


Genome-wide CRISPR screen identifies *ZIC2* as an essential gene that controls the cell fate of early mesodermal precursors to human heart progenitors

Jiejia Xu¹  | Chikai Zhou¹ | Kylie S. Foo² | Ran Yang² | Yao Xiao¹ |
Kristine Bylund² | Makoto Sahara^{1,2} | Kenneth R. Chien^{1,2}

¹Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm, Sweden

²Department of Medicine, Karolinska Institutet, Huddinge, Sweden

Correspondence

Kenneth R. Chien, MD, PhD, Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm SE-17177, Sweden.
Email: kenneth.chien@ki.se

Funding information

Forskningsrådet om Hälsa, Arbetsliv och Välfärd; Knut och Alice Wallenbergs Stiftelse; Swedish Research Council; European Research Council Advanced Research Grant Award, Grant/Award Number: AdG743225

[Funding information was added on 13 November 2020 after original online publication].

Abstract

Cardiac progenitor formation is one of the earliest committed steps of human cardiogenesis and requires the cooperation of multiple gene sets governed by developmental signaling cascades. To determine the key regulators for cardiac progenitor formation, we have developed a two-stage genome-wide CRISPR-knockout screen. We mimicked the progenitor formation process by differentiating human pluripotent stem cells (hPSCs) into cardiomyocytes, monitored by two distinct stage markers of early cardiac mesodermal formation and commitment to a multipotent heart progenitor cell fate: *MESP1* and *ISL1*, respectively. From the screen output, we compiled a list of 15 candidate genes. After validating seven of them, we identified *ZIC2* as an essential gene for cardiac progenitor formation. *ZIC2* is known as a master regulator of neurogenesis. hPSCs with *ZIC2* mutated still express pluripotency markers. However, their ability to differentiate into cardiomyocytes was greatly attenuated. RNA-Seq profiling of the *ZIC2*-mutant cells revealed that the mutants switched their cell fate alternatively to the noncardiac cell lineage. Further, single cell RNA-seq analysis showed the *ZIC2* mutants affected the apelin receptor-related signaling pathway during mesoderm formation. Our results provide a new link between *ZIC2* and human cardiogenesis and document the potential power of a genome-wide unbiased CRISPR-knockout screen to identify the key steps in human mesoderm precursor cell- and heart progenitor cell-fate determination during *in vitro* hPSC cardiogenesis.

KEYWORDS

CRISPR, differentiation, human embryonic stem cells, mesoderm

1 | INTRODUCTION

Cardiac progenitors emerge from the mesoderm precursor cells right after gastrulation. Other lineages, such as bone and muscle, are also generated from these mesoderm precursor cells. The regulation of the cell-fate commitment of these precursor cells holds the key to understand the development and differentiation to the specific cell lineage.

Cardiac progenitor formation is one of the earliest committed steps in human cardiogenesis and requires the cooperation of multiple gene sets governed by developmental signaling cascades. The cell-fate decisions are controlled by both the internal genetic factors and the external signaling cues from the environment. The *in vitro* model systems of human pluripotent stem cell (hPSC) differentiation enable us to control the external signaling cues, leaving the genetic factor as the

main component in cell-fate decision. Thus, a genetic screen approach on such a system may perhaps reveal the important genes that control the cell-fate decisions of both the mesodermal precursor cells and the cardiac progenitor cells during cardiogenesis.

CRISPR-knockout screen technology has been developed shortly after the establishment of CRISPR-Cas9 genome editing system in mammalian cells.^{1,2} Due to its feasibility and simplicity, the CRISPR-Cas9 system can be easily scaled up and integrated with an array synthesis system and/or mass parallel sequencing. CRISPR screen is a forward genetic approach, usually performed in pool screen manners. It is a versatile tool to discover genes associated with a specific phenotype in an unbiased fashion. CRISPR screens have been used in identifying genes for fitness of the cells,³ signaling transduction,⁴⁻⁷ survival, tumor metastasis,⁸ drug resistance,⁹ muscle formation,¹⁰ and cellular metabolism.^{11,12} Herein, we report a CRISPR-knockout screen to investigate the cell-fate decision and lineage specification in cardiogenesis. By combining a highly efficient cardiac differentiation protocol with the CRISPR screen, we have developed a two-stage genome-wide CRISPR-knockout screen system for identifying the key regulators in cardiac progenitor formation and commitment.

From the screen outputs, we compiled a list of 15 candidate genes. After validating seven genes of them, we identified *ZIC2* as a new essential gene for regulating cardiac progenitor formation. Importantly, our study provides a new link between *ZIC2* and human cardiogenesis and documents the potential power of the genome-wide unbiased CRISPR screen to identify the key steps in human mesoderm precursor cell- and heart progenitor cell-fate determination during human cardiogenesis, recapitulated in *in vitro* hPSC model systems.

2 | MATERIALS AND METHODS

2.1 | hESC culture and differentiation

Human pluripotent stem cells (ES03, NIH code: HES3) were maintained on Vitronectin (Thermo Fisher Scientific, Paisley, UK) coated plates in Essential 8 (Thermo Fisher Scientific) medium and passaged with Versene (Thermo Fisher Scientific). Both small molecule and growth factors based cardiac differentiation of hESCs were performed according to previously published growth factor based methods.¹³⁻¹⁶ Please see the supplemental method for details.

2.2 | Cell line generation

2.2.1 | Generation of Cas9 hPSC line

The lentiviral plasmid Cas9-NLS-FLAG-2A-Bsd was cotransfected into HEK293T cells with packaging plasmids psPAX2 and pMD2.G (Addgene, 12260 and 12259). HEK293T cells were cultured in high-glucose DME +10% FBS. Virus-containing media were collected at 48 and 72 hours after transfection and used for hPSC infection. Transduced cells were selected based on resistance to 1 µg/mL puromycin.

Significance statement

In order to unbiasedly uncover the regulators that control the formation of human cardiac progenitor, a genome-wide CRISPR-knockout screen based on cardiac differentiation from human embryonic stem cells (hESCs) was developed. From the screen output, *ZIC2* as an essential gene for cardiac progenitor formation was identified. The results of the study provide a new link between *ZIC2* and human cardiogenesis and document the potential power of genome-wide unbiased CRISPR-knockout screens to identify key steps in heart progenitor fate determination during human cardiogenesis with hESC model systems.

2.2.2 | Generation of *ZIC2* mutated hPSC lines

Three plasmids, pPB-CRISPR-*ZIC2*-g1/2 and pCyl43, were used for genome editing in hPSCs. Two million hPSCs, 1 µg pCyl43 and 4 µg pPB-CRISPR-*ZIC2*-g1/2 DNA were mixed in 100 µL nucleofection solution and then nucleofected with program B-16 using a Nucleofector 2b device (Lonza, Zurich, Switzerland). After 2 weeks of 0.5 µg/mL puromycin selection, cells were singularized and sparsely seeded to generate single cell derived clones.

Other mutants, *PLEKH02*, *CTSV*, *TTC39B*, *SEMA3B*, *MAGIX* were generated in the same way. *SMAD4* clones were reused from previous study and were generated in the same way.

2.3 | Generation of hPSC mutant libraries

The virus packaging and transfection was carried out in the virus lab. The library and two other lentivirus packaging plasmid psPAX2, and pMD2.G were transfected into 293TN cells with FuGENE HD (Promega, Madison, Wisconsin). The supernatants were collected at 48 and 72 hours after transfection and used to infection the Cas9 expressing ES03 cells. A 5 µg/mL Blasticidin selection was applied for 2 weeks. And the libraries were cultured for another 2 weeks before experiments.

2.4 | Mutation sequencing

Genomic DNA was extracted from putative *ZIC2* mutated hPSCs using Quick-gRNA MiniPrep kit (Zymo Research, Irvine, California, D3007). The targeted genome region was amplified primer pairs, forward: GGCGCAGAACGGCTTCGTTG and reverse: TGCTCAGTTGCTCGG GGTCCG. The amplified polymerase chain reaction (PCR) fragment was inserted into pCR4TOPO vector via the TOPO-TA method (Thermo Fisher Scientific). Plasmids from individual *Escherichia coli* clones were Sanger Sequencing using a T7 primer.

2.5 | Genomic sgRNA library preparation

Genomic DNA was extracted with Zymo-DNA Miniprep Kit (D3024) or Zymo FFPE DNA MiniPrep Kit (D3065) according to the manuals. Genomic DNA from the hESC libraries were extracted with Zymo-DNA Miniprep Kit. And the sorted cell population were fixed with 4% formaldehyde solution, so proteinase K digestion step were included, and Zymo FFPE DNA MiniPrep Kit were used to extract the genomic DNA in these cells.

A two-step PCR were used to prepare the sequencing library. In the first PCR reaction, the sgRNA containing region were amplify from the genomic DAN with lib-seq primers: lib-seq-fwr: ACACTCTTCCC TACACGACGCTCTCCGATCTgaggaTTGTGGAAAGACGAAACACCG and lib-seq-rvs: AGACGTGTGCTCTCCGATCTtactattctttcccctgcactgt.

In the second PCR reaction, sequencing adaptor and multiplex index were added by illumina's Nextra adaptor primers. The final PCR product were size-selected with E-Gel system and purified. The final libraries were qualified with bio-analyzer. And the finished libraries were send to SciLifeLab for sequencing.

2.6 | CRISPR screen data processing

Fastq files were fed into a perl script to extract the sgRNAs. The extracted sgRNA sequences were aligned to the human genome sequence with bowtie. Good alignments were annotated with a GTF file using bedtools. In the final step, a table was generated showing the sgRNA sequencing, genome coordinates, targeted gene information, and the count numbers. The count table is fed into edgeR to compute the "differential" genes between the different groups as the candidate genes.

2.7 | Antibodies

POU5F1 (sc-5279, Santa Cruz, Heidelberg, Germany); NANOG (4903S, Cell Signaling, Leiden, Netherlands); SOX2 (3579, Cell Signaling); PAX6 (DSHB, Iowa, Iowa); TBXT (AF2085, R&D system, Minneapolis, Minnesota); MESP1 (ab77013, Abcam, Cambridge, UK); ISL1 (39.4D5, DSHB); NKX2-5 (sc-14 033, Santa Cruz); TNNT2 (MS-295-P1, ThermoFisher Scientific); V5 (R960-25, ThermoFisher Scientific); ZIC2 (ab150404, Abcam).

2.8 | Flow cytometry

Cells were dissociated into single cells with Accutase (STEMCELL Technologies, Cambridge, UK) for 10 minutes and subsequently fixed with 1% paraformaldehyde for 20 minutes at room temperature and stained with primary and secondary antibodies in PBS with 0.2% Triton X-100 and 0.5% BSA. Data were collected on a FACSCaliber or FASCanto flow cytometer (Becton Dickinson, Franklin Lakes, New Jersey) and analyzed using FlowJo. Fluorescence-activated cell sorting (FACS) gating was based on the corresponding isotype or secondary only antibody control.

2.9 | Immunostaining

Cells were fixed with 4% paraformaldehyde for 15 minutes at room temperature and stained with primary and secondary antibodies in PBS plus 0.4% Triton X-100 and 5% nonfat dry milk. Nuclei were stained with DAPI. A ZEISS LSM 700 confocal microscope was used for imaging.

2.10 | RNA sequencing

Total RNA was isolated with Direct-Zol RNA Kits (Zymo Research). RNA quality was checked with BioAnalyzer 2100 (Agilent, Santa Clara, California). Library preparation type was Illumina TruSeq Stranded mRNA, Poly-A selection. Samples were sequenced on HiSeq2500 with a 2 × 126 setup using HiSeq SBS Kit v4 chemistry. Basic data processing was performed in the SciLifeLab facility. Briefly, the FASTQ files were mapped to Human Genome, GRCh37 with STAR aligner.¹⁷ The outputted binary version of a Sequence Alignment/Map (BAM) files were used for further analysis described later.

2.11 | RNA-Seq analysis

The expression count table was produced by feeding FeatureCounts¹⁸ with the BAM files produced in the data process step and annotation file gencode.v27lift37 from GENCODE. Differential gene expression analysis was performed with R package edgeR.¹⁹ Gene expression pattern analysis was performed with k-mean function in R. Analysis and graphs were produce with custom scripts available by request.

2.12 | Single cell RNA-seq

The cryo-preserved day 3 samples were thawed at 37°C and stained with Sytox Blue (Thermo Fisher Scientific). The Sytox blue negative population contained the live cells and were FACS sorted into 384-well plate, which contains the lysis buffer for single cell RNA-seq library preparation provided by the single cell RNA-seq facility in Integrated Cardio Metabolic Centre (ICMC), Department of Medicine, Huddinge, Karolinska Institutet. Strict gate had been applied to minimize the chance of more than one cell in each well. Three plates were sequenced, including one vial of wild type ES03, ZIC2 mutant clone 1 and clone 2. Single cell RNA libraries were prepared using the Smart-seq²⁰ protocol and sequenced with HiSeq2000 using 50 bp single read mode. The outputted FASTQ files were mapped to Human Genome, GRCh37 with HISAT2 aligner.²¹ The outputted BAM files were used for further analysis described later.

2.13 | Single cell RNA-seq analysis

The expression count table was produced by feeding FeatureCounts¹⁸ with the BAM files produced in the data process step and annotation file gencode.v27lift37 from GENCODE. The count table and

annotation information are integrated into SingleCellExperiment,²² an R object for single cell RNA-seq analysis. Cells with less than 7e3 genes expressed or total read count below 1.2e5 is filtered out. And genes that expressed in less than five cells are also excluded from the analysis. The number of cells passed final quality control (QC) is 1043, and the number of genes include in the analysis is 26 033. After the QC, we used R-package: scran,²³ to calculate the size factor for each cell according to the internal housekeeping genes and normalized the read counts for further analysis. T-SNE map is generated using t-SNE algorithm integrated in the SingleCellExperiment package. Cell clustering and marker gene identification are analyzed by sc3,²⁴ another R package for single cell RNA-seq analysis. The gene expression plot on the t-SNE map is plotted with customized R-script with ggplot2 package.

2.14 | Animal experiments

The animals' care was in accordance with the institutional guidelines of the Karolinska Institutet, and all animal experiments were approved by the local ethics committee (Stockholm, Sweden) in accordance with the Animal Protection Law, the Animal Protection Regulation, and the Regulation of the Swedish National Board for Laboratory Animals. Please refer the supplemental methods for the generation of mutant embryos with microinjection of zygotes.

3 | RESULTS

3.1 | Generation of hESC mutation libraries

Generation of a high-quality cell library is needed as the foundation of a genome-wide screen. We first engineered human embryonic stem cells (hESCs) constitutively expressing Cas9 using an ES03 cell line and a plasmid pLentiCas9Bsd.²⁵ After three passages of selection, as shown in Figure S1A, the cells expressed Cas9 uniformly, indicated by the expression of flag tags that attach at the C-terminal of Cas9. We further tested the differentiation efficiency of this cell line, and highly efficient differentiation could be eventually achieved (Figure S1B), making it suitable for the screen. Next, we transfected the cells with the sgRNA lentiviral library (see Methods). The transfection was carried out with an optimized low viral titer to avoid too many copies of insertions in one single cell. To increase the library coverage, three independent transfections were performed, and thus, three hESC libraries were generated (Figure S2B). These hESC libraries were separately maintained over 4 weeks to achieve the maximum on-target effect but inversely only the limited off-target effect.

Immediate after the transfection, some gene mutations may cause the cell death or slow growth, resulting in a dropout from the screen of CRISPR-knockout genes that potentially affect cardiac differentiation. After 4 weeks culture when the transduced hESC libraries should be stably settled down, we sequenced the genomic DNAs of the three libraries, respectively, and checked the sgRNAs they were carrying. In total, 47 million reads were obtained. We filtered out sgRNAs with total reads below 500, and found that the transduced sgRNAs finally

targeted around 6000 genes. Eighty percent of those sgRNAs were overlapped among the libraries, showing the reproducibility of generating the hESC mutant libraries (Figure S2C).

Since these libraries can self-renew and be differentiated into a wide variety of specific cell types, their applications are considered to be limitless. As one example, we performed a small screen to uncover genes responsible for the cell death induced by dissociation of hESCs. The cells carrying mutations in these genes might have a growth advantage over the other cell types. Identifying these genes likely helped to better interpret the screen outcome. The survival rate of hESCs after cell dissociation is usually around 1%, without using p160-Rho-associated coiled-coil kinase (ROCK) inhibitors. Even with a ROCK inhibitor, the survival rate increases but is only 27%.²⁶ If cells from the library carrying gene mutation that can disrupt the ROCK signaling pathway, they were likely to be naturally resistant to this sort of mechanical and chemical stress to cells and to have growth advantages rather than normal or other mutant cells. To prove this hypothesis, we pooled a small library from the three main libraries and passaged them without using a ROCK inhibitor. After two passages, only a few colonies survived. They could not only proliferate continuously, but also survive after the freeze-and-thawed stress without addition of a ROCK inhibitor. We extracted the genomic DNAs from these cells and amplified the sgRNAs integrated in their genome. With a simple TA-cloning method, we could separate the sgRNAs into individual plasmids, and 20 of them were sequenced. We found a list of genes whose mutations were responsible for the growth advantage of these cells. They involved *PMAIP1* (22%), *RYBP* (17% + 5%), *MYH9* (17%), *PCGF1* (5%), *ZNF418* (5%), *STOML1* (5%), *GLYATL2* (5%), and *B3GALT6* (5%) (Figure S2D). Taken together, we successfully created the three hESC mutant libraries for the CRISPR screen and managed to test the screen principle by identifying the genes associated with the potential cell growth advantage.

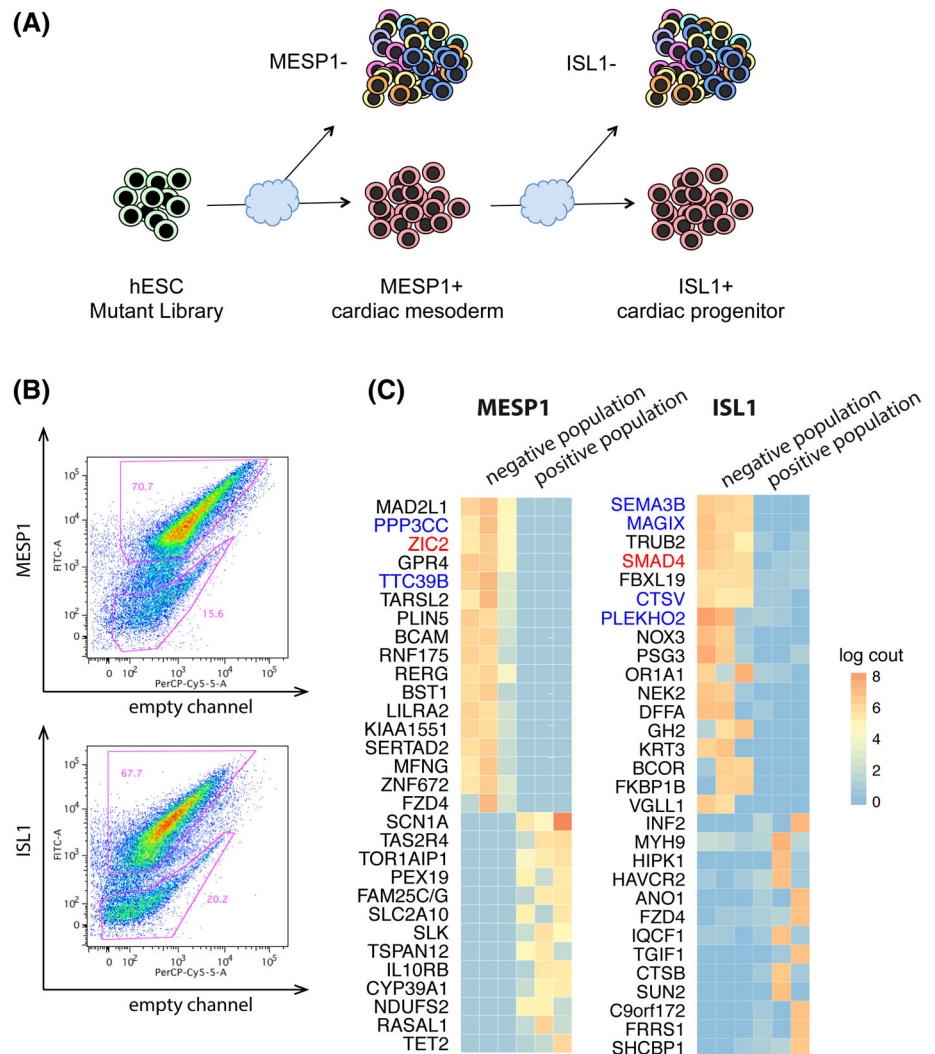
3.2 | Screen for essential genes for the formation of cardiac mesoderm and progenitors

We pooled the three hESC mutant libraries together to create one large genome-wide mutated library, covering more than 6000 genes. We differentiated the library and carried out the screening in the cardiac mesoderm and progenitor stages.

To screen the essential genes for the formation of cardiac mesodermal precursor cells, we used *MESP1* as a cardiac mesoderm marker. The cardiac mesoderm stage was corresponding to day 3 in cardiac differentiation in vitro, when more than 95% of the wild type hESC-derived cells were *MESP1* positive. We used a *MESP1* antibody to separate the mutant hESC-derived cells on day 3 into *MESP1*⁺ and *MESP1*⁻ populations via FACS. Theoretically, the cells with mutation of an essential gene for cardiac mesoderm formation failed to express *MESP1* and were enriched in the *MESP1*⁻ population. Therefore, by comparing integrated sgRNAs between *MESP1*⁺ and *MESP1*⁻ populations sorted on day 3, the sgRNA-targeted essential genes for cardiac mesoderm formation could be identified as genes enriched in the *MESP1*⁻ population but not found in the *MESP1*⁺ population (Figure 1A).

FIGURE 1 Cardiac mesoderm and progenitor stage screening. A, Schematic shows the screening strategy.

B, Examples of flow cytometry plots show the sorting of MESP1 or ISL1 positive and negative populations. In total, eight batches of mesoderm differentiation (collected on day 3), and 10 batches of progenitor differentiation (collected on day 6) were performed. The best three differentiation attempts were chosen for further sequencing analyses. C, Heat maps show the genes with sgRNAs enriched in marker negative and positive populations



Next, to screen the essential genes for the formation of cardiac progenitors, we used ISL1 as a cardiac progenitor marker. At the progenitor stage such as day 6 in cardiac differentiation, more than 90% of wild type hESC-derived cells expressed ISL1. The mutant hESC-derived cells were separated into ISL1⁺ and ISL1⁻ populations. Similar to the screening at cardiac mesoderm stage, cells with mutation of an essential gene for cardiac progenitor induction failed to express ISL1 and were enriched in the ISL1⁻ population on day 6 (Figure 1A).

To minimize the enrichment of negative population due to a low efficiency of cardiac differentiation, we differentiated the libraries multiple times, eight for cardiac mesoderm formation and 10 for cardiac progenitor induction, and the three attempts with the most efficient cardiac differentiation were chosen for further analyses, in which the highest percentages of MESP1 or ISL1 positive cells were obtained (Figure 1B). In total, six populations were collected, and genomic DNAs of those cells were extracted. The sgRNA regions integrated into the genome were amplified with a nest-PCR method. Adapters for massive parallel sequencing were added during the second PCR step (see Methods). HiSeq2500 (Illumina, San Diego, California) was used to sequence these sgRNA amplicons, which were extracted and aligned to the human genome. Then, sgRNA count tables were generated and fed into the

edgeR program to compare and detect the differential expression genes between the negative and positive populations for Mesp1 and Isl1.

3.3 | Validation of candidate genes essential for the formation of cardiac mesoderm and progenitors

As shown in Figure 1C, the well-known cardiac genes, such as *TBXT*, *MESP1*, *ISL1*, and *NKX2-5*, did not show on top of the list, suggesting these genes might not be as essential for the formation of cardiac mesoderm or progenitors as previously thought. Among the top candidates of essential genes for cardiac mesoderm or cardiac progenitor formation (Figure 1C), *SMAD4* is the most well-known gene, a key component in TGF β superfamily signaling pathways.²⁷ Its effect on cardiac differentiation was reported and thus predictable. Besides *SMAD4*, we compiled a list of 15 candidate genes. From the Mesp1⁺ cardiac mesoderm screening, *MAD2L1*, *PPP3CC*, *ZIC2*, *PODXL*, *NAPA*, *GPR4*, and *TTC39B* were identified as the top candidates. From the ISL1⁺ cardiac progenitor screening, *SEMA3B*, *MAGIX*, *TRUB2*, *SMAD4*, *FBXL19*, *CTSV*, *PLEKHO2*, *NOX3*, and *DFFA* were identified as the top candidates (Figure 1C).

3.4 | ZIC2 mutants reduce cardiac lineage commitment from early mesodermal precursors

Next, we investigated the mechanistic aspects underlying the failure of cardiomyocyte differentiation in the ZIC2 mutants. We first examined the appearances of ZIC2 mutants in each of the hESC differentiation stages. Although bright field images showed the colonies from ZIC2 mutants were not as compact as the wild type (Figure S3A), immunostaining images showed POU5F1 and NANOG were expressed in ZIC2 mutants as well as in the wild type (Figure S3B). This suggested that the ZIC2 mutants hold their pluripotency state, according to these markers' expression.

When the cardiac differentiation was initiated, the wild type hESCs would go through similar stages as embryonic development: on day 1, after 24 hours of stimulation by a GSK3 β inhibitor, hESCs differentiated into TBXT (also known as Brachyury T)-positive cells, resembling the *in vivo* cells locating at the primitive streak; on day 3, the cells

transformed into MESP1-positive cells, resembling the *in vivo* cardiac mesodermal precursor cells migrating through the primitive streak toward the anterior pole; and on day 6, ISL1-positive cardiac progenitors emerged from the cultured cells, indicating the formation of heart field progenitors. On the contrary, the ZIC2 mutants failed to express MESP1 at protein level on day 3 (Figure 3A), suggesting the failure of cardiac mesoderm formation. Interestingly, on day 6, the ZIC2 mutant could also express ISL1, a classical second heart field marker²⁸⁻³² (Figure 3B). On day 8, the progenitors from the wild type cells differentiated into cardiomyocytes expressing NKX2-5 and TNNT2 (Figure 3C). However, only few of the ZIC2 mutants were able to express cardiomyocyte markers on day 8. In fact, most of the ZIC2 mutants failed to express both NKX2-5 and TNNT2 (Figure 3C). These results suggested two things: first, ZIC2 is essential for the formation of MESP1⁺ cardiac mesodermal precursors that can commit to the cardiomyocyte lineage; and second, ISL1 expression is not fully controlled by the MESP1 signaling pathway and can be turned on irrespective of

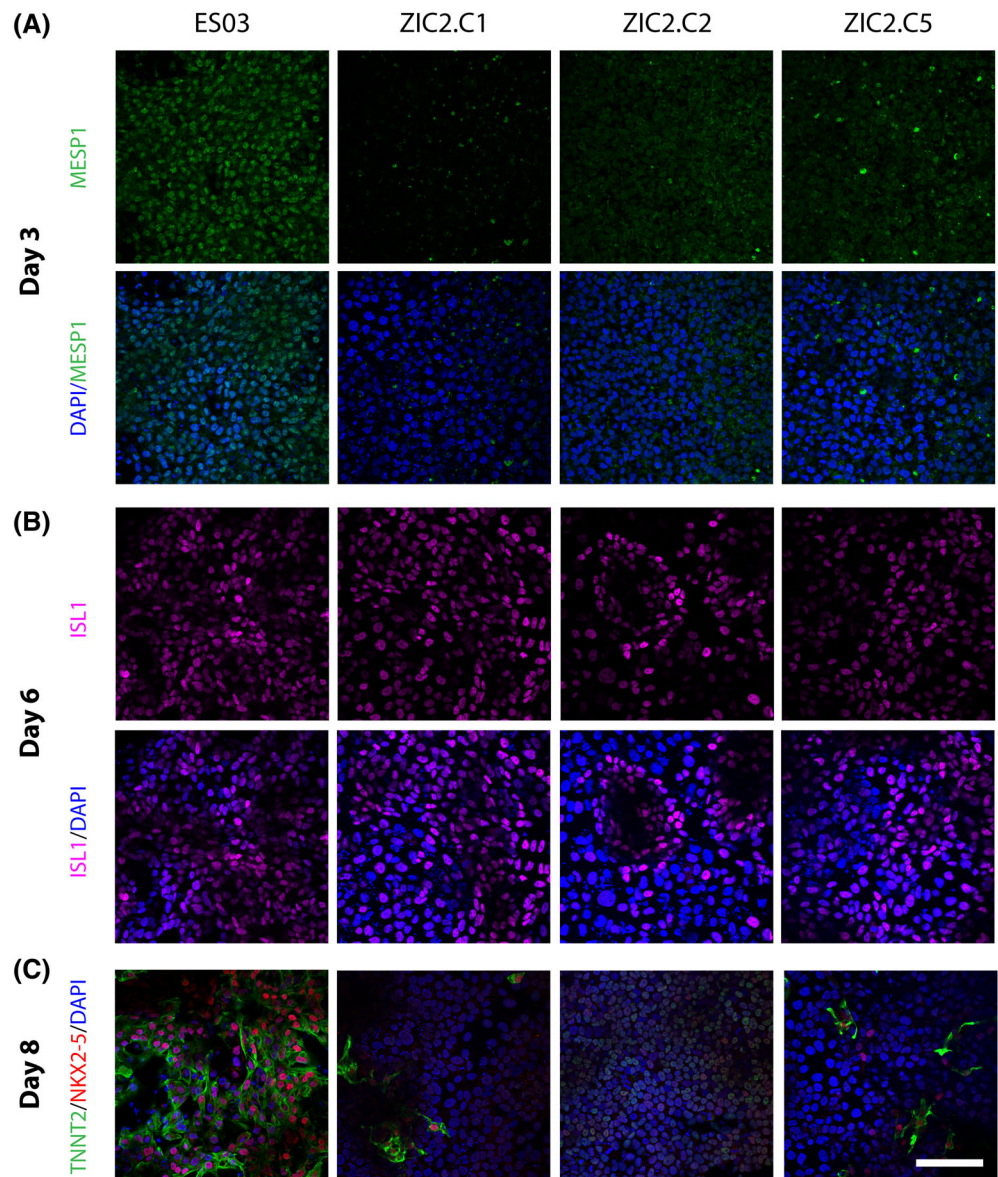


FIGURE 3 ZIC2 mutant reduce cardiac lineage commitment. A, ZIC2 mutants fail to express cardiac mesoderm marker MESP1 on day 3. B, ZIC2 mutants are able to express cardiac progenitor marker ISL1 on day 6. C, Few of the ZIC2 mutated progenitors can differentiate into cardiomyocyte fate. Scale bar = 100 μ m

precedent expression of *MESP1*, although the myogenic *ISL1*⁺ progenitors appear to be derived from the *MEPS1*⁺ lineage.

3.5 | *ZIC2* is essential for the formation of cardiac progenitors

To further explore the mechanism of *ZIC2*-mediated cardiac differentiation, we analyzed bulk RNA sequencing (RNA-seq) data to compare the transcriptome differences between wild type and the *ZIC2*-mutant hESC-derived cells at different stages in cardiac differentiation: day 0, day 1, day 2, day 3, and day 6. These time points correspond to distinct development stages established in previous study,²⁸ namely epiblast (day 0), primitive streak (day 1), early mesoderm (day 2), cardiac mesoderm (day 3), and cardiac progenitor (day 6) stages, respectively. Using multiple-dimensional scaling (MDS), we illustrated the relationship of these samples. They were surprisingly close to each other at the early stages. Both the mutant and the wild-type cells followed a similar trajectory (Figure 4A). However, the number of differentially expressed genes increased following the differentiation progress (Figure 4B), with day 1, 25 up, 25 down; day 2, 27 up, 39 down; day 3, 53 up, 69 down; and day 6, 126 up, 167 down (wild type vs *ZIC2* mutants). As shown in Figure 4C, by using classic cardiac differentiation markers, we found the *ZIC2* mutant cells were able to express early mesoderm precursor markers, such as *TBXT*, *EOMES*, and *MIXL1*. Conversely, the transcriptional similarities shown between the wild type and mutant cells seceded from the cardiac progenitor stage onward. Indeed, the *ZIC2* mutants failed to commit to the cardiomyocyte lineage, as indicated by the expression patterns of the classical cardiomyocyte markers such as *NKX2-5*, *MEF2C*, *TBX5*, *MYH6*, *TNNT2*, and *MYL7*, all of which were not induced in the *ZIC2* mutants (Figure 4C).

The similar transcriptome profiles between the wild type and *ZIC2* mutant cells at the early stage suggest the *ZIC2* mutants have adopted the early mesodermal formation, while the differences at later stage indicate the *ZIC2* mutants fails to become cardiac progenitors. These results suggest *ZIC2* may play an essential role in the cell-fate decision of the mesodermal precursor cells, which are considered to be indispensable for the following induction of cardiac progenitors.

3.6 | *ZIC2* mutants have switched the cell fate of early mesodermal precursors into the osteoblast lineage

To further decipher the molecular basis of causing the cell-fate switch of the mesodermal precursor cells, we focused on the three major developmental signaling pathways: TGF β superfamily, wingless/Int-1 (WNT), and fibroblast growth factor (FGFs) (Figure 4D). The major differences were highlighted in the FGF signaling pathways. The *ZIC2* mutant cells showed higher expression of *FGF8*, *FGF18*, and *FGF19* on day 1 than wild-type cells. Both *FGF8* and *FGF18* were reported as positive regulators for osteogenesis.³³ On the contrary, *FGF4* and

FGF13 had lower expression in the mutants. On day 6, *FGF10*, which is regulated by *ISL1*, was highly expressed in the wild-type cells, but not in the mutant cells. These observations suggest the incorrect FGF paracrine cues may be one of the reasons for the derailed differentiation program of the *ZIC2* mutants.

Furthermore, among the differentially expressed genes on day 6, several osteogenesis-related genes such as *RUNX2*, *DCN*, and *OGN* were upregulated in the *ZIC2* mutants. In contrast, the cardiogenic genes such as *NKX2-5*, *MEF2C*, *PLN*, and *MYH6* were downregulated in those cells (Figure 4E). By analyzing the protein-protein interaction networks in the STRING database, the skeletal development network was formed by the upregulated genes in the mutants on day 6 (Figure 4F). These results indicated the mutants might convert their cell fate from cardiac to the osteoblast lineage.

Additionally, the *ZIC2* mutants had more upregulated genes than the downregulated genes when compared with the wild type (Figure 4B), and among the upregulated genes, both cardiogenesis and osteogenesis genes were found. *MSX2*, another pro-osteogenesis gene,³⁴ was upregulated in the mutants. Cardiogenesis genes, such as *ISL1*, *PITX2*, were prematurely expressed in the mutant at cardiac mesoderm stage, while these genes remained inactive or at a low expression level in the wild types. On day 6, the cardiogenesis program was successfully executed in the wild type, while it was shut down in the *ZIC2* mutants. Among the upregulated genes in the mutants on day 6, an anticardiac factor *CDX2*³⁵ was found, suggesting the *ZIC2* mutants might not just adopt the osteogenic fate, but also activate molecular programs that suppress the induction of the cardiac cell fate. These results indicate one of the *ZIC2* functions previously unknown is to induce cardiac commitment and specification properly during the in vitro human cardiogenesis.

3.7 | Single-cell RNA-seq analysis reveals subpopulations emerging differently between the wild type and *ZIC2* mutant cells during mesoderm formation

To further understand the role of *ZIC2* that would play during the formation of cardiac mesoderm, we analyzed single-cell RNA-seq data of the wild type and *ZIC2* mutant cells during cardiac mesoderm formation on day 3, and sought to identify the subpopulations emerging differently from both cell types. On day 3 in cardiac differentiation, 384 cells from each group, such as wild type and *ZIC2* mutants (clones 1 and 2) were harvested for single-cell RNA-seq analysis. Eventually, 362 cells from the wild type, 339 cells from the *ZIC2* clone 1, and 342 cells from the *ZIC2* clone 2 finally passed the quality control test and were then used for the analysis. With a t-SNE approach, the analyzed cells were segregated into three large groups, as shown in Figure 5A. There was a group on the top, specifically composed by the wild-type cells. On the contrary, there was a different group on the bottom, composed by the two *ZIC2* mutant clones. One another group, mainly composed by the wild-type cells with a small fraction of the mutant cells, was also found on the right side of the t-SNE panel (Figure 5A).

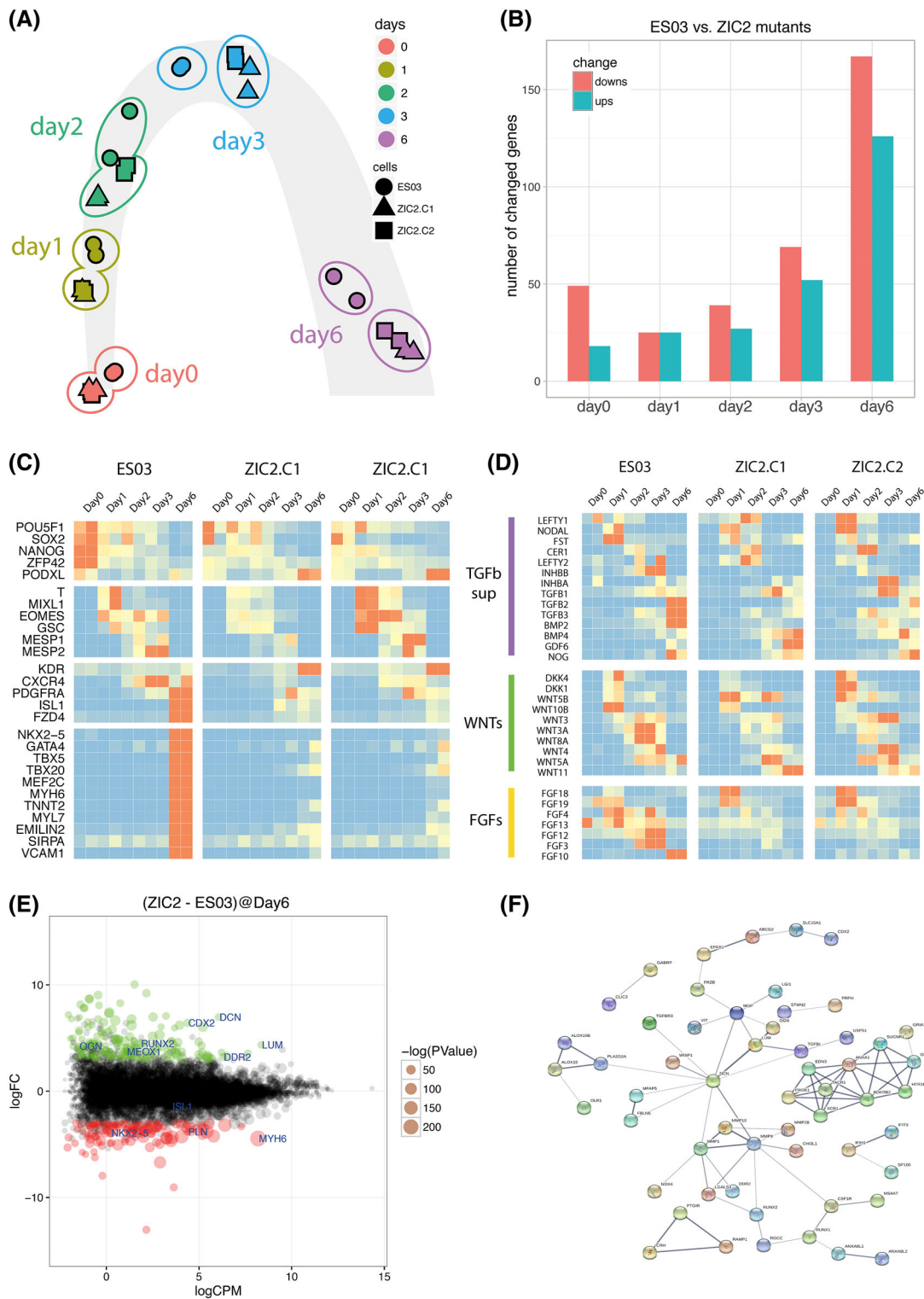


FIGURE 4 Profiling the differentiation of the ZIC2 mutants with RNA-Seq. A, MDS (multidimensional scaling) plot shows the transcription profile similarity between the wild type and ZIC2 mutants. B, The number of differentially expressed genes (DEGs) comparing the wild type and mutants at different stages. C, Heat map shows classic marker genes expressions. D, Heat map shows the dynamic expression of ligands from WNT, TGFβ, and FGF signaling pathways. E, MD (mean-differential) plot show the DEGs on day 6 between the mutant and the wild type. F, Protein-protein interaction network based on the DEGs on day 6 in the ZIC2 mutants suggests the mutants switch to osteoblast fate

We then examined expression patterns of the stage-specific markers in cardiac differentiation on the analyzed cells in Figure 5A. The pluripotency markers, *POU5F1*, *NANOG*, and *SOX2*,

were highly expressed in both the top and bottom groups, while expressed little or nothing (especially, *NANOG* and *SOX2*) in the right-most group (Figure 5B). The lower expression of *NANOG* and *SOX2* in

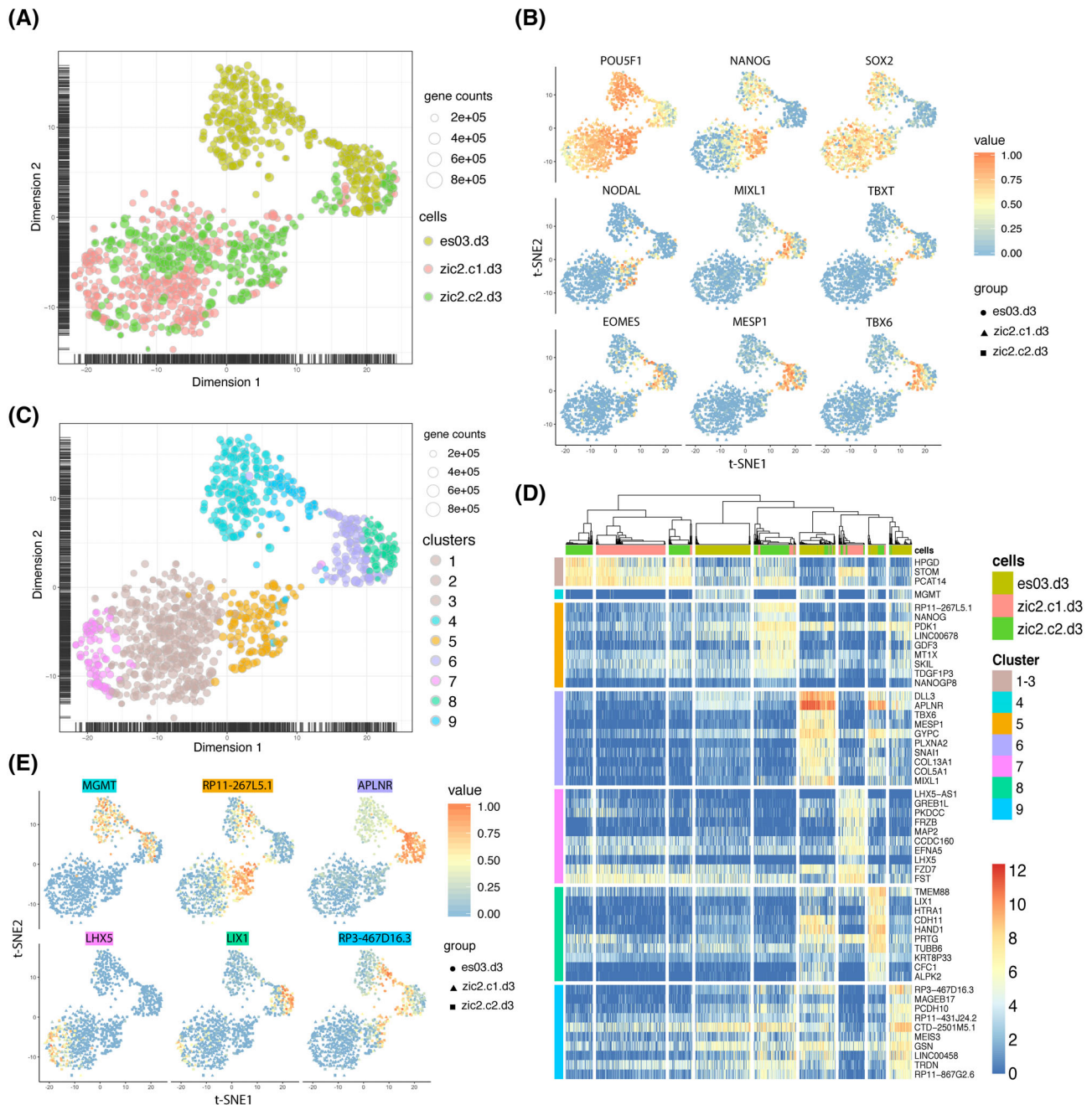


FIGURE 5 Single cell RNA-Seq reveal subpopulation difference between the wild type and ZIC2 mutants during mesoderm formation. A, tSNE representation of the transcription profiles of the day 3 cells from wild type and ZIC2 mutants hESCs. B, tSNE plots with the expression distribution of the classical developmental stage markers, including pluripotency markers: POU5F1, NANOG, SOX2; primitive markers: NODAL, MIXL1, T(TBXT); mesoderm stage markers: EOMES, MESP1, TBX6. C, Subpopulation clusters based on the transcription profile. There are in total nine clusters, but clusters 1-3 are very similar, so seven distinct populations can be identified. D, Heat map showing the marker genes of each cluster. Horizontal axis shows the composition of the cells from different groups. The vertical axis on the left shows the color code for each cluster, and the vertical axis on the right shows the names of marker genes. E, tSNE plots with the expression of example marker genes for different clusters

the rightmost group suggests this group is a more differentiated cell population than the top and bottom groups. The primitive streak-associated genes such as *NODAL*, *MIXL1*, and *TBXT*, which are normally highly expressed on day 1, were mainly expressed in the cells located in the middle among the three groups, indicating a developmental transition from the more immature top and bottom groups to the rightmost group. On the other hand, the cardiac mesoderm

formation markers such as *EOMES*, *MESP1*, and *TBX6*, which are normally highly expressed on day 3, were mainly expressed in the rightmost group, suggesting that this group is most likely composed of the cardiac mesoderm precursor cells, which the majority of the ZIC2 mutants failed to form (Figure 5B).

To proceed with a more detailed analysis, we adopted the k-means method to further cluster the analyzed cells according to their

gene expression patterns. After trying several parameters, we found that processing the data set with a parameter “ $k = 9$ ” yielded a reasonable result (Figure 5C). In this setting, the initial top and rightmost groups were subdivided into the two clusters, respectively. The top group contained the clusters 4 and 9, while the rightmost group contained the clusters 6 and 8. Since the bottom group involves the largest number of cells, it was subdivided into 5 clusters, namely clusters 1 to 3, 5, and 7. Cells in the clusters 1 to 3 in the bottom group shared very similar expression profiles, expressing the pluripotency markers and being composed of only the mutants, which suggests that these clusters were the *ZIC2* mutated ES-like cells. The cells in the cluster 9 of the top group were also expressing pluripotency marker genes, and thereby they were considered as wild-type ES-like cells. The clusters 4 and 5 were apparently intermittent cell populations transiting from the ES-like state to the mesoderm-like state. The primitive streak-associated genes were expressed in the cells of the clusters 4 and 5. Cells in the clusters 6 and 8 in the rightmost group were considered as the mesoderm-like ones, because of expressing the early and cardiac mesoderm formation marker genes. They formed a distinct population inside the rightmost group each other. Cells in the cluster 7, located on the leftmost edge of the bottom group and composed of only the *ZIC2* mutated cells, were considered as more differentiated cells that were completely different from the ES-like cells in the clusters 1 to 3 and 9. These cells in the cluster 7 showed low expression level of *NANOG*, but high-expression level of *LHX5*, suggesting they were nonmesodermal cells but likely to be differentiating into the neuronal lineage. A heat map image highlights the marker genes for each of the cell clusters (Figure 5D), and expression patterns of several newly identified marker genes are shown in Figure 5E. Among all of the clusters, the cardiac mesoderm-like clusters (clusters 6 and 8) are of most our interest. The markers for the two clusters are the apelin receptor gene (*APLNR*) and *LIX1* (Figure 5D,E). *APLNR* was not a very specific marker for the rightmost group, since the cells in the top group also expressed this gene. However, cells in the cluster 6 showed the highest expression level of *APLNR* (Figure 5E). Comparing the *APLNR* expression levels between the clusters 6 and 8 within the rightmost group, the latter of which contained more number of mutant cells than cluster 6, the cells in the cluster 8 showed a weaker expression level of *APLNR*. This observation suggests that the *APLNR*-related signaling pathway that was apparently activated in the wild-type cells during cardiac mesoderm formation was attenuated in the *ZIC2* mutants.

To further confirm the results analyzed earlier, we used alternative methods (ie, Seurat^{36,37} and U-map³⁸) to analyze the data set. In the end, similar results were obtained (Figure S4), and in them, *APLNR* could still mark the cardiac mesoderm populations.

Taken together, the single-cell RNA-seq analysis newly identified a distinct cardiac mesoderm forming population, which has the marker genes, such as *MESP1*, *TBX6*, *EOMES*, and *APLNR*. This population mainly contained the wild-type cells with only a small fraction of the *ZIC2* mutant-derived cells, suggesting that *ZIC2* may play an important role in the cell-fate decision of the early mesodermal precursors through the *APLNR*-related signaling pathway, at least in part, during

the in vitro human cardiogenesis. To validate the obtained single cell RNA-seq results, we examined in vitro protein expression of *APLNR* during cardiac differentiation of the wild type and the *ZIC2* mutated hESCs. As shown in Figure S5, protein expression of a pluripotency marker (*POUR5F1* [Oct3/4]; day 0) and a primitive mesoderm precursor marker (*Brachyury*; day 1-2) was comparable between the wild type and the *ZIC2* mutated hESC-derived cells, while protein expression of *APLNR* (day 1-3) was much attenuated in the *ZIC2* mutants, further supporting the single cell RNA-seq results.

3.8 | Differential roles of *ZIC2* in human and mouse cardiogenesis

Previous studies have shown that the *Zic2* gene point mutation in mouse models could recapitulate some of the developmental defects in humans, such as holoprosencephaly.³⁹⁻⁴¹ Cardiac defects, such as random looping and ventricular septal defect, were also observed in these mutant mice.^{39,42} However, if *Zic2* would play an essential role in murine cardiac progenitor formation, as seen in in vitro hESC differentiation, it can be hypothesized that abnormal cardiac phenotypes in mice with *Zic2* gene mutation would become severer than was thought before. Thus, as a proof-of-concept experiment to investigate a potential role of *Zic2* in murine cardiogenesis, we newly generated *Zic2*-mutated murine embryos using CRISPR-Cas9 system through zygote microinjection (see Methods). Two guide RNAs, approximately 500 bp apart from each other, were designed to target the exon 1 of the murine *Zic2* gene (Figure 6A). We have injected the guide RNAs into 50 murine zygotes. Among them, 26 injected zygotes turned into blastocysts, and only two developed into embryos on embryonic day 10.5 (E10.5). The results of PCR validation on the putative mutation regions indicated a high degree of mosaicism in the generated mutant mice (Figure 6B). Consistent with the previous reports,³⁹⁻⁴⁴ the E10.5 mutant embryos exhibited developmental defects in the heads. The two cerebral hemispheres were clearly visible in the wild-type embryo in the same stage, while these structures were atypical in the mutant embryos (Figure 6C). There were no folds at the tops of the wild type embryo, while the folds in the mutant embryos at the same region were clearly visible (Figure 6C). We next focused on the cardiac phenotypes in the mutant embryos. All the four chambers can be identified in the *Zic2* mutants (Figure 6D). From the left lateral view, the morphology of the left ventricle and atrium were similar between the wild type and mutant embryos. From the right lateral view, the right ventricle and atrium also emerged in both embryos. However, it appeared that the outflow tract region in the mutant embryos were flatter as compared with that in the wild type embryo (Figure 6D,E), suggesting that a slight outflow tract malformation might be detectable at this stage in the mutant embryos.

Together, the earlier and our studies have demonstrated that *Zic2* mutation in mice did not cause severe cardiac defect or malformation leading to cardiac death during early embryogenesis. These observations are contrasted with the results of the in vitro hESC studies described here, in which the *ZIC2* mutant hESCs showed almost

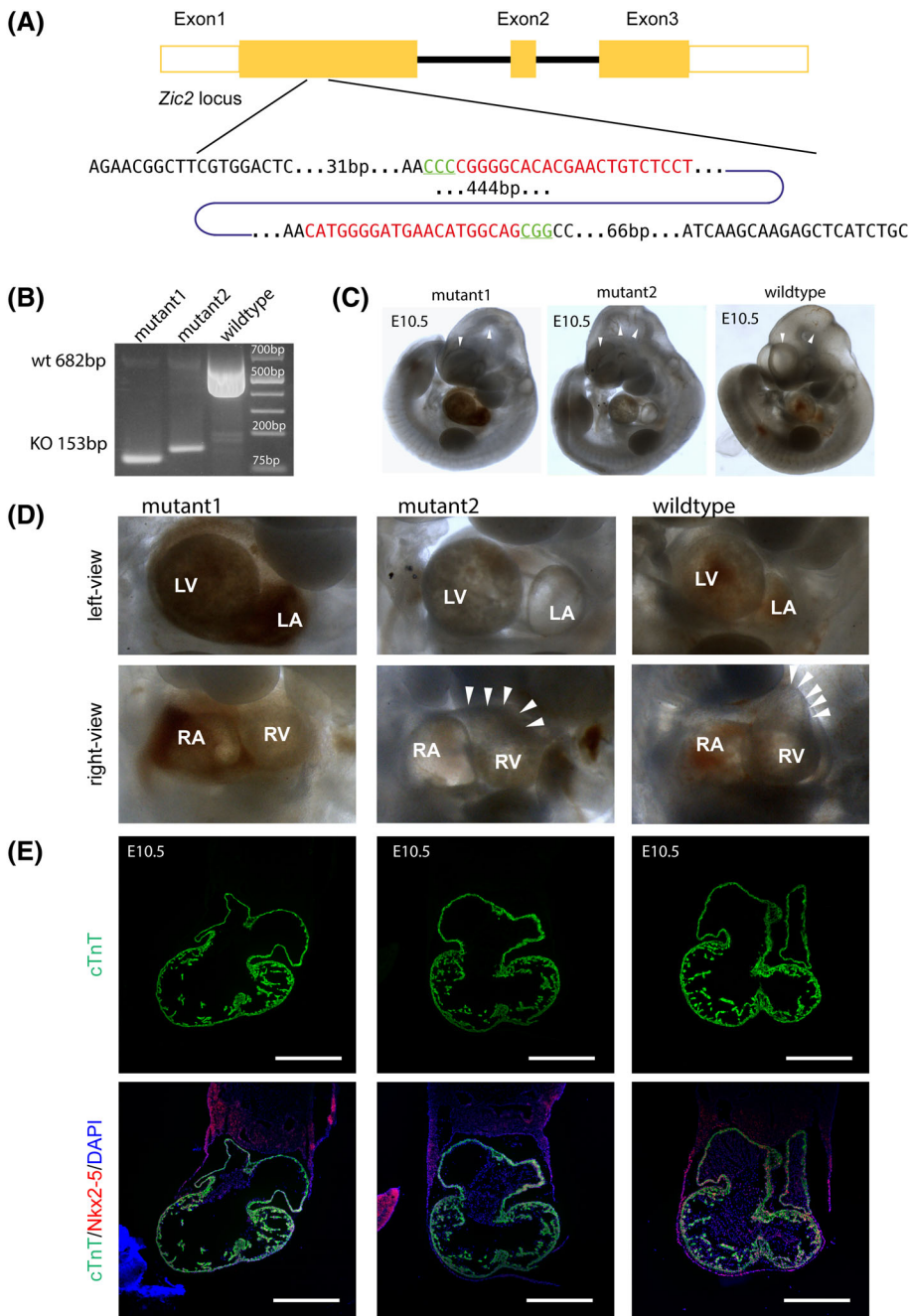


FIGURE 6 Cardiogenesis occurs in *Zic2* mutated mice embryos. **A**, Targeted sequences in *Zic2* loci. Two guide RNAs are labeled in red and their PAM site are colored green. The functional two guide RNAs will introduce a >500 bp deletion. **B**, Electrophoresis of the PCR product from the mutant and wild-type embryo tissue. The expected size of the wild type band is 682 bp and the expected size of the mutant band is 153 bp. **C**, E10.5 embryos images of mutant and wild-type mice. **D**, Zoom-in view of the embryo heart. LA, left atrium; LV, left ventricle; RA, right atrium; RV, left ventricle. Scale bar = 500 μ m

complete loss of the cardiomyocyte differentiation capabilities by shifting the cell fate of early mesodermal precursor cells into the noncardiac lineages. Collectively, our study identified the novel role of *ZIC2* as an essential regulator of early mesodermal precursors for cardiac commitment and differentiation in *in vitro* human cardiogenesis, and also highlights the potential species differences in regard to the *ZIC2*-related cardiogenesis between mice and humans.

4 | DISCUSSION

CRISPR-knockout screen is a powerful forward genetic tool to discover genes responsible for the phenotype of interest. We report here

a hPSC-based CRISPR screen for discovering the essential genes for cell-fate decision of early mesodermal precursors. hPSCs differentiation systems have for long time been used for modeling and studying the developmental process. The CRISPR-Cas9 genome editing provides a powerful tool for validating the function of a specific gene in the hPSC models. The screening methods we reported here involved the genome editing via the CRISPR-Cas9 technology and the development modeling of *in vitro* hPSC differentiation. As demonstrated here, hPSC mutant libraries are versatile, for they can be differentiated into multiple lineages and used to discover indispensable genes for the developmental commitment to specific cell lineages. For example, it can also be used for discovering the regulators for the pluripotency and proliferation of hPSCs, as shown in the current study. However,

this technology is still developing and has its limitations. First, the mutation outcome and efficiency of the nonhomolog repair from the CRISPR-Cas9 are still difficult to predict in spite of the large amount of efforts that have been put into the library designing. Second, the lack of strict selections usually causes high false positive rate. In the setting of pool screenings, even cells with the same genetic mutation might behave differently due to the local niche created by their neighbors. To filter out these levels of variabilities, strict selection should be applied in order to decrease the false positive rate. Otherwise, second round of screening or laborious individual validation is necessary. Third, the possibilities of potential off-target effects caused by recognition of the similar sequence sites to the target sequence on the human genome by guide RNAs cannot be completely excluded. About the two guide RNAs used for generation of the *ZIC2* mutated hESCs (Figure 2B), putative off-target sites with the highest similarity (two to three base pair mismatches in the PAM-distal part) to the target sequence were identified and sequenced; however, there were no mutations observed on those predicted off-target sites (Figure S6), suggesting that off-target cleavage did not appear to occur for the two guide RNAs.

In this study, we demonstrated the potential power of genome-wide unbiased CRISPR-knockout screens to identify the key steps in human mesodermal precursor cell- and heart progenitor cell-fate determination during human cardiogenesis with in vitro hPSC model systems. Through individual validations, we identified *ZIC2* as an essential gene that control the cardiac lineage commitment from early mesodermal precursor cells.

The zinc finger of the cerebellum (*ZIC*) genes are well known to have implication in neuroectoderm development and neural crest cell induction.⁴⁵ There are five *ZIC* genes in mammals, *ZIC1* to *ZIC5*. They all contain a five-Cys2His2-type zinc finger domain, which defines the family. *ZIC1* and *ZIC4* are associated with Dandy-Walker malformation (DWM).⁴⁶ Mutations in *ZIC2* cause a severe defect in forebrain development known as holoprosencephaly (HPE).⁴⁰ Mutations in *ZIC3* cause X-linked heterotaxy in addition to neural tube defects.^{47,48} No mutations have yet been described for *ZIC5* in humans. Mouse models have been used for the etiology study of these diseases. *Zic3* null mice exhibited approximately 50% embryonic lethality with additional 30% lethality in the perinatal period.⁴¹ The mutant embryos had defects in embryo turning, cardiac development, and neural tube closure.

Mice with one severe loss-of-function allele of *Zic2* is known as kumba. Unlike humans in which one allele mutation is sufficient to cause to disease, mice with one kumba allele have no describable defects, and the HPE phenotype presented only in the double mutant case. Studies in mutant mice showed that HPE is produced by a transient defect in the development of prechordal plate, a structure required for forebrain midline morphogenesis, validating the function of *Zic2* during organizer formation.⁴⁴ The *Zic2* null mice have cardiac situs problems, indicating a L-R axis patterning defect.⁴² By examining the embryos at later stages, more cardiovascular malformations have been found, such as abnormal ventricular topology, great artery transposition, double outlet right ventricle, ventricular septal defect, and rightward looped aortic arch.³⁹ Nevertheless, in our study, *Zic2*-

mutant murine embryos in the early stage (E10.5) had all four chambers of the heart, further illustrating the different roles of *ZIC2* required during cardiogenesis between mice and humans.

In terms of molecular functions, two functions of *ZIC2* have been previously reported: *ZIC2* can bind directly to TCF4 and inhibit the transcriptional activity of the β -catenin/TCF4 complex⁴⁹; and *ZIC2* can also function with Mbd3/NuRD in regulating the chromatin state and transcriptional outputs of genes linked to differentiation.⁵⁰ Based on these previous studies, the mechanism we proposed here is the inhibitory function of *ZIC2*, which allows the precise execution of specific molecular program for proper differentiation. During gastrulation, immature cells are migrating through the primitive streak and form a variety of mesodermal precursor cells. The activation signal is overwhelming. Without strong and proper inhibitory systems, multiple differentiation processes are likely activated at the same time and compete to take control of the cells. The inhibitory system can dampen spiking signaling and leaving the primary signals to convert the cells into the appropriate progenitors and their progenies in a spatiotemporal manner. By performing single-cell RNA-seq analysis of the differentiating cells on day 3, we have captured the transitional stage of early and cardiac mesoderm formation. By comparing the wild type and the *ZIC2* mutant cells, we found that unlike wild-type cells, *ZIC2* mutant cells could not enter the differentiation path into the cardiac mesoderm lineage whereas they often differentiated into other mesodermal lineages such as the osteoblast lineage, suggesting that *ZIC2* mutation switched the cell fate of the early mesoderm precursors. Further, although not frequently, some of the *ZIC2* mutant cells entered the differentiation path into the neuronal lineage while none of the wild-type cells differentiated into the neuronal lineage, suggesting that *ZIC2* mutation might also affect the cell-fate upstream the early mesodermal stage to some degree. Taken together, these observations indicate that the *ZIC2* mutant cells were somehow misinterpreting the environment signal and underwent un-directed differentiation into the noncardiac lineages. In addition, we found that in the mesoderm forming cluster (ie, cluster 6 in Figure 5C,E), which was almost all composed of wild-type cells, the *APLNR* expression was much higher than the rest of the clusters, for example, cluster 8 involving a lot of the *ZIC2* mutant cells (Figure 5C,E), suggesting the *APLNR*-related signaling pathway might be altered in the absence of the functional *ZIC2* gene. The *APLNR* signaling can modulate the NODAL signaling pathway, a key pathway for mesendoderm induction.^{27,51} Although it remains unclear whether *ZIC2* acts directly or indirectly on the *APLNR* signaling pathway, our study connects *ZIC2* with the *APLNR* signaling pathway in the early mesodermal precursor cell-fate decision and cardiac lineage formation during cardiogenesis.

5 | CONCLUSION

We have developed a genome-wide CRISPR-knockout screen in in vitro hPSC model systems to uncover crucial genes for the formation of human mesoderm precursors and heart progenitors, and identified *ZIC2* as an essential gene for in vitro human cardiomyocyte

commitment and differentiation. The transcriptome analysis and other experimental evidence revealed the essential role of ZIC2 in human cardiogenesis to control the cell fate of early mesodermal precursors so as to directly differentiate into cardiac progenitors and thereafter cardiomyocytes. These findings also provide new insights into the potential congenital heart defects in patients with mutation of the ZIC2 gene, emphasizing the importance of the cell-fate decision of the mesodermal precursors in the early embryonic stage for proper cardiogenesis in humans.

ACKNOWLEDGMENTS

We would like to acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure, NGI, and Uppmax for providing assistance in massive parallel sequencing and computational infrastructure. Single cell RNA-seq experiment is supported by the single cell facility in Integrated Cardio Metabolic Centre (ICMC), Department of Medicine, Huddinge, Karolinska Institutet. pLentiCRISPR was a gift from Feng Zhang at Massachusetts Institute of Technology. We thank Wellcome Trust Sanger Institute for making the PiggyBac material available. This work was supported by grants from the Swedish Research Council and the Knut och Alice Wallenbergs Stiftelse and the European Research Council Advanced Research Grant Award (AdG743225)

CONFLICT OF INTEREST

K.R.C. declared advisory role and research funding from EQT-AstraZeneca and stock ownership in SWIB-Procella-Smartwise-Moderna. The other authors declared no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

J.X.: conception and design, collection and assembly of data, data analysis and interpretation, manuscript writing; C.Z., K.S.F., R.Y., Y.X., K.B.: collection and/or assembly of data; M.S.: data analysis and interpretation, manuscript writing; K.R.C.: conception and design, financial support, data analysis and interpretation, manuscript writing, final approval of manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Jiejia Xu  <https://orcid.org/0000-0003-3885-3546>

REFERENCES

1. Shalem O, Sanjana NE, Hartenian E, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*. 2014;343:84-87.
2. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. *Science*. 2014;343:80-84.
3. Hart T, Chandrashekar M, Aregger M, et al. High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell*. 2015;163:1515-1526.
4. Parnas O, Jovanovic M, Eisenhaure TM, et al. A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell*. 2015;162:675-686.
5. DeJesus R, Moretti F, McAllister G, et al. Functional CRISPR screening identifies the ufmylation pathway as a regulator of SQSTM1/p62. *Elife*. 2016;5:e17290.
6. Wang T, Yu H, Hughes NW, et al. Gene essentiality profiling reveals gene networks and synthetic lethal interactions with oncogenic Ras. *Cell*. 2017;168:890-903. e815.
7. Pusapati GV, Kong JH, Patel BB, et al. CRISPR screens uncover genes that regulate target cell sensitivity to the morphogen sonic hedgehog. *Dev Cell*. 2018;44:271.
8. Chen S, Sanjana NE, Zheng K, et al. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell*. 2015;160:1246-1260.
9. Ruiz S, Mayor-Ruiz C, Lafarga V, et al. A genome-wide CRISPR screen identifies CDC25A as a determinant of sensitivity to ATR inhibitors. *Mol Cell*. 2016;62:307-313.
10. Bi P, Ramirez-Martinez A, Li H, et al. Control of muscle formation by the fusogenic micropeptide myomixer. *Science*. 2017;356:323-327.
11. Birsoy K, Wang T, Chen WW, Freinkman E, Abu-Remaileh M, Sabatini DM. An essential role of the mitochondrial electron transport chain in cell proliferation is to enable aspartate synthesis. *Cell*. 2015;162:540-551.
12. Arroyo JD, Jourdain AA, Calvo SE, et al. A genome-wide CRISPR death screen identifies genes essential for oxidative phosphorylation. *Cell Metab*. 2016;24:875-885.
13. Zhang J, Klos M, Wilson GF, et al. Extracellular matrix promotes highly efficient cardiac differentiation of human pluripotent stem cells: the matrix sandwich method. *Circ Res*. 2012;111:1125-1136.
14. Birket MJ, Casini S, Kosmidis G, et al. PGC-1alpha and reactive oxygen species regulate human embryonic stem cell-derived cardiomyocyte function. *Stem Cell Rep*. 2013;1:560-574.
15. Laflamme MA, Chen KY, Naumova AV, et al. Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. *Nat Biotechnol*. 2007;25:1015-1024.
16. Lian X, Hsiao C, Wilson G, et al. Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. *Proc Natl Acad Sci U S A*. 2012;109:E1848-E1857.
17. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15-21.
18. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923-930.
19. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26:139-140.
20. Picelli S, Faridani OR, Bjorklund AK, et al. Full-length RNA-seq from single cells using smart-seq2. *Nat Protoc*. 2014;9:171-181.
21. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12:357-360.
22. Korthauer ALDRK. SingleCellExperiment: S4 classes for single cell data. *Bioconductor Manual*. 2019. <https://bioconductor.org/packages/release/bioc/manuals/SingleCellExperiment/man/SingleCellExperiment.pdf>.
23. Lun AT, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res*. 2016;5:2122.
24. Kiselev VY, Kirschner K, Schaub MT, et al. SC3: consensus clustering of single-cell RNA-seq data. *Nat Methods*. 2017;14:483-486.
25. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods*. 2014;11:783-784.
26. Watanabe K, Ueno M, Kamiya D, et al. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat Biotechnol*. 2007;25:681-686.

27. Xu J, Gruber PJ, Chien KR. SMAD4 is essential for human cardiac mesodermal precursor cell formation. *STEM CELLS*. 2019;37:216-225.
28. Foo KS, Lehtinen ML, Leung CY, et al. Human ISL1(+) ventricular progenitors self-assemble into an in vivo functional heart patch and preserve cardiac function post infarction. *Mol Ther*. 2018;26:1644-1659.
29. Domian IJ, Chiravuri M, van der Meer P, et al. Generation of functional ventricular heart muscle from mouse ventricular progenitor cells. *Science*. 2009;326:426-429.
30. Bu L, Jiang X, Martin-Puig S, et al. Human ISL1 heart progenitors generate diverse multipotent cardiovascular cell lineages. *Nature*. 2009;460:113-117.
31. Moretti A, Caron L, Nakano A, et al. Multipotent embryonic isl1+ progenitor cells lead to cardiac, smooth muscle, and endothelial cell diversification. *Cell*. 2006;127:1151-1165.
32. Laugwitz KL, Moretti A, Lam J, et al. Postnatal isl1+ cardioblasts enter fully differentiated cardiomyocyte lineages. *Nature*. 2005;433:647-653.
33. Doerks T, Copley RR, Schultz J, Ponting CP, Bork P. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res*. 2002;12:47-56.
34. Satokata I, Ma L, Ohshima H, et al. Msx2 deficiency in mice causes pleiotropic defects in bone growth and ectodermal organ formation. *Nat Genet*. 2000;24:391-395.
35. Rao J, Pfeiffer MJ, Frank S, et al. Stepwise clearance of repressive roadblocks drives cardiac induction in human ESCs. *Cell Stem Cell*. 2016;18:341-353.
36. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36:411-420.
37. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. *Cell*. 2019;177:1888-1902. e1821.
38. Becht E, McInnes L, Healy J, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. 2018;37:38-44.
39. Dykes IM, Szumska D, Kuncheria L, et al. A requirement for Zic2 in the regulation of Nodal expression underlies the establishment of left-sided identity. *Sci Rep*. 2018;8:10439.
40. Elms P, Siggers P, Napper D, Greenfield A, Arkell R. Zic2 is required for neural crest formation and hindbrain patterning during mouse development. *Dev Biol*. 2003;264:391-406.
41. Purandare SM, Ware SM, Kwan KM, et al. A complex syndrome of left-right axis, central nervous system and axial skeleton defects in Zic3 mutant mice. *Development*. 2002;129:2293-2302.
42. Barratt KS, Glanville-Jones HC, Arkell RM. The Zic2 gene directs the formation and function of node cilia to control cardiac situs. *Genesis*. 2014;52:626-635.
43. Davies B, Davies G, Preece C, Puliyadi R, Szumska D, Bhattacharya S. Site specific mutation of the Zic2 locus by microinjection of TALEN mRNA in mouse CD1, C3H and C57BL/6J oocytes. *PLoS One*. 2013;8:e60216.
44. Warr N, Powles-Glover N, Chappell A, Robson J, Norris D, Arkell RM. Zic2-associated holoprosencephaly is caused by a transient defect in the organizer region during gastrulation. *Hum Mol Genet*. 2008;17:2986-2996.
45. Houtmeyers R, Souopgui J, Tejpar S, Arkell R. The ZIC gene family encodes multi-functional proteins essential for patterning and morphogenesis. *Cell Mol Life Sci*. 2013;70:3791-3811.
46. Grinberg I, Northrup H, Ardinger H, Prasad C, Dobyns WB, Millen KJ. Heterozygous deletion of the linked genes ZIC1 and ZIC4 is involved in Dandy-Walker malformation. *Nat Genet*. 2004;36:1053-1055.
47. Ware SM, Peng J, Zhu L, et al. Identification and functional analysis of ZIC3 mutations in heterotaxy and related congenital heart defects. *Am J Hum Genet*. 2004;74:93-105.
48. Gebbia M, Ferrero GB, Pilia G, et al. X-linked situs abnormalities result from mutations in ZIC3. *Nat Genet*. 1997;17:305-308.
49. Pourebrahim R, Houtmeyers R, Ghogomu S, et al. Transcription factor Zic2 inhibits Wnt/beta-catenin protein signaling. *J Biol Chem*. 2011;286:37732-37740.
50. Luo Z, Gao X, Lin C, et al. Zic2 is an enhancer-binding factor required for embryonic stem cell specification. *Mol Cell*. 2015;57:685-694.
51. Deshwar AR, Chng SC, Ho L, Reversade B, Scott IC. The Apelin receptor enhances Nodal/TGFbeta signaling to ensure proper cardiac development. *Elife*. 2016;5:e13758.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Xu J, Zhou C, Foo KS, et al. Genome-wide CRISPR screen identifies ZIC2 as an essential gene that controls the cell fate of early mesodermal precursors to human heart progenitors. *Stem Cells*. 2020;38:741-755. <https://doi.org/10.1002/stem.3168>