HHS Public Access

Author manuscript

Arterioscler Thromb Vasc Biol. Author manuscript; available in PMC 2024 February 01.

Published in final edited form as:

Arterioscler Thromb Vasc Biol. 2023 February ; 43(2): 312–322. doi:10.1161/ATVBAHA.122.318566.

BAV-associated Regulatory Regions Reveal *GATA4* Regulation and Function during hiPSC-based Endothelial-Mesenchymal Transition

Tingting Huang, BS[1,2,+], **Jiaxi Cheng, BS**[1,+], **Hao Feng, BS**[1,2], **Wei Zhou, PhD**[3], **Ping Qiu, PhD**[1], **Dong Zhou, BS**[1,2], **Dongshan Yang, PhD**[4], **Jifeng Zhang, PhD**[4], **Cristen Willer, PhD**[4], **Y. Eugene Chen, MD, PhD**[1,4], **Dogukan Mizrak, PhD**[1,*], **Bo Yang, MD, PhD**[1,*]

[1]Department of Cardiac Surgery, University of Michigan, Ann Arbor, MI, USA;

[2]Xiangya School of Medicine, Central South University, Changsha, China;

[3]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA;

[4]Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA.

Abstract

Background: The endothelial-mesenchymal transition (EndoMT) is a fundamental process for heart valve formation, and defects in EndoMT cause aortic valve abnormalities. Our previous genome-wide association study identified multiple variants in a large chromosome 8 segment as significantly associated with bicuspid aortic valve (BAV). The objective of this study is to determine the biological effects of this large noncoding segment in human induced pluripotent stem cell (hiPSC)-based EndoMT.

Methods: A large genomic segment enriched for BAV-associated variants was deleted in hiPSCs using two-step CRISPR/Cas9 editing. To address the effects of the variants on *GATA4* expression, we generated CRISPR repression hiPSC lines (CRISPRi) as well as hiPSCs from BAV patients. The resulting hiPSCs were differentiated to mesenchymal/myofibroblast-like cells through cardiovascular-lineage endothelial cells for molecular and cellular analysis. Single-cell RNA sequencing was also performed at different stages of EndoMT induction.

Results: The large deletion impaired hiPSC-based EndoMT in multiple biallelic clones compared with their isogenic control. It also reduced *GATA4* transcript and protein levels during EndoMT, sparing the other genes nearby the deletion segment. Single-cell trajectory analysis revealed the molecular reprogramming during EndoMT. Putative GATA binding protein targets

[*]Addresses for Correspondence: Bo Yang, MD, PhD, 1500 E Medical Center Dr. 5144 Frankel Cardiovascular Center, Ann Arbor, MI, 48109, boya@med.umich.edu, Dogukan Mizrak, PhD, 2800 Plymouth Road, NCRC-26, Room 263S, Ann Arbor, MI, 48109, dmizrak@med.umich.edu.
[+]These authors contributed equally to this work.

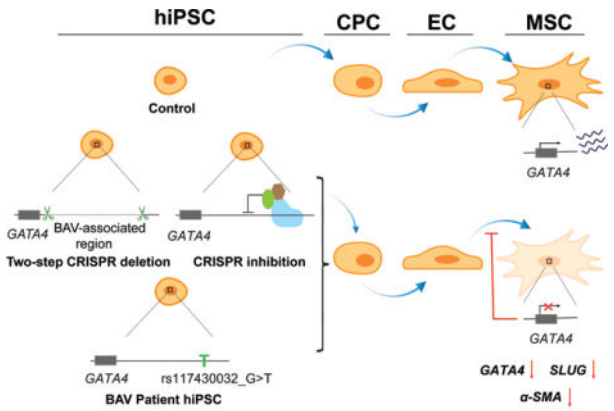
Disclosure
Dogukan Mizrak is a co-inventor on a patent application related to microwell technology for single-cell sequencing that was filed by Columbia University.

SUPPLEMENTAL MATERIALS
Supplemental Methods
Figures S1–S2
Table S1

during EndoMT were uncovered, including genes implicated in endocardial cushion formation and EndoMT process. Differentiation of cells derived from BAV patients carrying the rs117430032 variant as well as CRISPRi repression of the rs117430032 locus resulted in lower *GATA4* expression in a stage-specific manner. TWIST1 was identified as a potential regulator of *GATA4* expression, showing specificity to the locus tagged by rs117430032.

Conclusions: BAV-associated distal regions regulate *GATA4* expression during hiPSC-based EndoMT, which in turn promotes EndoMT progression, implicating its contribution to heart valve development.

Graphical Abstract



INTRODUCTION

Bicuspid aortic valve (BAV) is the most common heart malformation affecting 1–2% of all adults.[1] Patients with BAV have two aortic valve leaflets instead of the normal three, a defect formed during human aortic valve development. BAV patients are at a higher risk of developing aortic aneurysm and dissection eventually requiring surgery to replace the malfunctioning valve or to treat BAV-associated aortic aneurysm.[2] Despite the high frequency of BAV and the health risks associated with it, the genetic underpinnings of the disease are not fully understood.

Hereditary BAV is estimated to represent more than 20% of the cases.[3] Risk variants in *NOTCH1*, *SMAD6*, *MAT2A*, *ROBO4*, *CELSR1*, *GATA4*, *GATA5* and *GATA6* have been implicated in BAV formation.[4–11] Given the importance of these genes in broader heart development, BAV patients often have other heart defects. For example, GATA4 is a multifaceted transcription factor (TF) in heart development, and BAV patients carrying *GATA4* variants have higher incidences of Tetralogy of Fallot and ventricular septal defects.[12] Still, the majority of the BAV cases cannot be assigned to a single gene, and are likely a consequence of complex genetic interactions.[13]

At the early stages of human aortic valve development, myocardium induces endothelial cushion formation by promoting endothelial-mesenchymal transition (EndoMT), a complex process in which endothelial cells acquire a mesenchymal/ myofibroblast-like phenotype.[14] Regulated by blood flow and hemodynamic stress, the endocardial cushions undergo a

sequence of cell proliferation, differentiation, and death, and ultimately form distinct layers of aortic valves. EndoMT is a vital process for aortic valve modeling, and impaired EndoMT has been linked to BAV formation.[8,15,16] *SNAI1* (*SNAIL*), *SNAI2* (*SLUG*), and *TWIST1* are among the master regulators of EndoMT during valve formation.[17] Collectively, they repress the endothelial transcriptional program and promote the expression of mesenchymal/myofibroblast-like genes. Consistently, mutations in BAV-associated genes such as *NOTCH1*, *ROBO4*, and *GATA4* also impair EndoMT in animal models.[7,18,19] However, a mechanistic understanding of their role in EndoMT and their molecular interactions with key EndoMT TFs are yet to be determined.

In a genome-wide association study of a large cohort, we identified noncoding and protein-altering genetic variants at 8p23.1 associated with BAV.[8] Although we did not confirm the functional effects of these noncoding variants, chromatin conformation analysis implicated that the intergenic risk locus encompassing BAV variants physically interact with *GATA4* gene. In recent years, several variants located in noncoding distal elements have been linked to heritable disorders. rs2431697 is located in a distal enhancer of miR-146A and contributes to the pathogenesis of systemic lupus erythematosus by impairing miR-146A expression.[20] Parkinson's disease associated risk variant rs356168 tags a distal enhancer element of α-synuclein (SNCA) regulating its expression.[21] rs9349379 tags an enhancer signature regulating *EDN1* expression, which is associated with multiple vascular diseases.[22] However, noncoding risk variants contributing to BAV susceptibility have been unexplored.

In this study, we deleted a large genomic segment encompassing multiple BAV-associated variants in human induced pluripotent stem cells (hiPSCs) using two-step CRISPR/Cas9 editing. The large deletion reduced *GATA4* expression during hiPSC-based EndoMT, sparing the other genes neighboring the segment. The large deletion also impaired EndoMT by disrupting mesenchymal/myofibroblast-like cell output, and differentiation. Single-cell RNA sequencing (scRNA-seq) experiments revealed the temporal gene expression patterns highlighting upregulation of pluripotency, cytoskeleton remodeling, and TGFβ signaling dependent cell migration gene sets in the transformed cells. Motif analysis showed the enrichment of GATA binding motif around transcription start sites (TSS) of several late-stage EndoMT markers. CRISPRi repression experiments identified the rs117430032 locus in the large segment as a regulator of *GATA4* expression, which was also lower in cells derived from two different BAV patients carrying the rs117430032 variant. rs117430032 tags a putative binding site of master EndoMT regulators including TWIST1, which showed specificity to the locus. In sum, we provide evidence that BAV-associated distal elements regulate *GATA4* expression, which sustains hiPSC-based EndoMT.

METHODS

All experiments were performed according to the protocols approved by the Institutional Review Board at the University of Michigan. All materials can be made available from the corresponding authors on reasonable request. Detailed descriptions of reagents and experimental procedures are available in the Supplemental Material. The raw and processed single-cell RNA sequencing (scRNA-seq) data have been deposited to the Gene Expression Omnibus database under the accession number GSE193206.

EndoMT Induction

Endothelial cell (EC) differentiation through cardiovascular progenitor cell (CPC) was performed as described previously.[23] First, hiPSCs were dissociated into single cells using versene (Cat # 15040066, Gibco) and seeded on matrigel-coated plates at a density of 4 × 10^4 cells per cm^2 TesRE8 medium supplemented with 10 μM ROCK inhibitor Y2763217. After 24 hours, the medium was replaced with CPC differentiation medium, consisting of 1:1 mixture of DMEM-F12 (Gibco) and Neurobasal medium (Gibco), B27 supplement without vitamin A (Gibco), N2 supplement, penicillin–streptomycin, β-mercaptoethanol, 25 ng/ml BMP4 (PeproTech) and 8 μM CHIR99021 (PeproTech). After three days, CPC differentiation medium was replaced with EC induction medium, which contained Stempro34 medium (Gibco), Stempro 34 supplement (Gibco), 1x Glutamax (Gibco), 1x penicillin–streptomycin, 200 ng/ml VEGF (PeproTech) and 2 μM forskolin (Sigma-Aldrich). After EC induction for 2 days, immature EC were dissociated with Accutase (Gibco), and reseeded onto fibronectin (Cat 356008, Corning)-coated plates at a density of 2 × 10^4 per cm^2 in EC differentiation medium (Stempro34 medium, Stempro 34 supplement, Glutamax, penicillin–streptomycin, 50 ng/ml VEGF (PeproTech). EC differentiation medium was replenished daily for 5 days.

To induce EndoMT, EC differentiation medium was replaced with EndoMT induction medium at the end of EC differentiation. EndoMT induction medium contained Stempro34 medium, Stempro34 supplement, 1x Glutamax, 1x penicillin–streptomycin, 200 ng/ml BMP2 (PeproTech) and 50 ng/ml TGFβ2 (PeproTech). EndoMT induction was conducted for three days as previously described,[8] and the resulting cells were collected for the subsequent analysis.

Statistics

GraphPad Prism Software was used for the statistical analyses. All qRT-PCR, immunostainings, cell migration, apoptosis, CHIP-qPCR and enhancer testing data were presented as mean± standard deviation with at least three biological replicates. When analyzing more than two groups with five or less replicates, we conducted nonparametric Kruskal-Wallis test with multiple testing correction by controlling the False Discovery Rate (Two-stage step-up method of Benjamini, Krieger and Yekutieli). When analyzing more than two groups with six replicates that are normally distributed (Shapiro-Wilk normality test), we conducted one-way ANOVA analysis with multiple testing correction by controlling the False Discovery Rate (Two-stage step-up method of Benjamini, Krieger and Yekutieli). q value<0.05 was considered statistically significant, and individual p values were presented in the Figures up to the significance level p<0.001. When analyzing only two datasets, we used nonparametric two-tailed Mann Whitney test. p value<0.05 was considered statistically significant.

RESULTS

Deletion of a BAV-associated large genomic segment reduces GATA4 levels and impairs hiPSC-based EndoMT

Aortic valve development relies on EndoMT in which endocardial cells lose their endothelial identity and change toward a mesenchymal or myofibroblastic phenotype (Figure 1A).[24] Following EndoMT transition, transformed cells exhibit a migratory phenotype and a distinct morphology with high expression of *alpha smooth muscle actin* (*α-SMA*).[25] We previously identified a large genomic segment with a significant BAV association.[8] To understand the role of this large segment in hiPSC-based EndoMT, we deleted the region in hiPSCs using CRISPR/Cas9 editing (Figure 1B). We utilized a two-step guide RNA (gRNA) strategy to delete nearly 82 kilobase (kb) (Figure 1B). gRNA1 and gRNA2 target upstream and downstream regions of the 82 kb segment respectively. The first round of editing using gRNA1 and gRNA2 resulted in monoallelic deletion only. As the upstream gRNA target site was modified in the monoallelic clones, we used a different upstream gRNA (gRNA3) with the same downstream gRNA2 for the second round of CRISPR/Cas9 editing. This resulted in several biallelic deletion clones (Figure S1A). The deletion clones were first validated using two sets of primers, named L1+R1 and L2+R2 (Figure S1A). Karyotyping analysis also revealed no obvious chromosomal defects in the deletion lines (data not shown). To further confirm bi-allelic deletion, PCR products (L2+R2 primers) from three biallelic deletion clones (Δ−1, Δ−2 and Δ−3) were cloned into a TA cloning vector. Plasmid DNA was sequenced via Sanger sequencing, which indicated both alleles of the 82 kb segment were completely deleted in three clones with different indel patterns (Figure 1B).

To identify the effects of the deletion on *GATA4* expression, we differentiated the control and mutant hiPSCs to mesenchymal/myofibroblast-like cells (MSCs) through cardiovascular progenitor cell (CPC)-derived endothelial cells (ECs) using a previously established EndoMT protocol (Figure 1C).[8,23] First, we confirmed the generation of functional endothelial cells using the hiPSC model. Endothelial marker expression (*PECAM1* and *CDH5*) were hundreds to thousands of folds higher at the EC stage compared with the CPC stage, and hiPSC-derived ECs formed a continuous layer (Figures S1B–C). Importantly, *PECAM1* and *CDH5* levels were similar between hiPSC-derived ECs and primary human coronary artery endothelial cells (Figure S1B). Next, we conducted tube formation assay on hiPSC-derived ECs, which showed their ability to form capillary-like tubes (Figure S1D). Lastly, *PECAM1* and *CDH5* levels were significantly diminished in MSCs compared with ECs suggesting reduced endothelial identity at the MSC stage (Figure S1E).

To assay *GATA4* expression, we collected mRNA from three differentiation stages; CPCs, ECs and MSCs. Although *GATA4* was detectable in all three stages, we observed a significant elevation in *GATA4* levels at the MSC stage compared with CPC and EC stages in control cells suggesting additional *GATA4* transcriptional regulation during EndoMT (Figure 1C). Strikingly, *GATA4* transcript levels were significantly reduced in all three deletion clones at the MSC stage (Figure 1C). We also performed qRT-PCRs on other genes neighboring the 82 kb segment; *NEIL2*, *FDFT1*, *CTSB*, and *DEF134* at three differentiation

stages. Although we observed changes in these transcripts in individual deletion clones, we did not detect consistent changes in all three deletion clones (Figure S1F). Furthermore, 82 kb region harbors two genes; *DEFB135* and *DEFB136*. The expression levels of both genes were almost undetectable compared with *GAPDH* in MSCs suggesting that they are not actively transcribed during EndoMT (Figure S1G). We also performed immunostainings to quantify nuclear GATA4 protein levels at the MSC stage, which revealed significantly lower GATA4 levels in the deletion clones (Figure 1D). To confirm the *GATA4* elevation during EndoMT is not restricted to the hiPSC model, we examined *GATA4* levels in human umbilical vein endothelial cells before and after EndoMT induction. We observed a significant reduction in endothelial markers (*PECAM1* and *CDH5*) at the MSC stage while *GATA4* and *SLUG* expression were increased (Figure S1H).

Next, we measured the functional effects of the large deletion by quantifying the percentage of α-SMA+ cells in the control and mutant conditions on Day 3 of EndoMT. Controls MSCs exhibited strong α-SMA labeling with an enlarged morphology while we observed a dramatic reduction in α-SMA+ cell numbers in all three mutant conditions (Figure 1E). To examine the transcriptional defects caused by the large deletion, we performed qRT-PCRs to measure the levels for *α-SMA* and three key EndoMT TFs; *SNAIL*, *SLUG* and *TWIST1*. *SLUG* levels were significantly lower in the deletion clones while *SNAIL* and *TWIST1* levels were unaltered (Figure 1F).

Loss of cell-cell contact and acquisition of a migratory phenotype are hallmarks of MSCs undergoing EndoMT. To assess the migratory function of the mutant cells, we carried out a transwell migration assay.[26] Briefly, ten thousand cells were plated onto a chemotaxis insert and incubated for 24 hours at 37°C. The fluorescent signal in the migrated cells at the lower chamber were measured in the control and mutant conditions. The deletion clones exhibited reduced migratory ability at the MSC stage confirming hiPSC-based EndoMT defects (Figure 1G). EndoMT is also marked by a significant cell reduction particularly during the first two days of transition and MSCs settle for their final fate on Day 3. GATA4 is a critical TF for cell survival in heart development by regulating anti-apoptotic *BCLX* expression[27] and a loss of function GATA4 mutation causes hypocellularity in the heart valves.[25] When we measured *BCLX* levels in different MSC stages, we observed a *BCLX* peak on Day2 of EndoMT, which disappeared in the mutant MSCs (Figure S1I). We also measured the apoptotic rate by measuring Caspase-3 and Caspase-7 activities in the control and mutant cells between Day 1–2 of EndoMT induction. Mutant MSCs had significantly higher luminescence signal which is proportional to the apoptosis rate (Figure 1H). Overall, these data suggest that the deletion of an 82 kb region enriched for BAV variants impairs *GATA4* expression in a stage-specific manner sparing other neighbor genes, which in turn modulates cell differentiation, survival and migration properties during EndoMT.

Temporal Regulation of Molecular Reprogramming during EndoMT

To track the molecular changes during EC to MSC reprogramming, we performed scRNA-seq on cells isolated on Day 1 (D1) and Day 3 (D3) of EndoMT induction. The single-cell trajectory analysis using Monocle 3 R package revealed the order of gene expression changes (Figure 2A).[28] First, we identified pseudotime-dependent genes and colored their

expression on the trajectory. Endothelial markers such as *PECAM1*, *CDH5* and *NOTCH1* were highly expressed on D1 while classical EndoMT markers such as *α-SMA*, *COL1A1* and *COL3A1* as well as *GATA4* were enriched in D3 MSCs (Figures 2B–C). *SERPINE1 (PAI-1)* and *TFPI2* were enriched in a group of D1 cells that are more advanced in the pseudotemporal order. These transitory cells also had lower expression of endothelial markers suggesting that they are beginning to lose their endothelial identity while expressing immediate early genes (Figure 2B). Of note, TGFβ ligand is an essential factor of EndoMT and induces *SERPINE1* expression, which is a molecular switch controlling cardiac TGFβ signaling.[29,30]

To investigate the top molecular pathways co-regulated with *GATA4* expression, we identified the significantly enriched transcripts (247 transcripts, adjusted p value (p-adj<0.01, Fold-change>1.5, Supplemental Excel File I) in D3 MSCs compared with D1 cells, and performed integrated pathway analysis using Metacore from Genego (Figure S1J, Supplemental Excel File I). This analysis revealed enrichment of cytoskeleton and extracellular matrix (ECM) remodeling, self-renewal and pluripotency as well as TGFβ induced migration, fibrosis and epithelial-mesenchymal transition (EMT) gene sets in D3 MSCs confirming significant reprogramming during EndoMT process.

To reveal co-regulated TFs with *GATA4*, we calculated the co-expression probability ratios for all genes, and isolated the TFs expressed in more than 1% of all cells.[31] Key EndoMT-drivers such as *TWIST1* and *SNAIL* had high co-expression scores with *GATA4* while *GATA5* was the top co-expressed TF (Figure 2D). *GATA4* expression was mutually exclusive with the regulators of endothelial identity such as *SOX7*, *SOX17*, *SOX18* and *ERG* further confirming additional *GATA4* regulation during EndoMT.[32,33] Next, we aimed to identify the transcriptional regulators of significantly enriched genes in D3 MSCs. Using iRegulon motif discovery method, we detected enriched TF motifs in the 20 kb region centered around TSS of D3 MSC enriched genes, and highlighted optimal TFs targeting these sites (Figure 2E). The candidate TFs included Serum response factor (SRF), NF1, and PARP1 which have been implicated in EndoMT and EMT induction.[34–37] Strikingly, GATA binding motif was one of the top enriched motifs implying the importance of GATA4 in EndoMT process (Figure 2E). Potential GATA4 targets included *MYOCD*, as well as *VCAN*, *GPC6*, and *LTBP1* which are implicated endocardial cushion formation (Supplemental Excel File I).[38–40] In sum, these data suggest that GATA4 is a transcriptional regulator of several key transcripts enriched in D3 MSCs.

BAV Patient-derived cells and CRISPRi repression reveal a putative causal variant in the Large Genomic Segment and TWIST1-mediated *GATA4* expression

Given the EndoMT and *GATA4* expression defects caused by the large deletion clones, we aimed to identify individual variants contributing to the GATA4 regulation. The large genomic segment contains 97 common variants genome-wide-significantly associated with BAV including rs6601627.[8] To predict chromatin impacts of these variants, we conducted a DeepSea analysis and curated a short list of five variants with high DeepSea disease-impact scores or heart-specific scores for additional analysis.[41] To address the effects of these variants on *GATA4* expression, we generated CRISPR repression lines (CRISPRi) by

stably integrating dcas9-KRAB-MeCP2 in hiPSCs.[42] We then designed fluorescent protein tagged sgRNAs targeting each variant locus, and selected sgRNA expressing cells using fluorescent-associated cell sorting (Figures S1K and S2A). Following EndoMT induction, we measured *GATA4* levels in D3 MSCs. *GATA4* expression was unaltered in rs75747817i, rs6601627i, rs118065347i lines while it was significantly lower in one rs117157630i clone and two independent clones of rs117430032i (Figure 2F) compared with their isogenic control carrying dcas9-KRAB-MeCP2 (KRAB Control). The neighboring gene levels were not altered in both rs117430032i clones suggesting that *GATA4* is likely the primary target of rs117430032i (Figure S2B). We then conducted a small scale scRNA-seq experiment on D3 rs117430032i MSCs and performed single-cell differential expression analysis comparing rs117430032i sample to the D3 control (Supplemental Excel File I). To construct a high confidence list of GATA4 targets, we found the overlapping genes between putative targets identified in the motif analysis (Figure S2C) and significantly de-enriched genes in rs117430032i MSCs (p-adj<0.01, FC>1.5, Supplemental Excel File I). This list included *GPC6, LTBP1, VCAN* and *MYOCD*. We validated rs117430032i effects on the expression of these genes, which were significantly lower in rs117430032i clones suggesting them as GATA4 targets in hiPSC-based EndoMT assay (Figure S2C).

rs117430032 is located near the edge of the topologically associated domain spanning from hg19:chr8:11250000 to 11825000 defined by Hi-C in K562 and GM12878 cells, which suggested the region was brought in close proximity to *GATA4*.[8] The variant is in high linkage disequilibrium (r^2 = 0.79 in Europe samples of 1000 Genomes) with the most significant variant rs6601627 at the locus and is significantly associated with the risk of BAV (odds ratio=2.35, $P = 1.85 \times 10^{-8}$) (Figure S2D).[43] Next, we generated hiPSCs using peripheral blood mononuclear cells isolated from two male BAV patients carrying the rs117430032 variant (BAV-1^{rs117430032} and BAV-2^{rs117430032}), and a healthy male control. Sanger Sequencing verified the presence of the variant sequence in BAV Patient-derived hiPSCs, and its absence in the new male control and the isogenic control hiPSCs used in the deletion experiments (Figure S2E). Consistent with the isogenic control used in the large deletion experiments, the new control also had elevated *GATA4* expression at the MSC stage compared with CPC and EC stages indicating that the stage-specific *GATA4* enrichment in hiPSC model is independent of genetic background. Strikingly, *GATA4* expression was significantly reduced in BAV-1^{rs117430032} and BAV-2^{rs117430032} samples at the MSC stage (Figure 2G) confirming the disruption of *GATA4* regulation in BAV Patient-derived cells during hiPSC-based EndoMT. We then measured the EndoMT marker expression in both BAV Patient-derived cells and rs117430032i clones. Consistent with the large deletion phenotypes, both *α-SMA* and *SLUG* levels were diminished in the BAV Patient samples and the rs117430032i clones at the MSC stage further confirming the gene expression correlation between *GATA4* and EndoMT markers (Figure 2H).

Lastly, we investigated how rs117430032 variant could affect *GATA4* expression. rs117430032 marks an E-box (enhancer box) motif (5'-CAGGTG-3', alternate allele 5'-CAGTTG-3'), which is recognized by master EndoMT drivers with high specificity.[44,45] SNAIL and SLUG are classically categorized as transcriptional repressors, while TWIST1 exhibits bi-functional roles as both a repressor and an activator of gene transcription.[46] Furthermore, *TWIST1* has a high co-expression score with *GATA4* at the MSC stage

(Figure 2D), which we confirmed at the protein level by immunostainings (Figure S2F). To examine whether the master EndoMT drivers interacts with this E-box element, we designed an enhancer testing vector with a luminescent reporter. The putative enhancer sequence (~200 bp centered at the E-box element tagged by rs117430032) was placed upstream of a minimal promoter (miniCMV) which controls the luciferase reporter expression. We then co-transfected the enhancer testing vector and *SNAIL* or TWIST1 expressing plasmids in hiPSCs and measured reporter activity. *TWIST1* expression significantly elevated reporter activity while *SNAIL* expression had no effect (Figure 2I). Consistently, *TWIST1* expression in hiPSCs significantly increased *GATA4* transcription (Figure 2I). To assay TWIST1 interaction with the locus tagged by rs117430032, we performed Chromatin Immunoprecipitation followed by quantitative PCR analysis (ChIP-qPCR). In control MSCs, the target locus expression was enriched by nearly five-fold with TWIST1 antibody compared with the IgG control (Figure 2J). In the BAV patient sample carrying rs117430032, the enrichment was only modest (~two-fold enrichment on average) with TWIST1 antibody compared with the isotype control. Taken together, these data suggest that TWIST1 shows specificity for the E-box element tagged by rs117430032 and is a potential regulator of *GATA4* expression.

DISCUSSION

Here we identified a large BAV-associated genomic segment that regulates *GATA4* expression during hiPSC-based EndoMT. Our molecular and cellular characterization using hiPSC model revealed an enrichment of *GATA4* during EndoMT induction, and its involvement in sustaining MSC identity and output. Our findings are in line with animal model studies of *Gata4* in heart valve development. Conditional knockout of *Gata4* in endothelial-derived cells disrupts EndoMT, and causes hypocellular atrioventricular cushions.[25] Similarly, *Gata4*$^{G295ski/wt}$ mutation reduces aortic valve cushion volume at early embryonic days. Although the total cusp volume partially recovers at a later stage, non-coronary cusp volume remains statistically smaller.[18] In addition, GATA4 promotes postnatal cardiomyocyte cell survival by activating anti-apoptotic *BCLX* expression.[27] In this study, we primarily focused on the EndoMT stage as the deletion of BAV-associated segment does not appear to affect *GATA4* expression in earlier stages. Importantly, *Gata4* activity is also critical for endocardial cushion endothelial cells in animal models. Detailed profiling of human endocardial cushion endothelial cells during embryonic aortic valve development would be instrumental in developing better hiPSC models reflecting *in vivo* developmental stages. Overall, the pleiotropic EndoMT defects observed in the large deletion clones in hiPSC model are consistent with the versatile role GATA4 plays in heart development.[25]

Generation of large biallelic deletions using CRISPR/Cas9 editing is challenging. We achieved this using a dual-gRNA strategy previously utilized in other studies.[47] The deletion clones revealed GATA4 regulation by the distal elements within the large genomic segment. We also carried out CRISPRi transcriptional repression experiments to identify noncoding variants regulating *GATA4* expression. Although we do not rule out the possibility that there are multiple functional variants affecting *GATA4* expression in the large genomic segment, our data implies the involvement of the rs117430032 variant in the transcriptional regulation

of GATA4 during EndoMT. rs117430032 tags a potential E-box element recognized by master EndoMT drivers with high specificity.[44,45] Our data implies that TWIST1 binds to this locus and regulates *GATA4* expression. SNAIL and SLUG are classically associated with transcriptional repression while TWIST1 can function both as a transcriptional repressor and activator.[46] This is consistent with our finding that *SNAIL* transfection did not alter the reporter expression. Our data further revealed that *GATA4* and *TWIST1* have high co-expression probability scores suggesting that they are enriched at the same MSC stage. GATA4-TWIST1 signaling have also been implicated in atherosclerosis development by promoting EC permeability and proliferation as well as EndoMT.[48]

We used TGFβ2 and BMP2 ligands to induce EndoMT as previously described.[8] Integrated pathway analysis revealed the enrichment of several TGFβ induced pathways during EndoMT progression. Our single-cell trajectory analysis revealed a transitory cell type among D1 cells with high expression of canonical TGFβ signaling target genes including *SERPINE1* while *GATA4* expression peaks at later EndoMT stages. We also generated a short list of potential GATA binding protein targets during EndoMT by combining motif and differential gene expression analysis, validated some key targets including *MYOCD*; a coactivator of SRF,[49] the binding motif of which was highly enriched around the TSS of D3 MSC enriched genes. *MYOCD* regulation implicates the effects of GATA4 on cytoskeletal and ECM genes critical for the migratory and invasive EndoMT phenotype. Of note, the effects on *MYOCD* expression could also be indirect as the potent *MYOCD* regulator; *TEAD1* is also enriched in D3 MSCs (Supplemental Excel File I).[50] Importantly, the other putative GATA targets; *GPC6*, *LTBP1* and *VCAN*, have been implicated in endocardial cushion formation and EndoMT process.[38–40,51] We also observed a correlation between *GATA4* levels and *SLUG* expression which contributes to the migratory phenotype during EndoMT.[52] *SLUG* expression is strictly regulated by TGFβ signaling. GATA4 could also regulate *SLUG* expression through *LTBP1* regulation, which promotes TGFβ signaling.[53]

Access to human endocardial cushion cells during aortic valve development is very limited. We used the hiPSC model and EndoMT assay as a proxy to study the role of a BAV-associated large genomic segment in aortic valve formation. This highlights a major limitation of this study as our study material was limited to human-derived cells and two-dimensional cell culture assays. To examine *GATA4* regulation during EndoMT, we generated hiPSCs from BAV Patients and healthy controls. Although the hiPSC model poses limitations such as not thoroughly reflecting *in vivo* cell features, we believe that unlimited source of BAV Patient-derived hiPSCs could offer new avenues in understanding the complex BAV formation and its association with aortic aneurysm and dissection. Furthermore, several tissue-engineered heart valve methods have been developed, but their overall goal is to generate implantable heart valves for valve replacement surgeries. Therefore, existing tissue-engineered heart valve models need to be optimized to measure valve formation abnormalities.[54] In the future studies, better models are required to evaluate valve fusion defects caused by BAV variants including the large genomic segment characterized in this study.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

Sources of Funding

This study was supported by National Institutes of Health grants HL130614, HL141891, and HL151776 (B.Y.), HL159900, HL109946 and HL159871 (Y.E.C.), HL138139, HL153710 and GM122181 (J. Z.).

NON-STANDARD ABBREVIATIONS AND ACRONYMS

BAV	bicuspid aortic valve
EndoMT	endothelial-mesenchymal transition
scRNA-seq	single-cell RNA sequencing
hiPSC	human induced pluripotent stem cell
TF	transcription factor
TSS	transcription start site
gRNA	guide RNA
EC	endothelial cell
CPC	cardiovascular progenitor cell
MSC	mesenchymal/myofibroblast-like cell
TGFβ	transforming growth factor beta
ECM	extracellular matrix
α-SMA	alpha smooth muscle actin

REFERENCES

1. Martin LJ, Ramachandran V, Cripe LH, Hinton RB, Andelfinger G, Tabangin M, Shooner K, Keddache M, Benson DW. Evidence in favor of linkage to human chromosomal regions 18q, 5q and 13q for bicuspid aortic valve and associated cardiovascular malformations. Hum Genet. 2007;121:275–284 [PubMed: 17203300]
2. Braverman AC. Aortic involvement in patients with a bicuspid aortic valve. Heart. 2011;97:506–513 [PubMed: 21339321]
3. Bravo-Jaimes K, Prakash SK. Genetics in bicuspid aortic valve disease: Where are we? Prog Cardiovasc Dis. 2020;63:398–406 [PubMed: 32599026]
4. Garg V, Muth AN, Ransom JF, Schluterman MK, Barnes R, King IN, Grossfeld PD, Srivastava D. Mutations in notch1 cause aortic valve disease. Nature. 2005;437:270–274 [PubMed: 16025100]
5. Xu YJ, Di RM, Qiao Q, Li XM, Huang RT, Xue S, Liu XY, Wang J, Yang YQ. Gata6 loss-of-function mutation contributes to congenital bicuspid aortic valve. Gene. 2018;663:115–120 [PubMed: 29653232]

6. Guo DC, Gong L, Regalado ES, Santos-Cortez RL, Zhao R, Cai B, Veeraraghavan S, Prakash SK, Johnson RJ, Muilenburg A, et al. Mat2a mutations predispose individuals to thoracic aortic aneurysms. Am J Hum Genet. 2015;96:170–177 [PubMed: 25557781]
7. Gould RA, Aziz H, Woods CE, Seman-Senderos MA, Sparks E, Preuss C, Wünnemann F, Bedja D, Moats CR, McClymont SA, et al. Robo4 variants predispose individuals to bicuspid aortic valve and thoracic aortic aneurysm. Nat Genet. 2019;51:42–50 [PubMed: 30455415]
8. Yang B, Zhou W, Jiao J, Nielsen JB, Mathis MR, Heydarpour M, Lettre G, Folkersen L, Prakash S, Schurmann C, et al. Protein-altering and regulatory genetic variants near gata4 implicated in bicuspid aortic valve. Nat Commun. 2017;8:15481 [PubMed: 28541271]
9. Shi LM, Tao JW, Qiu XB, Wang J, Yuan F, Xu L, Liu H, Li RG, Xu YJ, Wang Q, et al. Gata5 loss-of-function mutations associated with congenital bicuspid aortic valve. Int J Mol Med. 2014;33:1219–1226 [PubMed: 24638895]
10. Gillis E, Kumar AA, Luyckx I, Preuss C, Cannaerts E, van de Beek G, Wieschendorf B, Alaerts M, Bolar N, Vandeweyer G, et al. Candidate gene resequencing in a large bicuspid aortic valve-associated thoracic aortic aneurysm cohort: Smad6 as an important contributor. Front Physiol. 2017;8:400 [PubMed: 28659821]
11. Theis JL, Niaz T, Sundsbak RS, Fogarty ZC, Bamlet WR, Hagler DJ, Olson TM. Celsr1 risk alleles in familial bicuspid aortic valve and hypoplastic left heart syndrome. Circ Genom Precis Med. 2022:CIRCGEN121003523
12. McCulley DJ, Black BL. Transcription factor pathways and congenital heart disease. Curr Top Dev Biol. 2012;100:253–277 [PubMed: 22449847]
13. Balistreri CR, Cavarretta E, Sciarretta S, Frati G. Light on the molecular and cellular mechanisms of bicuspid aortic valve to unveil phenotypic heterogeneity. J Mol Cell Cardiol. 2019;133:113–114 [PubMed: 31199951]
14. Wirrig EE, Yutzey KE. Conserved transcriptional regulatory mechanisms in aortic valve development and disease. Arterioscler Thromb Vasc Biol. 2014;34:737–741 [PubMed: 24665126]
15. Kostina AS, Uspensky V, Irtyuga OB, Ignatieva EV, Freylikhman O, Gavriliuk ND, Moiseeva OM, Zhuk S, Tomilin A, Kostareva A, et al. Notch-dependent emt is attenuated in patients with aortic aneurysm and bicuspid aortic valve. Biochim Biophys Acta. 2016;1862:733–740 [PubMed: 26876948]
16. Thomas PS, Sridurongrit S, Ruiz-Lozano P, Kaartinen V. Deficient signaling via alk2 (acvr1) leads to bicuspid aortic valve development. PLoS One. 2012;7:e35539 [PubMed: 22536403]
17. Welch-Reardon KM, Wu N, Hughes CC. A role for partial endothelial-mesenchymal transitions in angiogenesis? Arterioscler Thromb Vasc Biol. 2015;35:303–308 [PubMed: 25425619]
18. LaHaye S, Majumdar U, Yasuhara J, Koenig SN, Matos-Nieves A, Kumar R, Garg V. Developmental origins for semilunar valve stenosis identified in mice harboring congenital heart disease-associated gata4 mutation. Dis Model Mech. 2019;12
19. Timmerman LA, Grego-Bessa J, Raya A, Bertrán E, Pérez-Pomares JM, Díez J, Aranda S, Palomo S, McCormick F, Izpisúa-Belmonte JC, et al. Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. Genes Dev. 2004;18:99–115 [PubMed: 14701881]
20. Hou G, Harley ITW, Lu X, Zhou T, Xu N, Yao C, Qin Y, Ouyang Y, Ma J, Zhu X, et al. Sle non-coding genetic risk variant determines the epigenetic dysfunction of an immune cell specific enhancer that controls disease-critical microrna expression. Nat Commun. 2021;12:135 [PubMed: 33420081]
21. Soldner F, Stelzer Y, Shivalila CS, Abraham BJ, Latourelle JC, Barrasa MI, Goldmann J, Myers RH, Young RA, Jaenisch R. Parkinson-associated risk variant in distal enhancer of alpha-synuclein modulates target gene expression. Nature. 2016;533:95–99 [PubMed: 27096366]
22. Gupta RM, Hadaya J, Trehan A, Zekavat SM, Roselli C, Klarin D, Emdin CA, Hilvering CRE, Bianchi V, Mueller C, et al. A genetic variant associated with five vascular diseases is a distal regulator of endothelin-1 gene expression. Cell. 2017;170:522–533.e515 [PubMed: 28753427]
23. Patsch C, Challet-Meylan L, Thoma EC, Urich E, Heckel T, O'Sullivan JF, Grainger SJ, Kapp FG, Sun L, Christensen K, et al. Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. Nat Cell Biol. 2015;17:994–1003 [PubMed: 26214132]

24. Martin PS, Kloesel B, Norris RA, Lindsay M, Milan D, Body SC. Embryonic development of the bicuspid aortic valve. J Cardiovasc Dev Dis. 2015;2:248–272 [PubMed: 28529942]
25. Rivera-Feliciano J, Lee KH, Kong SW, Rajagopal S, Ma Q, Springer Z, Izumo S, Tabin CJ, Pu WT. Development of heart valves requires gata4 expression in endothelial-derived cells. Development. 2006;133:3607–3618 [PubMed: 16914500]
26. Evrard SM, Lecce L, Michelis KC, Nomura-Kitabayashi A, Pandey G, Purushothaman KR, d'Escamard V, Li JR, Hadri L, Fujitani K, et al. Endothelial to mesenchymal transition is common in atherosclerotic lesions and is associated with plaque instability. Nat Commun. 2016;7:11853 [PubMed: 27340017]
27. Aries A, Paradis P, Lefebvre C, Schwartz RJ, Nemer M. Essential role of gata-4 in cell survival and drug-induced cardiotoxicity. Proc Natl Acad Sci U S A. 2004;101:6975–6980 [PubMed: 15100413]
28. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014;32:381–386 [PubMed: 24658644]
29. Zhang J, Thorikay M, van der Zon G, van Dinther M, Ten Dijke P. Studying tgf-β signaling and tgf-β-induced epithelial-to-mesenchymal transition in breast cancer and normal cells. J Vis Exp. 2020
30. Flevaris P, Khan SS, Eren M, Schuldt AJT, Shah SJ, Lee DC, Gupta S, Shapiro AD, Burridge PW, Ghosh AK, et al. Plasminogen activator inhibitor type i controls cardiomyocyte transforming growth factor-β and cardiac fibrosis. Circulation. 2017;136:664–679 [PubMed: 28588076]
31. Mizrak D, Levitin HM, Delgado AC, Crotet V, Yuan J, Chaker Z, Silva-Vargas V, Sims PA, Doetsch F. Single-cell analysis of regional differences in adult v-svz neural stem cell lineages. Cell Rep. 2019;26:394–406 e395 [PubMed: 30625322]
32. Shah AV, Birdsey GM, Randi AM. Regulation of endothelial homeostasis, vascular development and angiogenesis by the transcription factor erg. Vascul Pharmacol. 2016;86:3–13 [PubMed: 27208692]
33. Yao Y, Yao J, Bostrom KI. Sox transcription factors in endothelial differentiation and endothelial-mesenchymal transitions. Front Cardiovasc Med. 2019;6:30 [PubMed: 30984768]
34. Zhao L, Zhao J, Wang X, Chen Z, Peng K, Lu X, Meng L, Liu G, Guan G, Wang F. Serum response factor induces endothelial-mesenchymal transition in glomerular endothelial cells to aggravate proteinuria in diabetic nephropathy. Physiol Genomics. 2016;48:711–718 [PubMed: 27565710]
35. Rodríguez MI, Peralta-Leal A, O'Valle F, Rodriguez-Vargas JM, Gonzalez-Flores A, Majuelos-Melguizo J, López L, Serrano S, de Herreros AG, Rodríguez-Manzaneque JC, et al. Parp-1 regulates metastatic melanoma through modulation of vimentin-induced malignant transformation. PLoS Genet. 2013;9:e1003531 [PubMed: 23785295]
36. Armstrong EJ, Bischoff J. Heart valve development: Endothelial cell signaling and differentiation. Circ Res. 2004;95:459–470 [PubMed: 15345668]
37. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol. 2014;15:178–196 [PubMed: 24556840]
38. Todorovic V, Finnegan E, Freyer L, Zilberberg L, Ota M, Rifkin DB. Long form of latent tgf-β binding protein 1 (ltbp1l) regulates cardiac valve development. Dev Dyn. 2011;240:176–187 [PubMed: 21181942]
39. Tenin G, Crozier A, Hentges KE, Keavney B. Glypican-6 deficiency causes dose-dependent conotruncal congenital heart malformations through abnormal remodelling of the endocardial cushions. bioRxiv. 2021:2021.2006.2028.450191
40. Kern CB, Twal WO, Mjaatvedt CH, Fairey SE, Toole BP, Iruela-Arispe ML, Argraves WS. Proteolytic cleavage of versican during cardiac cushion morphogenesis. Dev Dyn. 2006;235:2238–2247 [PubMed: 16691565]
41. Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning-based sequence model. Nat Methods. 2015;12:931–934 [PubMed: 26301843]

42. Yeo NC, Chavez A, Lance-Byrne A, Chan Y, Menn D, Milanova D, Kuo CC, Guo X, Sharma S, Tung A, et al. An enhanced crispr repressor for targeted mammalian gene regulation. Nat Methods. 2018;15:611–616 [PubMed: 30013045]
43. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, et al. A global reference for human genetic variation. Nature. 2015;526:68–74 [PubMed: 26432245]
44. Lee KW, Yeo SY, Sung CO, Kim SH. Twist1 is a key regulator of cancer-associated fibroblasts. Cancer Res. 2015;75:73–85 [PubMed: 25368021]
45. Nieto MA. The snail superfamily of zinc-finger transcription factors. Nat Rev Mol Cell Biol. 2002;3:155–166 [PubMed: 11994736]
46. Xu R, Won JY, Kim CH, Kim DE, Yim H. Roles of the phosphorylation of transcriptional factors in epithelial-mesenchymal transition. J Oncol. 2019;2019:5810465 [PubMed: 31275381]
47. Song Y, Yuan L, Wang Y, Chen M, Deng J, Lv Q, Sui T, Li Z, Lai L. Efficient dual sgrna-directed large gene deletion in rabbit with crispr/cas9 system. Cell Mol Life Sci. 2016;73:2959–2968 [PubMed: 26817461]
48. Mahmoud M, Souilhol C, Serbanovic-Canic J, Evans P. Gata4-twist1 signalling in disturbed flow-induced atherosclerosis. Cardiovasc Drugs Ther. 2019;33:231–237 [PubMed: 30809744]
49. Kumar MS, Owens GK. Combinatorial control of smooth muscle-specific gene expression. Arterioscler Thromb Vasc Biol. 2003;23:737–747 [PubMed: 12740224]
50. Creemers EE, Sutherland LB, McAnally J, Richardson JA, Olson EN. Myocardin is a direct transcriptional target of mef2, tead and foxo proteins during cardiovascular development. Development. 2006;133:4245–4256 [PubMed: 17021041]
51. Dupuis LE, Osinska H, Weinstein MB, Hinton RB, Kern CB. Insufficient versican cleavage and smad2 phosphorylation results in bicuspid aortic and pulmonary valves. J Mol Cell Cardiol. 2013;60:50–59 [PubMed: 23531444]
52. Uygur B, Wu WS. Slug promotes prostate cancer cell migration and invasion via cxcr4/cxcl12 axis. Mol Cancer. 2011;10:139 [PubMed: 22074556]
53. Tritschler I, Gramatzki D, Capper D, Mittelbronn M, Meyermann R, Saharinen J, Wick W, Keski-Oja J, Weller M. Modulation of tgf-beta activity by latent tgf-beta-binding protein 1 in human malignant glioma cells. Int J Cancer. 2009;125:530–540 [PubMed: 19431147]
54. Zhang BL, Bianco RW, Schoen FJ. Preclinical assessment of cardiac valve substitutes: Current status and considerations for engineered tissue heart valves. Front Cardiovasc Med. 2019;6:72 [PubMed: 31231661]
55. Zhou D, Feng H, Yang Y, Huang T, Qiu P, Zhang C, Olsen TR, Zhang J, Chen YE, Mizrak D, et al. Hipsc modeling of lineage-specific smooth muscle cell defects caused by tgfbr1(a230t) variant, and its therapeutic implications for loeys-dietz syndrome. Circulation. 2021;144:1145–1159 [PubMed: 34346740]
56. Carpentier G, Berndt S, Ferratge S, Rasband W, Cuendet M, Uzan G, Albanese P. Angiogenesis analyzer for imagej - a comparative morphometric analysis of “endothelial tube formation assay” and “fibrin bead assay”. Sci Rep. 2020;10:11568 [PubMed: 32665552]
57. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184:3573–3587 e3529 [PubMed: 34062119]
58. Janky R, Verfaillie A, Imrichova H, Van de Sande B, Standaert L, Christiaens V, Hulselmans G, Herten K, Naval Sanchez M, Potier D, et al. Iregulon: From a gene list to a gene regulatory network using large motif and track collections. PLoS Comput Biol. 2014;10:e1003731 [PubMed: 25058159]
59. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. Locuszoom: Regional visualization of genome-wide association scan results. Bioinformatics. 2010;26:2336–2337 [PubMed: 20634204]

HIGHLIGHTS

- Deletion of a large genomic segment encompassing BAV-associated variants reduced *GATA4* expression during hiPSC-based EndoMT, and impaired EndoMT progression.
- GATA4 is instrumental in sustaining MSC identity and output during hiPSC-based EndoMT.
- BAV Patient-derived cells carrying the rs117430032 variant have lower *GATA4* expression after EndoMT induction, and TWIST1 shows specificity to the locus tagged by rs117430032.

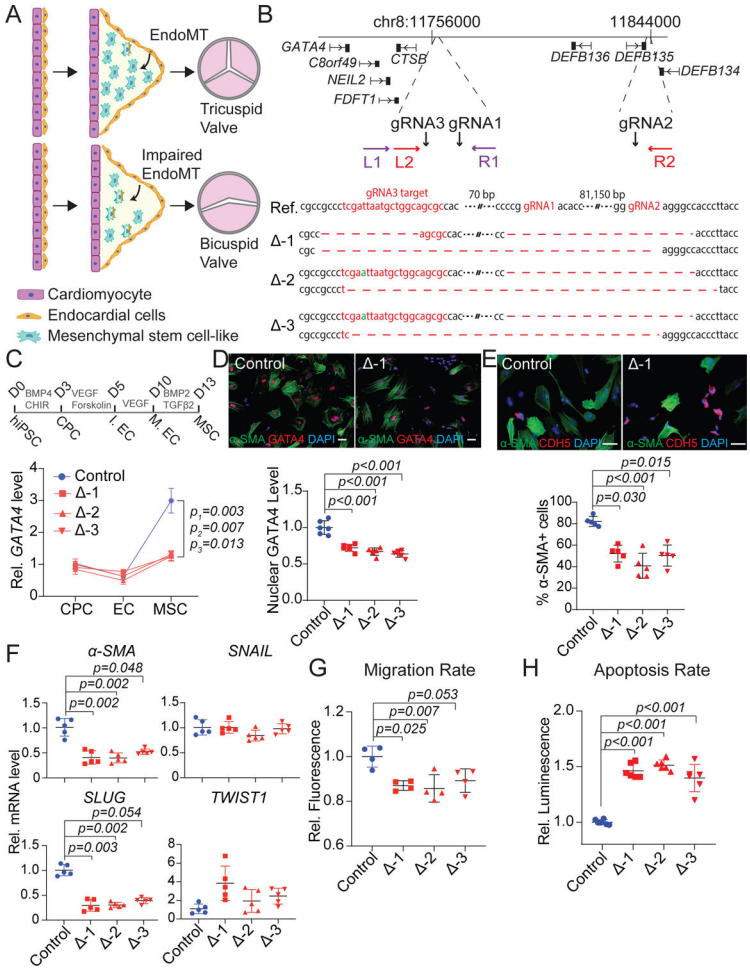


Figure 1. Deletion of a BAV-associated large genomic segment reduces GATA4 levels and impairs hiPSC-based EndoMT.

A) Diagram showing the role of EndoMT in endocardial cushion formation and human aortic valve development. B) Top: Strategy for the large segment deletion and the genes within or nearby the segment. gRNA1, gRNA2, gRNA3 indicate the gRNA target sites. L1, R1, L2 and R2 indicate PCR primer binding sites. Bottom: Sanger sequencing results for three biallelic deletion clones (Δ−1, Δ−2 and Δ−3) showing the large deletions on both alleles. Red and green letters indicate gRNA target sites and nucleotide insertions respectively, while the red dotted line indicate the deleted sequence. Ref.: Reference sequence. C) Top: Diagram of the differentiation process from hiPSC (human induced pluripotent stem cells) to MSC (Mesenchymal/Myofibroblast-like cells). D0: Day 0; D3: Day 3; D5: Day 5; D10: Day 10; D13: Day 13; CPC: Cardiovascular progenitor cells; I.EC: Immature endothelial cells; M.EC: Mature endothelial cells. Bottom: Relative *GATA4* levels in control and deletion clones at CPC, EC and MSC stages (N=5). The average expression level in control samples at the CPC stage were set to 1. D) Representative immunostaining images of α-SMA and GATA4 at the MSC stage, and the quantification of nuclear GATA4 staining intensity (N=6). Scale bars=50 μm. E) Representative immunostainings of α-SMA and CDH5 at the MSC stage, and the quantification of α-SMA+ cell percentage in each sample (N=5). Scale bars=100 μm. F) Relative mRNA levels of EndoMT markers (*α-SMA,*

SNAIL, SLUG, TWIST1) in different samples at the MSC stage (N=5). G) Cell migration rate in different samples at the MSC stage (N=4). The average fluorescence in the control was set to 1. H) Apoptosis rate in different samples at the MSC stage (N=6). The average luminescence in the control was set to 1. All data was presented as mean+/−standard deviation and the significance of the results was determined using either Kruskal-Wallis test (N≤5) or one-way ANOVA (N=6) with multiple testing correction by controlling the False Discovery Rate.

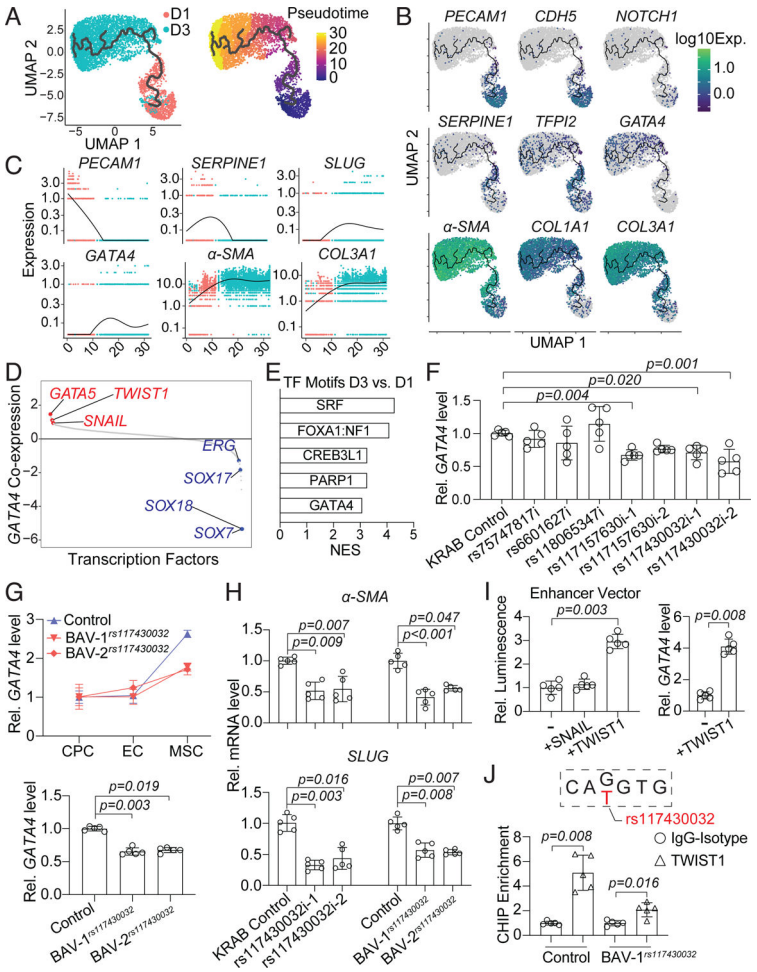


Figure 2. Stage-Specific Regulation and Function of GATA4 during hiPSC-based EndoMT.

A) Uniform Manifold Approximation and Projection (UMAP) visualization of Pseudotime trajectory of Day 1 (D1) and Day 3 (D3) cells during EndoMT colored by developmental stage. Left: D1 and D3 cells were colored in red and turquoise respectively. Right: The cells were colored by pseudotime. B) Expression of key endothelial (*PECAM1*, *CDH5* and *NOTCH1*), transitory (*SERPINE1* and *TFPI2*), canonical EndoMT markers (*α-SMA*, *COL1A1* and *COL3A1*) as well as *GATA4* on UMAP visualizations. C) Relative expression of genes across developmental pseudotime. D) Co-expression probabilities of different transcription factors (TF) with *GATA4*. E) iRegulon analysis revealing enriched TF motifs around transcription start sites of D3 MSC enriched genes. We highlighted an optimal TF for each motif. Motifs were ranked based on normalized enrichment score (NES). F) Relative *GATA4* expression in CRISPRi-repression lines (dcas9-KRAB-MeCP2+sgRNA) and their isogenic control (dcas9-KRAB-MeCP2) at the MSC stage (N=5). The average expression levels in the control were set to 1 and the statistically significant differences were highlighted. G) Top: Relative *GATA4* expression in control and BAV patient samples at CPC, EC and MSC stages (N=5). Bottom: Relative *GATA4* expression in BAV patient samples and rs117430032i cells at the MSC stage (N=5). H) Relative *α-SMA* and *SLUG* expression in BAV patient-derived and rs117430032i cells at the MSC stage (N=5). I) Left: Enhancer vector-based luminescence signal increases with TWIST1 overexpression

(N=5). Right: Consistently, TWIST1 overexpression increases *GATA4* expression (N=5). J) rs117430032 tags an E-box (enhancer box) motif. CHIP-qPCR results for Control and BAV patient-derived cells at the MSC stage (N=5). The average signal in their respective IgG1 isotype controls were set to 1. The quantifications in Figures 2F–2J were presented as mean+/−standard deviation. The significance of the results was determined using Kruskal-Wallis test (N≤5) with multiple testing correction by controlling the False Discovery Rate. When comparing only two conditions, the significance was determined using two-tailed Mann Whitney test.