Please cite this article in press as: Lu et al., Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming, Cell (2025), https://doi.org/10.1016/j.cell.2025.07.031

Cell

CellPress

Article

Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming

Jinlong Y. Lu,[1,4] William B. Tu,[1,4] Ronghui Li,[1,2] Mingxi Weng,[1] Bhargav D. Sanketi,[1] Baolei Yuan,[1] Pradeep Reddy,[1,2,3] Concepcion Rodriguez Esteban,[1,2,3] and Juan Carlos Izpisua Belmonte[1,2,5,*]

[1]San Diego Institute of Science, Altos Labs, San Diego, CA, USA
[2]Salk Institute for Biological Studies, La Jolla, CA, USA
[3]Senior author
[4]These authors contributed equally
[5]Lead contact
*Correspondence: jcbelmonte@altoslabs.com
https://doi.org/10.1016/j.cell.2025.07.031

SUMMARY

The loss of cellular and tissue identity is a hallmark of aging and numerous diseases, but the underlying mechanisms are not well understood. Our analysis of gene expression data from over 40 human tissues and 20 diseases reveals a pervasive upregulation of mesenchymal genes across multiple cell types, along with an altered composition of stromal cell populations, denoting a "mesenchymal drift" (MD). Increased MD correlates with disease progression, reduced patient survival, and an elevated mortality risk, whereas suppression of key MD transcription factors leads to epigenetic rejuvenation. Notably, Yamanaka factor-induced partial reprogramming can markedly reduce MD before dedifferentiation and gain of pluripotency, rejuvenating the aging transcriptome at the cellular and tissue levels. These findings provide mechanistic insight into the underlying beneficial effects of partial reprogramming and offer a framework for developing interventions to reverse age-related diseases using the partial reprogramming approach.

INTRODUCTION

The epigenetic landscape of aged cells reveals a diminished capacity to maintain cellular identity.[1–8] Previous work identified that cells undergoing replicative senescence exhibit features resembling a fibroblast-to-myofibroblast transition (FMT).[9] These studies suggest that a potential shift in cell state occurs during aging, often accompanied by changes in various molecular hallmarks.[4]

Yamanaka factors OCT4, SOX2, KLF4, and MYC (OSKM) can reprogram fully differentiated somatic cells into a pluripotent state.[10] OSKM-based cellular reprogramming has emerged as a promising approach to rejuvenate aging-related phenotypes, demonstrating effectiveness both *in vitro* and *in vivo*.[11–16] However, the precise mechanisms responsible for this rejuvenation effect remain only partially understood. Complete reprogramming of fibroblasts to a pluripotent state by OSKM initiates with a process known as mesenchymal-epithelial transition (MET), where the cells dedifferentiate from a mesenchymal state to an epithelial state.[17,18] An intriguing question arises regarding whether this shift in cellular state is essential for rejuvenation or if it could be a potentially detrimental byproduct of the reprogramming process that should be avoided.

Here, we analyzed human tissue biopsies and found widespread upregulation of mesenchymal genes and increased stromal cell populations during aging and in many age-related diseases, indicating a drift toward mesenchymal identity at the cellular and tissue levels. Intriguingly, Yamanaka factors can significantly mitigate MD by partially reprogramming cells *in vitro* and *in vivo*. These results provide conceptual advances in understanding the nature of aging and the mechanisms through which partial reprogramming rejuvenates organisms.

RESULTS

MD is a feature of human aging

To identify molecular signatures of aging, we analyzed gene expression data of 46 human tissues from the adult genotype-tissue expression (GTEx) database.[19] We calculated Spearman's ranked correlation coefficients between gene expression and age as input for gene set enrichment analysis (GSEA) to reveal noteworthy trends in gene expression (Figure 1A). Aging is consistently associated with downregulation of energy metabolism (oxidative phosphorylation, tricarboxylic acid (TCA) cycle, and mitochondria biogenesis) and cholesterol biosynthesis pathways,[20] alongside key cell cycle programs (mitotic checkpoint and sister chromatid separation), implying compromised proliferation linked to senescence.[21] Conversely, inflammatory pathways, including tumor necrosis factor alpha (TNF-α) signaling and cytokines, were broadly upregulated, consistent with "inflammaging," a phenomenon characterized by chronic

Please cite this article in press as: Lu et al., Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming, Cell (2025), https://doi.org/10.1016/j.cell.2025.07.031

inflammation throughout the aging process.[22] These enrichments were specific and not observed in random gene sets (Figure 1A).

Beyond these pan-tissue hallmarks, aging drove tissue-specific functional declines (Figure S1A).

Specialized pathways deteriorated in numerous tissues, including γ-aminobutyric acid (GABA) receptor activation and synaptic transmission in the brain, adipogenesis in adipose depots (normalized enrichment score [NES] = −3.7), meiosis in the testis (NES = −2.8), hemostasis in the liver (NES = −2.9), and keratinization and the formation of cornified envelope in the skin, bladder, vagina, and esophagus mucosa (NES = −3.4 to −2.5). Core processing machinery was also dysregulated. While small nuclear RNA assembly and transfer RNA processing pathways were downregulated in the stomach (NES $\leq$ −2.8), ribosome biogenesis showed a complex, divergent pattern: repressed in the kidney and pancreas (NES $\leq$ −2.6), but induced in the prostate and brain (NES $\geq$ 3.1), indicating a tissue-specific remodeling of protein synthesis capacity during aging.

Of particular interest, we observed a widespread upregulation of pathways related to epithelial-mesenchymal transition (EMT) in most tissues (Figure 1A), a finding that remained robust after adjusting for ischemia time and sex (Figure S1B). EMT represents a dynamic shift in cellular properties, involving the transition from an epithelial to a mesenchymal phenotype.[36] Notably, the regulatory networks associated with EMT exhibited significant overlap with other mesenchymal transitions or activation, such as endothelial-mesenchymal transition (EndoMT), FMT, pericyte-myofibroblast transition (PMT), macrophage-mesenchymal transition (MMT), fibro-adipogenic progenitors (FAPs) to myofibroblast transition, and potentially transitioned in other cell types.[37,38] To encapsulate this prevalent phenomenon across cell types, we introduce the term "mesenchymal drift" (MD), wherein cells undergo a partial or complete transition, leading to a compromised original cellular identity while acquiring new or intensifying existing mesenchymal traits, including altered extracellular matrix (ECM) production, cellular junction, tissue mechanical stiffening, and cytokine production. The prevalence of MD across aging tissues suggests a generalized "cell state drift" that contributes to organ dysfunction.

Prevalent MD in human diseases

We hypothesized that MD may also manifest in various diseases, especially those linked to aging. We prioritized examining changes in human disease by selecting *in vivo* biopsy data with substantial sample sizes for their physiological relevance and statistical power. While acknowledging that many forms of cancer are age-related diseases with EMT involvement in disease progression, we opted not to focus on this disease class in this study considering its extensive exploration in previous studies.[39]

In the lung, we analyzed gene expression data of 139 lung samples obtained from 14 non-diseased control (NDC) donors (mean age: 44.9 ± 12.9 years old [y.o.]) and 20 patients diagnosed with idiopathic pulmonary fibrosis (IPF) (62.2 ± 6.3 y.o.).[23] IPF, a condition primarily affecting elderly adults, is the most prevalent form of pulmonary fibrosis. Our analysis revealed a notable upregulation of MD genes in IPF patients (Figure 1B). Furthermore, we observed a similar, albeit less pronounced, degree of MD increase in two other lung disorders: interstitial lung disease (ILD) (53 ± 8.4 y.o.), characterized by progressive lung tissue scarring, and chronic lung allograft dysfunction (CLAD) (57.2 ± 6.6 y.o.), a condition recognized as the leading cause of mortality in lung transplant recipients beyond the first year post-transplantation (Figure 1B). Age-matched controls (donors aged >60 years, 63.2 ± 3.5 y.o.) also show significantly lower MD compared with disease samples (Figure S1C).

Next, we wanted to determine whether MD at the tissue level is through changes in cell composition and/or gene expression. In single-cell RNA sequencing (scRNA-seq) data comprising over 300,000 cells obtained from 32 IPF lungs (65.4 ± 5.4 y.o.), 18 chronic obstructive pulmonary disease (COPD) lungs (62.4 ± 4.9 y.o.), and 28 control donor lungs (45.6 ± 17.7 y.o.),[24] we observed a significant reduction in the number of alveolar type 1 (AT1) cells and alveolar type 2 (AT2) epithelial progenitor cells in IPF samples compared with control and COPD samples (Figure S1D; Table S1). Meanwhile, IPF samples tend to have an increased percentage of stromal cell types, including

Figure 1. Dysregulation of MD genes across aging and relevant diseases

(A) Top dysregulated pathways across 46 tissues during aging scored by normalized enrichment score (NES). Spearman's ranked correlation coefficient between age and gene expression was used as GSEA input.

(B–M) MD gene expression in different diseases and respective healthy controls in various tissues or cell types: (B and C) lung,[23,24] (D and E) liver,[25,26] (F and G) kidney,[27–29] (H) heart,[30] (I) blood vessel (endothelial cells),[31] (J) brain,[32] (K) skin,[33] (L) colon,[34] and (M) knee.[35]

Bulk expression data: (A), (B), (D), (F), (H), and (M); scRNA-seq: (C), (E), (G), (I), (J), (K), and (L).

Sample sizes are annotated, and all comparisons are made between control and disease groups. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, n.s. (no significant difference): $p \geq 0.05$. Statistical methods are described in STAR Methods.

Diseases: IPF: idiopathic pulmonary fibrosis; ILD: interstitial lung disease; CLAD: chronic lung allograft dysfunction; COPD: chronic obstructive pulmonary disease; MASLD: metabolic dysfunction-associated steatotic liver disease; MASH: metabolic dysfunction-associated steatohepatitis; eCLD: early chronic liver disease; ACLF: acute-on-chronic liver failure; DCM: dilated cardiomyopathy; HCM: hypertrophic cardiomyopathy; PPCM: peripartum cardiomyopathy; LN: lupus nephritis; DN: diabetic nephropathy; FSGS: focal segmental glomerulosclerosis; FGGS: focal global glomerulosclerosis; ESKD: end-stage kidney disease; AD: Alzheimer's disease; OA: osteoarthritis; AC: atherosclerotic core.

Cell types: Astro: astrocyte; AT1: alveolar type 1 cell; Chol: cholangiocyte; DKC: differentiated keratinocyte; Endo/EC: endothelial cell; Entero: enterocyte; Fib: fibroblast; Inf Fib: inflammatory fibroblast; Hep: hepatocyte; HSC: hepatic stellate cell; Inj PT: injured proximal tubule; LEC: lymphatic endothelial cell; Mac: macrophage; Myofib: myofibroblast; OPC: oligodendrocyte precursor cell; Peri: pericyte; PT: proximal tubule; T: T cell; Telo: telocyte; Tropho: trophocyte; UnKC: undifferentiated keratinocyte; VEC: vascular endothelial cell.

See also Figures S1–S4.

Please cite this article in press as: Lu et al., Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming, Cell (2025), https://doi.org/10.1016/j.cell.2025.07.031

myofibroblasts, pericytes, and smooth muscle cells (SMCs), as well as endothelial cells in the peribronchial regions (Figure S1D). The elevated proportion of peribronchial endothelial cells and SMCs may represent the cellular basis for bronchial vessel expansion, a characteristic feature of IPF.[40]

Within these remaining cells, AT1 and AT2 epithelial cells, as well as all stromal cells except SMCs and endothelial cells in lymphatic, arterial, and capillary regions, exhibited significant upregulations of MD gene expression (Figures 1C, S1E, and S1F). To ensure age-matched comparisons, we excluded control samples younger than 60 years, resulting in a filtered control group with an average age of 64.7 ± 2.1 y.o., comparable to the IPF and COPD cohorts. Similarly, we observed consistent upregulation of MD genes across AT1, AT2, fibroblasts, myofibroblasts, pericytes, and endothelial cells in capillary regions (Figure S1G). Thus, pulmonary fibrosis involves both an expanded mesenchymal cell population and a heightened mesenchymal state across multiple cell types, suggesting that upregulation of cell-intrinsic mesenchymal gene expression may contribute to changes in cell composition.

Next, we shifted our attention to liver diseases, particularly focusing on data from patients with metabolic dysfunction-associated steatotic liver disease (MASLD) and the more severe metabolic dysfunction-associated steatohepatitis (MASH), characterized by the accumulation of excessive fat in the liver, which can escalate to cirrhosis and liver failure.[25,41] While MD gene expression was similar between healthy controls and early-stage liver diseases (MASLD, early chronic liver disease [eCLD], and early MASH [F0-F1]), it significantly increased in advanced MASH (F2-F4), correlating with disease severity (Figures 1D and S2A). The mean age for patients with advanced fibrosis was 57.1 ± 10.1 years (Cluster A identified by Govaere et al.[25]), while for the remaining patients (Cluster B), it was 52.6 ± 12.4 y.o. In another independent dataset,[41] compared with control (64 ± 7 y.o.) and eCLD groups (58 ± 7 y.o.), similar increases were seen in late-stage decompensated cirrhosis (54 ± 18 y.o.) and acute-on-chronic liver failure (ACLF) (55 ± 6 y.o.), a condition linked to organ failure and high mortality (Figure S2A).

We subsequently examined single-cell transcriptomic data of over 100,000 cells characterizing cell types in age-matched healthy (57.4 ± 7.9 y.o.) and cirrhotic human livers (56.6 ± 5.8 y.o.).[26,42] Significant MD gene upregulation was observed in endothelial cells and hepatic stellate cells (HSCs) under both alcohol- and non-alcohol-associated liver disease conditions (Figure 1E). In addition, hepatocytes exhibited a pronounced upregulation of MD genes in alcohol-associated liver disease. By contrast, no consistent increase in MD gene expression was detected in cholangiocytes or Kupffer cells (Figures 1E and S2B). An independent single-cell dataset of steatotic and control human livers[43] revealed a similar pattern, with MD genes upregulated in endothelial cells, HSCs, and hepatocytes, but not in cholangiocytes or Kupffer cells (Figure S2C). These cell-specific findings align with the pivotal roles of endothelial cells and HSCs in the progression of MASH, where liver fibrosis typically originates in the perisinusoidal and periportal regions.[44,45] These data collectively suggest that MD plays a key role in liver disease progression.

Focusing on the kidney, we analyzed gene expression data from micro-dissected human kidney biopsy samples. These samples were obtained from both healthy living donors and patients with chronic kidney diseases (CKD), such as lupus nephritis (LN), diabetic nephropathy (DN), focal segmental glomerulosclerosis (FSGS), and focal global glomerulosclerosis (FGGS).[27] Comparing patient and healthy samples, we observed significantly elevated MD gene expression in both the glomeruli and tubules of CKD patients, with the tubules in FSGS showing a similar trend (Figures 1F and S2D). The insignificant change of MD in FSGS data is likely due to the limited sample size available ($n = 3$). Additionally, we noted a similar upregulation of MD genes in the basilic vein of the kidney in individuals with end-stage kidney disease (ESKD) (Figure 1F), the advanced phase of long-term CKD.[28]

We further analyzed publicly available scRNA-seq data from 73 human kidney tissue samples, comprising 35 healthy controls (mean age: 62.3 y.o.), 18 individuals with diabetic kidney disease (35.1 y.o.), and 20 individuals with hypertensive kidney disease (45.1 y.o.).[29] Compared with healthy controls, CKD samples had fewer proximal tubule epithelial cells and more injured counterparts (Figure S2E; Table S1). In comparison, average stromal cell percentage showed an approximate two-fold increase, with endothelial cells unchanged (Figure S2F). Significant upregulation of MD genes was noted in specific cell populations, including epithelial cells from the proximal tubule (PT) segment, thick ascending loop of Henle (TAL), distal convoluted tubule (DCT), and podocyte, as well as stromal cells such as fibroblasts, glomerulosclerosis (GS)-specific stromal cells, and mesangial cells. Endothelial cells from glomerular capillary tuft, lymphatic, and peritubular regions also showed a noticeable upregulation of MD genes (Figures 1G and S2G). Conversely, immune cells, except for $CD4^+$ T cells, did not exhibit obvious upregulation (Figure S2G). These findings illustrate disease-associated MD and altered cellular composition in the human kidney.

Next, we reanalyzed transcriptomic data from ventricles sourced from the Myocardial Applied Genomics Network (MAGNet),[30] including non-failing controls (55.9 ± 14.0 y.o.) and common types of cardiomyopathy, including dilated cardiomyopathy (DCM) (52.1 ± 10.6 y.o.), hypertrophic cardiomyopathy (HCM) (48.7 ± 12.6 y.o.), and peripartum cardiomyopathy (PPCM) (34.7 ± 10.3 y.o.).[46] Compared with individuals without heart failure, all the cardiomyopathies under investigation showed significant MD gene upregulation (Figure 1H).

Blood vessels, lined by endothelial cells, are key to circulatory function and barrier integrity, including the blood-brain barrier. Atherogenesis, aging-associated plaque formation in arteries, was explored by reanalysis of single-cell expression data of more than 50,000 cells from artery tissues in patients undergoing carotid endarterectomy.[31] Endothelial cells from calcified atherosclerotic core (AC) plaques exhibited significant upregulation of MD genes compared with proximal adjacent (PA) regions (Figure 1I). Macrophages displayed a less pronounced upregulation, while vascular SMCs (VSMCs) and T cells were not significantly altered (Figure S3A). Additionally, *PECAM1*, an adhesive protein preserving endothelial cell junction integrity, was downregulated in the endothelial cells from AC plaques compared with those from adjacent regions (Figure S3B). Notably, the plasma

Please cite this article in press as: Lu et al., Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming, Cell (2025), https://doi.org/10.1016/j.cell.2025.07.031

level of certain transforming growth factor beta (TGF-β)[47] members, which are potent inducers of EMT and EndoMT, was increased during human aging (Figure S3C). Exposure of endothelial cells to TGF-β and IL-1β[48] resulted in significant MD gene upregulation (Figure S3D), accompanied by a notable downregulation of *PECAM1* (Figure S3E). A similar pattern of MD gene upregulation was also observed in epithelial and fibroblast cells following TGF-β treatment (Figure S3F).[49,50] These observations are supported by a recently published study showing that EndoMT is associated with accelerated epigenetic aging of atherosclerotic plaques and a poor prognosis.[51] Together with prior findings in other organs, these results suggest that endothelial MD may contribute to vascular permeability and barrier dysfunction—a hallmark of many age-related disorders.

Aging also emerges as the primary risk factor for most neurodegenerative diseases, including Alzheimer's disease (AD). We reanalyzed single-nucleus (sn)RNA-seq data of brain tissues obtained from donors with AD (78.8 ± 8.3 y.o.) and unaffected controls (63.5 ± 13.7 y.o.).[32] Our investigation encompassed major cell types, spanning excitatory and inhibitory neurons, microglia, oligodendrocytes (OLGs), oligodendrocyte progenitor cells (OPCs), astrocytes, and pericytes. Our analysis revealed significant MD gene upregulation in astrocytes, oligodendrocytes, and pericytes, but not in OPCs, neurons, and microglia (Figures 1J and S3G). To ensure age-matched comparisons, we excluded control samples younger than 70 y.o., resulting in a filtered control group with an average age of 79 ± 4.4 y.o., comparable to the AD cohorts. Similarly, we observed consistent upregulation of MD genes across astrocytes, oligodendrocytes, and pericytes (Figure S3H).

Skin regeneration declines with age. From scRNA-seq data of over 40,000 cells from keloid and normal skin (age information not available),[52] we observed a substantial upregulation of MD genes in keratinocytes (KCs), fibroblasts, muscle fibers (MFs), and glandular cells (GCs), and less pronounced upregulation in endothelial cells and melanocytes (Figure S4A). The relatively modest changes in endothelial cells may be attributed to the expansion of newly formed endothelial cells, as keloids are known to exhibit increased angiogenesis.[53] Since the skin has been studied in the context of aging, we next analyzed scRNA-seq data from over 15,000 cells derived from skin samples taken from a defined, sun-protected area.[33] The samples were obtained from two young (25–27 y.o.) and three old (53–70 y.o.) male Caucasian donors. Our analysis revealed significant upregulation of MD genes in KCs, endothelial cells, pericytes, macrophages, and T cells (Figure 1K). Notably, MD gene upregulation was observed specifically in fibroblasts with higher mesenchymal potential, characterized by *ASPN*, but not in inflammatory fibroblasts characterized by *CCL19* expression. These results suggest that the diminished regenerative capacity of aged skin and the development of fibrotic scars are associated with significant upregulation of MD genes in key cell types.

In the eye, advanced age increases the risk of subretinal fibrosis in age-related macular degeneration (AMD). Our analysis revealed significant MD gene upregulation in human fetal retinal pigment epithelium (RPE) cells cultured or passaged under subconfluent conditions,[54] a state that serves as a proxy for the repeated wound stimulus occurring during aging (Figure S4B). This upregulation of MD genes coincided with a decrease in the expression of E-cadherin (encoded by *CDH1*), an epithelial marker, and an upregulation of N-cadherin (encoded by *CDH2*), a marker of EMT (Figure S4C).

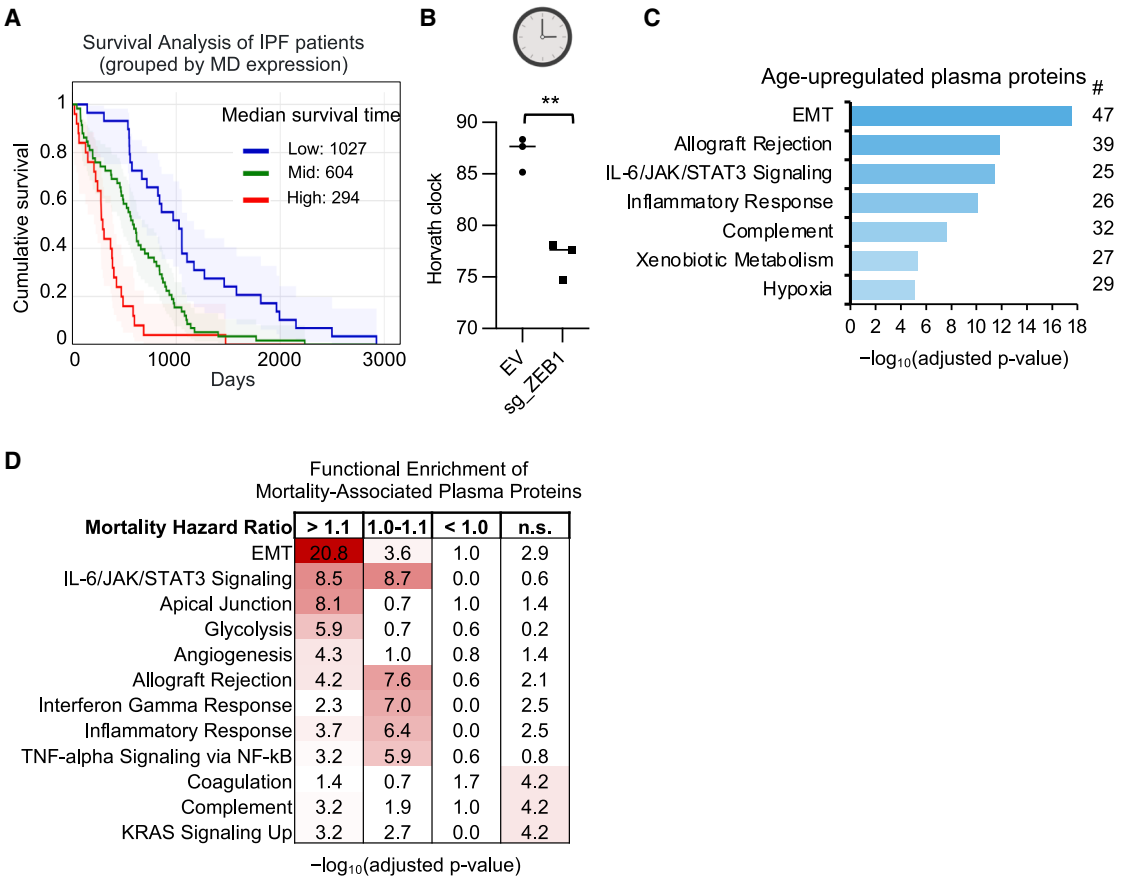
For diseases in the intestine, we reanalyzed data derived from patients with inflammatory bowel diseases (IBD), including Crohn's disease (CD)[55] and ulcerative colitis (UC).[34] Upon reanalysis of mucosal biopsies taken from the colons of 20 CD patients and controls matched in terms of age and sex, we found MD gene upregulation and *CDH1* downregulation (Figures S4D and S4E), implying a compromised intestinal barrier function. In a separate dataset of colonoscopic biopsies from 18 UC patients (inflamed and adjacent non-inflamed areas) and 12 healthy individuals, epithelial cells, including transit-amplifying (TA) cells, enterocytes, and goblet cells, were decreased in UC patients, while inflammatory fibroblasts, myofibroblasts, pericytes, and several immune populations (e.g., $CD4^+$ memory T cells, mast cells, inflammatory monocytes, and regulatory T cells) were increased (Table S1). MD genes were significantly upregulated in mesenchymal cell types (myofibroblasts, pericytes, $WNT2B^+$ sub-cryptal trophocytes, and $WNT5B^+$ sub-epithelial telocytes) in inflamed colon samples and at lower but still significant levels in adjacent non-inflamed areas (Figures 1L, S4F, and S4G). MD genes were also significantly upregulated in other non-mesenchymal populations of the colon, such as the capillary endothelium, glia, and multiple epithelial cell types, including enterocytes. Most immune cell types did not show significant alteration in MD genes except for $CD69^-$ mast cells and inflammatory monocytes, important effectors of the innate immune cascade (Figure S4H). These findings suggest the occurrence of widespread MD during IBD. The presence of MD in non-inflamed areas suggests that MD may precede tissue transition to a diseased state or that UC-affected regions exert a broader influence on neighboring non-inflamed areas, potentially contributing to disease progression.

In the joint, we conducted a gene expression analysis on biopsy data from knee articular cartilage.[35] The results indicated a substantial upregulation of MD genes in patients diagnosed with osteoarthritis (OA) (control: 36.6 ± 13.5; OA: 66 ± 7.3) (Figure 1M). A comparison of donors aged 50 to 60 y.o. between the control and OA groups showed a similar upregulation of MD genes in the OA group (Figure S4I).

MD is linked to disease progression, mortality risk, and epigenetic rejuvenation

Given the strong association between MD and disease progression, we hypothesize that MD may serve as a predictor of clinical outcomes. As a proof of concept, we analyzed gene expression data from bronchoalveolar lavage (BAL) cells of 132 patients with IPF and correlated MD scores with follow-up survival data.[56] Patients were divided into low, middle, and high MD gene expression groups, with median survival times of 1,027, 604, and 294 days, respectively (Figures 2A and S5A), showing an inverse correlation with their MD magnitude (Figure S5B). Remarkably, patients with the lowest and highest MD gene expression levels survived for 2,498 and 59 days, respectively, representing more than a 40-fold difference in survival time. These findings highlight

Please cite this article in press as: Lu et al., Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming, Cell (2025), https://doi.org/10.1016/j.cell.2025.07.031



Mortality Hazard Ratio	> 1.1	1.0-1.1	< 1.0	n.s.
EMT	20.8	3.6	1.0	2.9
IL-6/JAK/STAT3 Signaling	8.5	8.7	0.0	0.6
Apical Junction	8.1	0.7	1.0	1.4
Glycolysis	5.9	0.7	0.6	0.2
Angiogenesis	4.3	1.0	0.8	1.4
Allograft Rejection	4.2	7.6	0.6	2.1
Interferon Gamma Response	2.3	7.0	0.0	2.5
Inflammatory Response	3.7	6.4	0.0	2.5
TNF-alpha Signaling via NF-kB	3.2	5.9	0.6	0.8
Coagulation	1.4	0.7	1.7	4.2
Complement	3.2	1.9	1.0	4.2
KRAS Signaling Up	3.2	2.7	0.0	4.2

Figure 2. MD is linked to disease progression, mortality risk, and epigenetic rejuvenation

(A) Survival analysis of IPF patients with different levels of MD gene expression.

(B) Horvath DNA methylation clocks of SUM159 cells upon treatment of empty vector (EV) and ZEB1 repression. $^{**}p < 0.01$.

(C) Functional enrichment of upregulated plasma proteins during aging.

(D) Functional enrichment of mortality-associated plasma proteins.

See also Figure S5.

MD gene levels as strong prognostic markers and potential therapeutic targets in IPF.

To begin probing causality between MD and aging, we examined whether modulating MD regulators could reverse biological age. Using DNA methylation-based epigenetic clocks,[57] we analyzed public data following the repression of ZEB1, a key EMT transcription factor that suppresses epithelial markers like E-cadherin.[58] In SUM159 breast cancer cells, ZEB1 repression led to a rapid reduction in DNA methylation age, consistent across both Horvath's pan-tissue[59] and PhenoAge clocks[60] (Figures 2B and S5C), suggesting potential for MD-targeted rejuvenation.

We next explored potential drivers of MD across organs. Previous studies show that acute liver injury can trigger hepatocyte senescence and systemic multi-organ failure. Kiourtis et al. demonstrated that circulating TGF-β mediates this effect, and its inhibition prevents extrahepatic dysfunction.[61] This observation prompted us to explore plasma protein alterations during aging. Analyzing plasma data from Lehallier et al.,[47] we found that inflammation pathways were enriched in upregulated plasma proteins, consistent with previous reports.[22] Interestingly, the EMT pathway was significantly enriched among upregulated plasma proteins during aging (adjusted p value = 2.7×10^{-18}, n = 47), exhibiting markedly greater enrichment compared with downregulated proteins (adjusted p value = 0.05, n = 11) (Figures 2C and S5D; proteins listed in Table S2). For example, plasma levels of GDF15 (a member of the TGF-β superfamily), pleiotrophin (PTN), chemokine eotaxin-1 (CCL11), and beta-2 microglobulin (B2M) were significantly elevated during aging (Figure S5E) and have been reported to promote EMT.[62–65]

In addition to age-associated proteins, we extended our analysis to include mortality-associated plasma proteins identified from the UK Biobank dataset (n > 40,000 participants), adjusting for sex and chronological age.[66] We categorized the mortality-associated proteins into three groups: proteins strongly associated with increased mortality (mortality hazard ratio [HR] > 1.1; n = 360), proteins with a moderate association (1.1 > HR > 1.0; n = 530), and proteins associated with decreased mortality (HR < 1.0; n = 27). As a comparison,

Please cite this article in press as: Lu et al., Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming, Cell (2025), https://doi.org/10.1016/j.cell.2025.07.031

we also included 500 randomly selected plasma proteins that showed no significant association with mortality. We observed that proteins with a stronger mortality HR were significantly enriched in EMT pathways, a pattern that was either absent or notably less pronounced in the other categories (Figure 2D; Table S2). By contrast, proteins with a weaker association to increased mortality were predominantly associated with interferon, inflammatory, and TNF-α signaling pathways.

These findings highlight the significant association between MD regulators, aging, and increased mortality, suggesting that circulating factors may play a role in driving and propagating MD across multiple organs. Additionally, these results may offer insight into the rejuvenating effects observed with heterochronic parabiosis.[67–69]

Partial reprogramming mitigates MD in aged cells

We next examined whether MD in aged cells can be mitigated by epigenetic reprogramming using the Yamanaka factors, which have shown rejuvenating effects in aged cells and organisms.[11,12] During somatic reprogramming to induced pluripotent stem cells (iPSCs), differentiated cells undergo MET to successfully dedifferentiate to the iPSC state.[17,18] To understand the dynamics of the transition between mesenchymal and epithelial states during reprogramming in aged cells, we conducted scRNA-seq of human fibroblasts derived from an aged donor (96 y.o.) at different time points after OSKM induction (Figures 3A–3D and S6A). Trajectory analysis using Slingshot identified three paths from the starting fibroblasts (Figure 3D): (1) one trajectory is toward the emergence of the pluripotent state, as marked by canonical pluripotency genes *NANOG*, *DPPA4*, and *LEFTY2* (Figure 3E); (2) another is to the non-reprogrammed state that retains many of the starting fibroblast marker genes (Figure S6B), but it also expresses distinct markers perhaps due to viral infection and/or different culture conditions (Figure 3E); and (3) the third trajectory that we termed partially reprogrammed represents cells that exhibit features of rejuvenation we will elaborate on below.

In the partially reprogrammed state, we observed that mesenchymal genes, including core EMT transcription factors *SNAI1/2*, *TWIST1/2*, and *ZEB1/2*, were downregulated within the first 3 days of OSKM induction but were maintained in non-reprogrammed populations (Figure 3F). TGF-β genes were similarly downregulated in the partially reprogrammed state (Figure 3F). Conversely, reprogramming-associated epithelial genes *CDH1*, *EPCAM*, and *OCLN* were only beginning to be upregulated toward the pluripotency state (Figure 3F). These suggest that the early partial reprogramming window is critical to dampen MD through partial MET, prior to the stable upregulation of the epithelial state toward pluripotency.

To further characterize the partially reprogrammed state, we identified a set of genes that were upregulated and downregulated with age from a dataset of over 100 human dermal fibroblasts[70] (Table S3). The partially reprogrammed populations showed a robust reversal of transcriptomic aging changes, while non-reprogrammed populations maintained the aging transcriptome (Figure 3G). This pattern was accompanied by downregulation of the MD and TGF-β pathway scores in the partially reprogramming cells and their maintained expression in non-reprogrammed cells (Figure 3G). To further illustrate the effects of partial reprogramming by OSKM, we compared our observed cell states with signatures derived from an scRNA-seq atlas of human fibroblast populations in homeostatic and diseased states *in vivo*, including a universal, unactivated fibroblast state that may give rise to specialized states and activated myofibroblast subtypes that contribute to tissue fibrosis.[71] The partially reprogrammed state resembled the universal fibroblast state marked by *PI16*, while signatures of activated myofibroblasts marked by *LRRC15* and *COL3A1* were decreased (signature score: Figure 3H; marker gene expression: Figure S6C). As uncontrolled myofibroblast activation is associated with fibrotic diseases and aging, partial reprogramming may inhibit this process to mediate rejuvenation effects in tissues.

To compare the effects of OSKM in young and aged cells, we performed the same reprogramming scRNA-seq time course in human fibroblasts from a young donor (22 y.o.) and integrated the results with those from the 96-y.o. aged donor (Figures 4A–4C, S6D, and S6E). We annotated the integrated cell states based on gene expression signatures (top 200 genes) identified from the 96-y.o. dataset to fibroblast, partially reprogrammed, non-reprogrammed, early pluripotency, and pluripotency states (Figures 4B and S6F; Table S3). These populations express the top cell state marker genes identified previously (Figures 3E and S6G). The time points and cell states between the young and aged cells largely align across the reprogramming time course, demonstrating experimental and analytical consistency (Figures 4C, S6D, and S6E). We observed that the starting-aged fibroblasts exhibit higher MD and TGF-β expression programs compared with young fibroblasts (Figure 4D). Upon OSKM expression, cells from both ages undergo the initiation of MET during the partial reprogramming phase, as reflected by decreased MD gene program, but because of different starting points, aged cells exhibit a greater decrease (Figure 4D), suggesting OSKM more strongly suppresses the MD program in aged cells. Similarly, the starting-aged fibroblasts show elevated myofibroblast signatures that are more sharply inhibited with OSKM than young fibroblasts (Figure 4E). Conversely, young fibroblasts reprogram more efficiently toward pluripotency than aged fibroblasts (Figure S6H), as more young cells at

(D) Slingshot trajectory inference of three paths from fibroblast toward partial reprogramming, pluripotency, and non-reprogramming states.
(E) Marker gene expression of cell states during reprogramming.
(F) Expression of mesenchymal, epithelial, and TGF-β genes for each cell state during reprogramming.
(G) Module scores of age-upregulated and age-downregulated genes (defined from data in Fleischer et al.[70] as described in STAR Methods), MD genes, and TGF-β genes for each cell state during reprogramming.
(H) Module scores of *in vivo* human fibroblast subtypes (defined from Buechler et al.[71]) for each cell state during reprogramming.
See also Figure S6.

Please cite this article in press as: Lu et al., Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming, Cell (2025), https://doi.org/10.1016/j.cell.2025.07.031

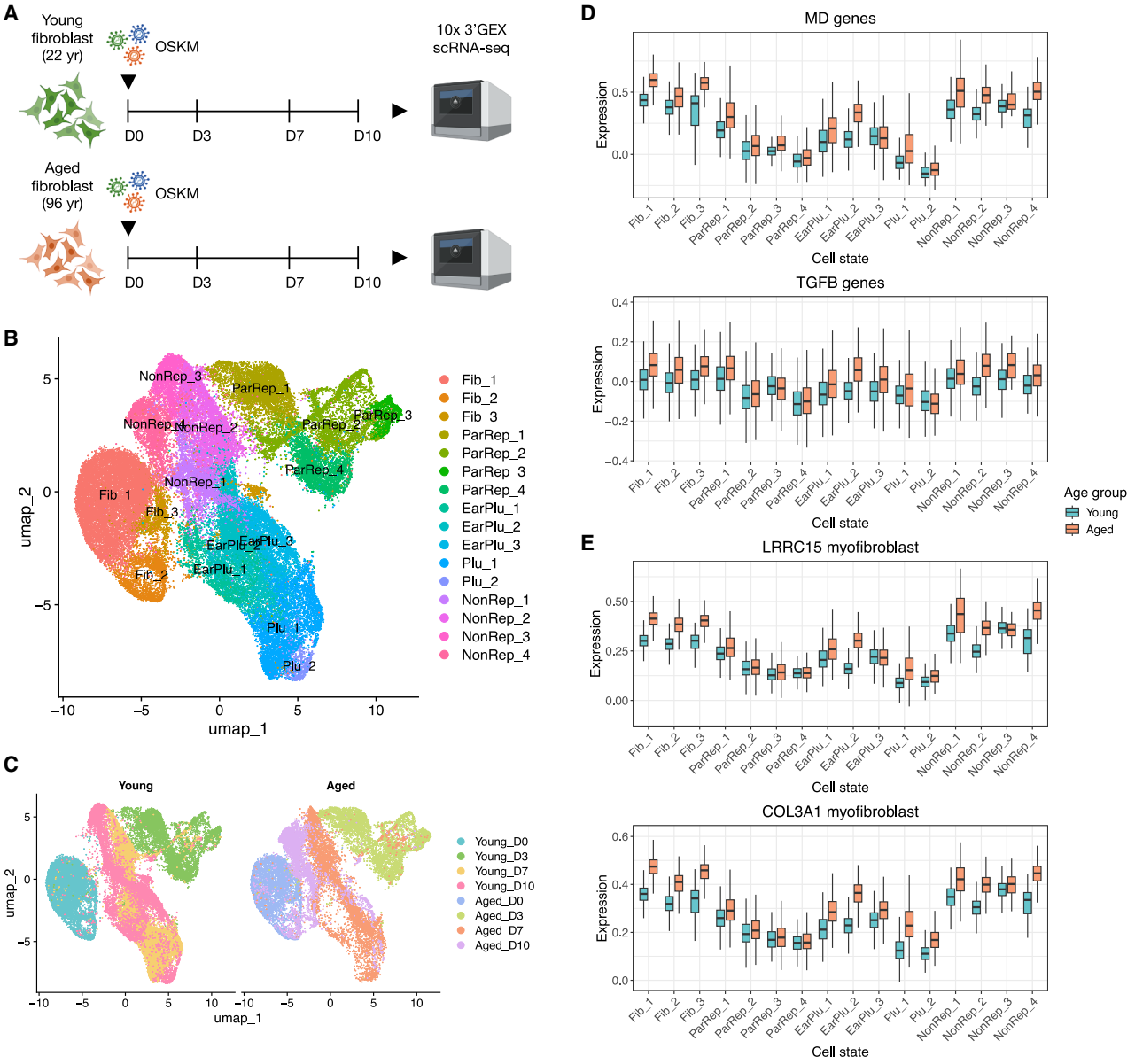


Figure 4. Yamanaka factor effect on MD in young and aged cells

(A) Schematic of scRNA-seq time course of OSKM reprogramming in young and aged fibroblasts.

(B) UMAP of Louvain clusters and their defined cell state in young and aged combined datasets.

(C) UMAPs of different time points split by age.

(D) Module scores of MD and TGF-β genes for each cell state split by age.

(E) Module scores of *in vivo* human myofibroblast subtypes for each cell state split by age.

Fib: fibroblast; ParRep: partial reprogramming; EarPlu: early pluripotency; Plu: pluripotency; NonRep: non-reprogramming.

See also Figure S6.

days 7–10 are represented in the early pluripotency and pluripotency states than aged cells (Figure S6E). These indicate that while Yamanaka factors have been characterized to induce MET to reprogram cells into iPSCs, they are particularly effective in dampening the mesenchymal program in aged cells as part of the partial MET effect that is not associated with progression to iPSCs.

To further disentangle the relationship between MD, MET, rejuvenation, and pluripotency, we investigated a Yamanaka factor mutant incapable of reprogramming cells into iPSCs: the OCT4 mutant (OCT4YR) that is defective in dimerization with SOX2.[72,73] In mouse embryonic fibroblasts (MEFs), Oct4YR with wild-type Sox2 and Klf4 suppress the MD gene program similar to wild-type OSK, without activating the

Please cite this article in press as: Lu et al., Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming, Cell (2025), https://doi.org/10.1016/j.cell.2025.07.031

pluripotency marker *Nanog* (Figure S6I). We extended this to aged human fibroblasts expressing wild-type OSK or OCT4YR with SK. Following a 4-day induction, we observed a robust downregulation of the MD gene program by both factor combinations (Figure S6J). OCT4YR with SK was able to reverse the transcriptomic aging signatures similar to OSK in aged human fibroblasts (Figure S6J), but this pattern is less apparent in embryonic-stage MEFs (Figure S6I). These findings support the notion that mitigating MD, independent of inducing pluripotency, is a potential mechanism in driving the rejuvenation process.

Partial reprogramming mitigates MD *in vivo*

We next sought to determine how MD is affected at the tissue level. In naturally aging mice, partial reprogramming by long-term systemic OSKM induction (7 months of 2-day on 5-day off cycles) significantly decreased MD gene expression in the kidney and liver, two organs where fibrosis prominently develops with age, and trended toward decrease in most other organs (Figure 5A), and this is consistent with the decreasing trend in epigenetic age by DNA methylation.[12] Short-term OSKM inductions (1 month of 2-day on and 5-day off) at an advanced age (25 months) showed a significant decrease of MD in the spleen and a trend in the liver (Figure 5A), and this regimen displayed less obvious epigenetic age change but nonetheless demonstrates that even short-term OSKM expression can affect MD in certain tissues. Recently, targeted expression of OSK under the *Cdkn2a* promoter was shown to extend lifespan and improve features of health in progeroid and natural aging mice.[74] Analysis of transcriptomics from progeroid mice (LAKI, carrying the *Lmna* gene with a G609G mutation) treated with adeno-associated virus (AAV)-GFP control or *Cdkn2a*-OSK for 75 days demonstrated a marked decrease in MD in the bone marrow compartment and a decreasing trend in the skin and spleen (Figure 5B). This is consistent with *Cdkn2a*-OSK-treated mice exhibiting a shift in hematopoietic cell populations toward those of young mice and decreased inflammatory characteristics in the bone marrow and spleen.[74]

To identify the cell types in which MD gene expression is affected by partial reprogramming *in vivo*, we examined scRNA-seq data in mouse intestine[75] and pancreas.[76] In the intestine, systemic OSKM induction for 4 days led to decreased MD gene expression in most cell types of the intestinal epithelium, including enterocytes, chief cells, crypt cells, Paneth cells, goblet cells, and enteroendocrine cells (Figure 5C). This may be linked to improved intestinal regenerative capacity.[75] In the pancreas, OSKM was expressed for a longer duration (7 days). While most pancreatic cell types, including subpopulations of pancreatic stellate, ductal, endocrine, endothelial, and immune cells, showed increased MD expression, a subpopulation of pancreatic stellate cells (pancreatic stellate 2) showed decreased MD expression (Figure 5D). This suggests that extended OSKM induction has differential effects depending on cell type. Further determining the duration of Yamanaka factor induction and limiting expression to specific cell types and states[74] may improve effects on MD mitigation and tissue health. As these scRNA-seq analyses were performed in young and adult mice, evaluating partial reprogramming effects at the single-cell level in aged and disease models will pinpoint cell populations that are affected.

Our findings revealed that the partially reprogrammed state exhibited a notable decrease in the expression of MD genes through partial MET. This observation corresponds to a reversal of aging signatures and inhibition of myofibroblast activation *in vitro* and improved health and tissue function *in vivo*.

DISCUSSION

In this study, we conducted a large-scale meta-analysis of human sequencing data to elucidate mechanisms underlying aging and age-related diseases. Our findings revealed a phenomenon of cell state drift toward a mesenchymal state across multiple organs (Figure 6). These changes are not merely due to an increase in mesenchymal populations but also encompass a transcriptomic shift across a diverse range of cell types, including epithelial, endothelial, glial, and immune cells.

Cellular dysregulation exhibits both consistent patterns and tissue- or disease-specific variations. For instance, parenchymal epithelial cells such as AT1 and AT2 cells in the lung, hepatocytes in the liver, KCs in the skin, renal tubular epithelial cells in the kidney, enterocytes in the colon, and RPE in the eye exhibit MD under disease conditions. Endothelial cells also show widespread dysregulation across several tissues, including lung, liver, kidney, blood vessels, colon, and aged skin. This aligns with recent findings showing that endothelial cells contribute to fibrosis in mice by inducing the expression of mesenchymal genes and the transcription factor SOX9.[77] Immune cells, particularly myeloid cells, exhibit context-dependent variability in their responses, which can have profound effects. In chronic allograft rejection biopsies, $CD68^+$/alpha smooth muscle actin $(\alpha\text{-SMA})^+$ cells, indicating macrophage-to-myofibroblast transition, comprised ~50% of myofibroblasts and correlated with allograft function and fibrosis severity.[78] A similar transition of macrophages into fibroblast-like cells is also observed during myocardial infarction healing.[79] Recent studies suggest that in the skin, immune cells such as neutrophils and macrophages can acquire a mesenchymal-like signature and even undergo myofibroblast transition in response to injury and wound healing.[80,81] The upregulation of MD genes in

Figure 5. Tissue-specific mitigation of MD by partial reprogramming *in vivo*

(A) MD gene expression of wild-type mice and systemic OSKM mice treated with doxycycline for 2 days on and 5 days off from 15 to 22 months of age (upper), and systemic OSKM mice treated with or without doxycycline for 2 days on and 5 days off from 25 to 26 months of age (lower). $p < 0.05$ indicates significance. Raw data from Browder et al.[12]

(B) MD gene expression of LAKI mice treated with AAV-GFP control or AAV-*Cdkn2a*-OSK for 75 days. $p < 0.05$ indicates significance. Raw data from Sahu et al.[74]

(C and D) MD gene signature score of scRNA-seq of mouse intestine (C) and pancreas (D) with OSKM treatment. Raw data from Kim et al.[75] and Chondronasiou et al.[76], respectively. ***$p < 0.001$, ****$p < 0.0001$.

See also Figure S6.

Please cite this article in press as: Lu et al., Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming, Cell (2025), https://doi.org/10.1016/j.cell.2025.07.031

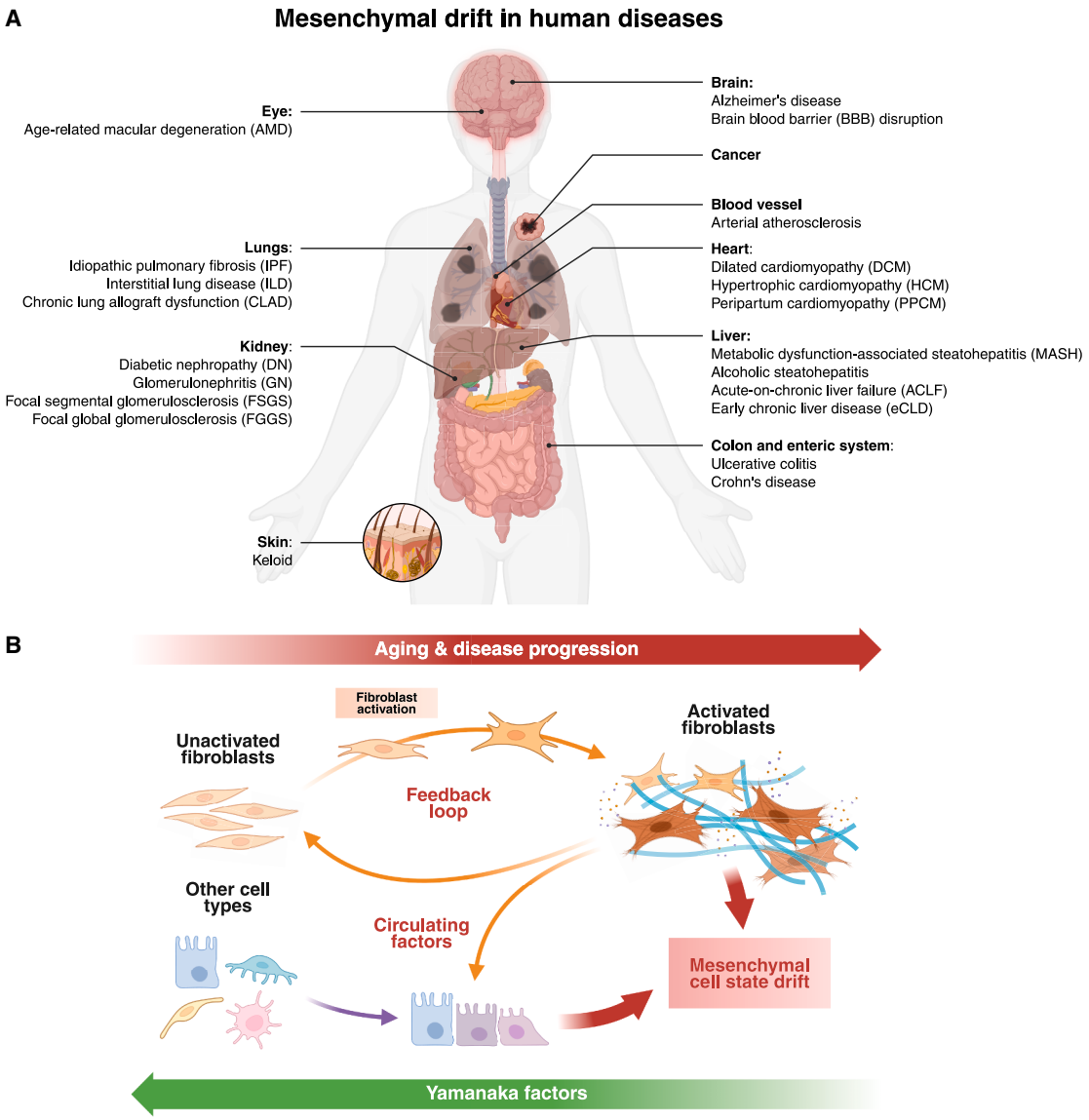


Figure 6. Model of pervasive MD in human aging and diseases

(A) Summary of evaluated disease types that showed MD dysregulation.

(B) Model of mesenchymal and other cell type contributions to MD and mitigation of MD with Yamanaka factors partial reprogramming.

immune cells of aged skin indicates a chronic state of injury and healing.

Basal mesenchymal gene expression levels differ between tissues. For example, lung fibroblasts tend to exhibit higher mesenchymal gene expression compared with kidney fibroblasts. This aligns with the observation that alveolar fibroblasts are more terminally differentiated[71] and show a stronger association with LRRC15^{+} myofibroblasts, a terminal state of myofibroblast differentiation.[82]

Although MD is associated with a range of diseases, further research is needed to clarify whether MD precedes the onset of these diseases and contributes causally to their development. MD can be detected at the early stages of MASH liver disease, with a positive correlation to disease severity (Figure 1D). MD has been observed in the non-inflammatory regions of the colon in UC, suggesting its broader role in disease progression (Figure 1L). In support of this notion, one recent study suggested that age-related EMT limits thymic function and regeneration.[83] Mesenchymal transitions or activations commonly occur during development and tissue repair, typically in a transient manner.[84] This observation underscores an intriguing convergence among aging, development, and tissue regeneration, raising the question of how mesenchymal transitions, normally transient during development and regeneration, become aberrantly sustained and widespread in aged tissues and age-associated pathologies.

Please cite this article in press as: Lu et al., Prevalent mesenchymal drift in aging and disease is reversed by partial reprogramming, Cell (2025), https://doi.org/10.1016/j.cell.2025.07.031

One possibility is that MD arises from repeated cycles of tissue injury and repair over the lifespan.[84] For instance, in an acute kidney injury model, *SOX9* is activated to repair the epithelium. However, a subset of epithelial cells, characterized by high expression of *SOX9* and an EMT marker, *CDH6*, are unable to return to their original state.[85] These cells activate adjacent fibroblasts, leading to tissue fibrosis. Comparable processes have been observed in endothelial cells across multiple organs.[77] In this scenario, epithelial and endothelial cells with mesenchymal features may accumulate after each injury. Additionally, aging induces significant damage to long-lived protein components of the ECM,[86] contributing to tissue fibrosis.[87] This detrimental process is partly driven by the excessive release of TGF-β family members and other growth factors and the nuclear translocation of the transcriptional co-activator with PDZ-binding motif (TAZ) and Yes-associated protein (YAP) transcription factors.[87] Increased ECM stiffness and the development of a pro-fibrotic niche further promote MD processes, including FMT and EMT.[88,89]

MD has the potential to spread from specific tissues to other organs, as demonstrated during liver injury.[61] In support of this, the pathway most significantly upregulated among plasma proteins during aging and mortality-associated proteins are associated with EMT (Figures 2C and 2D). Aging tissues undergoing MD may behave like dysfunctional endocrine organs, releasing excessive pro-aging and pro-EMT signals that contribute to systemic drift of cellular and tissue identity. This uncontrolled MD could be a fundamental mechanism driving aging and other aging-associated phenotypes. Importantly, MD could offer a potential target for the development of interventions to combat aging.

Fibroblast activation, transient ECM stiffening, and the secretion of inflammatory cytokines are necessary to protect wounded tissue from rupture and infection. These mechanical and cytokine changes can further activate normal fibroblasts, establishing a self-perpetuating fibrotic feedback loop. As the injury progresses into a chronic state, the fibrotic process persists and fails to resolve, leading to a shift toward a mesenchymal-favoring environment and an alteration in cell fate.[87] Here, we demonstrate that partial reprogramming of fibroblasts from aged individuals can reverse MD, resulting in an overall rejuvenated transcriptomic signature (Figures 3 and 4). This result is consistent with previous *in vivo* findings, which demonstrate that partial reprogramming leads to a reduction in fibrosis and scar formation, as well as enhanced regeneration.[90,91] In natural aging and progeroid mice and at the single-cell level in specific tissues, we show that a decrease in the MD program is associated with improved molecular and functional metrics of age. This systemic improvement may result from partial reprogramming transforming a stressed, pro-aging secretory profile into a more youthful state. Altogether, these results provide an explanatory framework in which partial reprogramming mitigates this broadly conserved process in age-related diseases and physiological aging.

Limitations of the study

Our study has several limitations. Although we found that Yamanaka factors suppress MD gene expression both *in vitro* and *in vivo*, additional epigenomic analyses will be necessary to clarify the chromatin mechanisms underlying this effect. While our analysis revealed age-associated differences in reprogramming dynamics and MD reversal, the limited number of donors may constrain the broader applicability of our findings. Future studies using a larger and more diverse cohort of fibroblasts and other cell types will be crucial to assess the robustness and inter-individual variability of MD dynamics during partial reprogramming. Such efforts may ultimately inform more personalized strategies for applying Yamanaka factor-based partial reprogramming therapies.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Juan Carlos Izpisua Belmonte (jcbelmonte@altoslabs.com).

Materials availability

This study did not generate new, unique reagents.

Data and code availability

Published data analyzed in this paper are listed in Table S1 and in the key resources table. Newly generated genomics data have been deposited in the Gene Expression Omnibus (GEO) repository under the accession numbers GEO: GSE297234 (scRNA-seq data) and GEO: GSE297233 (bulk RNA-seq). This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank Hal Barron, Rick Klausner, Steve Horvath, and all J.C.I.B. lab members for sharing ideas and feedback. We thank May Schwarz and Mariana Morales Valencia for administrative support.

AUTHOR CONTRIBUTIONS

Supervision, funding acquisition, resources, and project administration, C.R.E. and J.C.I.B.; conceptualization, investigation, formal analysis, visualization, and writing – original draft, J.Y.L., W.B.T., R.L., M.W., B.D.S., B.Y., P.R., C.R.E., and J.C.I.B.; writing – review & editing, J.Y.L., W.B.T., R.L., M.W., B.D.S., B.Y., P.R., C.R.E., and J.C.I.B. All the authors contributed to the discussion and provided feedback.

DECLARATION OF INTERESTS

The authors are employees of Altos Labs.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Cell lines
- METHOD DETAILS
  - Gene sets
  - GTEx data analysis
  - Gene set enrichment analysis
  - Gene expression analyses of human diseases
  - Single-cell RNA-seq analysis of publicly available data
  - Single-cell RNA-seq sample preparation
  - Single-cell RNA-seq analysis of newly generated data
  - Bulk RNA-seq cell culture conditions and sample preparation