctrl culture has numerous collagen-I fibrils, as pointed out by arrows. Yellow arrow points to a fibripositor, and green box is expanded to the left bottom corner, showing the distinct D-banding pattern of collagen-I fibril when observed with electron microscopy. VPSko clones all have fewer and thinner fibrils present in the culture (pointed out by red arrows). Representative of N = 3. Scale bar = 0.5 µm.

B. Fluorescence images of collagen-I (red) and DAPI counterstain in ctrl and VPSko iTTFs. Yellow arrows indicating collagen fibrils, and white arrows pointing to collagen-I presence in intracellular vesicles. Representative of N>6. Scale bar = 25 µm.

C. Matrix deposition by ctrl or VPSko iTTFs, after one week of culture. Left: decellularized matrix mass. N = 4, * p = 0.0299. Right: hydroxyproline content presented as a ratio between ctrl and VPSko cells. N = 4, *p = 0.0254. Ratio-paired t-test used.

D. Fluorescence images of collagen-I (red) and DAPI counterstain in ctrl and VPSoe iTTFs. Representative of N>6. Scale bar = 20 µm.

E. Matrix deposition by ctrl or VPSoe iTTFs, after one week of culture. Left: decellularized matrix mass, N = 4. Right: hydroxyproline content presented as a ratio between ctrl and VPSoe cells, N = 4. Ratio-paired t-test used.

F. Relative collagen fibril count in synchronized ctrl (black) and VPSoe (pink) iTTFs, corrected to the number of fibrils in ctrl cultures at start of time course. Fibrils scored by two independent investigators. Bars show mean± s.e.m. of N = 2 with n=6 at each time point.

G. Western blot analysis of conditioned media taken from ctrl and VPSoe iTTFs after 72 h in culture. Top: probed with collage-I antibody (ColI), bottom: counterstained with Ponceau (Pon) as control. Protein molecular weight ladders to the left (in kDa). Representative of N = 3.

Figure 6

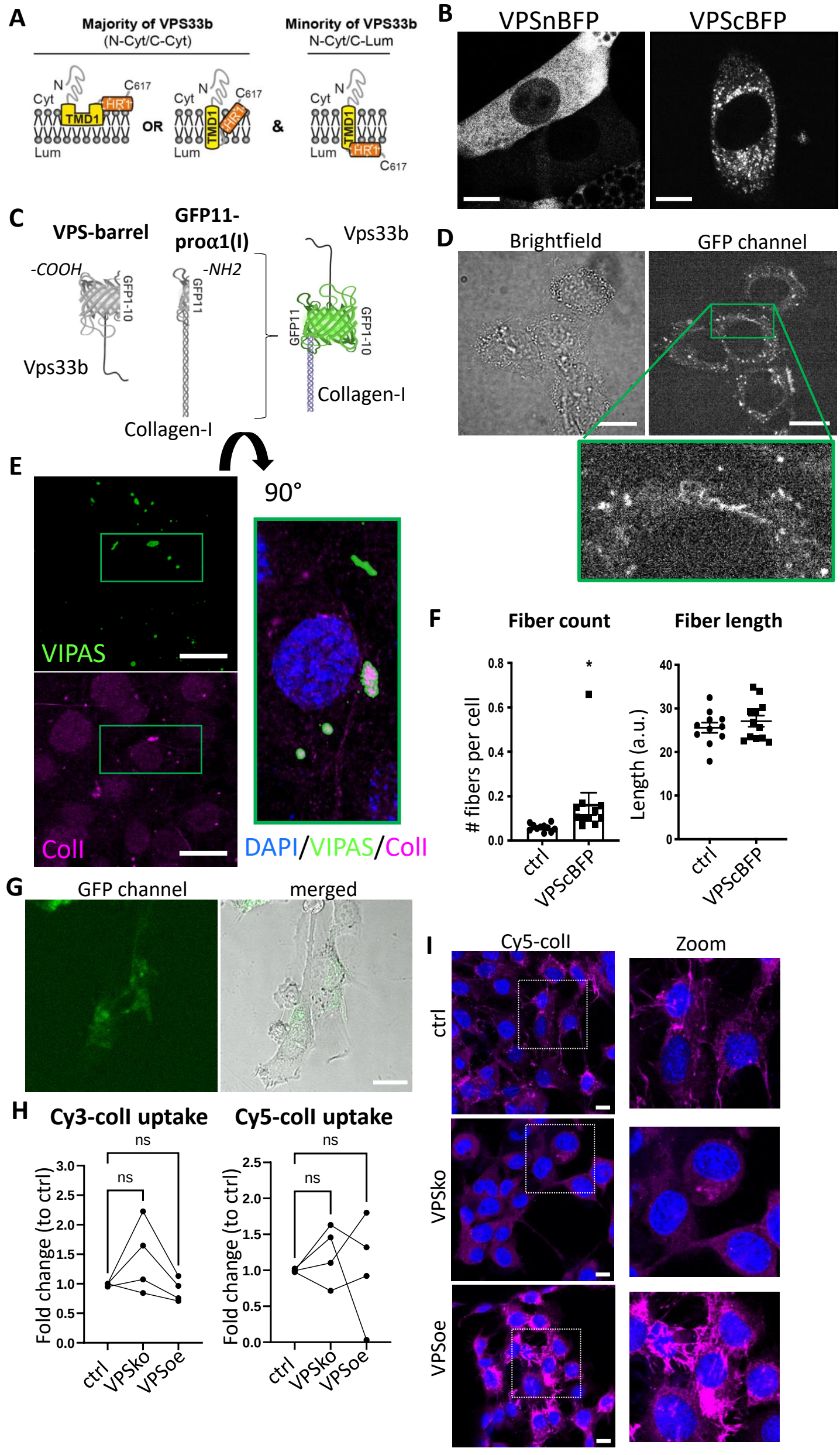

A
Majority of VPS33b
(N-Cyt/C-Cyt)
Minority of VPS33b
N-Cyt/C-Lum
N
C617
Cyt
HR1
TMD1
Lum
OR
&
B
VPSnBFP
VPScBFP
C
VPS-barrel
GFP11-
proα1(I)
Vps33b
-COOH
-NH2
GFP1-10
GFP11
Vps33b
Collagen-I
Collagen-I
90°
D
Brightfield
GFP channel
E
VIPAS
Coll
DAPI/VIPAS/Coll
F
Fiber count
Fiber length
fibers per cell
Length (a.u.)
ctrl
VPScBFP
G
GFP channel
merged
H
Cy3-coll uptake
Cy5-coll uptake
ns
Fold change (to ctrl)
ctrl
VPSko
VPSoe
I
Cy5-coll
Zoom
ctrl
VPSko
VPSoe

Figure 6: Procollagen-I and VPS33B localize to the same compartments

A. Schematic depicting the proposed membrane topologies of VPS33b.

B. iTTFs expressing BFP-tagged VPS33B. Left: BFP tagged on the N-terminal end of VPS33B (VPSnBFP), Right: BFP tagged on the C-terminal end (VPScBFP). Images taken in Airy mode. Representative of N>4. Scale bar = 10 µm.

C. Schematic of the split-GFP system. GFP1-10 barrel is introduced into VPS33B (VPS-barrel), and GFP11 to alpha-1 chain of Collagen-I (GFP11-proα1(I)). If the two tagged proteins co-localize (e.g. in a vesicle), a GFP signal will be emitted.

D. Brightfield (top) and fluorescence (middle) images of iTTFs expressing both VPS-barrel and GFP11-proα1(I) constructs. Representative of N=5. Green box is expanded to the bottom, to highlight the punctate fluorescence signals within intracellular vesicular structures, as well as fibril-like structures suggestive of fibril assembly sites. Scale bar = 20 µm.

E. Fluorescence images of VIPAS (green), collagen-I (red) and DAPI counterstain in iTTFs. Representative of N=3. Green box is expanded to the right (flipped 90°) to show strong VIPAS signal encasing collagen-I. Scale bar = 25 µm.

F. Quantification of average number of fibrils per cell (left) and average fibril length (right) in control endogenously-tagged Dendra-colI expressing 3T3 cells (ctrl) and Dendra-colI expressing 3T3 overexpressing VPScBFP (VPScBFP). >500 cells quantified per condition. N=12. *P=0.048.

G. Brightfield (left) and fluorescence (middle) images of iTTFs expressing VPS-barrel incubated with conditioned media containing Col1a1-GFP11 for 24hr. Scale bar = 25μm.

H. Line charts comparing the percentage of iTTFs that have taken up 5 µg/mL Cy3-colI (left) and Cy5-colI (right) after one hour incubation between control (ctrl), VPS33B-knockout (VPSko) and VPS33B-overexpressing (VPSoe) cells, corrected to control. RM one-way ANOVA analysis was performed. N = 4.

I. Fluorescence images of iTTFs of different levels of VPS33B expression, fed with Cy5-colI and further cultured for 72 h. Cultures were counterstained with DAPI. Box expanded to right of images to show zoomed-in images of the fibrils produced by the fibroblasts. Representative of N = 2.

Figure 7

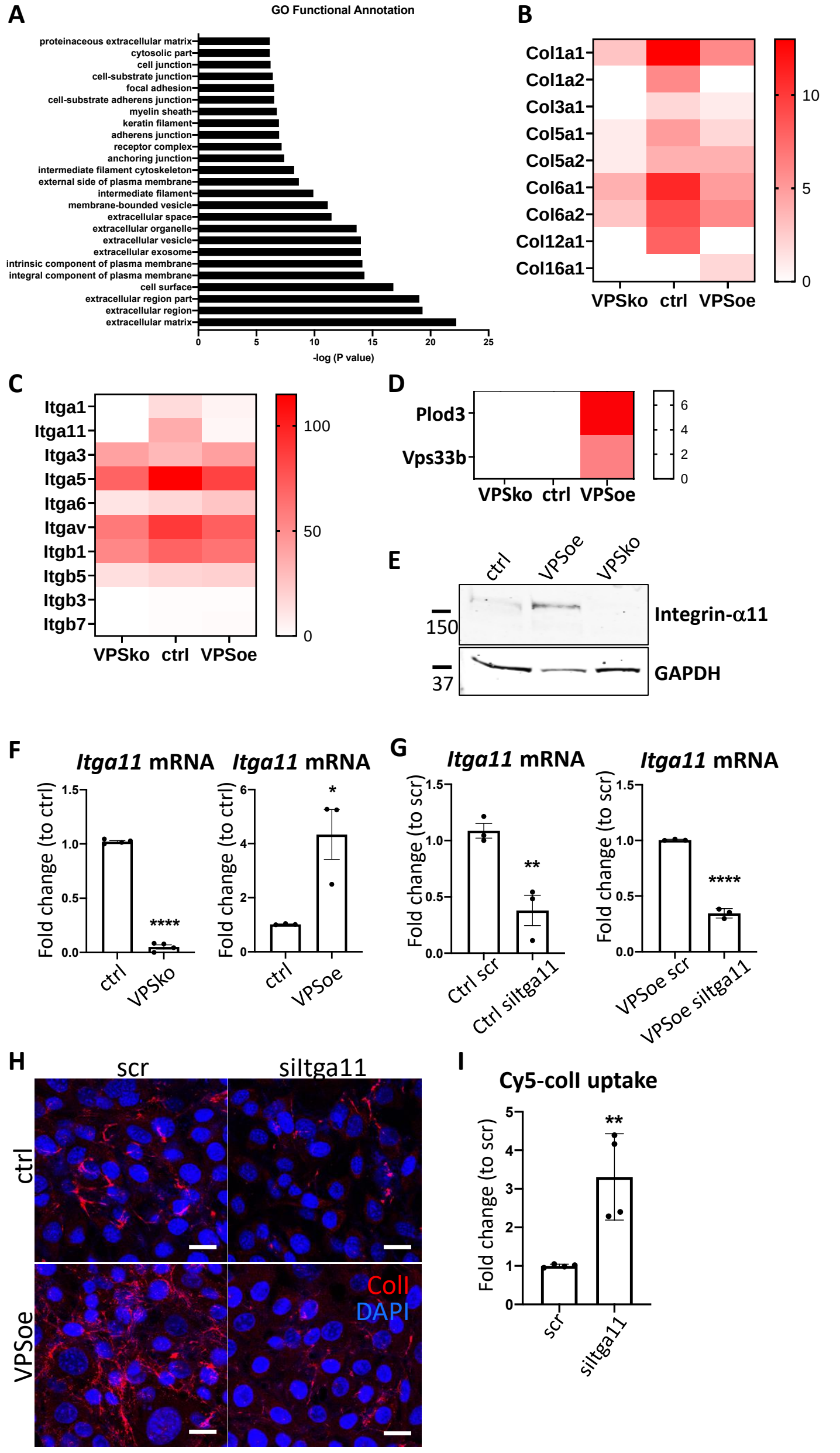

A
GO Functional Annotation
proteinaceous extracellular matrix
cytosolic part
cell junction
cell-substrate junction
focal adhesion
cell-substrate adherens junction
myelin sheath
keratin filament
adherens junction
receptor complex
anchoring junction
intermediate filament cytoskeleton
external side of plasma membrane
intermediate filament
membrane-bounded vesicle
extracellular space
extracellular organelle
extracellular vesicle
extracellular exosome
intrinsic component of plasma membrane
integral component of plasma membrane
cell surface
extracellular region part
extracellular region
extracellular matrix
0
5
10
15
20
25
-log (P value)
B
Col1a1
Col1a2
Col3a1
Col5a1
Col5a2
Col6a1
Col6a2
Col12a1
Col16a1
VPSko
ctrl
VPSoe
10
5
0
C
Itga1
Itga11
Itga3
Itga5
Itga6
Itgav
Itgb1
Itgb5
Itgb3
Itgb7
VPSko
ctrl
VPSoe
100
50
0
D
Plod3
Vps33b
VPSko
ctrl
VPSoe
6
4
2
0
E
ctrl
VPSoe
VPSko
150
37
Integrin-α11
GAPDH
F
Itga11 mRNA
Fold change (to ctrl)
1.5
1.0
0.5
0.0
****
ctrl
VPSko
Itga11 mRNA
Fold change (to ctrl)
6
4
2
0
*
ctrl
VPSoe
G
Itga11 mRNA
Fold change (to scr)
1.5
1.0
0.5
0.0
**
Ctrl scr
Ctrl siltga11
Itga11 mRNA
Fold change (to scr)
1.5
1.0
0.5
0.0
****
VPSoe scr
VPSoe siltga11
H
scr
siltga11
ctrl
VPSoe
Coll
DAPI
I
Cy5-coll uptake
Fold change (to scr)
5
4
3
2
1
0
**
scr
siltga11

Figure 7: Integrin α11 subunit mediates VPS33B-effects and is required for collagen-I fibrillogenesis

A. Top 25 Functional Annotation of proteins detected in biotin-enriched samples when compared to non-enriched samples based on p-values. Y-axis denotes the GO term, X-axis denotes –log (P value).

B. Heatmap representation of spectral counting of collagens detected in biotin-enriched surface proteins from control (ctrl), VPS33B-knockout (VPSko), and VPS33B-overexpressing (VPSoe) iTTFs. Scale denotes quantitative value as normalized to total spectra, as determined by Proteome Discoverer.

C. Heatmap representation of spectral counting of integrins detected in biotin-enriched surface proteins from control (ctrl), VPS33B-knockout (VPSko), and VPS33B-overexpressing (VPSoe) iTTFs. Scale denotes quantitative value as normalized to total spectra, as determined by Proteome Discoverer.

D. Heatmap representation of spectral counting of Plod3 and VPS33B detected in biotin-enriched surface proteins from control (ctrl), VPS33B-knockout (VPSko), and VPS33B-overexpressing (VPSoe) iTTFs. Scale denotes quantitative value as normalized to total spectra, as determined by Proteome Discoverer.

E. Western blot analysis of integrin α11 subunit levels in control (ctrl), VPS33B-overexpressing (VPSoe), VPS33B-knockout (VPSko) iTTFs. Top: probed with integrin α11 antibody, bottom: reprobed with GAPDH antibody. Protein molecular weight ladders to the left (in kDa). Representative of N = 3.

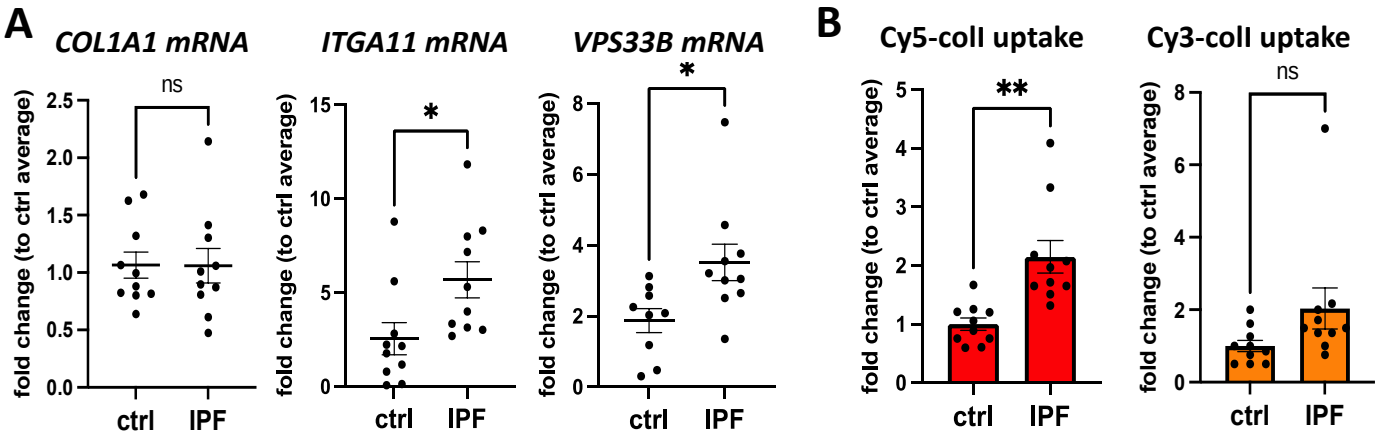
F. qPCR analysis of *Itga11* transcript levels in ctrl compared to VPSko iTTFs (left), and ctrl compared to VPSoe iTTFs (right). N>3, ****$P<0.0001$, *$P = 0.0226$.

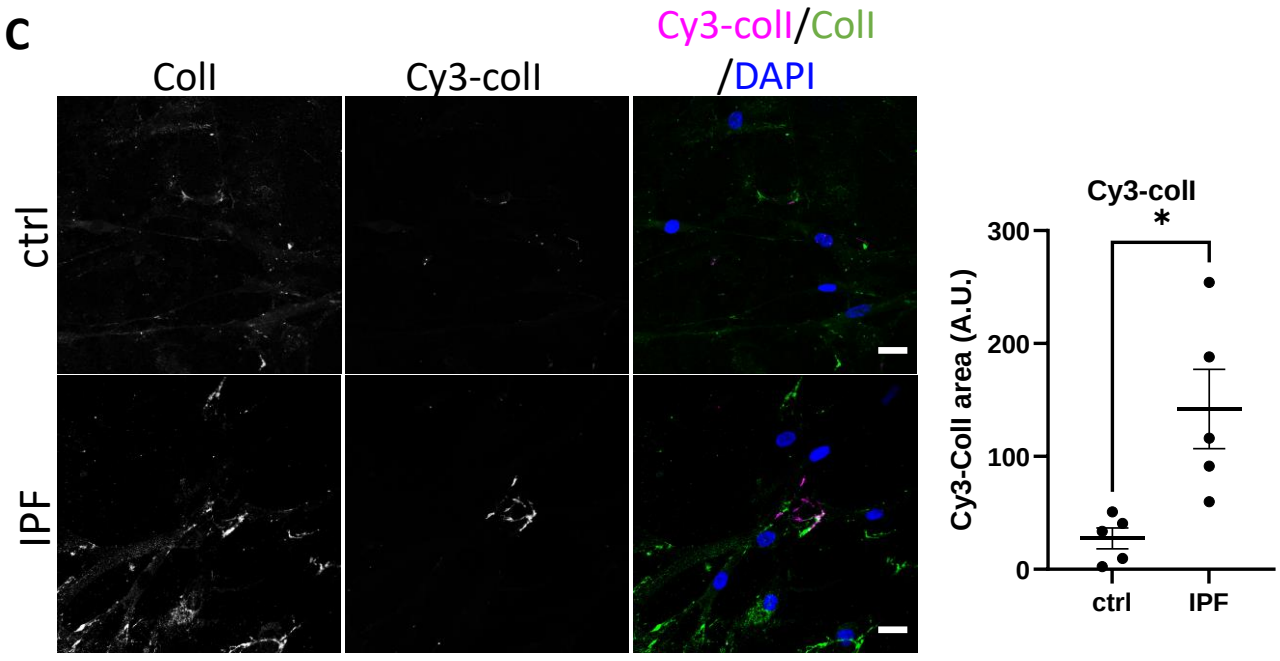
G. qPCR analysis of *Itga11* mRNA expression in ctrl (left) or VPSoe (right) iTTFs treated either with scrambled control (scr) or siRNA against Itga11 (siItga11), collected after 96 h. N=3, **$P=0.0091$, ****$P<0.0001$.

H. IF images of ctrl and VPSoe iTTFs treated either with control siRNA (scr) or siRNA again Itga11 (siItga11), after 72 h incubation; collagen-I (red) and DAPI (blue) counterstained. Representative of N = 3. Scale bar = 25 µm.

I. Bar chart comparing the percentage of iTTFs that have taken up 5 µg/mL Cy5-collI after one hour incubation between fibroblasts treated with scrambled control (ctrl) or siRNA against Itga11 (siItga11), corrected to scr. N=3. **$P=0.0062$.

Figure 8


A
COL1A1 mRNA
ITGA11 mRNA
VPS33B mRNA
ns
*
*
fold change (to ctrl average)
ctrl
IPF
B
Cy5-coll uptake
Cy3-coll uptake
**
ns
fold change (to ctrl average)
ctrl
IPF


C
Coll
Cy3-coll
Cy3-coll/Coll /DAPI
ctrl
IPF
Cy3-coll
*
Cy3-Coll area (A.U.)
ctrl
IPF

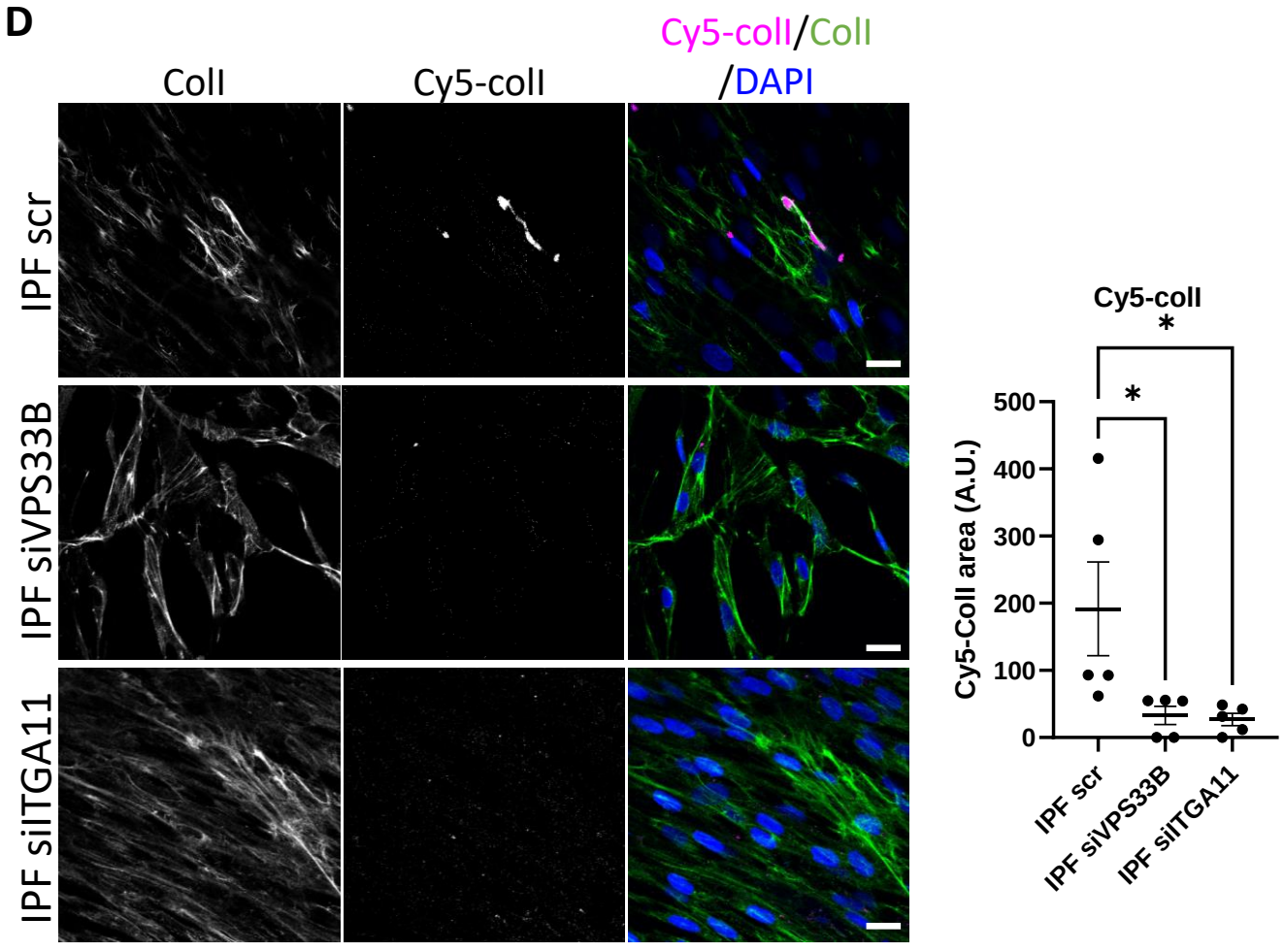

D
Coll
Cy5-coll
Cy5-coll/Coll /DAPI
IPF scr
IPF siVPS33B
IPF siITGA11
Cy5-coll
*
*
Cy5-Coll area (A.U.)
IPF scr
IPF siVPS33B
IPF siITGA11

Figure 8: Fibroblasts derived from IPF patients have higher collagen endocytic-recycling capacity that is mediated by VPS33B and ITGA11.

A. qPCR analysis of patient-derived fibroblasts isolated from control (ctrl) or IPF lungs. Bars showing mean± s.e.m, 5 patients in each group from 2 independent experiments (technical repeats not shown here). *Itga11, *p* = 0.0259; *VPS33B*, *p* = 0.0183.

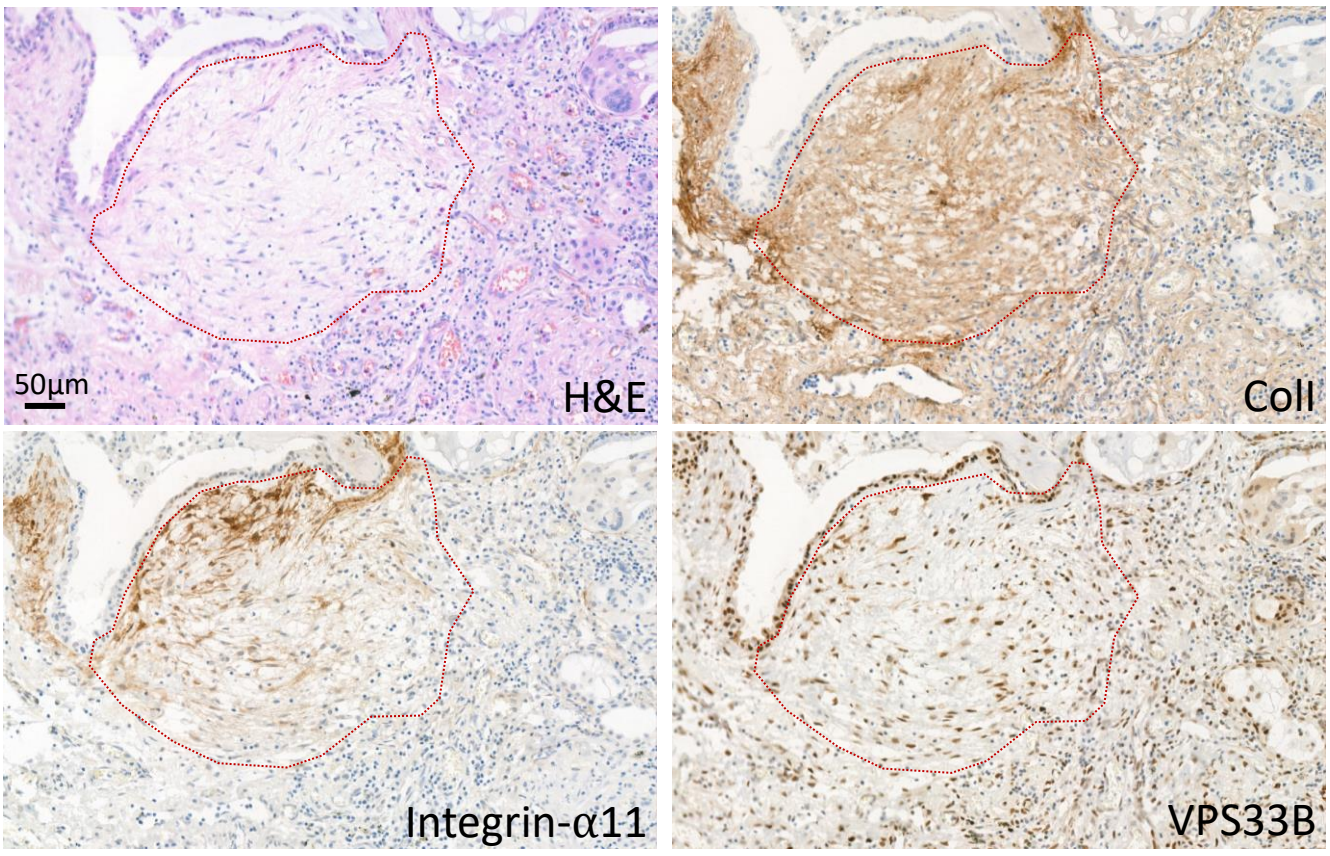
B. Fold change of percentage Cy5-colI (left) or Cy3-colI (right) taken up by ctrl or IPF lung fibroblasts, corrected to average of control fibroblasts. Bars showing mean± s.e.m., 5 patients in each group from 2 independent experiments (technical repeats not shown here). ***p* = 0.003.

C. Fluorescent images of ctrl or IPF lung fibroblasts that have taken up Cy3-colI (magenta), followed by further culture for 48 h in presence of ascorbic acid, before subjected to collagen-I staining (green). Labels on top denotes the fluorescence channel corresponding to proteins detected. Quantification of Cy3-colI signal to the right. IPF fibroblasts produced more Cy3-labelled fibrillar structures. Representative of N = 5. Scale bar = 20 µm. * *p* = 0.0135.

D. Fluorescent images of IPF lung fibroblasts treated with siRNA scrambled control (scr), siRNA against VPS33B (siVPS33B) or siRNA against ITGA11 (siITGA11) prior to uptake of Cy5-colI (magenta). This was followed by further culture for 48 h in presence of ascorbic acid, before subjected to collagen-I staining (green). Labels on top denotes the fluorescence channel corresponding to proteins detected. Quantification of Cy5-colI signal to the right. Both siVPS33B and siITGA11 significantly reduced recycled collagen signals. Representative of n = 5 across N = 2. Ordinary one-way ANOVA with multiple comparisions (to scr) was performed on quantification of Cy5-colI signal. siVPS33B, * *p* = 0.0341; siITGA11, * *p* = 0.0282.

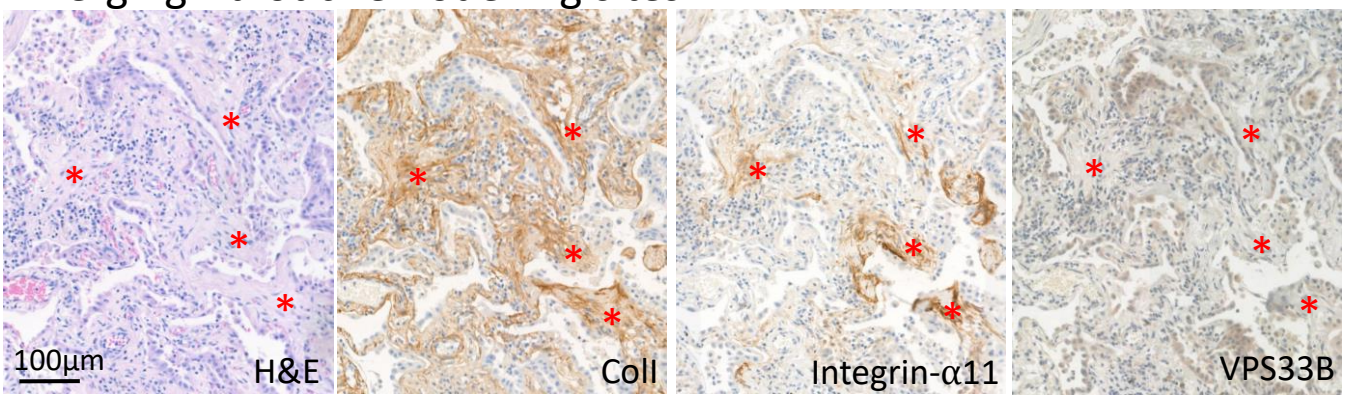
Figure 9

A



B

Emerging fibrotic remodelling sites



C

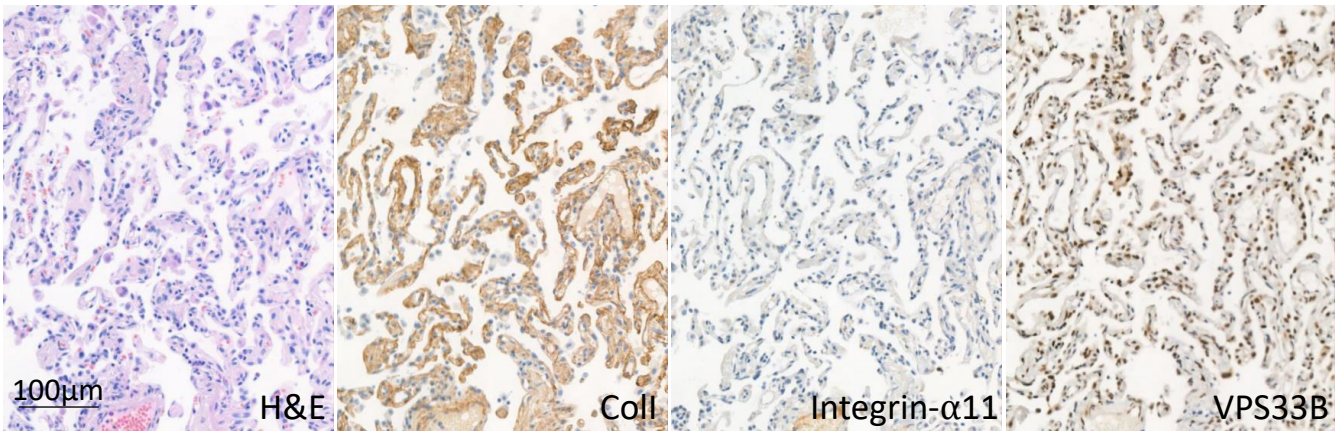


Figure 8: The IPF fibrotic focus is positive for integrin α11 subunit and VPS33B.

A. Immunohistochemistry of IPF patient (patient 1) with red dotted line outlining the fibroblastic focus, the hallmark lesion of IPF. Sections were stained with hematoxylin and eosin (H&E), collagen-I (ColI), integrin-α11, VPS33B. Scale bar = 50 µm.

B. Immunohistochemistry of IPF patient 4 showing regions of emerging fibrotic remodeling with evidence of fibroblastic foci formation (red asterisks). Sections were stained with H&E, ColI, integrin-α11, VPS33B. Scale bar = 100 µm.

C. Immunohistochemistry of 5 µm thick sequential lung sections taken from lungs classified as control (Control 1). Sections were stained with hematoxylin and eosin (H&E), collagen-I (ColI), integrin-α11, VPS33B. Scale bar = 100 µm.

Figure 10

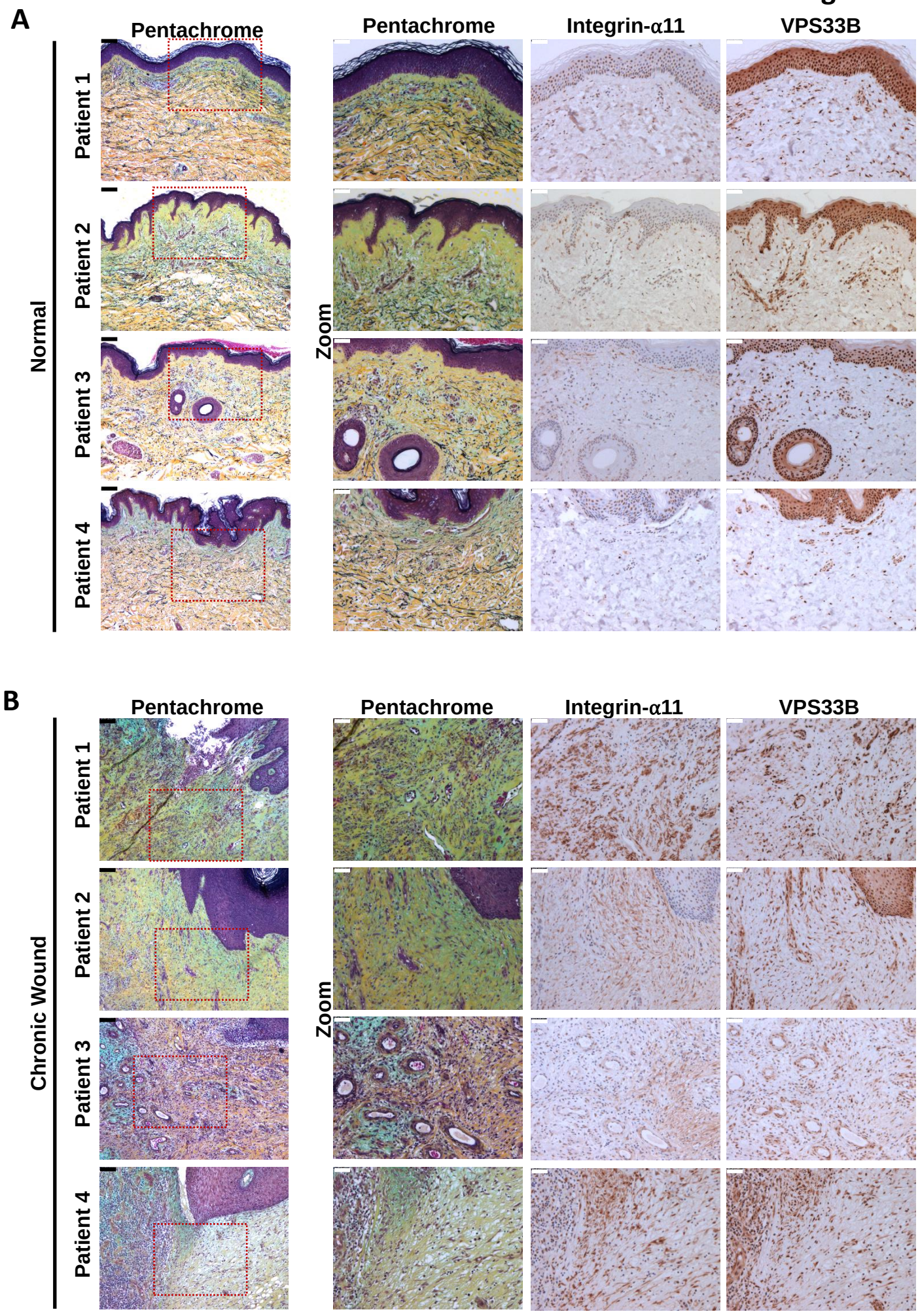

A
Pentachrome
Pentachrome
Integrin-α11
VPS33B
Normal
Patient 1
Patient 2
Patient 3
Patient 4
Zoom
B
Pentachrome
Pentachrome
Integrin-α11
VPS33B
Chronic Wound
Patient 1
Patient 2
Patient 3
Patient 4
Zoom

Figure 10: Proteins responsible for collagen fibrillogenesis are also co-localized to diseased areas of chronic skin wounds.

A. Immunohistochemistry of 5 µm thick sequential skin sections taken from normal skin regions of patients with chronic skin wounds (Patient 1, Patient 2, Patient 3, Patient 4). Sections were stained with pentachrome, integrin-α11, VPS33B. Scale bars positioned in top left corner: black (unzoomed pentachrome) = 100 µm, white (zoomed sections) = 50 µm.

B. Immunohistochemistry of 5 µm thick sequential skin sections taken from the chronic wound areas from patients with chronic skin wounds (Patient 1, Patient 2, Patient 3, Patient 4). Sections were stained with pentachrome, integrin-α11, VPS33B. Scale bars positioned in top left corner: black (unzoomed pentachrome) = 100 µm, white (zoomed sections) = 50 µm.

Figure 11

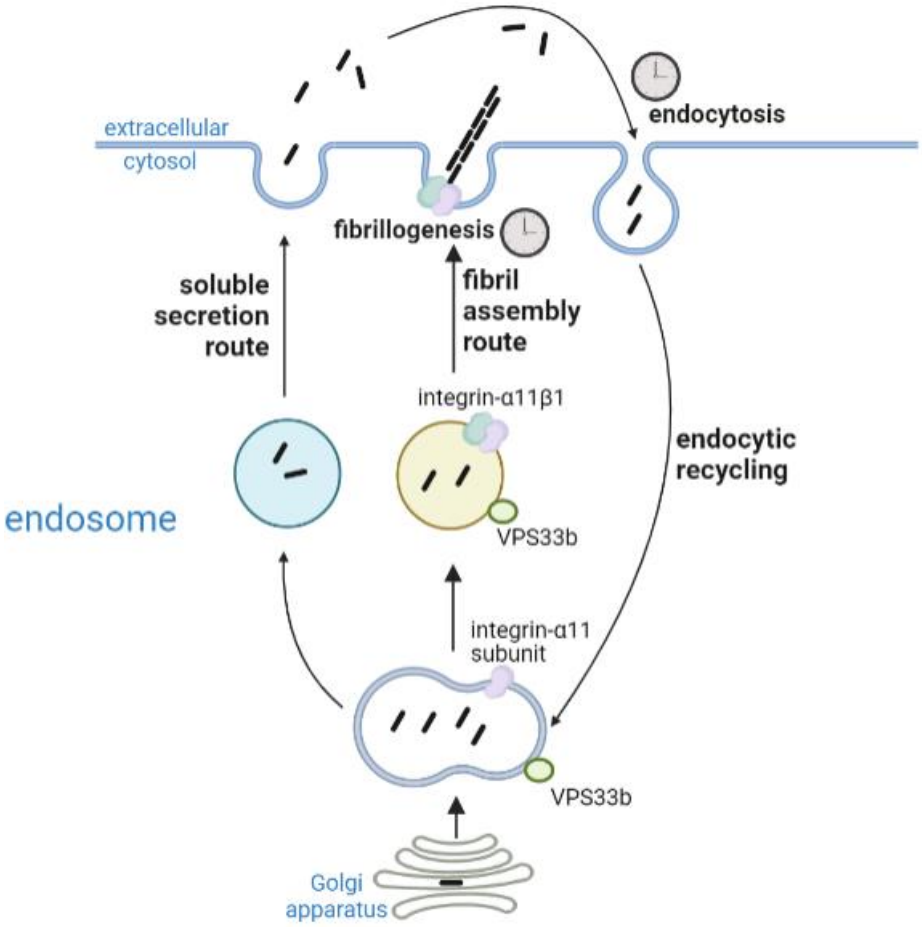

extracellular
cytosol
endocytosis
fibrillogenesis
soluble
secretion
route
fibril
assembly
route
integrin-α11β1
endocytic
recycling
endosome
VPS33b
integrin-α11
subunit
VPS33b
Golgi
apparatus

Figure 11: Proposed working model of collagen homeostasis in fibroblasts.

Endogenous collagen is either secreted as protomers (soluble secretion route, not circadian rhythmic) or made into fibrils (fibril assembly route, circadian rhythmic). Secreted collagen protomers can be captured by cells through endocytosis (circadian rhythmic) and recycled to make new fibrils. Integrin α11 and VPS33b directs collagen to fibril formation.