Supplementary Discussion

Supplementary Discussion 1. Heterogeneity of CF (Thy1+ non-immune, non-myocyte cells). Related to Fig 2a-b and Extended Data Fig. 6e-i. Unsupervised PCA and HC on data collected from control CFs identified five gene clusters: fibroblast/cardiac genes (e.g. *Pdgfrb*, *Tbx20*), myofibroblast/ fibroblast/smooth muscle genes (e.g. *Col1a1*, *Acta2*), cell cycle genes (e.g. *Ccnb1*, *Mki67*), epicardial genes (e.g. *Msln*, *Wt1*), and endothelial genes (e.g. *Pecam1*, *Cdh5*, Fig. 2a-b). Five cell populations based on the five gene clusters were identified by both HC and PCA (Fig. 2a-b) and the major population is composed of cells expressing high levels of fibroblast and myofibroblast/smooth muscle genes (CCI Fb and CCA Fb, 76.3% total). Two minor populations are epicardial-like cells (CCI Epi and CCA Epi, 15.8% total) and endothelial-like cells (Endo, CCI only, 7.9%). Interestingly, Epi showed a much higher CCA: CCI ratio than Fb and Endo (Fig 2b and Extended Data Fig. 6f). Immunostaining of the myofibroblast/smooth muscle marker αSMA (protein product of the *Acta2* gene) and the endothelial marker Cd31 (protein product of the *Pecam1* gene) followed by flow cytometry showed that ~72.6% of CF were αSMA+ and ~9% were CD31+ (Extended Data Fig. 6g-i). Both RNA and protein expression data suggest a relatively high percentage of CF expressing myofibroblast/smooth muscle markers and a relatively low percentage of CF expressing endothelial markers. Together, these data suggest that the heterogeneous Thy1+ non-immune, non-myocyte cells are composed of a major fibroblast population and minor epicardial-like and endothelial-like populations.

Supplementary Discussion 2. Molecular cascades underlying iCM induction. Related to Fig. 3a-g. Six clusters of genes were identified that significantly related to and showed similar trends over reprogramming (Fig. 3a). The largest cluster (cluster 1) showed a trend of immediate and continuous down-regulation of gene expression. The most enriched GO terms in this cluster are "ribonucleoprotein" and "translation" (corrected p values of 1.8e-54 and 2.7e-25, Fig. 3b). Genes in these two GO terms are mainly structural constituents of the ribosome such as *Rpl18*, suggesting an immediate and potent down-regulation of the cell's protein biosynthesis machinery upon the expression of reprogramming factors. In addition to ribosomal components, the cellular machineries for protein post-translational modification ("endoplasmic reticulum"), maturation

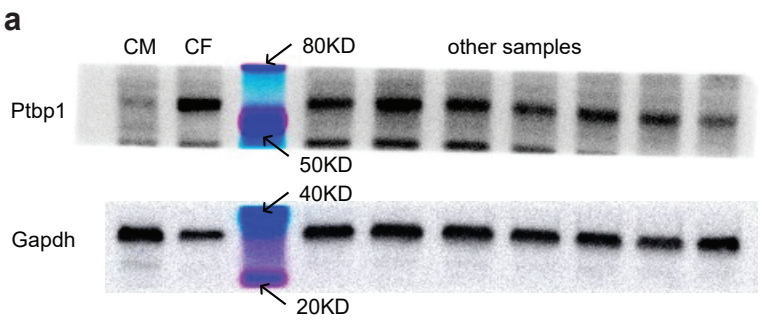
("protein folding"), trafficking ("protein transport"), and degradation ("proteasome") were all down-regulated. Among genes in another GO term enriched in cluster 1 – "transcription", there are many transcription factors including the CF lineage drivers *Tcf21* and *Sox9*, suggesting that the introduction of M, G, T may suppress lineages other than CM via down-regulation of non-CM transcription factors. Cell cycle genes are also enriched in cluster 1, consistent with our previous data in Fig. 1.

The second identified gene cluster showed a trend of expression that was initially unchanged and then dramatically decreased at three quarters of reprogramming pseudotime (Fig. 3a cluster 2). Among enriched GO terms are "endoplasmic reticulum", "aminoacyl-tRNA synthetase" and "protein folding", which are again related to protein synthesis and maturation likely for the same reasons as discussed above (Fig. 3c). In addition, the GO terms "extracellular matrix (ECM)", "glycoprotein", and "EGF-like domain" are also enriched in this cluster, suggesting a late suppression of fibroblast genes and growth factors, which is consistent with our previous data in Fig. 1 and 2 showing suppression of fibroblast markers from piCM to iCM. The third gene cluster showed an initial down-regulated and then temporarily up-regulated trend of expression (Fig. 3a cluster 3) and enriched GO terms include "mitochondrion", "electron transport", "hydrogen ion transport" and "histone H2B" (Fig. 3d). The fourth and fifth clusters are both up-regulated genes with cluster 4 showing continuous up-regulation and cluster 5 showing initial slight down-regulation and then continuous up-regulation (Fig. 3a cluster 4 and 5). Cluster 4 is enriched of GO terms "muscle protein", "cardiac muscle contraction", "GTPase activator activity" and "Wnt signaling pathway" (Fig. 3e) and cluster 5 is enriched of GO terms "oxidative phosphorylation", "cardiac muscle contraction" and "mitochondrion" (Fig. 3f), suggesting that cells are engaging in metabolic shift and structural changes towards a CM fate. The last identified cluster contained very few genes and showed only one significant GO term "glycoprotein" (Fig. 3a cluster 6 and Fig. 3g), which may worth follow-up studies focused on such proteins.

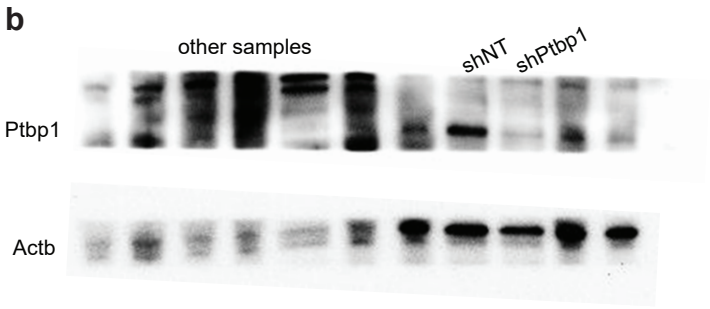
Supplementary Discussion 3. Transcriptome-repatterning at the transcription/gene expression level upon *Ptbp1* silencing. Ptbp1-mediated splicing pattern changes are a direct impact of *Ptbp1* knockdown. Yet results from these splicing pattern changes might be more profound and reflected at the gene expression level. Therefore, we further analyzed differentially

expressed genes (DEG) between MGT+shPtbp1 and MGT+shNT samples by DESeq2 (FDR < 0.05, fold change > 1.25, Supplementary Table 6). We found a very small overlap of genes that were alternatively spliced and differentially expressed upon *Ptbp1* knockdown, indicating the differences in direct and indirect consequences (Extended Data Fig. 9j). By comparing DEG between MGT vs lacZ and MGT+shPtbp1 vs MGT+shNT, we identified 28% overlapped genes (Extended Data Fig. 9k) and within the overlapped genes, 81% showed the same direction of changes (Extended Data Fig. 9l). GO analysis of overlapped genes revealed terms "ECM", "EGF-like domain", "myosin tail" and "potassium channel activity", most of which are cardiac or fibroblast fate-related (Extended Data Fig. 9n). These results suggest that *Ptbp1* knockdown enhanced the transcription changes induced by MGT, by further decreasing fibroblast gene expression and increasing cardiac gene expression (Fig. 3p). GO analysis of DEG identified only in MGT+shPtbp1 vs MGT+shNT showed enrichment of different sets of fibroblast and cardiac-related genes including increased cardiac structural genes (*Tpm1, Myl2,* and *Myh7*), fibrillogenesis inhibitor (*Fmod*), and proteases for ECM degradation (*Mmp11*), and decreased non-conventional myosins (*Myo1d, Myo1g,* and *Myo6*), fibroblast markers (*Postn, Loxl2*), and growth factors (*Tgfbr1,* and *Ctgf*) (Extended Data Fig. 9o, Fig. 3q). In mouse heart, α-Tropomyosin encoded by *Tpm1* is predominantly expressed compared to the β-Tropomyosin encoded by *Tpm2*; the ratio of α: β isoform increases from 5:1 in early embryonic life to 60:1 in the adult mice[1]. Intriguingly, depletion of Ptbp1 increased the ratio by upregulating *Tpm1* and downregulating *Tpm2*. Taken together, *Ptbp1* knockdown promoted CM fate by amplifying the MGT-induced expression changes of a subset of genes as well as regulating the expression of another different set of genes, both of which were cardiac and fibroblast fate-related.

1 Muthuchamy, M., Pajak, L., Howles, P., Doetschman, T. & Wieczorek, D. F. Developmental analysis of tropomyosin gene expression in embryonic stem cells and mouse embryos. *Molecular and cellular biology* **13**, 3311-3323 (1993).



Related to Extended Data Fig. 7h



Related to Extended Data Fig. 8b

Sup Fig 1