Genome-wide screens identify SEL1L as an intracellular rheostat controlling collagen turnover

Michael J. Podolsky[1,*], Benjamin Kheyfets[1], Monika Pandey[1], Afaq H. Beigh[1], Christopher D. Yang[2], Carlos Lizama Valenzuela[2], Ritwik Datta[2], Paul J. Wolters[3], Michael T. McManus[4], Ling Qi[5,6], Kamran Atabai[2,3,7,*]

*Co-corresponding author

1- Department of Medicine, Weill Cornell Medical College, New York, NY, USA.
2- Cardiovascular Research Institute, University of California, San Francisco, CA, USA.
3- Department of Medicine, University of California, San Francisco, CA, USA.
4- Department of Microbiology and Immunology and UCSF Diabetes Center, University of California, San Francisco, CA, USA.
5- Department of Internal Medicine, Division of Metabolism, Endocrinology and Diabetes, University of Michigan Medical School, Ann Arbor, MI, USA.
6- Department of Molecular and Integrative Physiology, University of Michigan Medical School, Ann Arbor, MI, USA.
7- Lung Biology Center, University of California, San Francisco, CA, USA.

The authors have declared that no conflict of interest exists.

Abstract

Accumulating evidence has implicated impaired extracellular matrix (ECM) clearance as a key factor in fibrotic disease. Despite decades of research elucidating the effectors of ECM

clearance, relatively little is understood regarding the upstream regulation of this process. Collagen is the most abundant constituent of normal and fibrotic ECM in mammalian tissues. Its catabolism occurs through extracellular proteolysis and cell-mediated uptake of collagen fragments for intracellular degradation. Given the paucity of information regarding the regulation of this latter process, we executed unbiased genome-wide screens to understand the molecular underpinnings of cell-mediated collagen clearance. Using this approach, we discovered a previously unappreciated mechanism through which collagen biosynthesis is sensed by cells internally and directly regulates clearance of extracellular collagen. The sensing mechanism is dependent on endoplasmic reticulum-resident protein SEL1L and occurs via a noncanonical function of SEL1L. This pathway functions as a homeostatic negative feedback loop that limits collagen accumulation in tissues. In human fibrotic lung disease, the induction of this collagen clearance pathway by collagen synthesis is impaired, thereby contributing to the pathological accumulation of collagen in lung tissue. Thus cell-autonomous, rheostatic collagen clearance is a previously unidentified pathway of tissue homeostasis.

Main Text

Fibrosis can be seen as the excess accumulation of extracellular matrix that interferes with tissue and organ function. Collagen is the most abundant constituent of fibrotic extracellular matrix, yet we still lack a complete understanding of the regulatory mechanisms governing collagen homeostasis. Whereas decades of research have elucidated pathways of collagen deposition in normal tissue and in fibrosis, relatively less is understood about collagen clearance. Clearance of extracellular collagen from tissue occurs via two pathways that can act in parallel and sequentially[1]: an extracellular proteolytic process and a process of intracellular uptake and degradation. This latter process is vitally important in controlling normal tissue

53 homeostasis and mitigating fibrosis as demonstrated in several recent publications that show
54 the effects of interfering with this process in animal models of fibrosis[2–5]. Although an impaired
55 collagen-degradative environment is a feature of fibrotic diseases such as Idiopathic Pulmonary
56 Fibrosis (IPF), Scleroderma, and cirrhosis[6–13], very little is known about the role of cell-based
57 collagen degradation in human disease.

58 A key mediator of cell-based collagen clearance is MRC2, a canonical collagen
59 endocytic receptor that binds to and internalizes collagen via a Fibronectin-2 domain, with
60 subsequent lysosomal degradation[14,15]. Earlier studies of this receptor showed that it mitigates
61 lung fibrosis[3] as well as liver and kidney models of fibrosis[16,17]. Our recent work added to this
62 literature by showing that MRC2-mediated collagen clearance is essential for resolution of
63 fibrosis (rather than development of fibrosis) – deletion of MRC2 is sufficient to impair resolution
64 of fibrosis in the bleomycin mouse model of lung fibrosis[18].

65 We and others have been studying the process of cellular uptake and degradation of
66 collagen fragments because of the accumulating evidence that this is an important homeostatic
67 mechanism *in vivo* and controls fibrosis resolution[2,16–19]. Previous attempts at uncovering novel
68 mediators of this process have met with some success, for example uncovering the role of
69 Flotillin proteins or the regulation by MRC2 of matrix proteases[20,21]. However, experimental
70 approaches have either been targeted based on known genes and pathways or have been
71 limited to a subset of potential genes because of limited technical ability to use genome-wide
72 discovery approaches. The advent of high throughput CRISPR-based screens that are done in
73 a pooled format and leverage high throughput sequencing[22] now permits the unbiased
74 evaluation of all potential genes for any role in upstream regulatory pathways that could be
75 contributing to this process. Therefore in this work we started with complementary CRISPRi and
76 CRISPRa screens to address this knowledge gap, followed by mechanistic work to characterize
77 a previously undescribed pathway that controls cell-based collagen clearance.

78

79

80 Results

81

82 We performed genome wide unbiased screens using CRISPR interference (CRISPRi)
83 and CRISPR activation (CRISPRa) to determine genes that regulate cellular uptake of
84 exogenous collagen (Fig. 1). In independent pooled screens, we used CRISPRi to silence every
85 gene in the genome, or CRISPRa to activate every gene in the genome along with use of
86 10,000 non-targeting guides. We then performed a phenotypic screen for cellular uptake of
87 fluorescent collagen fragments. The screens were carried out independently with entirely
88 separate guide libraries for CRISPRi and CRISPRa. Coverage was at least 350x throughout the
89 screening process for each screen. The screens were performed in U937 cells, a human
90 monocytic leukemia cell line that readily takes up collagen[23] and grows in suspension, which
91 facilitates the use of a large-scale screen platform requiring culturing of hundreds of millions of
92 cells.

93 We determined multiple genetic regulators of collagen uptake (Fig. 1B-C; Supp. Table
94 1). Some of these were previously known to have a role in this process, most importantly
95 *MRC2*, a canonical collagen endocytic receptor. Most of the other candidates were not
96 previously known to play a role in collagen turnover. Importantly, we found a negative
97 correlation when comparing phenotype scores for individual genes that were hits in both the
98 CRISPRi and CRISPRa screens (Fig. 1D). This suggests that at least some of these genes we
99 identified are both necessary and sufficient for cellular collagen clearance and that quantitative
100 control of collagen clearance by these genes might be tunable based on their expression levels.
101 To validate our findings, we re-screened the high and low uptake bins from the original CRISPR
102 screens individually and compared hits in these re-screens to the corresponding original
103 CRISPRi or CRISPRa screen (Supp. Fig. 1A-D). We found a marked correlation between the

104 original and repeat screens (Supp. Fig. 1C-D). To validate the performance of our pooled guide
105 library, we used arrayed delivery of CRISPRi reagents for 3 sgRNAs for multiple genes
106 identified as hits in either positive or negative directions in the original CRISPRi screens,
107 followed by collagen uptake assays. Phenotypic effects of arrayed delivery of these CRISPRi
108 sgRNAs were 77% concordant with original screen data (Supp. Fig. 1E).

109 We next performed bioinformatic analyses of hits from our CRISPR screens. We noticed
110 that multiple gene ontology (GO) groups related to collagen metabolism were top scoring GO
111 terms in our analysis using GSEA (Fig. 2A). This was true in all four screens, i.e. the original
112 CRISRPi and CRISPRa screens as well as the re-screens (Supp. Fig. 2). Intriguingly, this did
113 not only include collagen catabolism GO terms, but also collagen biosynthesis or collagen
114 production-related GO terms (Fig. 2A). This informatic analysis implied that collagen
115 biosynthetic activity may positively regulate cellular uptake of extracellular collagen fragments.
116 We compared the enrichment of these GO terms in our collagen uptake screens with data from
117 previously published CRISPR screens of phagocytosis of other cargo[24] (Fig. 2B). Only our
118 screens for collagen uptake exhibited uniform enrichment of collagen biosynthetic GO terms
119 (i.e. positive regulation of collagen uptake by collagen biosynthesis), compared with uptake of
120 other cargo (Fig. 2B). These data suggest that there is a biological basis for positive regulation
121 collagen uptake by collagen biosynthesis genes, and this is specific to uptake of collagen.

122 To validate whether collagen biosynthesis positively regulates collagen uptake, we
123 decided to evaluate individual collagen biosynthesis genes. A heat map of gene-level phenotype
124 scores from our original screens among top-enriched genes from the Collagen Biosynthesis GO
125 group is shown in Figure 3A. We also note that *TRAM2*, a gene previously described to be
126 critically important collagen biosynthesis[25] was a top hit in the CRISPRi screen as well (Fig. 1).
127 We used an orthogonal technique (shRNA) to silence TRAM2 in U937 cells which resulted in
128 decreased collagen uptake (Fig. 3B; Supp. Fig. 3A) and decreased MRC2 expression at the
129 protein and mRNA levels (Fig. 3C-D). To verify whether TRAM2-silencing indeed affects

130 collagen synthesis, we used MRC5 human fetal lung fibroblasts which express high levels of
131 collagen. Silencing of TRAM2 in fibroblasts led to decreased endogenous levels of collagen
132 (Fig. 3E-G) and decreased collagen uptake in fibroblasts as well (Fig. 3H). We next silenced
133 individual collagen genes that were seen in the original screen including type I collagens and
134 type V collagens and determined that silencing of these genes led to decreased collagen uptake
135 and a decrease in *MRC2* message in U937 cells (Fig. 3I-K). Silencing of type I collagen
136 synthesis in MRC5 fibroblasts also caused decreased collagen uptake (Fig. 3L) as well as
137 decreased MRC2 protein levels (Fig. 3M-N). By contrast, stimulating collagen expression by
138 pharmacologic means with TGFβ1 increased collagen uptake in U937 cells (Fig. 4A). Although it
139 is controversial whether leukocytes or myeloid cells produce significant quantities of collagen[26–
140 28], they do generate collagen mRNA and recently were found to contribute to collagen content *in*
141 *vivo* under certain conditions (including types I and IV collagens in zebrafish and mouse)[29]; a
142 Western blot verified that U937 cells do indeed make collagen, albeit in very low quantities (Fig.
143 4A; note 60 μg of protein needed to be loaded to visualize collagen from U937 cells, compared
144 with 20 μg loaded in most other Western blots in this manuscript; additionally, a more sensitive
145 total type I collagen antibody was used in this subpanel rather than a procollagen antibody that
146 is less sensitive but is specific for newly synthesized collagen, which we used in the other
147 subpanels of this figure). To corroborate these findings, we additionally stimulated collagen
148 synthesis in MRC5 fibroblasts and measured collagen uptake and *MRC2* expression.
149 Stimulation with TGFβ1 led to increased collagen uptake and increased *MRC2* expression (Fig.
150 4B-C). Using orthogonal pharmacologic stimulators of collagen synthesis, ascorbic acid and
151 LPA1 respectively, led to increasing *MRC2* message as well (Fig. 4D-E). These data together
152 provide additional evidence for a pathway by which collagen biosynthesis regulates *MRC2*
153 expression and collagen uptake.

154 Because the screens were based on flow cytometric sorting of individual cells and each
155 cell had a single individual genetic perturbation, phenotypes determined by our screens are
156 expected to be fundamentally cell-autonomous. Therefore we presumed that collagen synthesis
157 in cells must be sensed by some internal cell-autonomous mechanism. In examining hits in the
158 screen to determine whether there could be a cell-autonomous sensor of collagen synthesis, we
159 were intrigued to see that *SEL1L* was a top hit in our screens (Fig. 1). The gene product of
160 *SEL1L* is a single-pass, endoplasmic reticulum-resident transmembrane protein with its luminal
161 face containing a Fibronectin-2 (FN2) domain (Fig. 5A). FN2 domains are generally known to
162 bind collagen[30–32]. However the function of the FN2 domain in SEL1L has not previously been
163 characterized.
164 To validate the effect of *SEL1L* on collagen uptake, we used an inducible knockout
165 mouse embryonic fibroblast (MEF) cell line characterized previously[33]. We determined that
166 when *Sel1L* is deleted, there is reduced collagen uptake (Fig. 5B), reduced MRC2 protein
167 expression (Fig. 5C-D; Supp. Fig. 3B-C), reduced *Mrc2* mRNA (Fig. 5E; Supp. Fig. 3D), and
168 reduced MRC2 cell surface and total protein expression (Fig. 5F-G). By contrast overexpressing
169 SEL1L in fibroblasts leads to increased *MRC2* message and protein (Fig. 5H-I). Importantly
170 although SEL1L has chiefly been described as a member of the endoplasmic reticulum-
171 associated protein degradation (ERAD) quality control pathway[33–35], other members of the
172 ERAD pathway were not hits in our CRISPR screens of collagen uptake (Fig. 1; Supp. Table 1).
173 This included HRD1, which acts with SEL1L and is essential for ERAD[36–39] but was not a hit in
174 our screens. To test if SEL1L-mediated regulation of collagen uptake is ERAD-dependent, we
175 used cells null for SEL1L or HRD1 and determined that only SEL1L positively regulates collagen
176 uptake, MRC2 expression (Fig. 6A-C; Supp. Fig. 3E), and *MRC2* message production (Fig. 6D).
177 For these experiments and several following cell biology experiments, HEK293T cells were
178 chosen for the following reasons (1) efficiency of transfection and co-transfection; (2) conserved

179 positive regulation of MRC2 by SEL1L in this cell type as well; and (3) decreased proliferation of
180 MEFs after two passages following induction of complete *Sel1L* knockout, as described
181 previously[33]. However, where possible we have analyzed more than one cell as described
182 below.

183 We then tested whether MRC2 can be forcibly overexpressed in wildtype (WT), SEL1L-
184 knockout and HRD1-knockout cells. We were able to overexpress in all three conditions to
185 equal levels MRC2 that is functional and can take up exogenous fluorescent collagen fragments
186 (Fig. 6E-G; Supp. Fig. 3F). Since HRD1 is essential for ERAD as above, these data suggest
187 that the ERAD pathway is not necessary for MRC2 protein expression or maturation and argues
188 against ERAD itself being instrumental in regulating MRC2 expression. Furthermore, our data
189 indicate that *MRC2* is being regulated by SEL1L at the message level based on the data in
190 Figures 5 and 6, rather than protein level as may be expected if the process were ERAD-
191 dependent. Finally, though deficiency of SEL1L or HRD1 have both been shown to predispose
192 to ER stress[40,41], the mechanism by which deficiency of either of these molecules predisposes
193 to ER stress has been shown to be ERAD-dependent. Since the data in Figures 5 and 6 show
194 that HRD1 deficiency does not affect MRC2 and collagen uptake regulation, this argues against
195 a strong contribution of ER stress from SEL1L deficiency affecting this pathway via ERAD
196 impairment because HRD1 deficiency would be expected to have the same effect on ER stress.

197 To determine whether SEL1L could be a collagen sensor, we wanted to investigate
198 whether it binds collagen and whether its FN2 domain is important for its function of regulating
199 *MRC2* levels. Computational modeling predicted docking of collagen peptide in the FN2 domain
200 of SEL1L (Supp. Fig. 4A). In fibroblasts, we determined that SEL1L and collagen co-localize,
201 and this colocalization increases when collagen synthesis is induced by TGFβ1 (Fig. 7A; Supp.
202 Fig. 4B). Furthermore, there is a biochemical association between SEL1L and type I collagen,
203 as determined by co-immunoprecipitation followed by Western blotting (Fig. 7B). We next made
204 expression constructs of WT SEL1L vs. SEL1L with an in-frame deletion of its FN2 domain

205 (AA118-166; referred to as ΔFN2), each with C-terminal MYC tags, to examine the function of
206 this FN2 domain. Computational prediction suggested that this deletion should have minimal
207 overall effect on protein folding (Supp. Fig. 5A), since predicted full-length or ΔFN2 structures
208 had very predicted high alignment with Root Mean Squared Distance <1 Å. Transfection of
209 either construct led to localization of the MYC-tagged protein in the ER based on co-localization
210 with Calnexin (Supp. Fig. 5B). We also purified the WT and ΔFN2 proteins which migrated to the
211 expected sizes via SDS-PAGE (Supp. Fig. 5C) and analyzed the purified proteins with circular
212 dichroism (Supp. Fig. 5D); their spectra were similar, suggesting similar secondary structures.
213 Taken together, these data suggest the ΔFN2 mutant is fully translated, can be localized in the
214 correct subcellular compartment, and likely folds properly.
215 Next, we compared co-immunoprecipitation of the WT or ΔFN2 constructs with collagen
216 and show that the SEL1LΔFN2 mutant exhibited markedly diminished binding to procollagen
217 compared to WT SEL1L in lysates from co-transfected cells (Fig. 7C-D). To verify this is likely a
218 direct interaction, we incubated our purified protein constructs with rat-tail collagen and
219 demonstrate markedly reduced binding efficiency of the ΔFN2 construct, down to the
220 background levels of non-specific binding of collagen to beads (Supp. Fig. 5E). These data
221 indicate that the FN2 domain is essential for collagen binding by SEL1L. Importantly in both
222 fibroblasts and HEK293T cells, WT SEL1L drives MRC2 expression but SEL1LΔFN2 does not,
223 even though in these experiments these expression constructs are expressed at similar levels
224 (Fig. 7E-H). Furthermore, we evaluated the levels of OS9, a well-described ERAD substrate[33],
225 under conditions of SEL1L knockout or rescue with WT or ΔFN2 constructs (Fig. 7G). OS9
226 accumulated in the SEL1L knockout cells as expected, which was rescued with the WT
227 construct as well as the ΔFN2 mutant construct. Taken all together these data strongly suggest
228 that SEL1L is an internal sensor of collagen biosynthesis, and SEL1L is necessary for the effect
229 of collagen biosynthesis on *MRC2* message levels and hence the downstream phenotype of

230 collagen uptake. Consistent with our data in Figure 6 suggesting this phenomenon is ERAD-
231 independent, the FN2 domain appears to be essential for *MRC2* regulation but not for ERAD
232 based on the effects on OS9 levels. This effect on OS9 also validates the stability and function
233 of the ΔFN2 construct since it rescues ERAD functionality. Consistent with our data suggesting
234 that the FN2 domain is dispensable for ERAD, a comparison of protein domain structures of
235 SEL1L across diverse species demonstrates that although SEL1L and its ERAD function are
236 evolutionarily ancient and conserved down to yeast[42], the FN2 domain is only present in
237 vertebrates and higher organisms, consistent with the hypothesis that the collagen-sensing
238 function arose later and is likely to be independent of its original ERAD function (Fig. 7I). This
239 occurred evolutionarily after the appearance of the earliest collagens in sponges at the dawn of
240 metazoan[43–45].

241 To determine whether this pathway is relevant in human physiology and disease, we re-
242 analyzed single cell data from multiple published experiments including normal and fibrotic lung.
243 We found that in cells that express both *MRC2* and type I collagen, there is a positive
244 correlation between expression levels of these two mRNA species (Fig. 8A-D). However, this
245 positive correlation is diminished IPF lung cells (Fig. 8D). In fact, data from multiple independent
246 diverse datasets examining gene expression in different compartments (whole lung,
247 bronchoalveolar lavage, peripheral blood mononuclear cells) from control versus IPF patients
248 demonstrate that in all cases there is a positive correlation between collagen and *MRC2*
249 expression, but the relationship is impaired in IPF (Fig. 8E-G). These findings corroborate our *in*
250 *vitro* data showing collagen biosynthesis is positively associated with *MRC2* and collagen
251 uptake by cells. This diminished recruitment of *MRC2* could be a potential mechanism by which
252 fibrotic collagenous extracellular matrix fails to be cleared in diseases of persistent and
253 progressive fibrosis like IPF. By contrast in a spontaneously resolving model of fibrosis (the
254 bleomycin lung fibrosis model in C57Bl\6 mice) there is a positive correlation between collagen
255 and *Mrc2* expression in mouse lung and the relationship is not impaired under conditions of

256 fibrosis (Fig. 8H). Given that we have previously shown that MRC2 is necessary for resolution of
257 fibrosis[18], these data raise the possibility that impairment of biosynthesis-induced collagen
258 clearance via MRC2 could be an important pathway driving non-resolving fibrosis and IPF.
259 Finally, we examined SEL1L levels in lungs of pulmonary fibrosis patients versus control
260 lungs (see Supp. Table 3 for more details). We found that in some, but not all, fibrotic human
261 lung specimens there is diminished full-length SEL1L (Fig. 9A-B). Immunofluorescence in
262 fibrosis vs. control lungs corroborated this finding (Fig. 9C). Future work will investigate how this
263 process is taking place. Additional mechanisms by which collagen biosynthesis-induced
264 collagen clearance is impaired in IPF is also an area of study in our laboratory.
265
266
267 Discussion
268
269 In this work we have shown that collagen synthesis directly upregulates collagen
270 clearance by MRC2. Our screens identified uniquely important biological pathways controlling
271 collagen clearance that do not regulate phagocytic uptake of other cargo. This led us to identify
272 SEL1L as a critical regulator of collagen synthesis-induced collagen turnover. SEL1L is an
273 internal sensor of collagen biosynthesis and its collagen sensing function is necessary for the
274 homeostatic effect of collagen biosynthesis on upregulation of MRC2. SEL1L is known to be a
275 key member of the ERAD pathway, but our data indicate that the collagen-sensing function of
276 SEL1L is independent of its ERAD function, defining a novel non-canonical function for this
277 protein. Consistent with this concept is our observation that although SEL1L and the ERAD
278 pathways are both evolutionarily ancient with conservation to yeast[42], the collagen-sensing FN2
279 domain of SEL1L arose more recently with vertebrates after the evolutionary origin of collagen
280 proteins, which appeared concurrently with metazoa[43–45].

281 From a teleological perspective, priming the same cells that contribute to scar formation
282 to upregulate machinery integral for repair and clearance of that scar would be an efficient
283 means of reaching a goal of tissue homeostasis and regeneration of normal tissue. It is striking
284 that in IPF, a disease of persistent non-resolving fibrosis, the relationship between collagen
285 synthesis and clearance is uncoupled, in contrast to the preserved relationship of these two
286 processes in the spontaneously resolving model of bleomycin-induced lung fibrosis in C57Bl\6
287 mice[18,46,47]. These findings suggest this pathway could be ripe for therapeutic intervention in
288 IPF. Until now, it has not been apparent that this rheostatic collagen turnover pathway was
289 inadequately engaged because previous studies have shown a modest increase in *MRC2*
290 expression in IPF lung compared with normal lung[48–52]. However, this belies a dramatically
291 inappropriate recruitment of the homeostatic turnover mechanism necessary to resolve fibrosis
292 which would be expected commensurate with high levels of collagen synthesis in IPF.
293 In short, an imbalance of collagen production and degradation is necessary for the
294 hallmark accumulation of collagen in pulmonary fibrosis; our work here suggests a defective
295 cell-mediated pathway of collagen clearance, based on expression of MRC2 which correlates
296 closely to functional uptake of collagen[23], is a key part of this imbalance. It is well documented
297 that MRC2 mitigates fibrosis in *in vivo* models[3,16,17]. Direct cellular clearance of collagen by
298 MRC2 is likely the key function responsible for this phenotype. Collagen fragments are also
299 thought to be pro-inflammatory[53,54], so it is possible that unclear collagen fragments in the
300 setting of inadequate MRC2 activity could add to fibrogenesis. However, MRC2 has multiple
301 other functions besides collagen uptake such as clearance of collectins[55] and roles in apoptosis
302 regulation and other cellular processes[56,57]. The inadequate upregulation of MRC2 we have
303 described in Pulmonary Fibrosis could have multiple downstream consequences which we
304 intend to explore further in future work. Despite this, our gain-of-function data showing that
305 overexpression of SEL1L can lead to increased MRC2 production suggests this pathway could
306 be ripe for therapeutic targeting (i.e. to drive increased clearance of collagen to ameliorate

307 fibrosis). The decrement in SEL1L seen in a subset of fibrotic lungs in both Western blotting and
308 immunofluorescence lends further credence to this idea, since MRC2 is pro-resolving in
309 fibrosis[18].

310 Limitations of our study include not elucidating the specific cell type(s) in which SEL1L is
311 apparently diminished in pulmonary fibrosis lungs; this is an area of future study in our
312 laboratory. SEL1L is expressed in multiple cell types in the lung and may have other roles (most
313 notably its function in ERAD) that are or are not relevant in fibrosis biology. Since collagen
314 expression can be detected in other cell types besides fibroblasts, it is possible the pathway we
315 have described here is operational in other cell types as well. We also do not yet understand
316 how SEL1L is downregulated in fibrotic human lungs. Our data suggest this is one potential
317 mechanism by which there is incomplete recruitment of the homeostatic collagen clearance
318 mechanisms in IPF. How SEL1L itself is downregulated in fibrotic tissue is also the subject of
319 ongoing study in our laboratory. In addition, the mechanism by which SEL1L-mediated collagen
320 sensing transcriptionally regulates *MRC2* is not completely determined by this work. The
321 luciferase activity data we have produced suggest that *MRC2* transcript production may be a
322 downstream effect of SEL1L expression levels. The precise transcription factors or regulators
323 that are responsible for this are also the subject of ongoing study. Lastly, we cannot exclude the
324 possibility that MRC2 can also be an ERAD target, even though the mechanism by which
325 SEL1L positively regulates *MRC2* mRNA we have shown here is not ERAD-dependent; in fact,
326 the trend of increased MRC2 protein levels in the HRD1-KO cells may suggest MRC2 protein
327 could also be an ERAD target.

328 Despite these limitations, we demonstrate here that unbiased genome-wide approaches
329 based on phenotypic screening can yield new insights into cell and tissue-level biology. The
330 relationship between matrix production and degradation is turning out to be ever more complex
331 than previously appreciated; we believe understanding how this core relationship is

332 dysregulated in pulmonary fibrosis at the cellular level has the potential to help unlock insights
333 about this relentlessly progressive disease.
334
335
336 Methods
337
338 **CRISPR Screens**
339 Single cell clones of U937 cells expressing dCas9-KRAB-BFP (CRISPRi) or dCas9-VP64-BFP
340 (CRISPRa) were made as previously described[58]. Subsequently, independent guide libraries for
341 CRISPRi and CRISPRa were used as previously described[59]. Briefly, lentivirus particles
342 containing these guides were transduced into U937 clones above at low MOI (0.3), followed by
343 subsequent selection with puromycin (1 µg/ml). Puromycin-selected cells were amplified and the
344 underwent fluorescent collagen uptake assay (as below) in bulk (i.e. pooled format), with sorting
345 into high and low bins (top or bottom 10% of collagen uptake based on mean fluorescence
346 intensity), simultaneously on 4x FACSAria II devices, then re-cultured and harvested for
347 analysis or for repeat screening. Isolation of cell DNA, followed by amplification of the guide-
348 region of individual cells and high-throughput amplicon sequencing on an Illumina platform were
349 all done as previously described[59,60].
350
351 **Bioinformatics, Including Analysis of Screen Data**
352 Guide counts were extracted using Python v3 scripts previously published[22]. Analysis of screen
353 data was done via MAGeCK[61] using the lfc-mean method. Bioinformatic analysis was performed
354 using GSEA[62,63] on phenotype scores (defined as Log-fold change multiplied by the $-\log_{10}$(p-
355 value) as determined by the robust rank-aggregation method in MAGeCK). For comparison
356 across knockdown screens for different cargo, data from Haney MS et al. were used[24] and then
357 re-analyzed via GSEA along with our own collagen screens. Visualization of screen data was

358 done with ggplot2 package in R (version 4.1.1). Visualization of domain architecture of SEL1L
359 protein across species was performed using the OMA Orthology Browser[64]. For re-analysis of
360 single cell RNAseq data from multiple experiments (GSE136831, GSE135893, GSE121611,
361 GSE128033, GSE132771) were integrated and re-normalized via the Seurat R package version
362 4; cells with non-zero expression of *COL1A1* and *MRC2* were used for analysis in Fig. 7D.
363 Microarray data was re-analyzed as indicated using GEO2R.
364
365 **Quantitative real-time PCR**
366 Total RNA from tissue or sorted cells was isolated via Trizol (Invitrogen) or RNeasy kit (Qiagen).
367 Total RNA from each sample was reverse transcribed with a first-strand cDNA synthesis kit
368 (Quantabio or Thermo Fisher) according to the manufacturer's instructions. Quantitative real-
369 time PCR reactions were performed with different sets of primers and Sensifast SYBR (Bioline,
370 Taunton, MA) on a CFX96 or CFX384 Real-Time PCR Detection System (Bio-Rad) or Applied
371 Biosystems QuantStudio 5 using 20 ng cDNA/reaction. Each quantitative real-time PCR
372 reaction was performed at least twice and representative results are shown relative to a control
373 by the standard $2^{-\Delta\Delta Ct}$ method, where Ct represents the number of cycles required to reach
374 threshold for the target gene subtracted from the number of cycles required to reach threshold
375 for a control housekeeping gene (*GAPDH* or *ACTB*). All data are shown relative to the first
376 control condition in the given figure panel. Primer sequences are in Supplemental Table 2.
377
378 **Immunoblotting**
379 Cells or tissues were homogenized with cold lysis buffer (RIPA buffer: 50 mM Tris-HCl pH 7.4,
380 150 mM NaCl, 0.1% SDS, 0.5% DOC, and 1% Nonidet P-40; coIP buffer: 20mM Tris-HCl pH 8,
381 137mM NaCl, 2mM EDTA, 1% Nonidet P-40) supplemented with protease and phosphatase
382 inhibitor cocktail (Thermo Fisher Scientific). Lysates (20 μg except where otherwise noted) were
383 clarified by centrifugation at 14,000g for 15 minutes at 4°C, electrophoresed under reducing

384 conditions on 7.5% or 4-20% gradient precast gels (Bio-Rad), and transferred to PVDF
385 membranes (Bio-Rad and GVS Filter Technology). Membranes were blocked for 30 minutes in
386 5% Bovine Serum Albumin and incubated with primary antibodies against MRC2 (1:1,200: anti-
387 mouse: AF4789, 1μg/mL: anti-human: AF5770; R&D Systems, Bio-Techne), SEL1L (1:2,000;
388 Novus Biologicals: NBP2-93746; ABCAM: ab78298), collagen I (1:2,000; Southern Biotech:
389 1310-01), human procollagen I alpha 1 (1:1,500; Novus Biologicals: AF6220), MYC (1:5,000;
390 Cell Signaling Technology, 9B11), OS9 (1:2,000; Novus Biologicals: NB100-519), GAPDH
391 (1:4,000; Cell Signaling Technology, 14C10), or beta-Actin (1:10,000; Cell Signaling
392 Technology, 13E5) overnight at 4°C. Membranes were then incubated with horseradish
393 peroxidase–conjugated secondary antibodies (1:4000: Cell Signaling Technology; 1:2,000: R&D
394 Systems, Bio-Techne; Santa Cruz Biotechnology) for 1 hour at room temperature. Blots were
395 developed using enhanced chemiluminescence (Cytiva) by exposure onto film or ChemiDoc
396 XRS+ (Bio-Rad). Where indicated densitometry was performed using ImageJ (NIH), normalizing
397 the protein of interest to a loading control (GAPDH or beta-Actin), and comparisons were made
398 to the baseline condition. For immunoblots of FN2-deleted SEL1L, the Abcam antibody was
399 used because the immunogen is distal to the FN2 domain, whereas the immunogen for the
400 Novus antibody is the FN2 domain itself.
401
402 **Cell culture and molecular biology experiments**
403 U937 and MRC5 cells were obtained from ATCC. HEK293T cells that were WT, or null for
404 SEL1L or HRD1 were obtained from Dr. Qi and were described previously[33,65–67]. Immortalized
405 MEFs from WT (*Sel1L*$^{f/f;}$*ERcre*-) or littermate inducible KO KO (*Sel1L*$^{f/f}$*;ERcre*+) mice were
406 provided by Dr. Qi and were described previously[33]. MEFs were used at passage 4-16 and for
407 experiments in which Sel1L deletion was induced, MEFs were used within 1 passage after
408 addition of 4-hydroxytamoxifen (at 400 nM) vs. ethanol vehicle control as described previously[33].
409 Cells were grown in DMEM or RPMI (Corning) with 10% FBS (Gemini Bio-Products). ORFs

410 containing full length murine SEL1L or GAPDH were obtained from Origene and were
411 maintained in pCMV6 expression vectors and included an in-frame C-terminal MYC tag. FN2
412 deletion was accomplished using Q5 Site-Directed Mutagenesis kit (New England Biolabs)
413 according to the manufacturer's directions and in-frame deletion was verified with Sanger
414 sequencing. ORFs (including the MYC tag) were also subcloned into pHIV lentiviral transfer
415 plasmids using standard techniques and lentiviral particles were packaged in HEK293T cells
416 with pMD2G and psPAX2. For MRC2 overexpression, pCI-Mrc2 (murine Mrc2 cloned into the
417 pCI-Neo expression vector, from Promega) was used as we previously described[18]. All
418 transfections were done using Lipofectamine 2000 (Invitrogen) according to the manufacturer's
419 instructions. For shRNA experiments, lentiviral particles harboring shRNA machinery within
420 pSicoR-puromycin vectors were obtained from Sigma-Aldrich (scramble control: SHC002V;
421 shRNA sequences can be found in Supplemental Table 2). Lentiviral infection was done in
422 these experiments at 10 multiplicity of infection (MOI) in cells overnight in complete medium
423 after addition of polybrene (8 µg/mL), and medium was replaced with fresh complete medium
424 the next day. The infected cells were allowed to grow and then selected by resistance to
425 puromycin (1 µg/mL Sigma-Aldrich). Puromycin-selected cells were collected for experiments as
426 indicated. Cells were infected at passages <10, and cells were collected for experiments
427 between passages 5 and 16. For the luciferase assay, a 625bp fragment around the MRC2
428 transcriptional start site in a Firefly luciferase vector with a minimal promoter element (Promega)
429 was used as previously described[18]. The Stop&Glo® (Promega) system was used to read
430 luciferase activity in cells on a luminometer (Synergy H1). All firefly luciferase values were
431 normalized to luciferase levels of a constitutive Renilla for a transfection control.
432
433 **Co-immunoprecipitation**
434 Proteins were isolated in non-denaturing coIP lysis buffer (see above). Proteins were incubated
435 with fast-flow protein G- or protein A- Dynabeads (Invitrogen) for 2 hours and centrifuged to

436 eliminate nonspecifically bound proteins. The concentrations of the precleared proteins were
437 estimated by Bradford's assay. Primary antibodies against Procollagen (Novus) or SEL1L
438 (Novus) or a non-specific IgG control antibody were incubated with protein G- or protein A-
439 Dynabeads for 4 hours. Then, 200-400 µg pre-cleared protein were incubated overnight with the
440 antibody-bound beads. After washing, attached proteins were then eluted from the beads in 1%
441 SDS buffer by heating at 60˚C and vortexing followed by centrifugation to collect the
442 supernatants as the eluates. Equal volumes of the eluates were used for immunoblotting.
443 Immunoprecipitation with protein isolation buffer without lysates was run as a negative control
444 for the beads. Proteins immunoprecipitated using non-specific IgG (NS IgG) were negative
445 control for specific antibodies. Immunoblotting with total pre-cleared protein lysates were used
446 as input samples. For purification of the WT and ΔFN2 proteins, MYC-tagged protein was
447 purified from transduced cells (see above) with Myc-trap agarose (Chromotek), a nanobody
448 based system, according to the manufacturer's instructions. For co-IP experiments with purified
449 collagen, purified WT or ΔFN2 proteins were incubated with rat tail collagen (EMD Millipore) and
450 then co-IP was carried out as described above.
451
452 **Circular Dichroism**
453 WT or ΔFN2 purified protein samples were diluted to final concentration of ~0.35 mg/mL in a
454 Tris-Glycine buffer. CD spectra were taken with Chirascan V100 CD Spectrometer (Applied
455 Photophysics) using a 0.5-mm quartz cuvette between 200 and 300 nm at a set temperature of
456 25 °C.
457
458 **Flow cytometry and flow cytometric-based collagen uptake assay**
459 Flow cytometric analysis was done using a FACSVerse, FACSAriaII, or Accuri c6 (BD
460 Biosciences) and sorting was done using FACSAriaII devices (BD Biosciences) at the UCSF

461 Parnassus Flow Cytometry Core. For the collagen uptake assay, cells were incubated in fresh
462 media with Oregon Green 488-conjugated gelatin (10 μg/ml final concentration; Life
463 Technologies) for 1 hour at 37°C. Cells were then harvested (for suspension cells) or trypsinized
464 and harvested for analysis. To assess only internalized fluorescent collagen (i.e. to ignore
465 bound but not internalized collagen), cells were either washed before flow cytometry or Trypan
466 blue was added to quench bound but not internalized fluorochromes as we have done
467 previously[21,58]. For cell-surface staining, cells were stained using a standard FACS staining
468 protocol. Antibodies included: anti-MRC2 (as above) as a primary antibody with a secondary of
469 Donkey anti-sheep conjugated to a fluorophore (Invitrogen). Cells were then analyzed and/or
470 sorted on one of the above devices, with sorting and analysis gates based on fluorescence-
471 minus-one controls after compensation was performed with single-stain controls. For the
472 collagen uptake assay. Data were analyzed with FlowJo, version 7.6.1 (Tree Star, Ashland,
473 OR).
474
475 **Immunofluorescence Imaging**
476 Cells plated on coverslips were fixed with 4% PFA in PBS. Coverslips were blocked in
477 Phosphate-buffered saline (PBS) with 0.5% Bovine serum albumin, 0.1% Triton X-100 and 5%
478 donkey serum. Antibodies against Type I collagen (Southern Biotech, as above), MRC2 (R&D,
479 as above), Calnexin (Cell Signaling Technology, C5C9), MYC (Santa Cruz Biotechnology,
480 9E10), or SEL1L (Novus, as above for Fig. 6A; for Fig. 6D, anti-SEL1L was provided as a gift by
481 Dr. Qi[33,65–67] and was used since it can recognize SEL1L without its FN2 domain) was applied at
482 1:100 dilution followed by secondary antibody conjugated to Alexa Fluor 488 or 594 (Invitrogen)
483 at 1:100. Coverslips were mounted in Vectashield (Vector Laboratories). Human lung fragments
484 were fixed in 4% PFA in PBS and then dehydrated in sucrose and embedded in OCT and
485 frozen in chilled isopropanol on dry ice. Frozen sections (30 μm) were prepared. Sections were
486 blocked in Phosphate-buffered saline (PBS) with 0.5% Bovine serum albumin, 0.1% Triton X-

487 100 and 5% donkey serum. Antibody against SEL1L (Novus, as above) and Collagen I
488 (Rockland, 600-401-103) were applied at 1:100 dilution followed by secondary antibody
489 conjugated to Alexa Fluor 488 or 594 or 647 (Invitrogen) at 1:100. Sections were mounted in
490 Vectashield (Vector Laboratories).
491 Confocal images were captured using a Leica TCS SPE confocal microscope with an ACS APO
492 40x or 63x Oil CS objective lens at room temperature with Leica Type F Immersion Liquid
493 (11513859) or with a Leica Stellaris 8 confocal microscope with a 10x air or 40x Oil objective.
494 Wide-field epifluorescence images were captured by an Olympus DP70 CCD using a standard
495 Olympus BX51 upright microscope with an Olympus UplanFL 40x objective at room
496 temperature or with a Nikon TE2000-U microscope and Ds-Qi2 camera. Images were
497 processed with ImageJ (National Institutes of Health, Bethesda, Maryland, USA). All imaging
498 was done on the same platform within any given experiment. Images were processed with
499 ImageJ, in which any brightness/contrast was changed uniformly across all conditions and
500 pseudocoloring was applied. Single optical slices at Airy 1 are shown for confocal images.
501
502 **Human lung tissue**
503 Normal human lung tissue from deceased donors of different ages was obtained from lungs not
504 used by the Northern California Transplant Donor Network according to an IRB-exempted
505 protocol led by Dr. Wolters; IPF lung specimens were obtained from explanted lungs removed
506 during lung transplantation, in an IRB-approved protocol led by Dr. Wolters (see Supp. Table 3
507 for more details). After harvest, lung tissue was directly snap-frozen in liquid nitrogen and
508 protein was extracted as above or 1cm fragments were fixed in 4% PFA in PBS and processed
509 as described above for imaging.
510
511 ***In Silico* protein structure prediction**

512 Docking of the FN2 domain of SEL1L (Uniprot Q9UBV2; AlphaFoldDB structural prediction) with
513 a collagen peptide sequence (PGASGPMGPRGPPGPPGKNG) was done using MDockPeP[68,69]
514 and visualized with PyMOL (version 2.3; Schrödinger). For prediction of the structures of full
515 length SEL1L the ΔFN2 mutant, the Robetta online platform (which relies on RoseTTAFold
516 prediction algorithm[70]) was used to predict the structure of both the full-length sequence
517 (Uniprot Q9Z2G6) and the sequence without AA118-166 (the FN2 domain). Alignment and
518 visualization were performed in PyMOL.
519
520 **Statistics**
521 Data were evaluated with Graphpad Prism software (San Diego, CA): for two sample
522 comparisons, by 2-tailed Student's *t*-test or Mann-Whitney *U* test if data did not pass a test of
523 normality; for multiple comparisons, by ANOVA with post-hoc Bonferroni testing if the ANOVA
524 was significant; or by Pearson correlation as indicated. A p-value of 0.05 or less was considered
525 statistically significant with *=p<0.05, **=p<0.01, ***=p<0.001, ****=p<0.0001, NS=Not
526 Significant. Data are presented as mean ± SEM.
527
528 **Study approval**
529 All experiments using human tissue were approved by the Weill Cornell Medicine or UCSF
530 Committee on Human Research. Informed consent was obtained for study participants or
531 waived for human tissue from deceased donors in accordance with the UCSF Committee on
532 Human Research.
533
534
535 **Data Availability**
536
537 Source data are provided with this paper.

538

539 **Code Availability**

540

541 The codes used in CRISPR screen analysis are described above and are available from the

542 authors upon reasonable requests.

543

544 **Acknowledgments**

545

546 This research was supported by US National Institutes of Health grants K08 HL145015 (MJP)

547 and R01 HL136377 (KA), a Grant-in-aid from The Stony Wold-Herbert Fund (MJP), and a

548 Sponsored Research Agreement with Celgene (KA). We acknowledge the UCSF Parnassus

549 Flow CoLab (RRID:SCR_018206) for assistant generating flow cytometry data, supported in

550 part by Grant NIH P30 DK063720 and by the NIH S10 Instrumentation Grant S10

551 1S10OD021822-01. We acknowledge and thank the Precision Biomolecular Characterization

552 Facility at Columbia University for technical support and access to the Circular Dichroism

553 spectrometer, which is supported by NIH Award 1S10OD025102-01. We would like to thank S.

554 Layer for ongoing inspiration.

555

556

557 **Figure Legends**

558

559 Figure 1. CRISPR screens identify mediators of cellular collagen uptake. (A) Schematic of

560 experimental approach for flow cytometry-based screening for uptake of collagen fragments. (B)

561 Volcano plot of gene-level statistics of CRISPRi (inhibition) screen. Most significant gene hits

562 are labeled and colored in red. RRA = Robust Rank Aggregation score; LFC = Log-fold change.

563 (C) Volcano plot of gene-level statistics of CRISPRa (activation) screen. (D) Comparison of
564 overall phenotype score between CRISPRi and CRISPRa screens.
565
566 Figure 2. Collagen-biosynthetic pathways are regulators of collagen uptake uniquely compared
567 to screens for uptake of other cargoes. (A) Ranked plot of Normalized Enrichment Scores
568 derived from Gene-Set Enrichment Analysis (GSEA) of CRISPRi phenotype scores, using
569 10,561 Gene Ontology gene sets. Collagen-related gene sets are labeled and displayed in red.
570 (B) Comparison of GSEA statistics for labeled collagen-related gene sets across screens for
571 uptake of various cargoes[24]; screen method is indicated in parentheses. KD = Knock-down.
572
573 Figure 3. Collagen biosynthesis positively regulates cell-based collagen clearance and *MRC2*
574 expression. (A) Heatmap of overall phenotype scores from original screens as indicated for top
575 genes from 'Collagen Biosynthetic Process' GO gene-set (displayed genes had most extreme
576 'Score' values in one or more screen result). (B) Flow-cytometry based collagen-uptake assay in
577 U937 cells after shRNA treatment as indicated vs. Scramble control. MFI = Mean Fluorescence
578 Intensity. N=4 per group. (C) Western blot of U937 cells after shRNA treatment as indicated vs.
579 Scramble control, representative of N=4 independent experiments. (D) Q-RT-PCR in U937 cells
580 after shRNA treatment as indicated vs. Scramble control. N=4 per group. (E-F) Western blot
581 and densitometry of MRC5 fibroblasts after shRNA treatment as indicated vs. Scramble control,
582 representative of N=3 independent experiments. (G) Representative confocal
583 immunofluorescence images of MRC5 fibroblasts after shRNA treatment as indicated vs.
584 Scramble control. (H) Flow-cytometry based collagen-uptake assay in MRC5 cells after shRNA
585 treatment as indicated vs. Scramble control. N=4 per group. (I) Flow-cytometry based collagen-
586 uptake assay in U937 cells after shRNA treatment as indicated vs. Scramble control. N=4-6 per
587 group. (J-K) Q-RT-PCR in U937 cells after shRNA treatment as indicated vs. Scramble control.
588 N=4-5 per group. (L) Flow-cytometry based collagen-uptake assay in MRC5 fibroblasts after

589 shRNA treatment as indicated vs. Scramble control. N=6 and 4 in Scramble and shCOL1A1
590 groups respectively, N=2 in shMRC2 group. (M-N) Western blot and densitometry of MRC5
591 fibroblasts after shRNA treatment as indicated vs. Scramble control, representative of N=4
592 independent experiments. Statistics: (B-K) ANOVA with post-hoc Bonferroni testing; (L-N)
593 Student's *t*-test. * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$.
594
595 Figure 4. Pharmacological stimulation of collagen synthesis is positively associated with
596 collagen uptake and *MRC2* expression. (A-B) Flow-cytometry based collagen-uptake assay in
597 U937 (A) or MRC5 (B) cells after drug treatment as indicated vs. vehicle control. N=4 per group.
598 Western blot of corresponding cells and conditions shown below, each representative of N=3-4
599 independent experiments. (C-E) Q-RT-PCR in MRC5 cells after drug treatment as indicated vs.
600 vehicle control. N=3-5 per group. Western blot of corresponding cells and conditions shown
601 below, each representative of N=3 independent experiments. Statistics: (A-E) Student's *t*-test. *
602 $p < 0.05$, ** $p < 0.01$.
603
604 Figure 5. *Sel1L* positively regulates collagen uptake and *Mrc2* expression. (A) Schematic of
605 domain structure of SEL1L protein. FN2 = Fibronectin-2. (B) Flow-cytometry based collagen-
606 uptake assay in MEF cells (mouse embryonic fibroblasts) after treatment with 4-
607 hydroxytamoxifen (OHT) as indicated vs. vehicle control. WT = *Sel1L*$^{f/f;}$*ERcre*-, iKO =
608 *Sel1L*$^{f/f}$*;ERcre*+. N=3-4 per group. (C-D) Western blot and densitometry of MEFs after OHT
609 treatment as indicated vs. control, representative of N=4 independent experiments. (E) Q-RT-
610 PCR in MEFs after OHT treatment for genes as indicated. N=3 per group. (F) Flow-cytometry
611 based measurement of MRC2 surface expression in MEFs after OHT treatment. N=4 per group.
612 (G) Representative wide-field immunofluorescence images of MEFs after OHT treatment. (H) Q-
613 RT-PCR of MRC5 cells transfected with overexpression vectors containing genes as indicated
614 vs. vehicle control (lipofectamine only). N=4 per group. (I) Western blot of MRC5 cells

615 transfected with overexpression vectors containing genes as indicated vs. vehicle control
616 (lipofectamine only). N=2. Statistics: (B, H) ANOVA with post-hoc Bonferroni testing; (D - right
617 panel, E, F) Student's *t*-test; (D - left panel) Mann-Whitney *U* test. * $p < 0.05$, ** $p < 0.01$, *** $p <$
618 0.001, **** $p < 0.0001$, ns = not significant.
619
620 Figure 6. SEL1L, but not HRD1, positively regulates MRC2. (A) Flow-cytometry based collagen-
621 uptake assay in 293T of different genotypes (compared with baseline control without fluorescent
622 collagen). N=4-5 per group. (B-C) Western blot and densitometry of 293T cells of different
623 genotypes, representative of N=5 independent experiments. (D) Firefly luciferase assay of
624 *MRC2* promoter activity, normalized to constitutive Renilla luciferase in 293T cells of different
625 genotypes. N=3 per group. (E) Western blot of HEK293T cells of different genotypes transfected
626 with MRC2 overexpression vector vs. vehicle control, representative of N=6 independent
627 experiments. (F) Representative wide-field immunofluorescence images of 293T cells of
628 different genotypes transfected with MRC2 overexpression vector vs. vehicle control. (G) Flow-
629 cytometry based collagen-uptake assay in HEK293T of different genotypes transfected with
630 MRC2 overexpression vector (compared with baseline control without fluorescent collagen; note
631 different scale compared with panel A). N=3-4 per group. Statistics: (A, C, D, G) ANOVA with
632 post-hoc Bonferroni testing. * $p < 0.05$ ** $p < 0.01$, ns = not significant.
633
634 Figure 7. SEL1L binds collagen as it is being synthesized via its FN2 domain and this interaction
635 is necessary for regulation of *MRC2* expression. (A) Representative confocal
636 immunofluorescence images of MRC5 fibroblasts treated as indicated or vehicle control. (B)
637 Western blot of MRC5 fibroblasts after immunoprecipitation with anti-SEL1L antibody (vs.
638 controls as indicated), representative of N=2 independent experiments. (C) Western blot after
639 immunoprecipitation with anti-procollagen type I antibody (with input controls as indicated) of
640 HEK293T cells transfected with collagen and WT or mutant SEL1L constructs as indicated,

641 representative of N=3 independent experiments. (D) Representative confocal
642 immunofluorescence images of HEK293T cells transfected with collagen and WT or mutant
643 SEL1L constructs as indicated. (E) Western blot of MEFs after treatment with WT or mutant
644 SEL1L constructs as indicated, representative of N=3 independent experiments. (F) Flow-
645 cytometry based measurement of MRC2 surface expression on MEFs after treatment with WT
646 or mutant SEL1L constructs as indicated, vs. Empty Vector (EV). N=6 per group. (G) Western
647 blot of HEK293T cells either WT or KO for SEL1L as indicated after treatment with EV (empty
648 vector), WT or mutant SEL1L constructs as indicated, representative of N=3 independent
649 experiments. (H) Q-RT-PCR of SEL1L-KO HEK293T cells after transfection with WT or mutant
650 SEL1L as indicated vs. empty vehicle control. N=3 per group. (I) Comparison of protein domain
651 structures of Sel1L orthologs across phylogenetically diverse organisms; FN2 = Fibronectin-2
652 domain, tpr = Tetratricopeptide repeat domain. Statistics: (F-G) ANOVA with post-hoc
653 Bonferroni testing. * $p < 0.05$, ** $p < 0.01$, ns = not significant.
654
655 Figure 8. Collagen biosynthesis and *MRC2* expression are positively correlated in lung tissue,
656 but this relationship is impaired in human fibrotic lung disease. (A) UMAP plot with cell type
657 annotation of scRNAseq data from multiple available datasets of human lung control and IPF
658 tissue (GSE136831, GSE135893, GSE121611, GSE128033, GSE132771). N=184,672 single
659 cells. (B-C) Gene expression of individual genes as indicated in UMAP as generated in panel
660 (A). (D) Comparison of gene expression in all cells from data as in panel (A) that had non-zero
661 expression of *COL1A1* or *MRC2*. (E) Comparison of gene expression in whole human lung
662 microarray data from GSE110147 analyzed via GEO2R in pulmonary fibrosis lungs vs. control
663 specimens. N=37 vs. 11 per group. (F) Comparison of microarray gene expression in human
664 lung bronchoalveolar lavage specimens from GSE70867 analyzed via GEO2R from Idiopathic
665 Pulmonary Fibrosis (IPF) lungs vs. control specimens. N=112 vs. 20. (G) Comparison of
666 microarray gene expression in peripheral blood mononuclear cell (PBMC) specimens from

667 GSE37858 analyzed via GEO2R from Idiopathic Pulmonary Fibrosis (IPF) lungs vs. control
668 specimens. N=70 vs. 35. (H) Comparison of gene expression in whole mouse lung microarray
669 data from GSE40151 analyzed via GEO2R from Bleomycin-treated mice vs. Saline control
670 mice. N=56 vs. 55. Statistics: ANCOVA, p-values indicated in figures. Pearson correlations are
671 displayed in panel (D).
672
673 Figure 9. SEL1L protein levels are reduced in the lungs of patients with Pulmonary Fibrosis. (A-
674 B) Western blots of whole human lung lysates from Pulmonary Fibrosis lungs vs. control (donor
675 lungs deemed not suitable for transplant), with corresponding densitometry. N=13 in Control
676 and 11 in Pulmonary Fibrosis group. Arrowheads correspond to two lung lysates that were used
677 in both Western blots for normalization across blots, otherwise all other lanes are independent
678 samples. (C) Representative low-power wide-field immunofluorescence images of human lungs
679 (Pulmonary Fibrosis vs. control). Each row comes from an individual donor or Pulmonary
680 Fibrosis lung, with two different fields per lung (i.e. N=3 individual lungs per group). Statistics:
681 (B, D) Mann-Whitney *U* test. * $p < 0.05$.
682
683
684 Supplemental Figures
685
686 Supplemental Figure 1. Re-screens are correlated to original CRISPR screen results. (A)
687 Volcano plot of CRISPRi re-screen (Screen 2). (B) Volcano plot of CRISPRa re-screen (Screen
688 2). (C) Comparison of gene-level overall phenotype scores between CRISPRi screens 1 and 2.
689 (D) Comparison of gene-level overall phenotype scores between CRISPRa screens 1 and 2. (E)
690 Comparison of arrayed log-fold change in arrayed validation vs. original pooled screen log-fold
691 change; each point represents unique guide sequence. 65/84 guides demonstrated
692 concordance with original pooled screen data (77%). Pearson correlation coefficient is shown.

693

694 Supplemental Figure 2. Collagen-related gene ontology terms are enriched as regulators of
695 collagen uptake in CRISPRi and CRISPRa screens. The top twenty positive and negatively
696 performing gene sets are indicated with associated statistics for each screen as indicated,
697 among all 10,561 Gene Ontology gene sets. Collagen-related gene sets are highlighted with red
698 boxes (i.e. those with 'Collagen' in the GO gene set name). NES = Normalized Enrichment
699 Score.

700

701 Supplemental Figure 3. Validation of different tools used in this manuscript. (A) Q-RT-PCR in
702 U937 cells after shRNA treatment as indicated vs. Scramble control. N=4 per group. (B-C)
703 Additional densitometry of data from Figure 5C. N=4 per group. (D) Q-RT-PCR in MEFs after
704 OHT treatment for genes as indicated. N=3 per group. (E) Additional densitometry of data from
705 Figure 6B. N=4 per group. (F) Densitometry of the left side of Figure 6E. N=6 per group.
706 Statistics: (A, E, F) ANOVA with post-hoc Bonferroni testing; (B-D) Student's *t*-test. * $p < 0.05$, **
707 $p < 0.01$, **** $p < 0.0001$, ns = not significant.

708

709 Supplemental Figure 4. Additional support of SEL1L and collagen interaction. (A) Predicted
710 docking of collagen peptide with FN2 domain of SEL1L. (B) Additional representative confocal
711 immunofluorescence image of MRC5 fibroblasts at baseline (as in Figure 7A), with decreased
712 PMT gain in the Collagen channel, along with magnified inset to highlight co-localization.

713

714 Supplemental Figure 5. FN2 domain deletion mutant interferes with collagen binding but is
715 otherwise stable. (A) Alignment of predicted structures of full-length SEL1L and ΔFN2 SEL1L.
716 Two different views and alignments are shown. Root-mean squared distance (RMSD) is
717 displayed. Yellow arrow indicates the FN2 domain (in the pink structure on the left; in the green
718 structure on the right). (B) Representative confocal immunofluorescence images of HEK293T

719 cells transfected with MYC-tagged WT or ∆FN2 SEL1L constructs (at low transfection
720 efficiency), co-stained with ER marker Calnexin. (C) Western blot of purified WT or ∆FN2 SEL1L
721 proteins compared with input cell lysates, representative of N=3 independent experiments. (D)
722 Circular Dichroism spectra of purified WT or ∆FN2 proteins. (E) Western blot of co-incubated
723 purified either WT or ∆FN2 SEL1L proteins with rat tail collagen, after immunoprecipitation with
724 anti-MYC antibody (with input and additional controls as indicated), representative of N=3
725 independent experiments. 'Collagen only' indicates incubation of rat tail collagen with IP beads
726 without co-incubation with purified MYC-tagged protein.
727
728
729 Supplemental Tables
730
731 Supplemental Table 1. Gene-level data from CRISPRi and CRISPRa screens of collagen
732 uptake as analyzed via MAGeCK.
733
734 Supplemental Table 2. Primer and shRNA sequences used in this manuscript.
735
736 Supplemental Table 3. Demographic and clinical data on the sources of human lungs used in
737 this manuscript.
738
739
740 References
741
742 1. McKleroy W, Lee T-H, Atabai K. Always cleave up your mess: targeting collagen
743 degradation to treat tissue fibrosis. *AJP: Lung Cellular and Molecular Physiology*.
744 2013;304(11):L709–L721.

745 2. Atabai K, Jame S, Azhar N, Kuo A, Lam M, McKleroy W, DeHart G, Rahman S, Xia DD,
746 Melton AC, Wolters P, Emson CL, Turner SM, Werb Z, Sheppard D. Mfge8 diminishes the
747 severity of tissue fibrosis in mice by binding and targeting collagen for uptake by macrophages.
748 *Journal of Clinical Investigation*. 2009.
749 3. Bundesmann MM, Wagner TE, Chow Y-H, Altemeier WA, Steinbach T, Schnapp LM.
750 Role of urokinase plasminogen activator receptor-associated protein in mouse lung. *American*
751 *journal of respiratory cell and molecular biology*. 2012;46(2):233–9.
752 4. Cui H, Jiang D, Banerjee S, Xie N, Kulkarni T, Liu R-M, Duncan SR, Liu G. Monocyte-
753 derived alveolar macrophage apolipoprotein E participates in pulmonary fibrosis resolution. *JCI*
754 *Insight*. 2020;5(5).
755 5. Fan M-H, Zhu Q, Li H-H, Ra H-J, Majumdar S, Gulick DL, Jerome JA, Madsen DH,
756 Christofidou-Solomidou M, Speicher DW, Bachovchin WW, Feghali-Bostwick C, Pur Eacute E.
757 Fibroblast Activation Protein (FAP) Accelerates Collagen Degradation and Clearance from Lung
758 in Mice. *The Journal of biological chemistry*. 2015.
759 6. Montaño M, Ramos C, González G, Vadillo F, Pardo A, Selman M. Lung collagenase
760 inhibitors and spontaneous and latent collagenase activity in idiopathic pulmonary fibrosis and
761 hypersensitivity pneumonitis. *Chest*. 1989;96(5):1115–9.
762 7. Selman M, Ruiz V, Cabrera S, Segura L, Ramírez R, Barrios R, Pardo A. TIMP-1, -2, -3,
763 and -4 in idiopathic pulmonary fibrosis. A prevailing nondegradative lung microenvironment?
764 *American journal of physiology. Lung cellular and molecular physiology*. 2000;279(3):L562-74.
765 8. Montfort I, Pérez-Tamayo R. Collagenase in experimental carbon tetrachloride cirrhosis
766 of the liver. *The American journal of pathology*. 1978;92(2):411–20.
767 9. Okazaki I, Maruyama K. Collagenase activity in experimental hepatic fibrosis. *Nature*.
768 1974;252(5478):49–50.
769 10. Brady AH. Collagenase in scleroderma. *The Journal of clinical investigation*.
770 1975;56(5):1175–80.
771 11. Mahalanobish S, Saha S, Dutta S, Sil PC. Matrix metalloproteinase: An upcoming
772 therapeutic approach for idiopathic pulmonary fibrosis. *Pharmacological Research*.
773 2020;152:104591.
774 12. Costanzo L, Soto B, Meier R, Geraghty P. The Biology and Function of Tissue Inhibitor
775 of Metalloproteinase 2 in the Lungs. *Pulmonary Medicine*. 2022;2022:1–12.
776 13. Inoue R, Yasuma T, Fridman D'Alessandro V, Toda M, Ito T, Tomaru A, D'Alessandro-
777 Gabazza CN, Tsuruga T, Okano T, Takeshita A, Nishihama K, Fujimoto H, Kobayashi T,
778 Gabazza EC. Amelioration of Pulmonary Fibrosis by Matrix Metalloproteinase-2
779 Overexpression. *International Journal of Molecular Sciences*. 2023;24(7):6695.
780 14. Engelholm LH, List K, Netzel-Arnett S, Cukierman E, Mitola DJ, Aaronson H, Kjøller L,
781 Larsen JK, Yamada KM, Strickland DK, Holmbeck K, Danø K, Birkedal-Hansen H, Behrendt N,
782 Bugge TH. uPARAP/Endo180 is essential for cellular uptake of collagen and promotes fibroblast
783 collagen adhesion. *The Journal of cell biology*. 2003;160(7):1009–15.
784 15. Yuan C, Jürgensen HJ, Engelholm LH, Li R, Liu M, Jiang L, Luo Z, Behrendt N, Huang
785 M. Crystal structures of the ligand-binding region of uPARAP: effect of calcium ion binding. *The*
786 *Biochemical journal*. 2016;473(15):2359–68.
787 16. Madsen DH, Jürgensen HJ, Ingvarsen S, Melander MC, Vainer B, Egerod KL, Hald A,
788 Rønø B, Madsen CA, Bugge TH, Engelholm LH, Behrendt N. Endocytic collagen degradation: A
789 novel mechanism involved in protection against liver fibrosis. *Journal of Pathology*.
790 2012;227(1):94–105.
791 17. López-Guisa JM, Cai X, Collins SJ, Yamaguchi I, Okamura DM, Bugge TH, Isacke CM,
792 Emson CL, Turner SM, Shankland SJ, Eddy AA. Mannose receptor 2 attenuates renal fibrosis.
793 *Journal of the American Society of Nephrology : JASN*. 2012;23(2):236–51.

794 18. Podolsky MJ, Yang CD, Valenzuela CL, Datta R, Huang SK, Nishimura SL, Dallas SL,
795 Wolters PJ, Le Saux CJ, Atabai K. Age-dependent regulation of cell-mediated collagen turnover.
796 *JCI Insight*. 2020;5(10).
797 19. Melander MC, Jürgensen HJ, Madsen DH, Engelholm LH, Behrendt N. The collagen
798 receptor uPARAP/Endo180 in tissue degradation and cancer (Review). *International Journal of*
799 *Oncology*. 2015;47(4):1177–1188.
800 20. Messaritou G, East L, Roghi C, Isacke CM, Yarwood H. Membrane type-1 matrix
801 metalloproteinase activity is regulated by the endocytic collagen receptor Endo180. *Journal of*
802 *cell science*. 2009;122(Pt 22):4042–4048.
803 21. Lee T-H, McKleroy W, Khalifeh-Soltani A, Sakuma S, Lazarev S, Riento K, Nishimura
804 SL, Nichols BJ, Atabai K. Functional genomic screen identifies novel mediators of collagen
805 uptake. *Molecular biology of the cell*. 2014;25(5):583–93.
806 22. Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C,
807 Panning B, Ploegh HL, Bassik MC, Qi LS, Kampmann M, Weissman JS. Genome-Scale
808 CRISPR-Mediated Control of Gene Repression and Activation. *Cell*. 2014;159(3):647–661.
809 23. Madsen DH, Ingvarsen S, Jürgensen HJ, Melander MC, Kjøller L, Moyer A, Honoré C,
810 Madsen CA, Garred P, Burgdorf S, Bugge TH, Behrendt N, Engelholm LH. The non-phagocytic
811 route of collagen uptake: A distinct degradation pathway. *Journal of Biological Chemistry*. 2011.
812 24. Haney MS, Bohlen CJ, Morgens DW, Ousey JA, Barkal AA, Tsui CK, Ego BK, Levin R,
813 Kamber RA, Collins H, Tucker A, Li A, Vorselen D, Labitigan L, Crane E, et al. Identification of
814 phagocytosis regulators using magnetic genome-wide CRISPR screens. *Nature genetics*.
815 2018;50(12):1716–1727.
816 25. Stefanovic B, Stefanovic L, Schnabl B, Bataller R, Brenner DA. TRAM2 protein interacts
817 with endoplasmic reticulum Ca2+ pump Serca2b and is necessary for collagen type I synthesis.
818 *Molecular and cellular biology*. 2004;24(4):1758–68.
819 26. Bucala R, Spiegel LA, Chesney J, Hogan M, Cerami A. Circulating fibrocytes define a
820 new leukocyte subpopulation that mediates tissue repair. *Molecular medicine (Cambridge,*
821 *Mass.)*. 1994;1(1):71–81.
822 27. Phillips RJ, Burdick MD, Hong K, Lutz MA, Murray LA, Xue YY, Belperio JA, Keane MP,
823 Strieter RM. Circulating fibrocytes traffic to the lungs in response to CXCL12 and mediate
824 fibrosis. *The Journal of clinical investigation*. 2004;114(3):438–46.
825 28. Kleaveland KR, Velikoff M, Yang J, Agarwal M, Rippe RA, Moore BB, Kim KK.
826 Fibrocytes are not an essential source of type I collagen during lung fibrosis. *Journal of*
827 *immunology (Baltimore, Md. : 1950)*. 2014;193(10):5229–39.
828 29. Simões FC, Cahill TJ, Kenyon A, Gavriouchkina D, Vieira JM, Sun X, Pezzolla D,
829 Ravaud C, Masmanian E, Weinberger M, Mayes S, Lemieux ME, Barnette DN, Gunadasa-
830 Rohling M, Williams RM, et al. Macrophages directly contribute collagen to scar formation
831 during zebrafish heart regeneration and mouse heart repair. *Nature Communications*.
832 2020;11(1):600.
833 30. Ozhogina OA, Trexler M, Bányai L, Llinás M, Patthy L. Origin of fibronectin type II (FN2)
834 modules: structural analyses of distantly-related members of the kringle family idey the kringle
835 domain of neurotrypsin as a potential link between FN2 domains and kringles. *Protein science :*
836 *a publication of the Protein Society*. 2001;10(10):2114–22.
837 31. Bányai L, Patthy L. Evidence for the involvement of type II domains in collagen binding
838 by 72 kDa type IV procollagenase. *FEBS letters*. 1991;282(1):23–5.
839 32. Pickford AR, Potts JR, Bright JR, Phan I, Campbell ID. Solution structure of a type 2
840 module from fibronectin: implications for the structure and function of the gelatin-binding
841 domain. *Structure*. 1997;5(3):359–370.
842 33. Sun S, Shi G, Han X, Francisco AB, Ji Y, Mendonça N, Liu X, Locasale JW, Simpson
843 KW, Duhamel GE, Kersten S, Yates JR, Long Q, Qi L. Sel1L is indispensable for mammalian

844 endoplasmic reticulum-associated degradation, endoplasmic reticulum homeostasis, and
845 survival. *Proceedings of the National Academy of Sciences*. 2014;111(5).
846 34. Christianson JC, Olzmann JA, Shaler TA, Sowa ME, Bennett EJ, Richter CM, Tyler RE,
847 Greenblatt EJ, Wade Harper J, Kopito RR. Defining human ERAD networks through an
848 integrative mapping strategy. *Nature Cell Biology*. 2012;14(1):93–105.
849 35. Mueller B, Lilley BN, Ploegh HL. SEL1L, the homologue of yeast Hrd3p, is involved in
850 protein dislocation from the mammalian ER. *The Journal of cell biology*. 2006;175(2):261–70.
851 36. Gardner RG, Swarbrick GM, Bays NW, Cronin SR, Wilhovsky S, Seelig L, Kim C,
852 Hampton RY. Endoplasmic Reticulum Degradation Requires Lumen to Cytosol Signaling.
853 *Journal of Cell Biology*. 2000;151(1):69–82.
854 37. Hampton RY, Gardner RG, Rine J. Role of 26S proteasome and HRD genes in the
855 degradation of 3-hydroxy-3-methylglutaryl-CoA reductase, an integral endoplasmic reticulum
856 membrane protein. *Molecular Biology of the Cell*. 1996;7(12):2029–2044.
857 38. Bordallo J, Plemper RK, Finger A, Wolf DH. Der3p/Hrd1p Is Required for Endoplasmic
858 Reticulum-associated Degradation of Misfolded Lumenal and Integral Membrane Proteins.
859 *Molecular Biology of the Cell*. 1998;9(1):209–222.
860 39. Plemper RK, Bordallo J, Deak PM, Taxis C, Hitt R, Wolf DH. Genetic interactions of
861 Hrd3p and Der3p/Hrd1p with Sec61p suggest a retro-translocation complex mediating protein
862 transport for ER degradation. *Journal of Cell Science*. 1999;112(22):4123–4134.
863 40. Francisco AB, Singh R, Li S, Vani AK, Yang L, Munroe RJ, Diaferia G, Cardano M,
864 Biunno I, Qi L, Schimenti JC, Long Q. Deficiency of Suppressor Enhancer Lin12 1 Like (SEL1L)
865 in Mice Leads to Systemic Endoplasmic Reticulum Stress and Embryonic Lethality. *Journal of*
866 *Biological Chemistry*. 2010;285(18):13694–13703.
867 41. Yagishita N, Ohneda K, Amano T, Yamasaki S, Sugiura A, Tsuchimochi K, Shin H,
868 Kawahara K, Ohneda O, Ohta T, Tanaka S, Yamamoto M, Maruyama I, Nishioka K, Fukamizu
869 A, et al. Essential Role of Synoviolin in Embryogenesis. *Journal of Biological Chemistry*.
870 2005;280(9):7909–7916.
871 42. Biunno I, Castiglioni B, Rogozin IB, DeBellis G, Malferrari G, Cattaneo M. Cross-species
872 conservation of SEL1L, a human pancreas-specific expressing gene. *Omics : a journal of*
873 *integrative biology*. 2002;6(2):187–98.
874 43. Exposito J-Y, Cluzel C, Garrone R, Lethias C. Evolution of collagens. *The Anatomical*
875 *Record*. 2002;268(3):302–316.
876 44. Exposito JY, le Guellec D, Lu Q, Garrone R. Short chain collagens in sponges are
877 encoded by a family of closely related genes. *Journal of Biological Chemistry*.
878 1991;266(32):21923–21928.
879 45. Fidler AL, Vanacore RM, Chetyrkin S v, Pedchenko VK, Bhave G, Yin VP, Stothers CL,
880 Rose KL, McDonald WH, Clark TA, Borza D-B, Steele RE, Ivy MT, Aspirnauts, Hudson JK, et al.
881 A unique covalent bond in basement membrane is a primordial innovation for tissue evolution.
882 *Proceedings of the National Academy of Sciences of the United States of America*.
883 2014;111(1):331–6.
884 46. Hecker L, Logsdon NJ, Kurundkar D, Kurundkar A, Bernard K, Hock T, Meldrum E,
885 Sanders YY, Thannickal VJ. Reversal of persistent fibrosis in aging by targeting Nox4-Nrf2
886 redox imbalance. *Science Translational Medicine*. 2014;6(231).
887 47. Peng R, Sridhar S, Tyagi G, Phillips JE, Garrido R, Harris P, Burns L, Renteria L, Woods
888 J, Chen L, Allard J, Ravindran P, Bitter H, Liang Z, Hogaboam CM, et al. Bleomycin induces
889 molecular changes directly relevant to idiopathic pulmonary fibrosis: a model for “active”
890 disease. *PloS one*. 2013;8(4):e59348.
891 48. Morse C, Tabib T, Sembrat J, Buschur KL, Bittar HT, Valenzi E, Jiang Y, Kass DJ,
892 Gibson K, Chen W, Mora A, Benos P v, Rojas M, Lafyatis R. Proliferating SPP1/MERTK-
893 expressing macrophages in idiopathic pulmonary fibrosis. *The European respiratory journal*.
894 2019;54(2).

895 49. Reyfman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S,
896 Fernandez R, Akbarpour M, Chen C-I, Ren Z, Verma R, Abdala-Valencia H, Nam K, Chi M, Han
897 S, et al. Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the
898 Pathobiology of Pulmonary Fibrosis. *American journal of respiratory and critical care medicine*.
899 2019;199(12):1517–1536.
900 50. Habermann AC, Gutierrez AJ, Bui LT, Yahn SL, Winters NI, Calvi CL, Peter L, Chung M-
901 I, Taylor CJ, Jetter C, Raju L, Roberson J, Ding G, Wood L, Sucre JMS, et al. Single-cell RNA
902 sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in
903 pulmonary fibrosis. *Science advances*. 2020;6(28):eaba1972.
904 51. Adams TS, Schupp JC, Poli S, Ayaub EA, Neumark N, Ahangari F, Chu SG, Raby BA,
905 DeIuliis G, Januszyk M, Duan Q, Arnett HA, Siddiqui A, Washko GR, Homer R, et al. Single-cell
906 RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary
907 fibrosis. *Science advances*. 2020;6(28):eaba1983.
908 52. Tsukui T, Sun K-H, Wetter JB, Wilson-Kanamori JR, Hazelwood LA, Henderson NC,
909 Adams TS, Schupp JC, Poli SD, Rosas IO, Kaminski N, Matthay MA, Wolters PJ, Sheppard D.
910 Collagen-producing lung cell atlas identifies multiple subsets with distinct localization and
911 relevance to fibrosis. *Nature communications*. 2020;11(1):1920.
912 53. Weathington NM, van Houwelingen AH, Noerager BD, Jackson PL, Kraneveld AD, Galin
913 FS, Folkerts G, Nijkamp FP, Blalock JE. A novel peptide CXCR ligand derived from extracellular
914 matrix degradation during airway inflammation. *Nature Medicine*. 2006;12(3):317–323.
915 54. Riley DJ, Berg RA, Soltys RA, Kerr JS, Guss HN, Curran SF, Laskin DL. Neutrophil
916 Response following Intratracheal Instillation of Collagen Peptides into Rat Lungs. *Experimental*
917 *Lung Research*. 1988;14(4):549–563.
918 55. Jürgensen HJ, Nørregaard KS, Sibree MM, Santoni-Rugiu E, Madsen DH, Wassilew K,
919 Krustrup D, Garred P, Bugge TH, Engelholm LH, Behrendt N. Immune regulation by fibroblasts
920 in tissue injury depends on uPARAP-mediated uptake of collectins. *Journal of Cell Biology*.
921 2019;218(1):333–349.
922 56. Li L, Chen X, Zhang H, Wang M, Lu W. MRC2 Promotes Proliferation and Inhibits
923 Apoptosis of Diabetic Nephropathy. *Analytical Cellular Pathology*. 2021;2021:1–10.
924 57. Onursal C, Dick E, Angelidis I, Schiller HB, Staab-Weijnitz CA. Collagen Biosynthesis,
925 Processing, and Maturation in Lung Ageing. *Frontiers in Medicine*. 2021;8.
926 58. Podolsky MJ, Gupta D, Ha A, Ta R, Khalifeh-Soltani A, McKleroy W, Datta R, Sheppard
927 D, Atabai K. Cell division cycle 7 kinase is a negative regulator of cell-mediated collagen
928 degradation. *American Journal of Physiology - Lung Cellular and Molecular Physiology*.
929 2018;315(3):L360–L370.
930 59. Boettcher M, Tian R, Blau JA, Markegard E, Wagner RT, Wu D, Mo X, Biton A, Zaitlen
931 N, Fu H, McCormick F, Kampmann M, McManus MT. Dual gene activation and knockout screen
932 reveals directional dependencies in genetic networks. *Nature biotechnology*. 2018;36(2):170–
933 178.
934 60. Boettcher M, Covarrubias S, Biton A, Blau J, Wang H, Zaitlen N, McManus MT. Tracing
935 cellular heterogeneity in pooled genetic screens via multi-level barcoding. *BMC genomics*.
936 2019;20(1):107.
937 61. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, Irizarry RA, Liu JS, Brown M, Liu XS.
938 MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9
939 knockout screens. *Genome biology*. 2014;15(12):554.
940 62. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich
941 A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a
942 knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of*
943 *the National Academy of Sciences of the United States of America*. 2005;102(43):15545–50.
944 63. Mootha VK, Lindgren CM, Eriksson K-F, Subramanian A, Sihag S, Lehar J, Puigserver
945 P, Carlsson E, Ridderstråle M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub

946 TR, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately
947 downregulated in human diabetes. *Nature genetics*. 2003;34(3):267–73.
948 64. Altenhoff AM, Train C-M, Gilbert KJ, Mediratta I, Mendes de Farias T, Moi D, Nevers Y,
949 Radoykova H-S, Rossier V, Warwick Vesztrocy A, Glover NM, Dessimoz C. OMA orthology in
950 2021: website overhaul, conserved isoforms, ancestral gene order and more. *Nucleic acids*
951 *research*. 2021;49(D1):D373–D379.
952 65. Shrestha N, Liu T, Ji Y, Reinert RB, Torres M, Li X, Zhang M, Tang C-HA, Hu C-CA, Liu
953 C, Naji A, Liu M, Lin JD, Kersten S, Arvan P, et al. Sel1L-Hrd1 ER-associated degradation
954 maintains β cell identity via TGF-β signaling. *The Journal of clinical investigation*.
955 2020;130(7):3499–3510.
956 66. Zhou Z, Torres M, Sha H, Halbrook CJ, van den Bergh F, Reinert RB, Yamada T, Wang
957 S, Luo Y, Hunter AH, Wang C, Sanderson TH, Liu M, Taylor A, Sesaki H, et al. Endoplasmic
958 reticulum–associated degradation regulates mitochondrial dynamics in brown adipocytes.
959 *Science*. 2020;368(6486):54–60.
960 67. Bhattacharya A, Sun S, Wang H, Liu M, Long Q, Yin L, Kersten S, Zhang K, Qi L.
961 Hepatic Sel1L-Hrd1 ER-associated degradation (ERAD) manages FGF21 levels and systemic
962 metabolism via CREBH. *The EMBO Journal*. 2018;37(22).
963 68. Yan C, Xu X, Zou X. Fully Blind Docking at the Atomic Level for Protein-Peptide
964 Complex Structure Prediction. *Structure*. 2016;24(10):1842–1853.
965 69. Xu X, Yan C, Zou X. MDockPeP: An *ab-initio* protein–peptide docking server. *Journal of*
966 *Computational Chemistry*. 2018;39(28):2409–2413.
967 70. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong
968 Q, Kinch LN, Schaeffer RD, Millán C, Park H, Adams C, Glassman CR, DeGiovanni A, et al.
969 Accurate prediction of protein structures and interactions using a three-track neural network.
970 *Science*. 2021;373(6557):871–876.
971

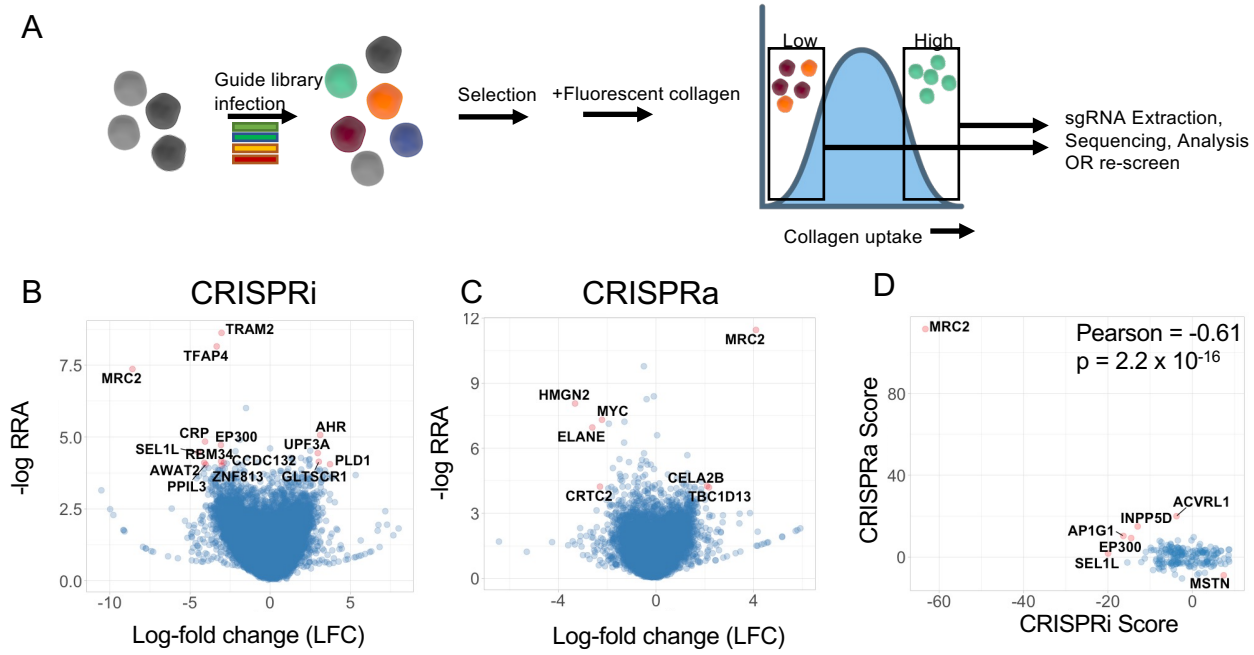
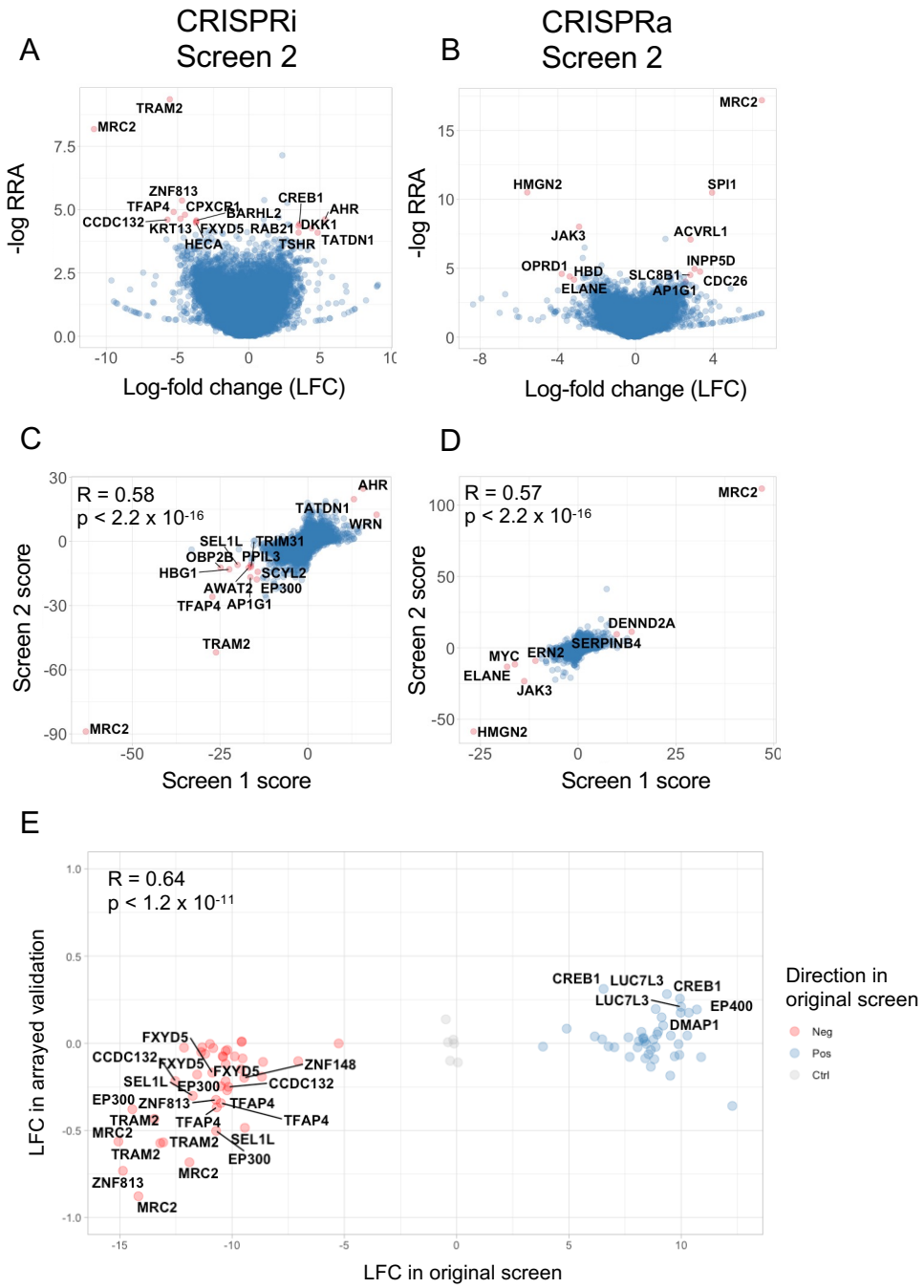


Figure 1. CRISPR screens identify mediators of cellular collagen uptake. (A) Schematic of experimental approach for flow cytometry-based screening for uptake of collagen fragments. (B) Volcano plot of gene-level statistics of CRISPRi (inhibition) screen. Most significant gene hits are labeled and colored in red. RRA = Robust Rank Aggregation score; LFC = Log-fold change. (C) Volcano plot of gene-level statistics of CRISPRa (activation) screen. (D) Comparison of overall phenotype score between CRISPRi and CRISPRa screens.



Supplemental Figure 1. Re-screens are correlated to original CRISPR screen results. (A) Volcano plot of CRISPRi re-screen (Screen 2). (B) Volcano plot of CRISPRa re-screen (Screen 2). (C) Comparison of gene-level overall phenotype scores between CRISPRi screens 1 and 2. (D) Comparison of gene-level overall phenotype scores between CRISPRa screens 1 and 2. (E) Comparison of arrayed log-fold change in arrayed validation vs. original pooled screen log-fold change; each point represents unique guide sequence. 65/84 guides demonstrated concordance with original pooled screen data (77%). Pearson correlation coefficient is shown.

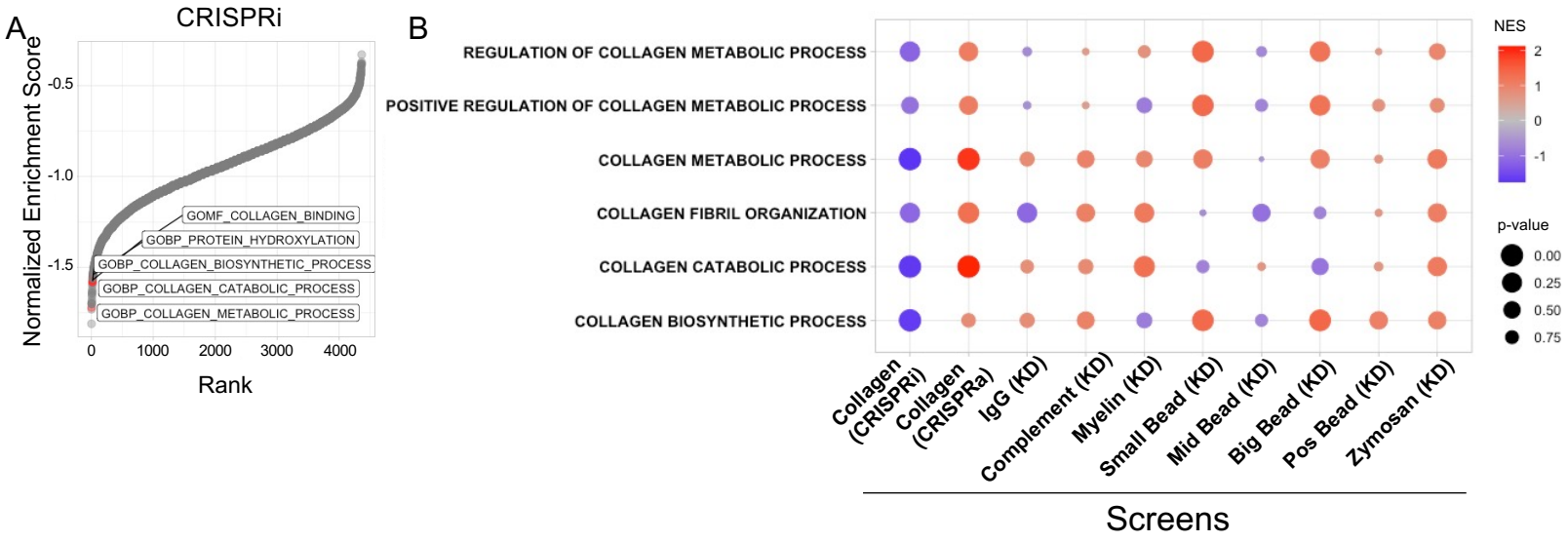
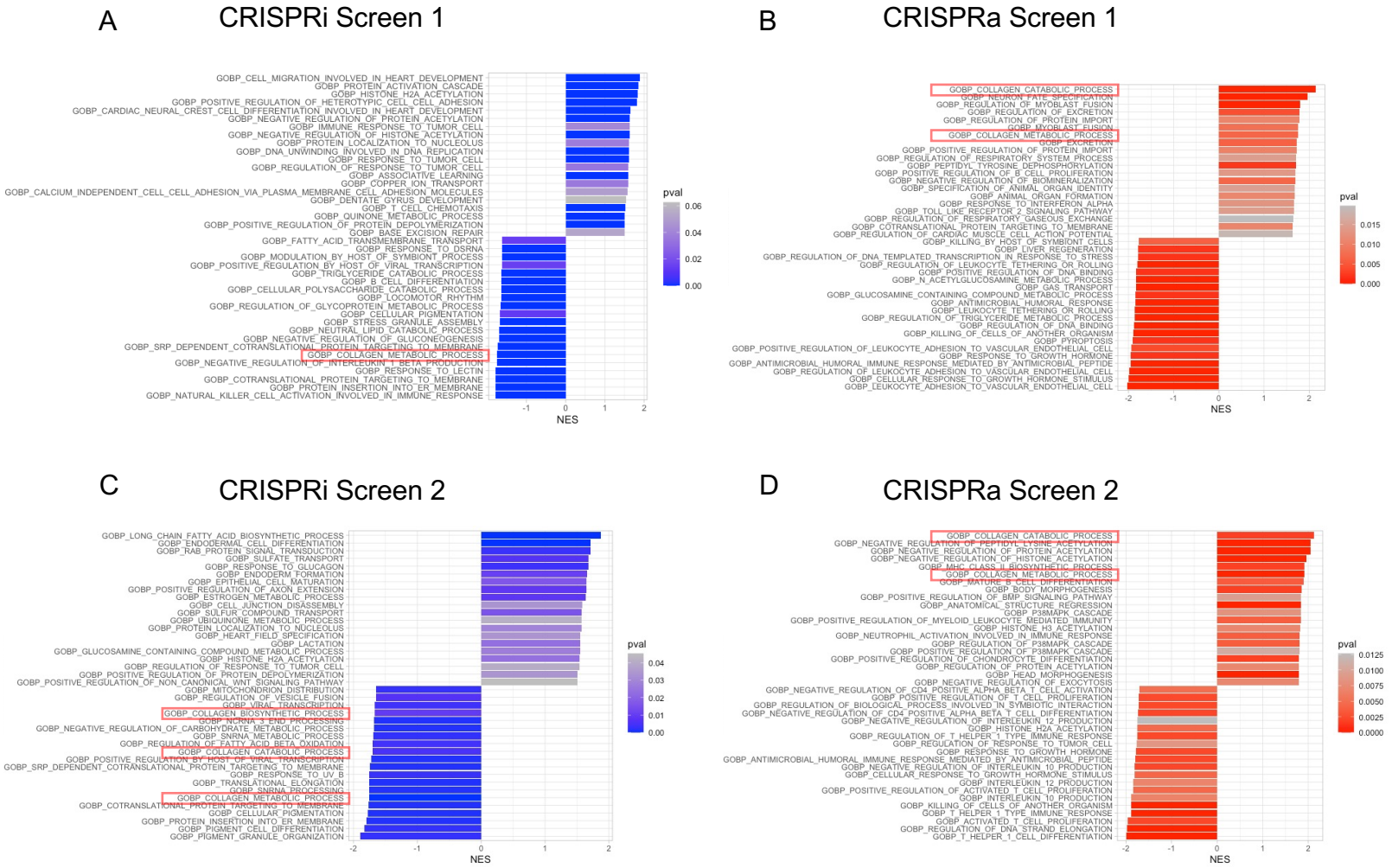


Figure 2. Collagen-biosynthetic pathways are regulators of collagen uptake uniquely compared to screens for uptake of other cargoes. (A) Ranked plot of Normalized Enrichment Scores derived from Gene-Set Enrichment Analysis (GSEA) of CRISPRi phenotype scores, using 10,561 Gene Ontology gene sets. Collagen-related gene sets are labeled and displayed in red. (B) Comparison of GSEA statistics for labeled collagen-related gene sets across screens for uptake of various cargoes[22]; screen method is indicated in parentheses. KD = Knock-down.



Supplemental Figure 2. Collagen-related gene ontology terms are enriched as regulators of collagen uptake in CRISPRi and CRISPRa screens. The top twenty positive and negatively performing gene sets are indicated with associated statistics for each screen as indicated, among all 10,561 Gene Ontology gene sets. Collagen-related gene sets are highlighted with red boxes (i.e. those with ‘Collagen’ in the GO gene set name). NES = Normalized Enrichment Score.

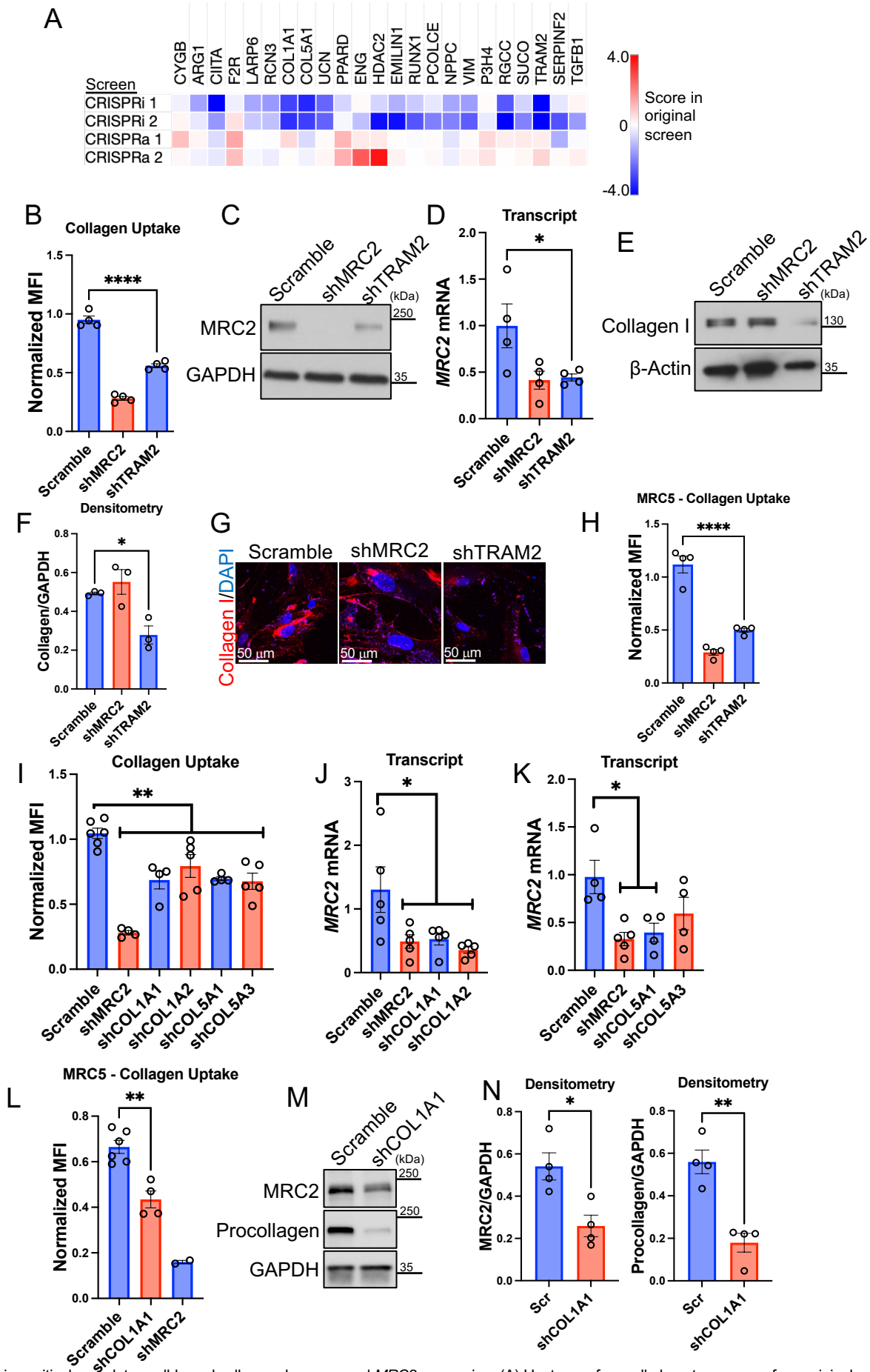
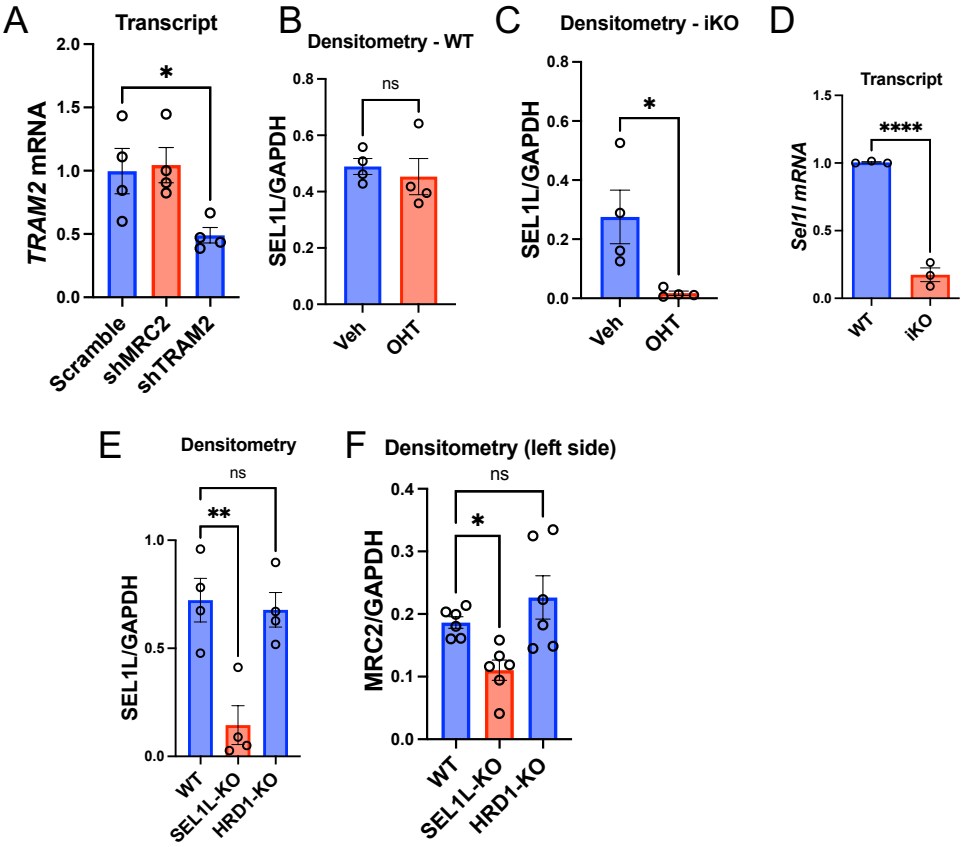


Figure 3. Collagen biosynthesis positively regulates cell-based collagen clearance and *MRC2* expression. (A) Heatmap of overall phenotype scores from original screens as indicated for top genes from 'Collagen Biosynthetic Process' GO gene-set (displayed genes had most extreme 'Score' values in one or more screen result). (B) Flow-cytometry based collagen-uptake assay in U937 cells after shRNA treatment as indicated vs. Scramble control. MFI = Mean Fluorescence Intensity. N=4 per group. (C) Western blot of U937 cells after shRNA treatment as indicated vs. Scramble control, representative of N=4 independent experiments. (D) Q-RT-PCR in U937 cells after shRNA treatment as indicated vs. Scramble control. N=4 per group. (E-F) Western blot and densitometry of MRC5 fibroblasts after shRNA treatment as indicated vs. Scramble control, representative of N=3 independent experiments. (G) Representative confocal immunofluorescence images of MRC5 fibroblasts after shRNA treatment as indicated vs. Scramble control. (H) Flow-cytometry based collagen-uptake assay in MRC5 cells after shRNA treatment as indicated vs. Scramble control. N=4 per group. (I) Flow-cytometry based collagen-uptake assay in U937 cells after shRNA treatment as indicated vs. Scramble control. N=4-6 per group. (J-K) Q-RT-PCR in U937 cells after shRNA treatment as indicated vs. Scramble control. N=4-5 per group. (L) Flow-cytometry based collagen-uptake assay in MRC5 fibroblasts after shRNA treatment as indicated vs. Scramble control. N=6 and 4 in Scramble and shCOL1A1 groups respectively, N=2 in shMRC2 group. (M-N) Western blot and densitometry of MRC5 fibroblasts after shRNA treatment as indicated vs. Scramble control, representative of N=4 independent experiments. Statistics: (B-K) ANOVA with post-hoc Bonferroni testing; (L-N) Student's *t*-test. * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$.



Supplemental Figure 3. Validation of different tools used in this manuscript. (A) Q-RT-PCR in U937 cells after shRNA treatment as indicated vs. Scramble control. N=4 per group. (B-C) Additional densitometry of data from Figure 5C. N=4 per group. (D) Q-RT-PCR in MEFs after OHT treatment for genes as indicated. N=3 per group. (E) Additional densitometry of data from Figure 6B. N=4 per group. (F) Densitometry of the left side of Figure 6E. N=6 per group. Statistics: (A, E, F) ANOVA with post-hoc Bonferroni testing; (B-D) Student's *t*-test. * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$, ns = not significant.

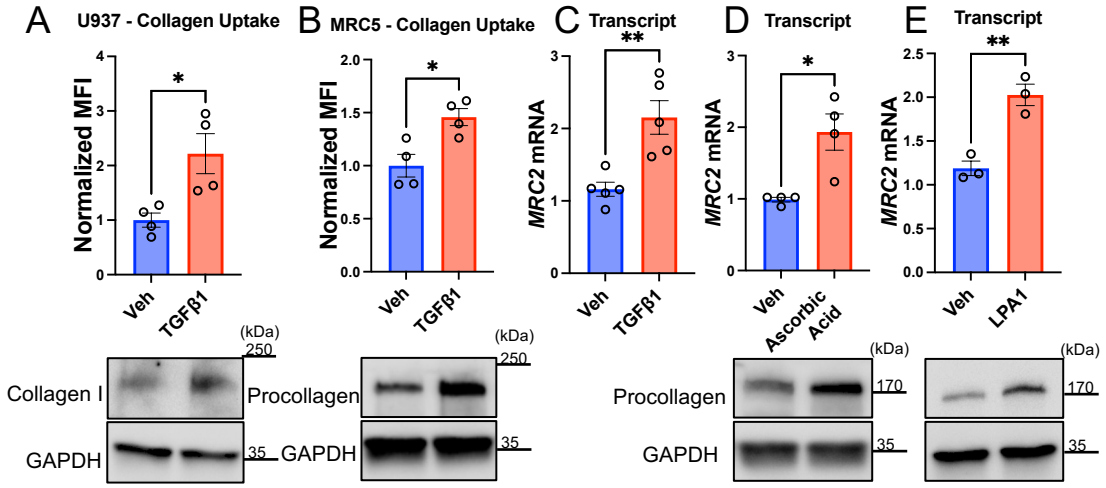


Figure 4. Pharmacological stimulation of collagen synthesis is positively associated with collagen uptake and *MRC2* expression. (A-B) Flow-cytometry based collagen-uptake assay in U937 (A) or MRC5 (B) cells after drug treatment as indicated vs. vehicle control. N=4 per group. Western blot of corresponding cells and conditions shown below, each representative of N=3-4 independent experiments. (C-E) Q-RT-PCR in MRC5 cells after drug treatment as indicated vs. vehicle control. N=3-5 per group. Western blot of corresponding cells and conditions shown below, each representative of N=3 independent experiments. Statistics: (A-E) Student's *t*-test. * $p < 0.05$, ** $p < 0.01$.

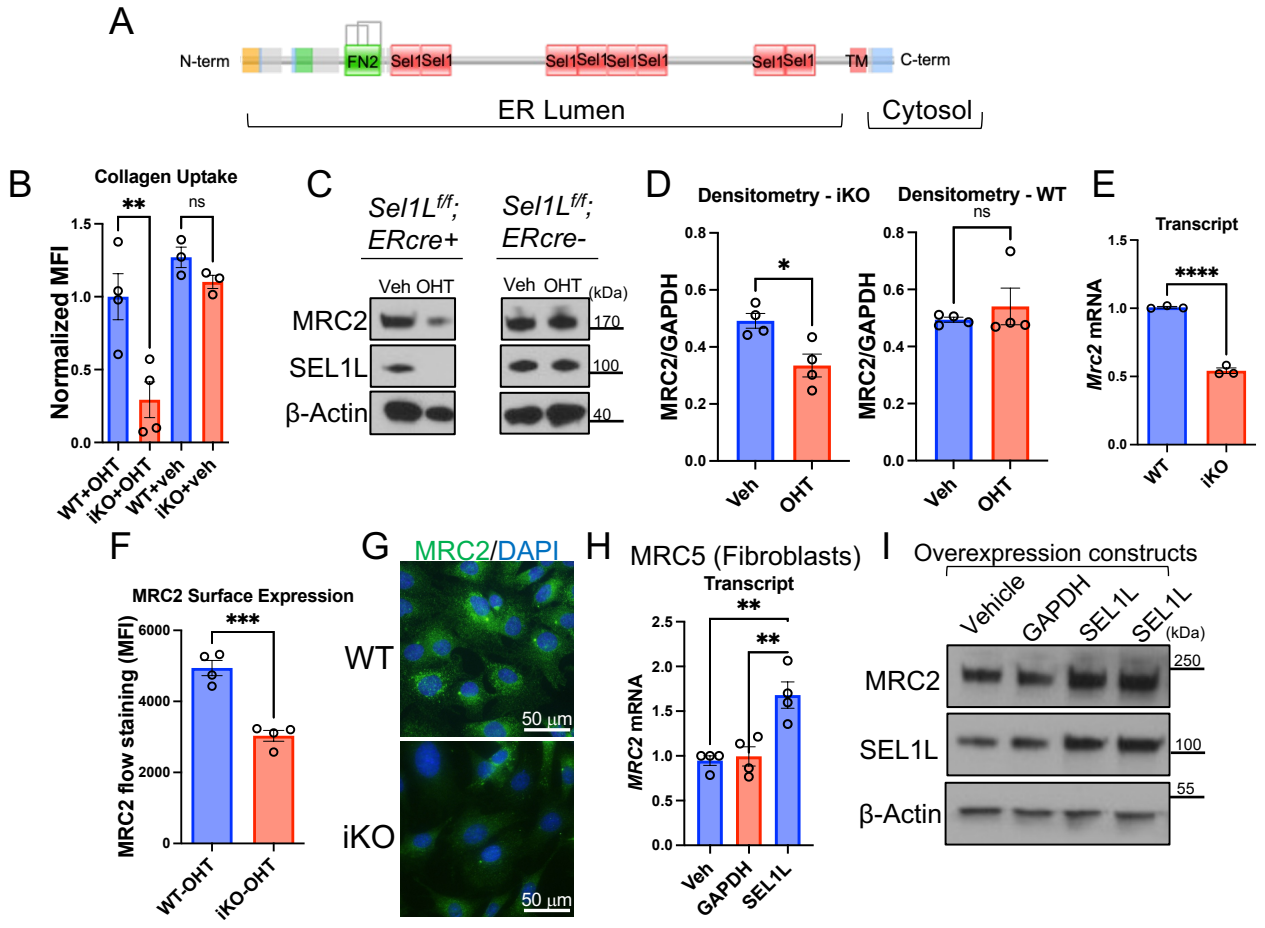


Figure 5. *Sel1L* positively regulates collagen uptake and *Mrc2* expression. (A) Schematic of domain structure of SEL1L protein. FN2 = Fibronectin-2. (B) Flow-cytometry based collagen-uptake assay in MEF cells (mouse embryonic fibroblasts) after treatment with 4-hydroxytamoxifen (OHT) as indicated vs. vehicle control. WT = *Sel1L*$^{f/f;}$*ERcre*-, iKO = *Sel1L*$^{f/f}$*;ERcre*+. N=3-4 per group. (C-D) Western blot and densitometry of MEFs after OHT treatment as indicated vs. control, representative of N=4 independent experiments. (E) Q-RT-PCR in MEFs after OHT treatment for genes as indicated. N=3 per group. (F) Flow-cytometry based measurement of MRC2 surface expression in MEFs after OHT treatment. N=4 per group. (G) Representative wide-field immunofluorescence images of MEFs after OHT treatment. (H) Q-RT-PCR of MRC5 cells transfected with overexpression vectors containing genes as indicated vs. vehicle control (lipofectamine only). N=4 per group. (I) Western blot of MRC5 cells transfected with overexpression vectors containing genes as indicated vs. vehicle control (lipofectamine only). N=2. Statistics: (B, H) ANOVA with post-hoc Bonferroni testing; (D - right panel, E, F) Student's *t*-test; (D - left panel) Mann-Whitney *U* test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, ns = not significant.

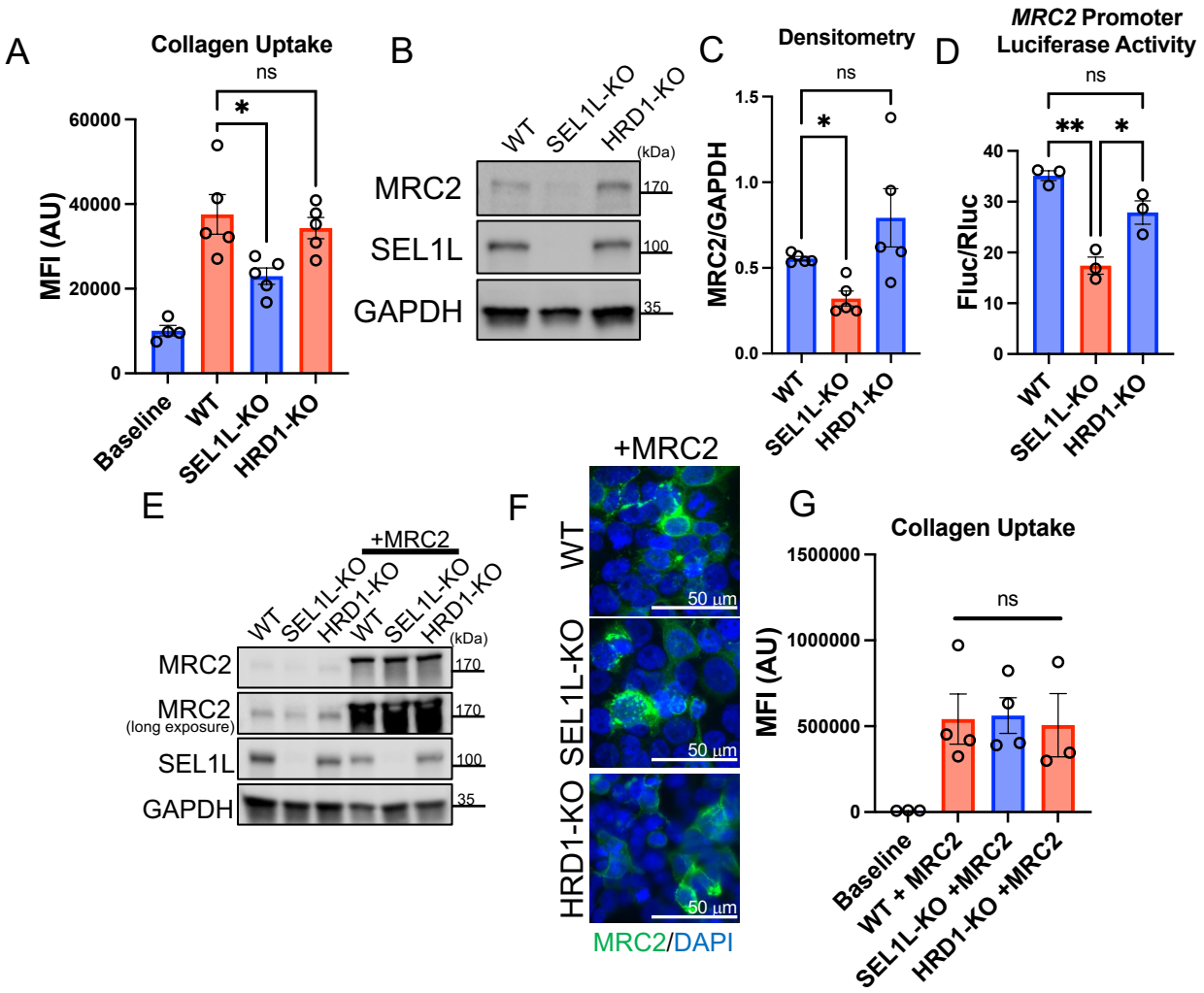
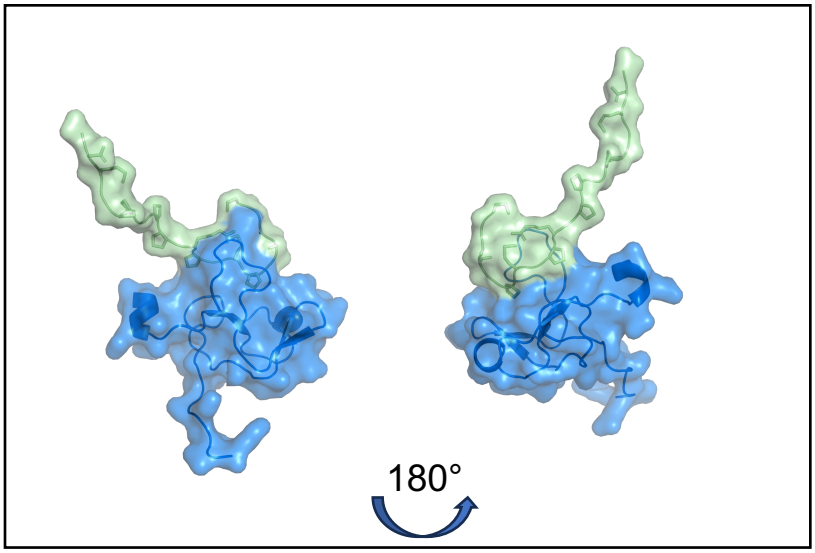


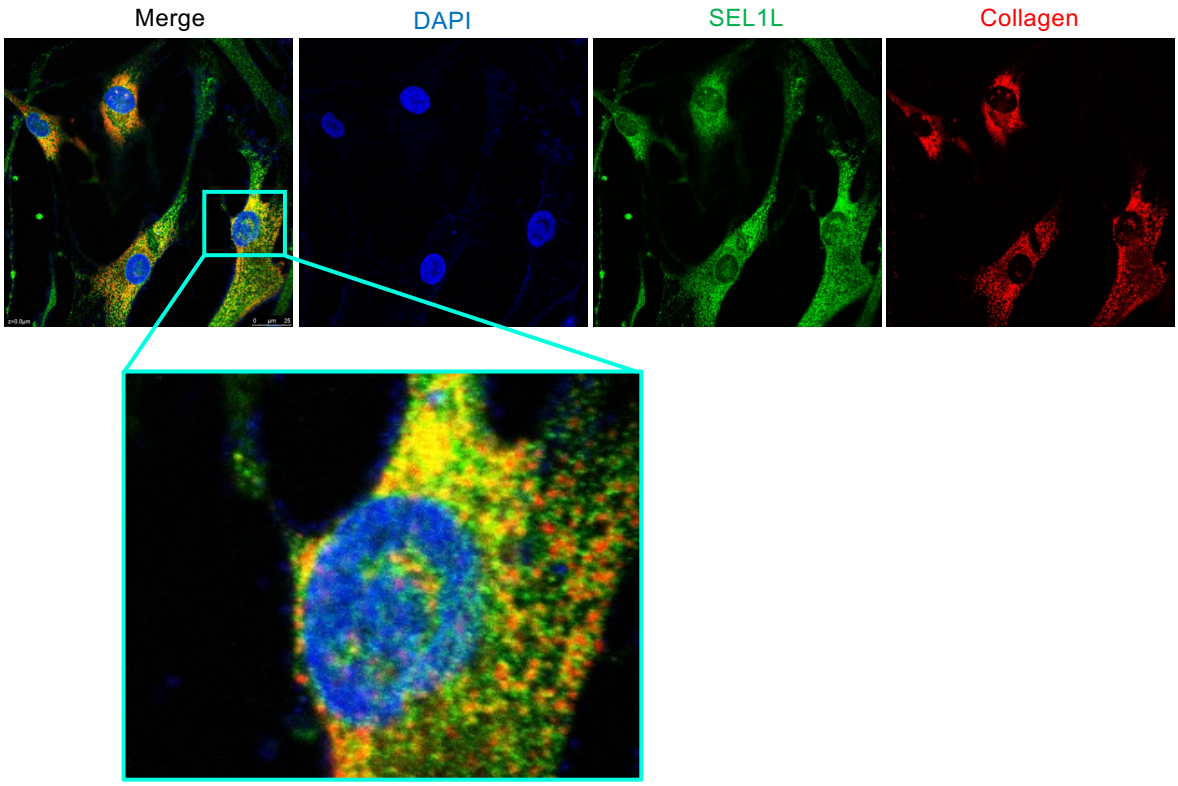
Figure 6. SEL1L, but not HRD1, positively regulates MRC2. A) Flow-cytometry based collagen-uptake assay in 293T of different genotypes (compared with baseline control without fluorescent collagen). N=4-5 per group. (B-C) Western blot and densitometry of 293T cells of different genotypes, representative of N=5 independent experiments. (D) Firefly luciferase assay of *MRC2* promoter activity, normalized to constitutive Renilla luciferase in 293T cells of different genotypes. N=3 per group. (E) Western blot of HEK293T cells of different genotypes transfected with MRC2 overexpression vector vs. vehicle control, representative of N=6 independent experiments. (F) Representative wide-field immunofluorescence images of 293T cells of different genotypes transfected with MRC2 overexpression vector vs. vehicle control. (G) Flow-cytometry based collagen-uptake assay in HEK293T of different genotypes transfected with MRC2 overexpression vector (compared with baseline control without fluorescent collagen; note different scale compared with panel A). N=3-4 per group. Statistics: (A, C, D, G) ANOVA with post-hoc Bonferroni testing. * $p < 0.05$ ** $p < 0.01$, ns = not significant.

A



Green: collagen peptide; Blue: FN2 domain of SEL1L

B



Supplemental Figure 4. Additional support of SEL1L and collagen interaction. (A) Predicted docking of collagen peptide with FN2 domain of SEL1L. (B) Additional representative confocal immunofluorescence image of MRC5 fibroblasts at baseline (as in Figure 7A), with decreased PMT gain in the Collagen channel, along with magnified inset to highlight co-localization.

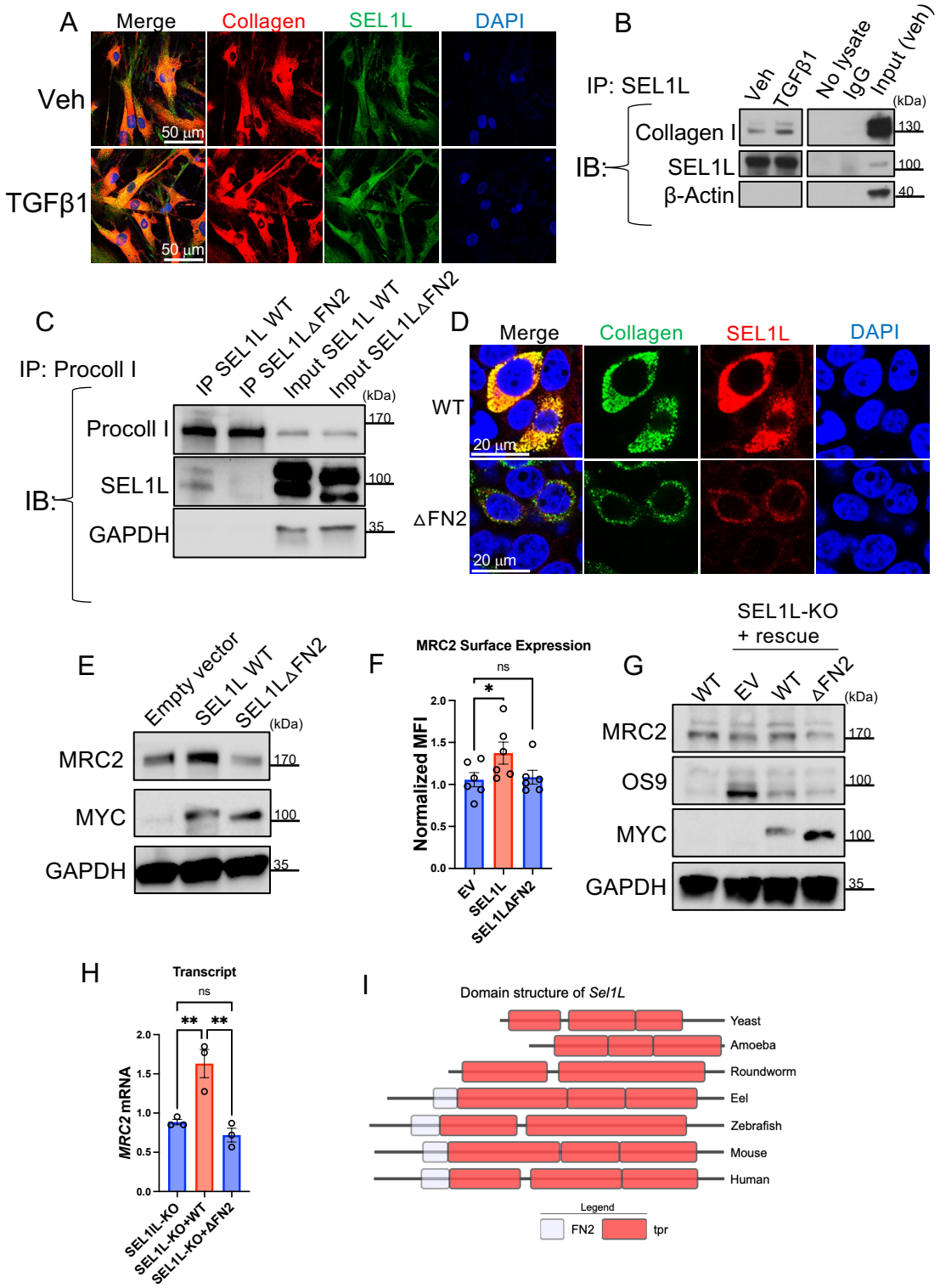
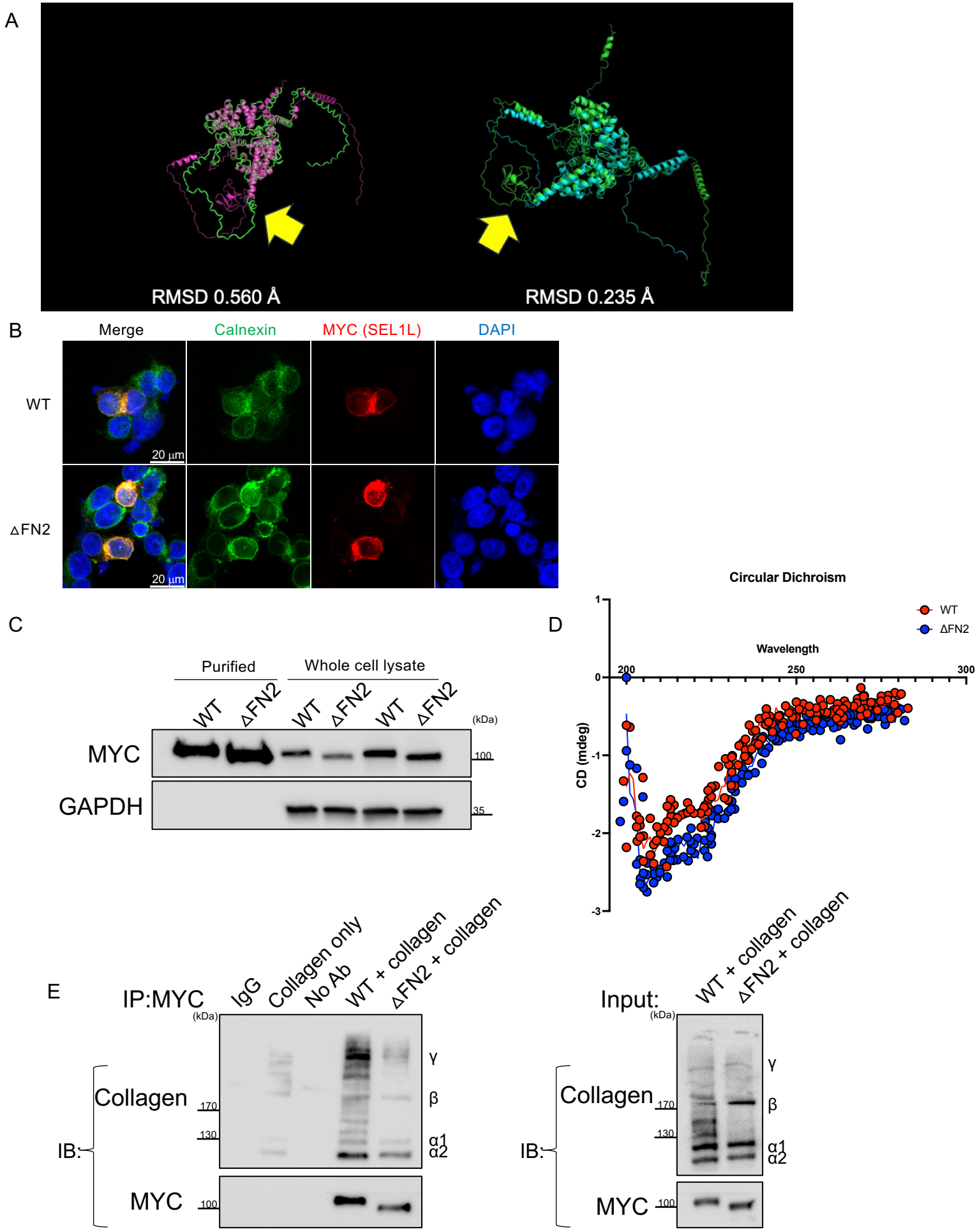


Figure 7. SEL1L binds collagen as it is being synthesized via its FN2 domain and this interaction is necessary for regulation of *MRC2* expression. (A) Representative confocal immunofluorescence images of MRC5 fibroblasts treated as indicated or vehicle control. (B) Western blot of MRC5 fibroblasts after immunoprecipitation with anti-SEL1L antibody (vs. controls as indicated), representative of N=2 independent experiments. (C) Western blot after immunoprecipitation with anti-procollagen type I antibody (with input controls as indicated) of HEK293T cells transfected with collagen and WT or mutant SEL1L constructs as indicated, representative of N=3 independent experiments. (D) Representative confocal immunofluorescence images of HEK293T cells transfected with collagen and WT or mutant SEL1L constructs as indicated. (E) Western blot of MEFs after treatment with WT or mutant SEL1L constructs as indicated, representative of N=3 independent experiments. (F) Flow-cytometry based measurement of MRC2 surface expression on MEFs after treatment with WT or mutant SEL1L constructs as indicated, vs. Empty Vector (EV). N=6 per group. (G) Western blot of HEK293T cells either WT or KO for SEL1L as indicated after treatment with EV (empty vector), WT or mutant SEL1L constructs as indicated, representative of N=3 independent experiments. (H) Q-RT-PCR of SEL1L-KO HEK293T cells after transfection with WT or mutant SEL1L as indicated vs. empty vehicle control. N=3 per group. (I) Comparison of protein domain structures of Sel1L orthologs across phylogenetically diverse organisms; FN2 = Fibronectin-2 domain, tpr = Tetratricopeptide repeat domain. Statistics: (F-G) ANOVA with post-hoc Bonferroni testing. * $p < 0.05$, ** $p < 0.01$, ns = not significant.



Supplemental Figure 5. FN2 domain deletion mutant interferes with collagen binding but is otherwise stable. (A) Alignment of predicted structures of full-length SEL1L and ΔFN2 SEL1L. Two different views and alignments are shown. Root-mean squared distance (RMSD) is displayed. Yellow arrow indicates the FN2 domain (in the pink structure on the left; in the green structure on the right). (B) Representative confocal immunofluorescence images of HEK293T cells transduced with MYC-tagged WT or ΔFN2 SEL1L constructs, co-stained with ER marker Calnexin. (C) Western blot of purified WT or ΔFN2 SEL1L proteins compared with input cell lysates, representative of N=3 independent experiments. (D) Circular Dichroism spectra of purified WT or ΔFN2 proteins. (E) Western blot of co-incubated purified either WT or ΔFN2 SEL1L proteins with rat tail collagen, after immunoprecipitation with anti-MYC antibody (with input and additional controls as indicated), representative of N=3 independent experiments. 'Collagen only' indicates incubation of rat tail collagen with IP beads without co-incubation with purified MYC-tagged protein.

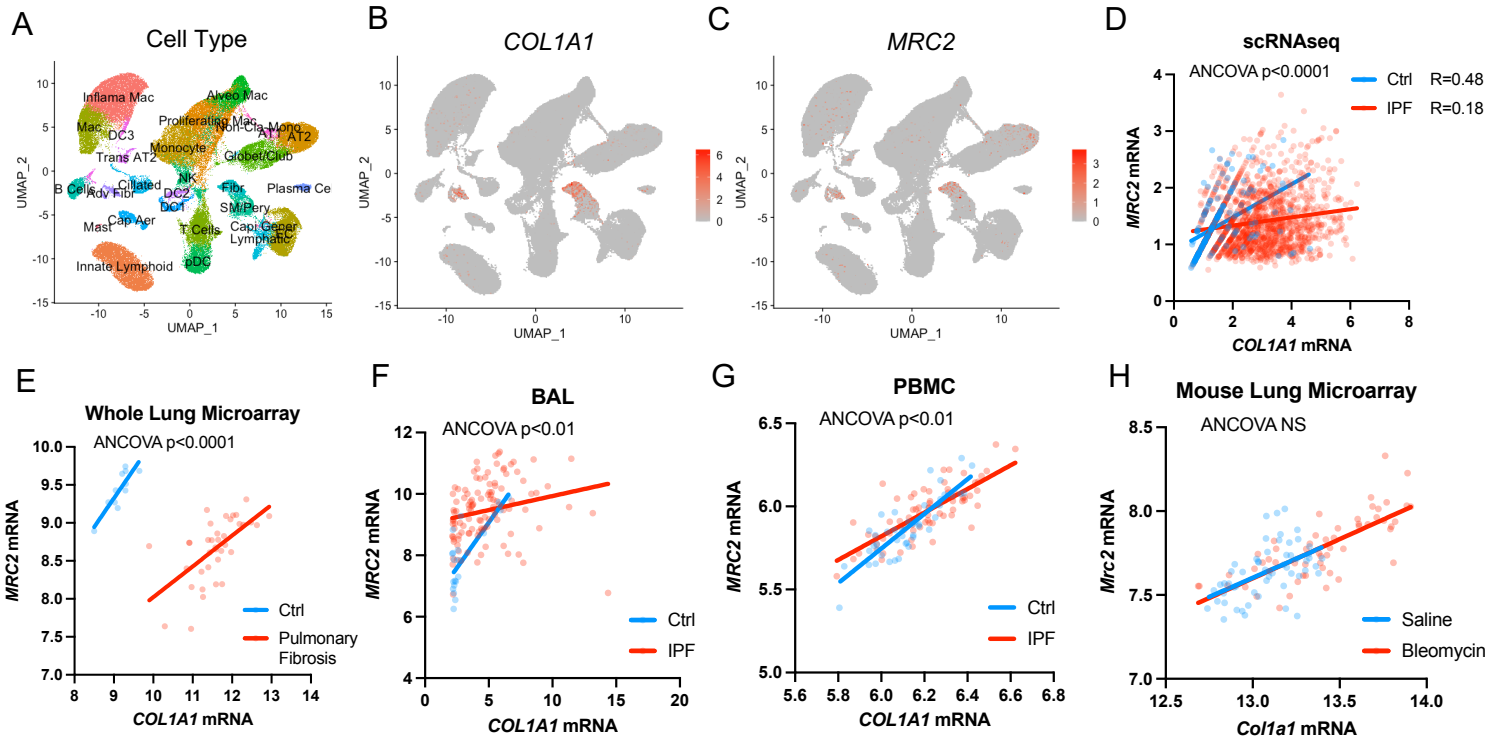


Figure 8. Collagen biosynthesis and *MRC2* expression are positively correlated in lung tissue, but this relationship is impaired in human fibrotic lung disease. (A) UMAP plot with cell type annotation of scRNAseq data from multiple available datasets of human lung control and IPF tissue (GSE136831, GSE135893, GSE121611, GSE128033, GSE132771). N=184,672 single cells. (B-C) Gene expression of individual genes as indicated in UMAP as generated in panel (A). (D) Comparison of gene expression in all cells from data as in panel (A) that had non-zero expression of *COL1A1* or *MRC2*. (E) Comparison of gene expression in whole human lung microarray data from GSE110147 analyzed via GEO2R in pulmonary fibrosis lungs vs. control specimens. N=37 vs. 11 per group. (F) Comparison of microarray gene expression in human lung bronchoalveolar lavage specimens from GSE70867 analyzed via GEO2R from Idiopathic Pulmonary Fibrosis (IPF) lungs vs. control specimens. N=112 vs. 20. (G) Comparison of microarray gene expression in peripheral blood mononuclear cell (PBMC) specimens from GSE37858 analyzed via GEO2R from Idiopathic Pulmonary Fibrosis (IPF) lungs vs. control specimens. N=70 vs. 35. (H) Comparison of gene expression in whole mouse lung microarray data from GSE40151 analyzed via GEO2R from Bleomycin-treated mice vs. Saline control mice. N=56 vs. 55. Statistics: ANCOVA, p-values indicated in figures. Pearson correlations are displayed in panel (D).

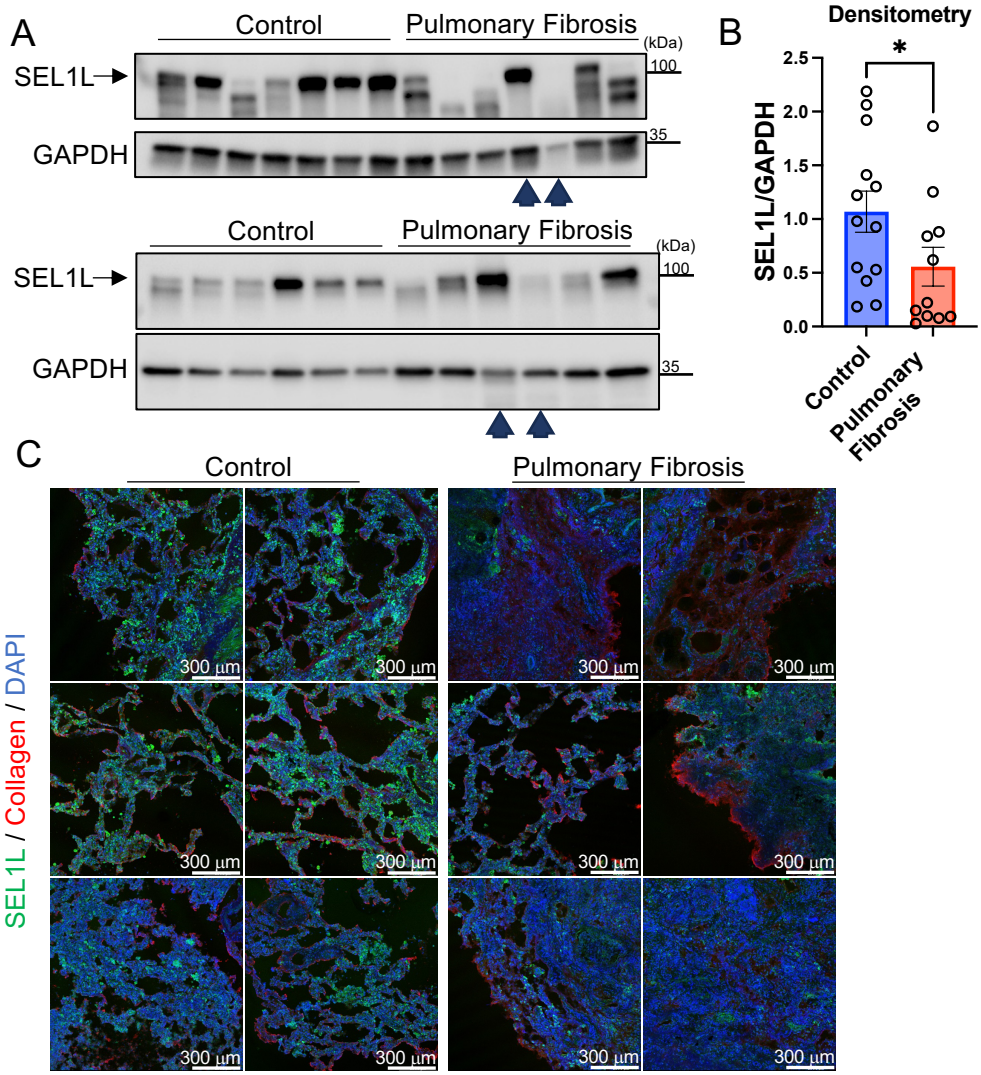


Figure 9. SEL1L protein levels are reduced in the lungs of patients with Pulmonary Fibrosis. (A-B) Western blots of whole human lung lysates from Pulmonary Fibrosis lungs vs. control (donor lungs deemed not suitable for transplant), with corresponding densitometry. N=13 in Control and 11 in Pulmonary Fibrosis group. Arrowheads correspond to two lung lysates that were used in both Western blots for normalization across blots, otherwise all other lanes are independent samples. (C) Representative low-power wide-field immunofluorescence images of human lungs (Pulmonary Fibrosis vs. control). Each row comes from an individual donor or Pulmonary Fibrosis lung, with two different fields per lung (i.e. N=3 individual lungs per group). Statistics: (B, D) Mann-Whitney *U* test. * $p < 0.05$.