Supplemental Materials For:

Embryologic Origin Influences Smooth Muscle Cell Phenotypic Modulation Signatures in Murine Marfan Syndrome Aortic Aneurysm

Albert J. Pedroza MD[1], Alex R. Dalal MD[1], Rohan Shad MD[1], Nobu Yokoyama MD[1], Ken Nakamura MD[1], Paul Cheng MD PhD[2], Robert C. Wirka MD[3], Olivia Mitchel BS, Michael Baiocchi PhD[4], William Hiesinger MD[1], Thomas Quertermous MD[2], Michael P. Fischbein MD PhD[1]

[1]Department of Cardiothoracic Surgery, Stanford University School of Medicine. Stanford CA, USA
[2]Division of Cardiovascular Medicine, Stanford University School of Medicine. Stanford CA, USA
[3]Division of Cardiology, UNC School of Medicine, Chapel Hill NC, USA
[4]Department of Epidemiology and Population Health, Stanford Unviersity School of Medicine. Stanford CA, USA

Contents:

Supplemental Methods:

Concurrent scRNAseq/scATACseq Data Generation

Mouse aortic root/ascending aortic tissue was surgically dissected from 16-week-old mice after cervical dislocation and whole-body perfusion via left ventricular puncture. Tissues from four age- and sex-matched animals for each genotype were pooled, finely minced, and digested in enzymatic mix as described previously.[10] Following digestion, cell suspensions were washed in Hank's balanced salt solution. Cells were then sorted based on *tdTomato* fluorescence via fluorescence-activated cell sorting to segregate SHF-derived cells from non-lineage-traced cells. The resulting sorted cell pools were then split into two partitions for RNA and ATAC sequencing. RNA samples (8) were processed immediately using the Next GEM Single Cell 3' v3.1 reagent kit. To ensure adequate nuclei yield for ATACseq capture, separate cell suspensions from male and female cohorts were combined and centrifuged to pellet the cells. Cell lysis, washing, and nuclear counting were performed according to supplier protocols (10X Genomics, Pleasanton, CA). Nuclei were immediately processed for scATAC according to supplier specifications using the scATAC v1.1 reagent kit. RNA libraries were sequenced in parallel on a NovaSeq 6000 S1 flow cell, and ATAC libraries were sequenced in parallel on a NovaSeq 6000 SP flow cell. All library preparation and sequencing steps were performed by the Stanford Genome Sequencing Service Center.

scRNAseq Data Analysis

Raw sequencing data were processed into gene expression matrices using Cell Ranger Single-cell software suite 3.1.0 on the Stanford Sherlock High Performance Computing Cluster as described previously[10]. Base-call files were demultiplexed to generate individual samples FASTQ files using the cellranger mkfastq pipeline. Reads were then aligned to the mouse transcriptome (mm 10–3.0.0), cell barcodes and unique molecular identifiers were filtered and corrected using the cellranger count pipeline. Samples were similarly aligned to a custom transcriptome file incorporating the *tdTomato* transcript provided by Dr. Wirka to validate efficient lineage tracing via *tdTomato* gene counts in lineage-positive sorted cells (**Figure 2D**). The final output filtered expression matrices were imported into the Seurat package in R and built into Seurat objects using the *CreateSeuratObject* function. Thresholds for individual cell read counts (nCount_RNA) and genes (nFeature_RNA) were determined for each data set by excluding indices (cells) with counts outside of normal distributions for these variables. Cells with >5% mitochondrial gene content were excluded. Data normalization, scaling, and regression by mitochondrial content were then performed using the *SCTransform* command under default settings in Seurat. Principal component analysis and nonlinear dimensional reduction using uniform manifold approximation and projection (UMAP) was performed. Cell clustering was then assessed across a range of predetermined resolution scales to ensure separation of known major aortic cell types without excessively subclustering. The *FindMarkers* function in Seurat was applied to perform nonparametric Wilcoxon rank-sum test where differential expression testing is noted in **Results**. After preliminarily evaluating gene expression profiles in the disease-associated modulated SMC cluster (**Supplemental Figure 4**), cells from male and female mice were analyzed jointly, limiting the assessment of sex dimporphism in subsequent analyses.

scATAC Data Analysis

Raw base-call files from the NovaSeq 6000 sequencer were demultiplexed to generate FASTQ files using the cellranger-atac mkfastq pipeline (v 1.2.0). The reads were aligned to the mouse atac-transcriptome (mm-10–1.2.0), following which cell barcodes and unique molecular identifiers were filtered and corrected using the cellranger-atac count pipeline, yielding peak matrices and fragments files for each sample. The runs were normalized to sequencing depth and aggregated to generate a single peak-barcode matrix via the cellranger-atac aggr pipeline. Peak matrix and fragment files from individual samples were then integrated into a single chromatin assay within the Signac package in R. The dataset was filtered by multiple QC metrics (cells with >60% of reads in peaks, nucleosome signal < 2) followed by normalization, linear dimensional reduction and clustering. Differential peak accessibility testing was performed using the *FindMarkers* function in Signac, applying nonparametric Wilcoxon rank-sum test. An RNA expression matrix was built from the DNA peaks assay using the

GeneActivity function and integrated label transfer from the scRNA dataset was performed using the *FindTransferAnchors* and *TransferData* functions. Differentially accessible peaks were analyzed using Genomic Regions Enrichment of Annotations (GREAT) using the complete GRCm38 genome as background region and assigning each input peak to the single nearest gene transcription start site. Resulting enriched GO biological processes with FDR Q-value < 0.05 were considered significant. Transcription factor motif accessibility was performed using the chromVAR package in R.

Lentiviral-mediated *TWIST1* overexpression

Lentiviral plasmids for Myc-DDK-tagged TWIST1 (NM_000474, Cat# RC202920L3, Origene), TWIST1 specific targeting shRNA (Cat# TL308556, Origene) and control scrambled shRNA (Cat# TR30021, Origene) were co-transfected to HEK293T cells (Sigma-Aldrich) with packaging plasmid and transfection reagent (TR30037, Origene) according to supplier's instructions. Fresh DMEM supplemented with 10% FBS was replaced 6-8 hours after transfection. Viral-supernatants were collected 48 and 72 hours after transfection and then filtrated with 0.45um PES filter (Cat# 6896-2504, Sigma-Aldrich) and stored at -80°C.

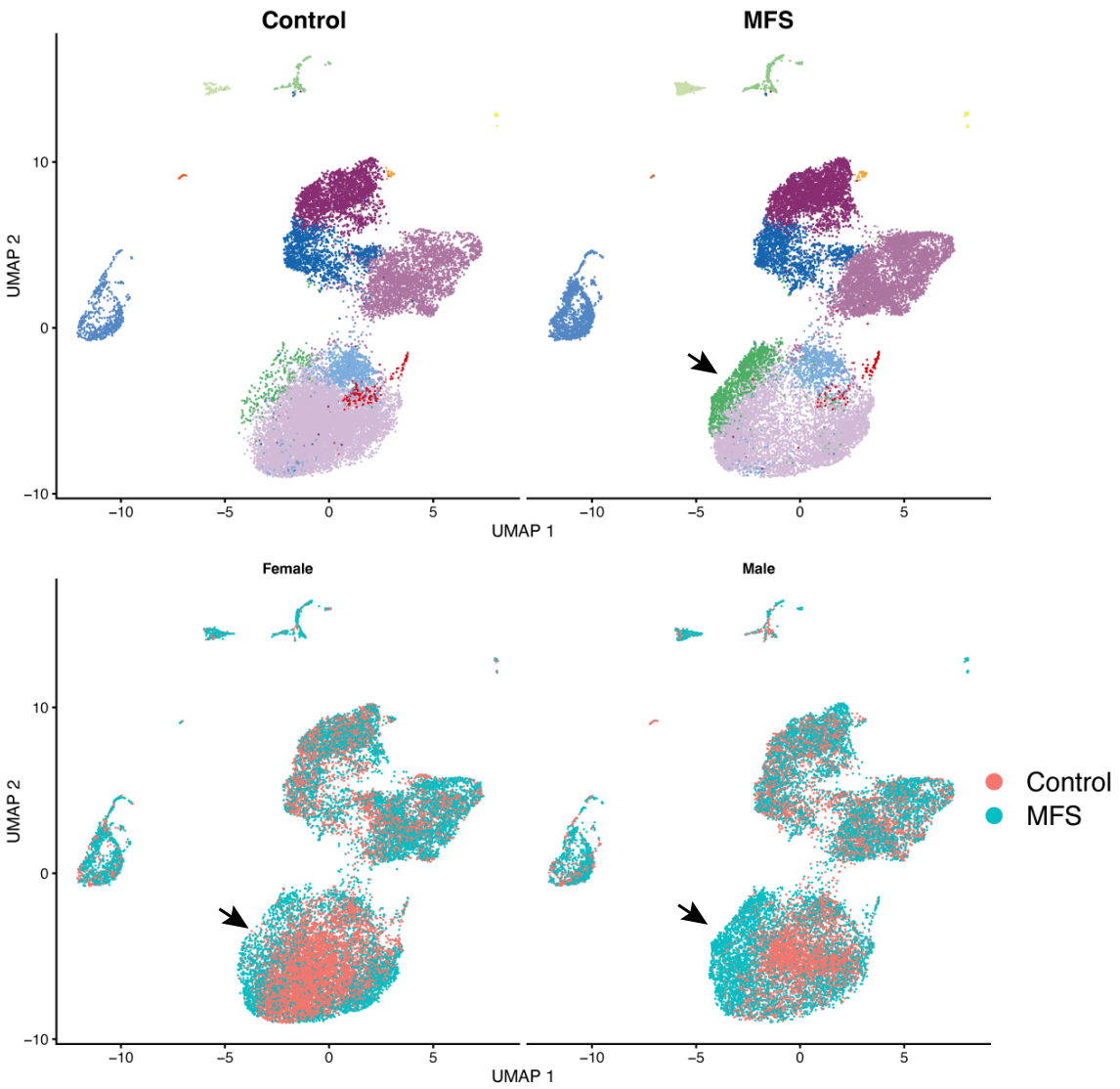
Aortic SMCs were split into 6-well plate at a density of $1.5x10^5$ cells per well. The day after splitting, viral-supernatants mixed with 0.8ug/ml polybrene (Cat# TR-1003-G, Sigma-Aldrich) were added to the wells and centrifuged 1000xg, 1 hour at 32°C. Fresh aortic SMC medium was replaced 12 hours after transduction. 48 hours after transduction, transduced cells were selected with puromycin (1ug/ml) for 5 days. Stably transduced cells were then passaged into new 6-well plates at a density of $1.0x10^5$ cells per well. Two days after splitting, cells were processed for RNA extraction or ATAC sequencing

RT-qPCR

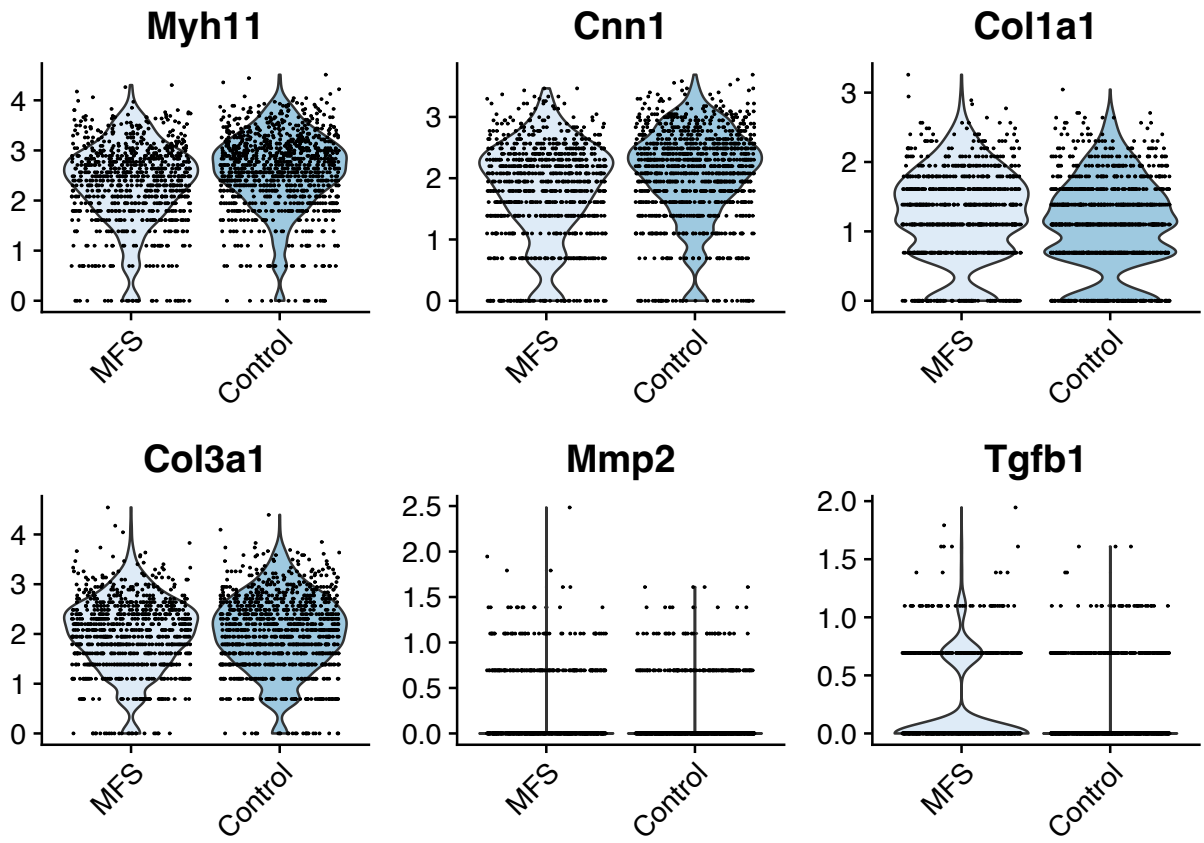
Cell lysis and RNA purification was performed using the QuickRNA miniprep kit (#R1054, Zymo Research) according to the supplier's instructions. 500ng of RNA was used for cDNA synthesis using the Maxima First Strand cDNA Synthesis kit (#K1641, ThermoFisher). cDNA was diluted ten-fold and 2uL of cDNA was used for duplexed RT-qPCR using pre-designed TaqMan assays (ThermoFisher) for specified genes of interest and *GAPDH* (housekeeping gene).

Bulk ATACseq

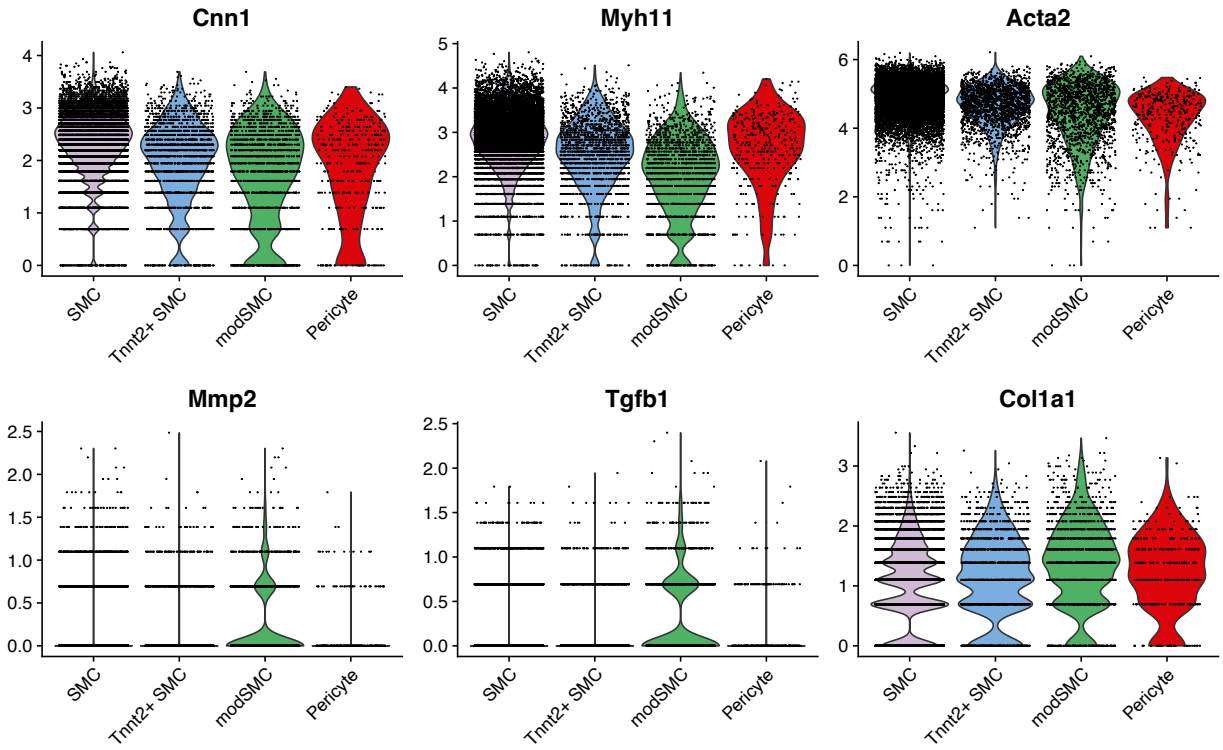
Stably transduced lines were prepared using the Active Motif ATAC-Seq kit (#53150, Active Motif) according to the provided instructions. Cells were trypsinized into single cell suspension, washed in PBS, and counted. $2x10^6$ cells from each sample were then pelleted any lysed with ATAC lysis buffer, followed by Tagmentation for 30 minutes at 37 degrees. Tagmented DNA fragments were then purified with the provided spin columns and PCR amplified with the provided i5/i7 indexed primers. The PCR product was purified with SPRI beads. Purified libraries were sequenced (150bp paired-end reads) on a NovaSeq 6000 device. Data were analyzed using a built-in pipeline comprising peak-calling using MACS2 followed by genome annotation (hg19) and enrichment analysis using Basepair software (Basepair Tech, New York, NY) via the online interface.



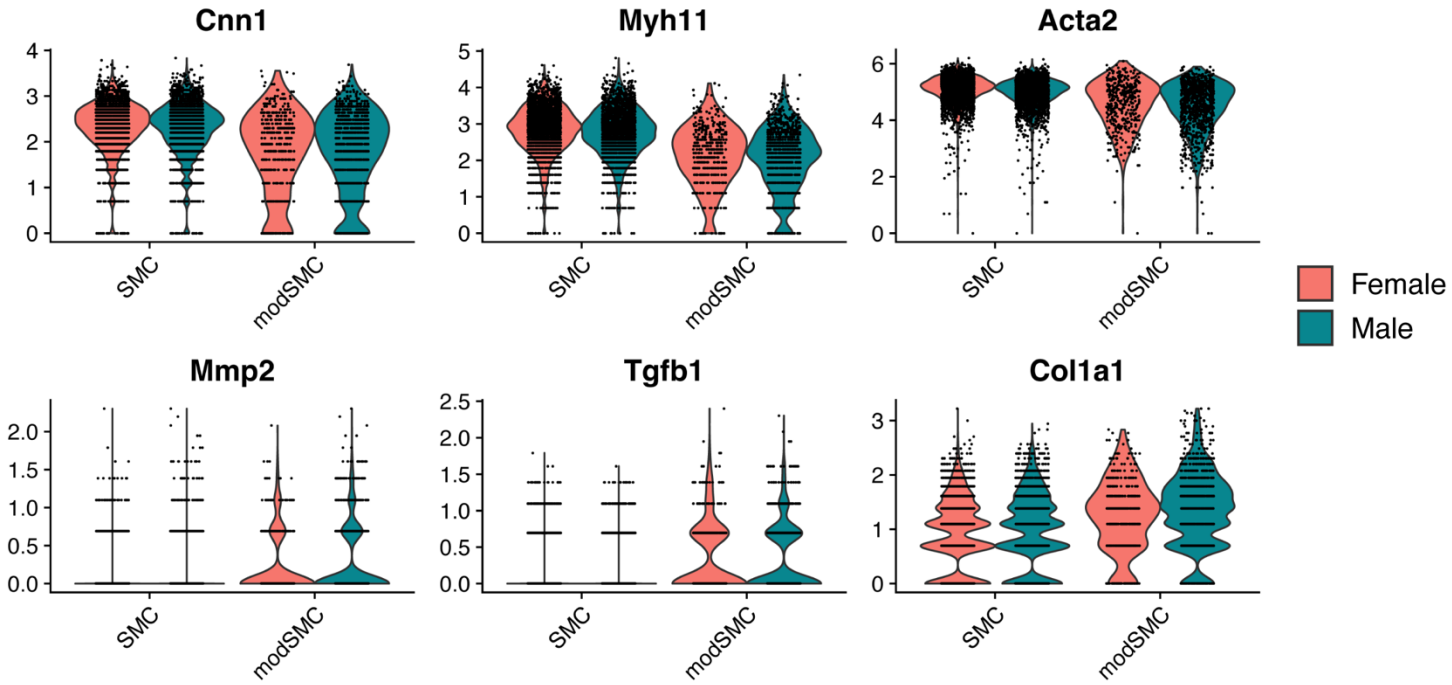
Supplemental Figure 1: UMAP projections for total scRNAseq dataset stratified by genotype with unsupervised clustering (top panel) and split by sex (bottom) grouped by genotype. Enrichment of the disease-associated cluster (green in top panel, marked by black arrow) is consistent in both male and female MFS mice.



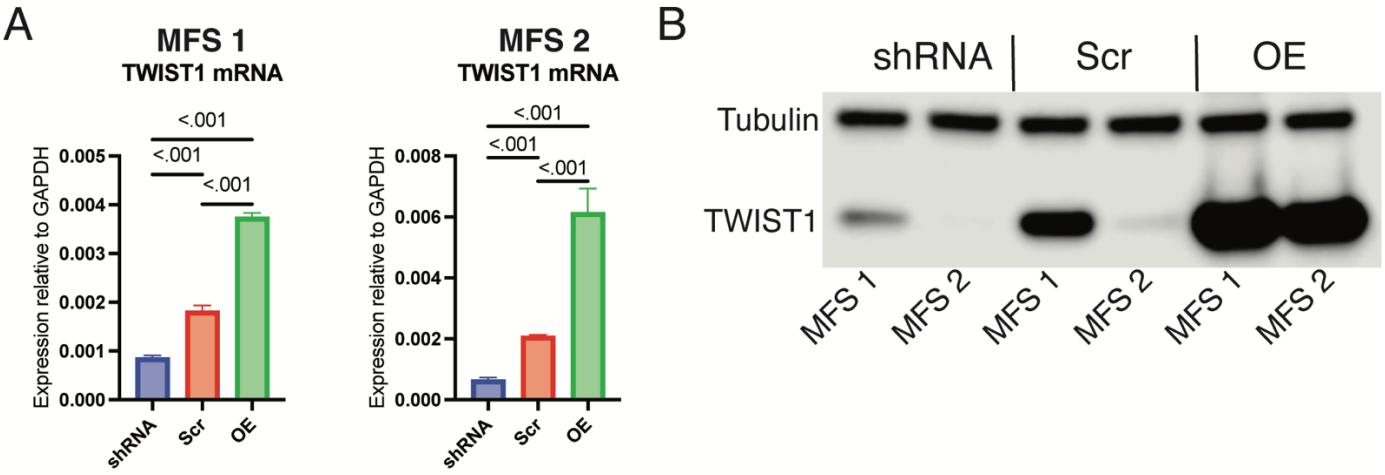
Supplemental Figure 2: Violin gene expression plots for typical markers of quiescent, contractile smooth muscle cells (*Myh11* and *Cnn1*) and Marfan aortic aneurysm modulated smooth muscle cell markers (*Col1a1, Col3a1, Mmp2, Tgfb1*) for cells from the *Tnnt2*+ SMC subset for both Marfan (MFS) and littermate control mice showing no significant difference in expression of these genes between genotypes.



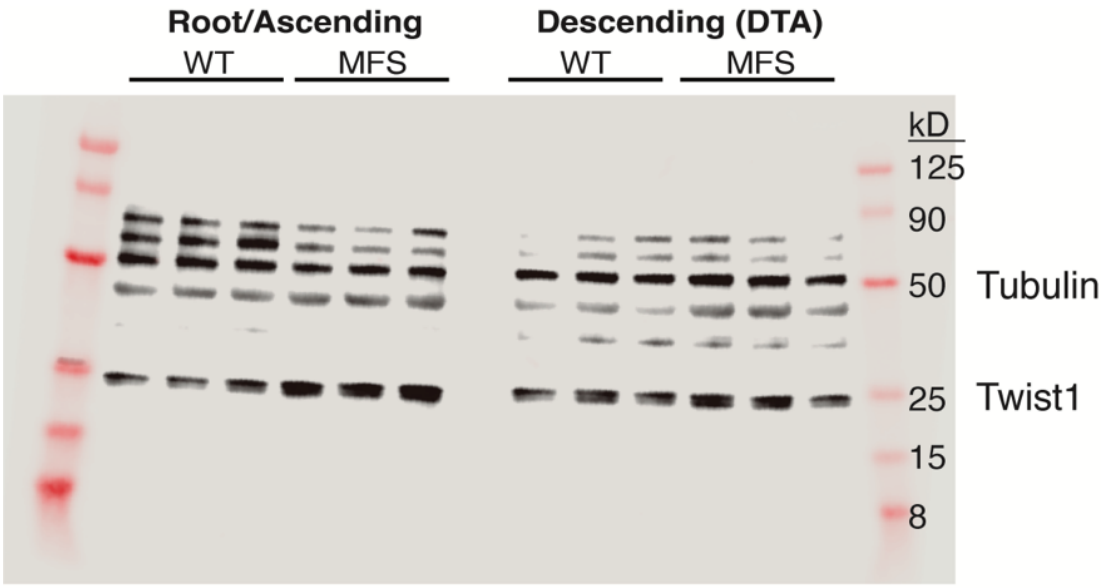
Supplemental Figure 3: Violin gene expression plots for representative smooth muscle cell contractile markers with significantly reduced expression in modSMC cluster compared to all other clusters (*Cnn1, Myh11* and *Acta2*) and modSMC cluster enriched genes (*Mmp2, Tgfb1, Col1a1*) for all four identified SMC subsets in the complete dataset (includes both SHF and non-traced lineages and Marfan/control cells).



Supplemental Figure 4: Violin gene expression plots for representative differentially expressed genes between SMC and modSMC clusters stratified by sex. None of the denoted genes were differentially expressed between male and female MFS mice.



Supplemental Figure 5: *TWIST1* overexpression in n=2 MFS aortic root SMC lines. **A** RT-qPCR data for n=4 technical replicate subcultures following lentiviral-mediated transduction of *TWIST1*-specific shRNA, scrambled shRNA, or *TWIST1* overexpression vectors and puromycin selection to establish stably transduced lines. p values indicate results of Brown-Forsythe and Welch ANOVA testing followed by Dunnett T3 multiple comparisons test. **B** Western blot demonstrating qualitative TWIST1 protein expression following transduction with indicated lentiviral constructs.



Supplemental Figure 6: Uncropped western blot image corresponding to Figure 6D in manuscript. Red ladder bands depict visible bands in 700nm channel overlaid with chemiluminescent signal.

Major Resources Table

Genetically Modified Animals

Strain	Species	Vendor or Source	Background Strain	Other Information	Persistent ID / URL
Nkx2-5$^{IRES-Cre}$	Mus musculus	The Jackson Laboratory	C57BL/6J	Second heart field (SHF) lineage-traced *Cre* recombinase	https://www.jax.org/strain/024637
Ai9 (Rosa^{tdTomato})	Mus musculus	The Jackson Laboratory	C57BL/6J	Cre-inducible fluorescent reporter line	https://www.jax.org/strain/007909
Fbn1$^{C1041G/+}$	Mus musculus	The Jackson Laboratory	C57BL/6J	Marfan mouse strain	https://www.jax.org/strain/024637

RNA *in situ* Hybridization Probes

Target mRNA	Vendor or Source	Catalog #	Working concentration	Persistent ID / URL
tdTomato (Mus musculus)	Advanced Cell Diagnostics	#317041-C2	Proprietary (as supplied)	https://acdbio.com/catalog-probes
Igfbp2 (Mus musculus)	Advanced Cell Diagnostics	#405951	Proprietary (as supplied)	https://acdbio.com/catalog-probes
Tnnt2 (Mus musculus)	Advanced Cell Diagnostics	#418681	Proprietary (as supplied)	https://acdbio.com/catalog-probes

Antibodies

Target protein	Vendor or Source	Catalog #	Working concentration	Persistent ID / URL
Twist1 (Mus musculus, Homo sapiens)	Proteintech	26465-1-AP	1:500 (1.2ug/mL)	https://www.ptglab.com/products/TWIST1-Antibody-25465-1-AP.htm
Beta-tubulin (Mus musculus, Homo sapiens)	Cell Signaling Technology	#2146	1:500	https://www.cellsignal.com/products/primary-antibodies/b-tubulin-antibody/2146

cDNA Clones

Clone Name	Vector	Source / Repository	Persistent ID / URL
TWIST1 Human Tagged ORF Clone	pLenti-C-Myc-DDK-P2A-Puro	Origene	https://www.origene.com/catalog/cdna-clones/expression-plasmids/rc202920l3/twist-twist1-nm_000474-human-tagged-orf-clone
TWIST1 Human shRNA	pGFP-C-shLenti	Origene	https://www.origene.com/catalog/rnai/shrna-plasmids/tl308556/twist-twist1-human-shrna-plasmid-kit-locus-id-7291
Scrambled shRNA Control	pGFP-C-shLenti	Origene	https://www.origene.com/catalog/rnai/shrna-plasmids/tr30021/scrambled-shrna-control-in-pgfp-c-shlenti-shrna-vector

Cultured Cells

Name	Vendor or Source	Sex (F, M, or unknown)	Persistent ID / URL
MFS 1	Clinical sample	Male	n/a
MFS 2	Clinical sample	Female	n/a

Data & Code Availability

Description	Source / Repository	Persistent ID / URL
Aortic single-cell RNA sequencing data from second heart field lineage traced mice with Fbn1C1041G/+ and control mice.	NCBI Gene Expression Omnibus (GEO)	GSE186845