

1 **A transposable element into the human long noncoding RNA *CARMEN* is a switch**
2 **for cardiac precursor cell specification**

3
4 Isabelle Plaisance^a, Panagiotis Chouvardas^b, Yuliangzi Sun^c, Mohamed Nemir^a, Parisa Aghagolzadeh^a,
5 Farhang Aminfar^a, Sophie Shen^c, Woo Jun Shim^c, Francesca Rochais^d, Rory Johnson^{b,e}, Nathan Palpant^c,
6 Thierry Pedrazzini^{a, f}

7
8 ^aExperimental Cardiology Unit, Division of Cardiology, University of Lausanne Medical School, Lausanne,
9 Switzerland; ^bDepartment of Medical Oncology, Inselspital, University of Bern, Bern, Switzerland; ^cInstitute
10 for Molecular Bioscience, University of Queensland, Brisbane, Australia; ^dAix Marseille University,
11 INSERM, MMG, U1251, Marseille, France; ^eSchool of Biology and Environmental Science, University
12 College Dublin, Dublin, Ireland;

13
14
15 ^fCorrespondence to:
16 Thierry Pedrazzini, Experimental Cardiology Unit, Division of Cardiology, Department of Cardiovascular
17 Medicine, University of Lausanne Medical School, Lausanne, Switzerland

18 Email: thierry.pedrazzini@chuv.ch

19 Phone: +41 21 314 0765

20

21

22 **Short title:** A molecular switch encoded in a long noncoding RNA

23

24 **Original article:** 8672 words

1 **ABSTRACT**

2

3 **Aims:** The major cardiac cell types composing the adult heart arise from common multipotent precursor
4 cells. Cardiac lineage decisions are guided by extrinsic and cell-autonomous factors, including recently
5 discovered long noncoding RNAs (lncRNAs). The human lncRNA *CARMEN*, which is known to dictate
6 specification towards the cardiomyocyte (CM) and the smooth muscle cell (SMC) fates, generates a
7 diversity of alternatively spliced isoforms.

8

9 **Methods and Results:** The *CARMEN* locus can be manipulated to direct human primary cardiac
10 precursor cells (CPCs) into specific cardiovascular fates. Investigating *CARMEN* isoform usage in
11 differentiating CPCs represents therefore a unique opportunity to uncover isoform-specific function in
12 lncRNAs. Here, we identify one *CARMEN* isoform, *CARMEN-201*, to be crucial for SMC commitment.
13 *CARMEN-201* activity is encoded within an alternatively-spliced exon containing a MIRc short
14 interspersed nuclear element. This element binds the transcriptional repressor REST (RE1 Silencing
15 Transcription Factor), targets it to cardiogenic loci, including *ISL1*, *IRX1*, *IRX5*, and *SFRP1*, and thereby
16 blocks the CM gene program. In turn, genes regulating SMC differentiation are induced.

17

18 **Conclusions:** These data show how a critical physiological switch is wired by alternative splicing and
19 functional transposable elements in a long noncoding RNA. They further demonstrated the crucial
20 importance of the lncRNA isoform *CARMEN-201* in SMC specification during heart development.

21

1 TRANSLATIONAL PERSPECTIVE

2
3 LncRNAs regulate cell commitment and differentiation during development. Taking advantage of cardiac
4 precursor cells isolated from the human fetal and adult heart, we identify a novel mechanism mediated by
5 a specific isoform of the lncRNA *CARMEN*, which controls specification into the smooth muscle cell vs. the
6 cardiomyocyte fate. These results have direct implication in cell therapy for heart disease. Moreover,
7 *CARMEN* is associated with pathological states and represents an interesting biomarker for assessing the
8 extent of damage in the cardiovascular system. Our data propose therefore a mean to control cardiac cell
9 identity and behavior during heart development and disease.

10

11

12 KEY WORDS

13 Cardiac precursor cells; smooth muscle cells; long noncoding RNAs; Splicing; Transposable elements

1 INTRODUCTION

2 The mammalian heart is composed of several cell types that derives from mesodermal progenitor
3 cells of the first and second heart field ¹. The distinct lineages arise from multipotent cardiovascular
4 precursors ²⁻⁴. In this framework, very few studies have evaluated transcriptional regulation in primary
5 human cardiac precursor cells (CPCs). We have previously derived clonal populations of CPCs from the
6 fetal and the adult human heart ^{5, 6}. Their comparison allows for the dissection of the molecular
7 mechanisms controlling cardiac specification and differentiation. Understanding the processes regulating
8 cardiac cell programming and reprogramming provides also novel insights for the treatment of
9 cardiovascular disease.

10 Next generation sequencing coupled to assessment of the epigenetic landscape has shown that
11 mammalian genomes produce thousands of noncoding transcripts. Of these, long noncoding (lnc)RNAs
12 represent the most heterogeneous and diverse class of RNA molecules. lncRNAs can be multiexonic,
13 spliced, capped and polyadenylated ⁷. Current estimates predict the existence of approximately 200,000
14 lncRNAs in human, of which very few have been fully characterized. lncRNAs are implicated in a variety
15 of functions that define cell identity and behavior. They exert *Cis*- and *Trans*-regulatory functions,
16 controlling chromatin remodeling and transcription. *Cis*-acting lncRNAs operate at a close vicinity to their
17 site of transcription ⁸. On the other hand, *Trans*-acting lncRNAs leave their site of transcription and exert
18 functions at remote locations in the genome ⁹. In this case, lncRNAs partner with proteins such as
19 chromatin remodelers to modify the local chromatin environment at target locations. In the cardiovascular
20 system, several lncRNAs have been identified as key players of cell differentiation and homeostasis.
21 *Braveheart*, *Fendrr* and *Meteor* have been involved in mesoderm and cardiac differentiation ¹⁰⁻¹². *Myheart*
22 controls CM hypertrophy, and *Wisper* is a regulator of cardiac fibroblasts, critical for the development of
23 fibrosis ^{13, 14}. Then, *SMILR* and *SENCR* have been associated with SMC proliferation and differentiation ^{15,}

24 ¹⁶.
25 The way in which functions are encoded in lncRNAs' sequences remains enigmatic. It is thought
26 that lncRNAs comprise modular assemblages of functional elements, composed of structural motifs that
27 interact with proteins or other nucleic acids ¹⁷. Recent studies have implicated repetitive transposable
28 elements as a source of functional lncRNA domains ^{18, 19}. In this context, lncRNA loci can produce a

1 variety of transcripts through alternative splicing ^{7, 20}. It has been speculated that these alternative
2 isoforms, by containing different combinations of exon sequences, thereby exert diverse functions. Some
3 years ago, we identified *CARMEN*, a conserved lncRNA essential for cardiogenesis ²¹. *CARMEN* has
4 been previously referred to as *MIR143HG* because the locus hosts *MIR-143* and *MIR-145*, two
5 microRNAs (miRNAs) important for SMC differentiation ²². Nevertheless, the miRNA precursor represents
6 only one *CARMEN* isoform among several others. All other isoforms are lncRNA splice variants, with no
7 apparent coding potential, for which functions remain to be fully defined. We showed previously that three
8 isoforms were involved in cardiac specification in human fetal CPCs ²¹. Moreover, we reported that one
9 isoform, referred earlier as to *CARMEN-7* (formally *CARMEN-201* or ENST00000505254), was
10 differentially expressed in CPCs committing to the CM vs. the SMC lineage ⁶. Interestingly, the *CARMEN*
11 locus can be manipulated to direct CPCs into specific cardiovascular fates. Investigating *CARMEN*
12 isoform usage in differentiating CPCs represents therefore a unique opportunity to uncover isoform-
13 specific function in lncRNAs. Here, we demonstrate that *CARMEN-201* controls specification into the SMC
14 lineage in human CPCs, and that this function is encoded within an alternatively-spliced exon containing a
15 MIRc short interspersed nuclear element.

16

17

18 **METHODS**

19

20 Methods are described in details in the Supplementary Information available online.

21

22 **Human cardiac precursor cells (CPCs)**

23 Fetal heart biopsies were collected at 5 weeks of gestation following abortion, and adult atrial appendages
24 were obtained from patients undergoing cardiac surgery. CPCs were isolated by enzymatic digestion as
25 previously described ^{5, 6}.

26

27 **Human plasma samples**

1 Plasma samples were collected from patients admitted to the Lausanne University Hospital with a
2 diagnosis of myocardial infarction.

3

4 **Study approval**

5 The study was approved by the Lausanne University Hospital Ethics Committee and the Swiss Ethics
6 Committee (Human cardiac precursor cells: Protocols 22/03 and 178/09; Human plasma from patients
7 with myocardial infarction: Protocol PB_2018-00231 and 94/15), and was conducted according to the
8 Declaration of Helsinki. Written informed consent was obtained from all patients included in the study.

9

10 **GapmeR-mediated knockdown**

11 *CARMEN-201* silencing was obtained by adding *CARMEN-201*-specific GapmeRs targeting Exon 2 to
12 CPCs at a final concentration of 20 nM.

13

14 **siRNA-mediated knockdown**

15 siRNA transfection was performed using RNAiMax (ThermoFisher). CPCs were transfected with the
16 indicated siRNA at a final concentration of 10 nM.

17

18 **RNA extraction, RT-PCR and real time PCR**

19 Total RNA from plasma and from cultured cells was extracted using the miRNeasy Serum/Plasma
20 Advanced kit and the miRNeasy kit (Qiagen).

21

22 **Absolute quantification of CARMEN isoform expression**

23 pBluescript SK+ plasmids containing *CARMEN-201*, *CARMEN-205* or *CARMEN-217* cDNA were
24 synthesized (GenScript, USA). Data were converted into transcript copy per cell, assuming 100%
25 efficiency in conversion of RNA into cDNA.

26

27

28

1 **Subcellular fractionation**

2 Cells were harvested and lysed. The lysate was centrifuged at 3800g for 2 min. The supernatant
3 represented the cytoplasmic fraction. The pellet was used to produce the nuclear fraction.

5 **CRISPR/Cas9-mediated Exon 2 deletion**

6 A CRISPR/Cas9-D10A Nickase-based strategy was used to delete Exon 2 in the *CARMEN* gene.

8 **CRISPR-On activation**

9 We used the CRISPR/dCas9-based Synergistic Activator Mediator (SAM) gain of function system to
10 activate *CARMEN* isoform expression in fetal CPCs²³.

12 **RNA sequencing**

13 Sequencing libraries were prepared according to Illumina RNA Seq library kit instructions. Libraries were
14 sequenced with the Illumina HiSeq2000 (100bp; PE).

16 **TRIAGE analysis**

17 The Transcriptional Regulatory Inference Analysis from Gene Expression (TRIAGE) analysis, providing a
18 means to identify cell type-specific regulatory genes has been described²⁴.

20 **Uniform Manifold Approximation and Projection (UMAP) density plots**

21 We used the data provided by Asp et al., 2019 (accession number is European Genome-phenome
22 Archive (EGA): EGAS00001003996). The UMAP were generated using *Nebulosa* package.

24 **Lentiviral vectors**

25 The SIN-cppt-CMV-EGFP-WHV plasmid was a kind gift of Dr. Nicole Deglon (University of Lausanne,
26 Lausanne, Switzerland). EGFP sequences were replaced by either the wild-type *CARMEN*-201 Exon 2 or
27 a mutated Exon 2 containing scrambled MIRc sequences.

28

1 **RNA Pulldown and identification of *CARMEN-201* protein partners**

2 The pBluescript SK+ plasmids containing *C-201* Ex2 or *C-201* mutEx2 were used to synthesize
3 biotinylated probes. Precleared lysate was incubated with either no probe, biotinylated *C-201* Ex2 or
4 biotinylated mut*C-201* Ex2. Proteins were loaded on 12% SDS-PAGE.

6 **Tandem mass spectrometry**

7 For identifying *CARMEN-201* protein partners, tryptic peptide mixtures were injected on an Ultimate RSLC
8 3000 nanoHPLC system (Dionex, Sunnyvale, CA, USA).

10 **Western blotting**

11 Proteins associated to biotinylated *C-201* transcript were resolved by SDS-PAGE and electroblotted onto
12 PVDF membranes (GE Healthcare).

14 **RNA Immunoprecipitation (RIP)**

15 The RIP experiment was conducted as described¹⁴.

17 **RIP following by sequencing**

18 RIP was performed using anti-REST IgG. REST-associated transcripts were purified using the RNeasy
19 isolation kit (Qiagen). Sequencing libraries were prepared according to Illumina RNA Seq library kit
20 instructions. Libraries were sequenced with the Illumina HiSeq2000 (100bp; PE).

22 **Chromatin immunoprecipitation followed by real-time quantitative PCR (ChIP-qPCR)**

23 ChIP-qPCR was performed using the Pierce magnetic ChIP Kit (Thermo Scientific) and the ChIPAb+
24 REST Kit (Millipore) according to the manufacturers' instructions.

25
26
27
28

1 **Immunohistochemistry**

2 Cells were fixed in 2% paraformaldehyde, permeabilized in 0.3% Triton-X100 and processed for
3 immunostaining using appropriate antibodies. Coverslips were mounted with VECTASHIELD Antifade
4 Mounting medium with DAPI (VECTOR LABORATORIES).

6 **RNA BaseScope in situ hybridization**

7 RNA BaseScope in situ hybridization was used to detect *CARMEN-201* expression in adult CPCs in
8 culture or in sections of human hearts. In situ detection was performed the BaseScope kit (ACD biotech,
9 323900).

11 **Masson's trichrome staining**

12 Paraffin tissue sections were also processed for Masson's trichrome staining and analyzed with a Zeiss
13 Axioscan Z1 (Carl Zeiss).

15 **Statistics**

16 All data were collected from at least 3 independent experiments, performed at least in triplicates. Data
17 throughout the paper are expressed as mean \pm SEM. Statistical analysis: ANOVA with post-hoc Tukey.

20 **RESULTS**

21 *CARMEN* isoforms are differentially expressed in CPCs committing to the CM vs. the SMC fate.

22 To study the importance of *CARMEN* isoforms in cardiac differentiation, we took advantage of
23 primary CPCs isolated from the human heart. We isolated CPCs from the fetal heart at 5 weeks of
24 gestation, hereafter referred to as fetal CPCs ⁵. Adult CPCs were isolated from atrial appendages of
25 cardiac patients ⁶. Fetal CPCs have a high propensity to differentiate into CMs whereas adult CPCs
26 preferentially produce SMCs. Indeed, seven days after inducing differentiation, CMs, expressing *ACTN2*,
27 *ACTC1*, *MYH6*, *MYH7* and *TNN1*, were readily detected in fetal CPC cultures. In contrast, adult CPCs
28 differentiated into SMCs expressing *ACTA2*, *CALD1*, *CNN1* and *MYH11* (Fig. 1a and b; Supplemental Fig.

1 S1a). We next analyzed *CARMEN* expression under these two experimental conditions. Using capture
2 long-read sequencing²⁵, we detected seven annotated isoforms, i.e. *CARMEN (C)-201*, *-202*, *-205*, *-210*,
3 *-212*, *-215*, and *-217* (GENCODE v33; Fig. 1c). *C-215* is the precursor of *MIR-143* and *MIR145* whereas
4 all other isoforms represent lncRNA transcripts whose expression terminates upstream of the miRNAs. All
5 isoforms were significantly expressed in adult CPCs during SMC differentiation. In contrast, one isoform,
6 *C-201*, was downregulated during differentiation of fetal CPCs into CMs (Supplemental Fig. S1b).
7 Absolute quantification confirmed *C-201* represented the main isoform in adult CPCs but the least
8 abundant in fetal CPCs (Fig. 1d), suggesting its involvement in SMC differentiation. All *CARMEN* isoforms
9 but the miRNA precursor (*C-215*) were more abundant in the nucleus than the cytoplasm, a feature
10 compatible with the postulated function of lncRNAs as regulators of gene expression (Fig. 1e;
11 Supplemental Fig. S1c).

13 *CARMEN-201* controls SMC specification via its second exon

14 We next evaluated the involvement of *C-201* in SMC differentiation using a knockdown approach.
15 Antisense oligonucleotides (GapmeRs) were designed to target the *C-201* second exon, uniquely present
16 in this isoform (Fig. 1c). *C-201* was downregulated in adult CPCs following GapmeR transfection
17 (Supplemental Fig. S2a). The anti-*C-201* GapmeRs affected no other isoforms, demonstrating the
18 specificity of the approach. *C-201* silencing had no effects on differentiation of fetal CPCs into CMs but
19 completely blocked the capacity of adult CPCs to produce SMCs (Supplemental Fig. S2b and c). To
20 explore the importance of the *C-201* second exon in SMC specification, we produce adult CPCs lacking
21 the exon using CRISPR/Cas9 deletion. Guide (g)RNAs were designed to remove the second exon without
22 affecting any other exons (Supplemental Fig. S3a). *C-201* Exon 2-deleted adult CPC clones were derived.
23 Importantly, the *C-201* isoform was still expressed in adult CPCs lacking *C-201* Exon 2. However, the
24 transcript was reduced by the size of the second exon (Supplemental Fig. S3b). Endogenous *C-201*
25 expression was similarly induced in both wild-type and deleted adult CPC ($\Delta 201\text{Ex}2$) clones during
26 differentiation as detected by using a primer pair amplifying the 3' end of the transcript (Fig. 1f; primer pair
27 P1). The deletion of the second exon was verified using two primer pairs spanning the exon (i.e. P2; P3).
28 Next, we evaluated the capacity of deleted adult CPCs to produce SMCs. Measurement of marker gene

1 expression as well as immunostaining demonstrated that adult CPCs lacking *C-201* Exon 2 lost their
2 ability to differentiate into SMCs (Fig. 1f-h).

3

4 *CARMEN-201* induction is sufficient to trigger a SMC gene program in undifferentiated fetal CPCs

5 To evaluate the capacity of *C-201* to redirect fetal CPCs into the SMC lineage, we used a
6 CRISPR-On approach. We targeted transcriptional activators (dCas9-VP64; MS2-p65-HSF1), 200 bp
7 upstream of the *C-201* transcriptional start site (TSS) via expression of a modified gRNA containing MS2
8 aptamers (SAM system; ²³). Endogenous *C-201* expression was downregulated in fetal CPCs in the
9 absence of gRNA but markedly increased when the gRNA was expressed (Fig. 2a). Compared to the
10 large *C-201* induction, the other *CARMEN* isoforms, as well as the two hosted miRNAs, were marginally
11 activated (Fig. 2a; Supplemental Fig. S3c). We tested therefore the effects of *C-201* manipulation on the
12 fate of normally cardiogenic fetal CPCs. Strikingly, cells with forced *C-201* transcription produced large
13 amount of SMCs, indicating that *C-201* expression was sufficient to adopt a SMC fate (Fig. 2b-c).
14 Production of CMs was minimally affected, likely reflecting differentiation of untransfected fetal CPCs
15 (Supplemental Fig. S3d-e). Globally, expression of *C-201* was found to be necessary and sufficient for
16 inducing SMC specification in CPCs.

17

18 The *CARMEN-201* second exon contains a functional transposable element that drives SMC commitment

19 The results above prompted us to evaluate the role of the *C-201* second exon in SMC
20 specification. We looked at its primary structure (397 nucleotides) and detected a short interspersed
21 nuclear element (SINE), which was identified as a 126 nucleotide-long Mammalian-wide Interspersed
22 Repeat (MIR)c element. Of note, the exon is highly conserved in primates but not found in other species
23 (Fig. 2d). Intriguingly, MIRc is part of a catalog of predicted Repeat Insertion Domains of LncRNAs
24 (RIDLs) promoting nuclear localization ¹⁸, a feature consistent with the pronounced nuclear enrichment of
25 *C-201* (Fig. 1e). The exon also contains a partial ALU sequence.

26 To study the possible role of the MIRc element in determining SMC specification, we produced
27 two lentiviral vectors for overexpressing the entire Exon 2 in fetal CPCs (Supplemental Fig. S4a). The first
28 version contained wild-type MIRc sequences (*C-201* Ex2) whereas, in the second vector, the whole MIRc

1 sequence was scrambled (*C-201* mutEx2), thus maintaining length and sequence composition
2 (Supplemental Fig. S4a). Endogenous *C-201* expression was downregulated in differentiating fetal CPCs
3 as expected (Fig. 2e; P1). However, significant wild-type and mutated *C-201* Exon 2 expression was
4 measured in the respective transduced groups as judged by using Exon 2-specific primers (P3). We next
5 evaluated SMC and CM differentiation (Fig. 2f-g; Supplemental Fig. S4b-d). Non-transduced fetal CPCs
6 differentiated into CMs. In sharp contrast, overexpression of wild-type *C-201* Exon 2 forced fetal CPCs to
7 adopt a SMC fate, as evidenced by expression of SMC markers and immunostaining. Importantly, when
8 the MIRc element was mutated (*C-201* mutEx2), SMC differentiation was not observed, supporting a
9 critical role for this transposable element in the capacity of *C-201* to direct CPCs into the SMC lineage.

10

11 Identification of upstream regulators of CM and SMC specification in human CPCs

12 To better understand the processes leading to a switch in cell identity, we profiled the CPC
13 transcriptomes under different experimental conditions. Principal component analysis (PCA) was
14 conducted to evaluate differentiation when *C-201* Exon 2 was expressed and not expressed (Fig. 3a).
15 Samples of differentiating fetal CPCs (None) revealed temporal changes (d0 / Expansion; d1; d7)
16 characterizing CM specification. In contrast, fetal CPCs overexpressing wild-type *C-201* Exon 2 deviated
17 significantly from the original differentiation track. Importantly, CPC samples with overexpression of the
18 MIRc-mutated *C-201* Exon 2 were transcriptionally similar to untransfected samples, indicating again that
19 the transposable element was necessary for SMC commitment.

20 We next analyzed the transcriptomic data in details (Fig. 3b; Supplemental Fig. S4e). Although
21 fetal CPC differentiation was associated with early expression of cardiac TFs (e.g. *ISL1*; *TBX5*; *GATA4*;
22 *NKX2-5*) and late expression of genes associated with CM excitation-contraction coupling (e.g. *MYH6*;
23 *MYH7*; *SCN5A*; *TNNT2*), *C-201* Exon 2-overexpressing cells induced genes characterizing epicardial (e.g.
24 *WT1*; *TCF21*; *TBX18*), pericyte (e.g. *ACTA2*; *PDGFRB*; *NTS5E*) and SMC lineages (e.g. *TAGLN*; *CNN1*;
25 *MYH11*; *MYLK*). The two distinct transcriptomes were enriched with relevant terms in a gene ontology
26 analysis (Supplemental Fig. S4f). Several factors characterizing the second heart field (SHF) and the
27 outflow tract (OFT) were differentially expressed under the two experimental conditions (e.g. *GATA6*;
28 *HAND2*; *ISL1*; *MEIS1*; *MEIS2*; *PITX2*), suggesting a developmental origin for the fetal and adult CPCs.

1 Interestingly, undifferentiated fetal CPCs (Expansion) expressed factors from the Iroquois family of TFs
2 (*IRX1; IRX3; IRX4; IRX5*), which are known to be activated in mesodermal tissues, in particular the dorsal
3 mesoderm from which the heart derives ²⁶.

4 To identify regulators of specification during either CM or SMC differentiation, we took advantage
5 of TRIAGE (Transcriptional Regulatory Inference Analysis of Gene Expression), a novel metric for
6 inferring genes orchestrating cell identity ²⁴. TRIAGE is based on the observation that broad H3K27me3
7 occupancy at promoters enriches for genes driving cell fates. TRIAGE calculates a repressive score for
8 each gene, which can be combined with any type of genome-wide sequencing data to predict genes
9 governing cell differentiation. We therefore identified upstream regulators of fetal CPC specification with
10 and without *C-201* Exon 2 overexpression, 1 and 7 days after induction of differentiation. The results of
11 this inference analysis are presented in a way to compare rank orders based on input gene expression
12 (right column) vs. TRIAGE ranking (left column) (Fig. 3c and d). The top ranked TRIAGE candidates
13 revealed that key mesodermal and cardiac TFs were involved in the commitment of untransfected fetal
14 CPCs into the CM fate whereas determinants of pericyte and SMC identity were activated following *C-201*
15 Exon 2 expression. The complete lists of TRIAGE regulators identified under the two different
16 experimental conditions are presented in Supplemental Table S1. Analysis of cell identity (ARHS4; ²⁷)
17 and associated biological pathways (GO BP) using all identified TRIAGE candidates validated their
18 functional roles in CM and SMC commitment (Supplemental Fig. S5a).

19 To further investigate the association of *CARMEN* expression with blood vessel development in
20 vivo, we used data reporting the comprehensive transcriptional analysis of the embryonic human heart at
21 the single-cell level at different stages of gestation ²⁸. To understand how *CARMEN* expression was
22 related to gene network changes, we performed a correlation analysis of *CARMEN* against all genes in all
23 cells across all developmental time points (Fig. 3e). *CARMEN* abundance was associated preferentially
24 with pericyte and SMC gene programs, and not with the expression of CM or endothelial cell markers. We
25 next reanalyzed data generated at 6.5 weeks post-conception ²⁸, an important point in development
26 corresponding to the formation of the cardiac vasculature. Newly generated Uniform Manifold
27 Approximation and Projection (UMAP) density plots revealed lineage-specific markers, reflecting the
28 cellular diversity of the embryonic human heart including epicardial cells (*WT1; TCF21; TBX18*), pericytes

1 (*MCAM*; *CSPG4*; *ACTA2*; *PDGFRB*), and SMCs (*GATA6*; *TAGLN*; *CNN1*; *MYH11*) (Fig. 3f). *CARMEN*
2 was found expressed preferentially in these cells, substantiating an important role for this locus in pericyte
3 and SMC specification from the epicardium during development of the human heart. In contrast, CMs and
4 endothelial cells, marked by *TBX5*; *GATA4*; *NKX2-5*; *MYH6* and *PECAM1*; *KDR*, did not expressed
5 *CARMEN* at this developmental stage (Supplemental Fig. S5b). Interestingly, *ALDH1A2* and the COUP
6 transcription factor *NR2F2*, critical for atrial identity, were also expressed in *CARMEN*-expressing cells.

7

8 The MIRc element is a binding module for the RE1 Silencing Transcription Factor

9 Many lncRNAs function through interacting with proteins. To identify *C-201* protein partners, we
10 performed a RNA pulldown assay. Biotinylated *C-201* Exon 2 was used as a bait to purify *C-201*-
11 associated proteins from adult CPC lysates. As control, we used a mutated *C-201* Exon 2 lacking the
12 MIRc element. Proteins were identified by mass spectrometry. Three proteins were detected as
13 significantly associated to *C-201* Exon 2, namely the transcriptional repressor RE1 Silencing Transcription
14 Factor (REST; aka Neuron Restrictive Silencer Factor), the RNA methyltransferase NOP2/Sun RNA
15 Methyltransferase 6 (NSUN6), and the Replication Protein A1 (RPA1), a protein implicated in stabilization
16 of single-stranded DNA (Supplemental Fig. S6a; Supplemental Table S2). We first confirmed the
17 association of each protein with *C-201* by pulling down the full *C-201* transcript and quantifying the
18 amount of bound proteins by Western blotting (Fig. 4a-c; see full unedited gels in Supplementary
19 Information). An antisense *C-201* transcript was used as control. The results demonstrated the specific
20 interaction of *C-201* with REST, NSUN6 and RPA1 in proliferating and differentiating adult CPCs. We next
21 performed a RNA immunoprecipitation assay (RIP) using antibodies directed against REST, NSUN6 and
22 RPA1 respectively (Fig. 4d-f). Quantitative measurement of bound *CARMEN* isoforms confirmed the
23 interaction of *C-201* with REST, NSUN6 and RPA1 during CPC expansion and differentiation. No other
24 *CARMEN* isoforms were found associated to REST (Supplemental Fig. S6b). In contrast, small amounts
25 of *C-205* were detected as bound to NSUN6, and *C-217* appeared to interact with both NSUN6 and RPA1
26 (Supplemental Fig. S6c-d). Finally, to determine whether REST possessed intrinsic propensity to bind
27 MIRc-containing transcripts, we performed a REST RIP coupled to RNA profiling. REST-bound transcripts
28 were found significantly enriched in sequences containing a MIRc element as compared to transcripts not

1 bound by REST with similar length distribution and orientation. Nevertheless, the enrichment is also
2 higher when exploring repeat-containing genes in general, suggesting that global REST binding to RNA
3 molecules could require additional transposable elements (Supplemental Fig. S6e; Supplemental Table
4 S3).

5 To evaluate the role of REST in *C-201*-mediated CPC specification, we first used a knockdown
6 approach. Fetal CPCs were transfected with control or REST siRNA, and induced to differentiate into
7 SMC following *C-201* Exon 2 overexpression (Fig. 4g-h). Under control conditions, *C-201* Exon 2
8 expression forced CPCs to adopt a SMC specification. In sharp contrast, REST knockdown abolished the
9 capacity of *C-201*-expressing CPCs to commit to the SMC lineage. To confirm these results, we tested the
10 effects of REST silencing in adult CPCs spontaneously differentiating into SMC (Supplemental Fig. S6f).
11 Again, in the absence of REST, adult CPCs were unable to produce a SMC progeny. Interestingly, *C-201*
12 appeared downregulated in differentiating REST-deficient CPCs. This was also confirmed using RNA
13 BaseScope in situ hybridization in adult CPCs (Supplemental Fig. S6g). The resetting of *C-201* expression
14 following REST knockdown mimicked therefore what observed in differentiating fetal CPCs (Supplemental
15 Fig. S1a-b). Moreover, the cellular distribution of the *C-201* isoform was modified following REST silencing
16 (Supplemental Fig. S6h). Significant cytoplasmic enrichment was evident in the absence of REST, in
17 contrast to what measured under basal conditions. This observation suggested therefore that REST,
18 which carries a nuclear localization signal, might contribute to retain *C-201* in the nucleus via its capacity
19 to bind the MIRc element in the second exon.

21 *CARMEN-201* inhibits the CM fate via REST-mediated repression of cardiogenic transcription factor 22 expression

23 The association of *C-201* with REST suggested a mechanism involving the targeting of the
24 repressor to important regulatory loci to control cell fate in differentiating CPCs. In addition, RPA1, a *C-*
25 *201* protein partner, has been implicated in RNA:DNA triple helix stabilization²⁹, indicating that *C-201*
26 could interact with DNA sequences at target promoters. Thus, we took advantage of Triplex Domain
27 Finder (TDF), an application developed to detect DNA binding domains (DBDs) in lncRNAs³⁰. TDF
28 identifies also the DNA regions bound by the selected lncRNAs, i.e. gene promoters containing binding

1 sites for the lncRNA DBDs. Because REST had been associated with repression, we sought to identify C-
2 201 target promoters within the list of downregulated genes following C-201 Exon 2 overexpression.
3 Several DBDs were predicted in C-201, in particular in the sequences spanning the second exon (Fig. 5a).
4 As control, we performed a similar analysis for C-205 and C-217. The C-205 transcript was found to
5 contain distinct DBDs (Fig. 5a) whereas C-217 was not predicted to contain significant DBDs (not shown).
6 In total, 447 gene promoters were identified as potentially bound by C-201, and 387 by C-205
7 respectively. Among those, 337 were uniquely associated with C-201. These genes were related to GO
8 Biological Processes defining striated muscle contraction (Fig. 5b). In order to identify relevant targets of
9 C-201/REST action, we crossed the list of C-201-bound genes as predicted by TDF with the list of
10 TRIAGE candidates and the list of validated cardiac genes (Human Protein Atlas - ENSEMBL) (Fig. 5c;
11 Supplemental Table S4). Hypergeometric tests explored the significance of the overlaps and revealed four
12 primary candidates: *IRX1*; *IRX5*; *ISL1* and *SFRP1*. *ISL1* is a member of the LIM homeodomain family of
13 transcription factor, crucial for the development of the SHF. The two Iroquois homeobox transcription
14 factors *IRX1* and *IRX5* have been involved in developmental patterning in the embryonic heart. Finally,
15 *SFRP1* is a modulator of the WNT pathway that plays important roles in cardiac specification and
16 differentiation. Importantly, single-cell analysis demonstrated that these factors were not expressed in
17 *CARMEN*-expressing cells in the embryonic human heart at 6.5 weeks of gestation (Supplemental Fig.
18 S5b). Importantly, all four genes contained REST binding sites as determined by chromatin
19 immunoprecipitation followed by sequencing (ChIP-Seq) in a study interrogating REST binding in various
20 cell types (Supplemental Fig. S7a; Gene Expression Omnibus: GSM803369; GSM1010735;
21 GSM1010804)³¹. We evaluated therefore REST occupancy at the promoters of *IRX1*, *IRX5*, *ISL1* and
22 *SFRP1* in differentiating fetal and adult CPCs by ChIP-quantitative real-time (q)PCR. As expected, REST
23 occupied the promoter of the different candidate gene solely in adult CPCs expressing C-201, and not in
24 fetal CPCs (Supplemental Fig. S7b). In these experiments, we used *GAPDH* as a negative control (REST
25 occupancy in neither fetal nor adult CPCs) and *SYN1* as a positive control (REST occupancy in both fetal
26 and adult CPCs). Then, to formally demonstrate the dependence of REST targeting to *IRX1*; *IRX5*; *ISL1*
27 and *SFRP1* on C-201 action, we performed an additional experiment in adult CPCs with or without C-201

1 silencing. Consistently, REST occupancy at candidate promoters in adult CPCs was blunted following
2 GapmeR-mediated *C-201* knockdown (Figure 5d).

3 To validate the relevance of *ISL1*, *IRX1*, *IRX5* and *SFRP1* in CPC specification, we first
4 determined expression in fetal CPCs induced to differentiate into SMC following forced *C-201* expression
5 using CRISPR-On (Fig. 5e). Each candidate was downregulated after induction of *C-201* expression. We
6 then measured expression in adult CPC clones lacking *C-201* exon 2 ($\Delta 201\text{Ex}2$), i.e. not able to activate a
7 SMC gene program. *ISL1*, *IRX1*, *IRX5* and *SFRP1* expression was restored in these cells during
8 differentiation as compared to what observed in wild-type cells (Fig. 5f). Furthermore, we evaluated
9 expression after manipulating REST. We observed re-expression of the four factors in differentiating adult
10 CPCs after REST knockdown (Supplemental Fig. S7c). Moreover, REST silencing allowed also re-
11 expression of *ISL1*, *IRX1*, *IRX5* and *SFRP1* in fetal CPCs overexpressing *C-201* Exon 2 (Fig. 5g).
12 Altogether, these findings supported that the four candidates were under control by *C-201* via REST-
13 mediated repression.

14

15 *ISL1*, *IRX1*, *IRX5* and *SFRP1* silencing promotes the SMC fate

16 We next proceeded to validate the importance of *ISL1*, *IRX1*, *IRX5* and *SFRP1* in controlling
17 specification into the CM vs. the SMC lineage. To mimic REST-mediated repression, we used a siRNA
18 approach to knockdown each candidate in fetal CPCs normally committing to the CM lineage. We first
19 evaluated the effects of individual factor silencing on the capacity of fetal CPCs to produce a functional
20 progeny. Knocking down either *IRX1*, *IRX5* or *SFRP1* did not affect other candidate gene expression but
21 *ISL1* knockdown slightly decreased *IRX1*, *IRX5* and *SFRP1* levels (Supplemental Fig. S8a-d), suggesting
22 *ISL1* lies upstream of these factors in the cardiac regulatory network, in accordance with its role as a
23 pioneer transcription factor in the developing heart³². We then investigated re-specification of fetal CPCs
24 into the SMC fate following siRNA-mediated silencing of each factor individually or in combination. In
25 differentiating fetal CPCs, knocking down either *IRX1*, *IRX5*, *ISL1* or *SFRP1* restored epicardial gene
26 expression, i.e. *WT1*, *TCF21* and *TBX18*, confirming that downregulation of these critical cardiogenic
27 factors was a mandatory step in reprogramming CPCs into the smooth muscle fate (Fig. 6a-d). In addition,
28 individual factor silencing also resulted in the reexpression of *GATA6*, *HAND2*, *PDGFRA* and *TBX20*,

1 (Supplemental Fig. S8e-h), which are known to mark bipotential cardiac precursors giving raise to CMs
2 and SMCs³³. Nevertheless, manipulating each candidates separately has little impact on SMC gene
3 expression (Supplemental Fig. S8i-l). In fact, *ISL1* knockdown had even a negative effect on late SMC
4 maker expression. This suggested that *ISL1* operated at the onset of CPC specification but was also
5 necessary for late-stage differentiation.

6 We tested therefore combinations of siRNAs targeting *IRX1*, *IRX5* and *SFRP1* (Fig. 6e). As a
7 positive control for the activation of the SMC gene program, *C-201* Exon 2-overexpressing fetal CPCs
8 were included in the experiment. Each combination was associated with a large induction of epicardial
9 and SMC gene expression as compared to individual knockdown, with maximal impact achieved when all
10 three factors were downregulated simultaneously. This manipulation was as potent as *C-201* Exon 2
11 overexpression in inducing SMC genes in normally cardiogenic CPCs. Accordingly, massive SMC
12 differentiation was observed by immunostaining following *IRX1*, *IRX5* and *SFRP1* knockdown (Fig. 6f).
13 Interestingly, commitment occurred at the expense of the cardiogenic lineage (Fig. 6e and f) but had no
14 impact on endothelial cell production (not shown). Our data indicated therefore that *IRX1*, *IRX5* and
15 *SFRP1* downregulation was sufficient to redirect fetal CPCs into the epicardial and the SMC lineages.
16 Nevertheless, as mentioned above, *ISL1* expression appeared necessary during SMC differentiation. To
17 formally demonstrate this point, we performed an additional experiment in which the four factors were
18 silenced together (Supplemental Fig. S8m). In this case, *ISL1* silencing produced a slight negative effect
19 on SMC marker expression induced by combined *IRX1*, *IRX5* and *SFRP1* knockdown, sustaining a role
20 for *ISL1* in the late stage of SMC differentiation.

21 Interestingly, manipulating *IRX1*, *IRX5* or *SFRP1* had a striking effect on *CARMEN* isoform
22 expression (Fig. 6g). Indeed, while endogenous *C-201* was downregulated during specification in
23 differentiating fetal CPCs, its expression was reactivated after *IRX5* and *SFRP1* knockdown, and even
24 more so when *IRX5* and *SFRP1* were silenced together, suggesting *C-201* was negatively regulated by
25 the two cardiogenic factors. Remarkably, *C-217* expression demonstrated a mirror image, consistent with
26 coordinated regulation of the two isoforms and suggesting a switch might operate during SMC
27 specification. *C-205* was not modulated under these different conditions.

28

1 CARMEN-201 expression is increased in response to myocardial infarction in humans

2 In an attempt to evaluate the relevance of our findings in disease, we queried the association of
3 *CARMEN* with cardiovascular traits using CTG-VIEW (<https://view.genoma.io>). We identified important
4 phenotypes related to cardiovascular conditions as strongly associated with *CARMEN* (Fig. 7a). This
5 prompted us to investigate whether *C-201* was differentially expressed in the damaged myocardium. We
6 first used RNA BaseScope in situ hybridization to localize *C-201* expression in the human heart. Samples
7 were collected from explanted hearts of transplant patients, and expression was compared in CMs vs.
8 mural cells (Fig. 7b-e). *C-201* was found uniquely expressed in mural cells of large coronary vessels.
9 Immunostaining for VIMENTIN (marking endothelial cells and fibroblasts) and smooth-muscle myosin
10 heavy chain (SMMHC; marking SMCs) supported *C-201* expression being associated primarily with SMCs
11 (Fig. 7f; Supplemental Fig. S8n). *C-201* expression seemed equally distributed in the ventricular and atrial
12 vasculature. Next, we measured *C-201* expression in the blood of patients experiencing acute coronary
13 syndrome, with no prior history of cardiac disease. Plasma samples were obtained during angioplasty that
14 took place less than 1 hours after myocardial infarction, and at 24 and 48 hours thereafter (Figure 7g).
15 Individuals were classified based on the presence or absence of ST elevation, namely STEMI and non-
16 STEMI (NSTEMI). *C-201* was not expressed immediately after infarction, suggesting that the transcript
17 was not induced under basal conditions. However, the amounts of transcript dramatically increased after
18 one and two days. Importantly, circulating *C-201* concentrations were more elevated in STEMI vs. non-
19 STEMI patients. Altogether, it suggested that *C-201* was expressed in large vessels of the heart and
20 responded acutely to hemodynamic stress with an expression being proportional to the severity of the
21 disease.

24 **DISCUSSION**

25 In this study, we characterized for the first time the role of lncRNA isoforms in cell fate
26 determination through a systematic examination of the human *CARMEN* locus. In primary human CPCs
27 committing to the SMC lineage, the *C-201* isoform associates with REST via its MIRc element, targets the
28 repressor to important cardiogenic loci, namely *IRX1*, *IRX5*, *SFRP1* and *ISL1*, represses their expression

1 and promotes SMC specification (see Graphical Abstract). Conversely, in CPCs adopting a CM fate, *C-*
2 *201* is not expressed, allowing *IRX1*, *IRX5*, *ISL1* and *SFRP1* expression and the subsequent activation of
3 the cardiogenic program. Importantly, the MIRc-containing exon in *C-201* is found in primates but not in
4 other mammals, suggesting that MIRc-mediated functions controlling commitment into the SMC lineage
5 has been integrated in the *CARMEN* locus only recently in evolution. The *CARMEN* locus has been
6 involved in cardiogenesis, implicating however other isoforms than *C-201*²¹. Our data suggest that
7 coordinate regulation of *C-201* and *C-217* expression takes place during specification, providing a
8 plausible mechanism for controlling specification. *CARMEN* isoforms appear to share a single promoter.
9 Yet, additional transcription start sites have been recently detected in the *C-201* isoform, suggesting
10 transcriptional regulation might control *C-201* expression in differentiating SMCs³⁴.

11 Multipotent cardiovascular precursor cells expressing *ISL1* give rise to both CMs and SMCs²⁻⁴. In
12 this context, our TRIAGE analysis identifies key cardiac TFs as regulators of fetal CPC differentiation. On
13 the other hand, CPCs respecified into the SMC lineage after *C-201* Exon 2 overexpression express a
14 different gene program. Induction of *GATA6*, *HAND2*, *PDGFRA* and *TBX20* in CPCs is a characteristic
15 feature of cardiovascular intermediates capable of producing CMs and SMCs³³. We show also that
16 commitment to the SMC lineage is characterized by the stepwise expression of markers of epicardium-
17 derived cells (EPDCs) such as *WT1*, *MEOX1*, *KRT19* and *TBX18*, and pericytes such as *MCAM*, *CSPG4*,
18 *ACTA2* and *PDGFRB*. During development, EPDCs establish the subepicardial mesenchyme, then
19 migrate into the myocardium. These cells represent a known source of pericytes and SMCs for the
20 forming coronary vasculature³⁵. In addition, genetic tracing experiments suggest that epicardial cells can
21 also give rise to a myocardial progeny³⁶⁻³⁸. Along the same line, a recent single-cell analysis identifies the
22 juxta-cardiac field (JCF) contributing to both EPDCs and CMs³⁹. Trajectory analysis revealed a link
23 between precursors from the JCF and the posterior SHF, supporting the postulated developmental origin
24 of *CARMEN*-expressing CPCs. Of note, TRIAGE identifies *IRX1* and *IRX5* as important regulators of
25 cardiogenesis. *IRX1* is detected in the trabeculated and compact myocardium of the developing
26 ventricular septum whereas *IRX5* demonstrates a subendocardial to subepicardial gradient of expression
27⁴⁰. Consistently, our experiments show that knocking down *IRX1* and *IRX5* in fetal CPCs allows
28 reexpression of epicardial and SMC markers, suggesting expression of the two factors is sufficient to

1 maintain a cardiogenic identity in committed precursors. *C-201* targets also *SFRP1*, a known WNT
2 antagonist. Downregulation of the WNT pathway is an important step in establishing cardiac fates ⁴¹.

3 In specifying CPCs, *C-201* acts via REST-mediated repression. A role for *C-201* in blocking REST
4 activity during SMC determination, for instance via sequestering REST, is unlikely since REST silencing
5 abolishes SMC commitment in *C-201*-expressing CPCs. Consistently, REST acts as a transcriptional
6 repressor in the developing heart, where it is thought to repress adult cardiac gene expression ⁴²⁻⁴⁴.
7 Accordingly, blockade of REST in the heart leads to cardiac dysfunction ⁴⁵. In addition, our results suggest
8 that temporal REST expression during the development of the heart also reflects the role of REST in cell
9 fate determination. REST binds *C-201* but not *C-205* and *C-217*, further substantiating the importance of
10 the *C-201*/REST complex for SMC differentiation. *C-201* appears to be also indirectly under control by
11 REST. *C-201* relocalizes into the cytoplasm upon REST silencing. Therefore, the nuclear enrichment of *C-*
12 *201* could depend in part on its binding to REST. In this vein, REST binds *C-201* via the MIRc repeat in
13 the second exon, which was recently demonstrated to be associated with transcript nuclear localization ¹⁸.

14 Then, *C-201* associates with NSUN6 and RPA1. A recent study demonstrated a role for NSUN6 in
15 methylating mRNAs and lncRNAs, such as *MALAT1*, *NEAT1* AND *XIST* ⁴⁶. Mechanistically, lncRNA m5C
16 modification could be involved in transcript structure and stability ⁴⁷. *C-201* associates also to RPA1.
17 Importantly, RPA1 binds RNA with high affinity and promotes R-loop formation with homologous DNA ²⁹.
18 RNA-DNA hybrids initiate cellular processes regulating transcription and genome dynamics, two important
19 determinants of cell specification ^{48, 49}. Thus, RPA1 might contribute to effective targeting of REST at
20 regulatory loci via its capacity to stabilize *C-201*/promoter association, a feature consistent with the
21 predicted DNA binding domains in *C-201*.

22 *CARMEN* is expressed in adult tissues, particularly in the heart and the vasculature, reflecting
23 expression in CMs and SMCs ⁵⁰. An increasing body of evidence suggests *CARMEN* is associated with
24 pathological states in the cardiovascular system ²². Relevant to the present work, *CARMEN* was recently
25 demonstrated to regulate SMC differentiation and proliferation in atherosclerotic plaques ³⁴. Unstable
26 regions, in which high proliferation of dedifferentiated SMCs is observed, were characterized with
27 decreased *CARMEN* levels. Consistently, SMCs adopting a synthetic phenotype characterized *Carmen*
28 knockout mice. We have demonstrated previously that *CARMEN* is induced in the stressed mouse and

1 human hearts ²¹. Interestingly, human *CARMEN* isoforms were found differentially expressed depending
2 on the cardiac pathology, exemplifying again the complexity of the regulation of the *CARMEN* locus. Here,
3 we show that *C-201* levels increase in the blood during the acute phase of myocardial infarction. The likely
4 source of circulating *C-201* is the damaged heart. However, we cannot rule out the possibility that
5 hemodynamic stress also stimulates release from the peripheral vasculature. Nevertheless, assuming a
6 cardiac origin for *C-201*, its expression in CPCs could be part of the healing process initiated following
7 injury. In this scenario, CPCs expressing *C-201* would be specified preferentially into the SMC lineage.
8 Increased myocardial tissue perfusion has been reported in cell-based regenerative therapies for heart
9 disease. However, clinical trials failed to demonstrate functional improvement. This can be expected if
10 precursors are diverted from the cardiogenic lineage secondary to *C-201* expression. Our data propose
11 therefore a mean to improve CM production via modulating *C-201* expression in cardiac precursors.
12 Finally, *C-201* could represent an interesting biomarker for assessing the extent of cardiovascular damage
13 in various pathological situations.

14 Altogether, this work demonstrates how a biological switch is encoded in lncRNA sequence to
15 regulate cardiovascular specification. We have linked two key phenomena, namely alternative splicing and
16 the presence of deeply-conserved transposable elements. LncRNAs display far greater levels of
17 alternative splicing, although it has not been clear whether this reflects relaxed constraint or regulated
18 production of isoforms with distinct functions ²⁵. Here, we have shown an example where these two
19 processes converge to produce functional transcript isoforms, and provided the first physiological role for
20 a transposable element acting via a lncRNA during heart development.

21 22 23 **FUNDING**

24 This work was supported by grants from the Swiss National Science Foundation (T.P.; Grant No. CRSII5-
25 1_173738 and 31003A_182322).

26
27
28

1 AUTHORS CONTRIBUTION

2 T.P. conceived the project, designed experiments and wrote the paper. I.S. designed and
3 performed all wet-lab experiments, with help from M.N., P.A. and F.A.. P.C. and R.J. conducted the
4 bioinformatic analyses. Y.S., S.S., W.J.S. and N.P. performed the TRIAGE, GWAS and single-cell
5 analyses. F.R. provided critical human material for RNA FISH experiment, and her expertise in heart
6 development.

7

8

9 ACKNOWLEDGEMENTS

10 We express our gratitude to Dr Gabriel Cuellar Partida, University of Queensland, Brisbane,
11 Australia, for his help in analyzing the GWAS data. We are grateful to Dr Nicole Deglon and Dr Maria Del
12 Rey, University of Lausanne Medical School, Switzerland, for the plasmids used to prepare lentiviral
13 vectors. We thank the Genomic Technologies Facility, the Protein Analysis Facility, the Cellular Imaging
14 Facility and the Mouse Pathology Facility at the University of Lausanne, Switzerland, for providing
15 expertise in transcriptomics, proteomics, imaging and histology respectively.

16

17

18 CONFLICT OF INTEREST

19 T.P. is co-founder of Haya Therapeutics, Epalinges, Switzerland

20

21

22 DATA AVAILABILITY

23 All transcriptomic data has been deposited to GEO with the identifier GSE199930.

24

1 **REFERENCES**

2

3 1. Meilhac SM, Buckingham ME. The deployment of cell lineages that form the mammalian heart.
4 *Nat Rev Cardiol* 2018;**15**:705-724.

5 2. Kattman SJ, Huber TL, Keller GM. Multipotent flk-1+ cardiovascular progenitor cells give rise to
6 the cardiomyocyte, endothelial, and vascular smooth muscle lineages. *Dev Cell* 2006;**11**:723-732.

7 3. Moretti A, Caron L, Nakano A, Lam JT, Bernshausen A, Chen Y, Qyang Y, Bu L, Sasaki M,
8 Martin-Puig S, Sun Y, Evans SM, Laugwitz KL, Chien KR. Multipotent embryonic isl1+ progenitor
9 cells lead to cardiac, smooth muscle, and endothelial cell diversification. *Cell* 2006;**127**:1151-
10 1165.

11 4. Wu SM, Fujiwara Y, Cibulsky SM, Clapham DE, Lien CL, Schultheiss TM, Orkin SH.
12 Developmental origin of a bipotential myocardial and smooth muscle cell precursor in the
13 mammalian heart. *Cell* 2006;**127**:1137-1150.

14 5. Gonzales C, Ullrich ND, Gerber S, Berthonneche C, Niggli E, Pedrazzini T. Isolation of
15 cardiovascular precursor cells from the human fetal heart. *Tissue Eng Part A* 2012;**18**:198-207.

16 6. Plaisance I, Perruchoud S, Fernandez-Tenorio M, Gonzales C, Ounzain S, Ruchat P, Nemir M,
17 Niggli E, Pedrazzini T. Cardiomyocyte lineage specification in adult human cardiac precursor cells
18 via modulation of enhancer-associated long noncoding RNA expression. *JACC BST* 2016;**1**:472-
19 493.

20 7. Uszczynska-Ratajczak B, Lagarde J, Frankish A, Guigo R, Johnson R. Towards a complete map
21 of the human long non-coding RNA transcriptome. *Nat Rev Genet* 2018;**19**:535-548.

22 8. Gil N, Ulitsky I. Regulation of gene expression by cis-acting long non-coding RNAs. *Nat Rev*
23 *Genet* 2020;**21**:102-117.

24 9. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nat Rev*
25 *Genet* 2016;**17**:47-62.

26 10. Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, Ding H,
27 Butty VL, Torrey L, Haas S. Braveheart, a long noncoding RNA required for cardiovascular
28 lineage commitment. *Cell* 2013;**152**:570-583.

- 1 11. Grote P, Wittler L, Hendrix D, Koch F, Währisch S, Beisaw A, Macura K, Bläss G, Kellis M,
2 Werber M. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall
3 development in the mouse. *Dev Cell* 2013;**24**:206-214.
- 4 12. Alexanian M, Maric D, Jenkinson SP, Mina M, Friedman CE, Ting CC, Micheletti R, Plaisance I,
5 Nemir M, Maison D, Kernen J, Pezzuto I, Villeneuve D, Burdet F, Ibberson M, Leib SL, Palpant
6 NJ, Hernandez N, Ounzain S, Pedrazzini T. A transcribed enhancer dictates mesendoderm
7 specification in pluripotency. *Nat Commun* 2017;**8**:1806.
- 8 13. Han P, Li W, Lin CH, Yang J, Shang C, Nurnberg ST, Jin KK, Xu W, Lin CY, Lin CJ, Xiong Y,
9 Chien HC, Zhou B, Ashley E, Bernstein D, Chen PS, Chen HS, Quertermous T, Chang CP. A long
10 noncoding RNA protects the heart from pathological hypertrophy. *Nature* 2014;**514**:102-106.
- 11 14. Micheletti R, Plaisance I, Abraham BJ, Sarre A, Ting CC, Alexanian M, Maric D, Maison D, Nemir
12 M, Young RA, Schroen B, Gonzalez A, Ounzain S, Pedrazzini T. The long noncoding RNA Wisper
13 controls cardiac fibrosis and remodeling. *Sci Transl Med* 2017;**9**.
- 14 15. Bell RD, Long X, Lin M, Bergmann JH, Nanda V, Cowan SL, Zhou Q, Han Y, Spector DL, Zheng
15 D, Miano JM. Identification and initial functional characterization of a human vascular cell-enriched
16 long noncoding RNA. *Arterioscler Thromb Vasc Biol* 2014;**34**:1249-1259.
- 17 16. Ballantyne MD, Pinel K, Dakin R, Vesey AT, Diver L, Mackenzie R, Garcia R, Welsh P, Sattar N,
18 Hamilton G, Joshi N, Dweck MR, Miano JM, McBride MW, Newby DE, McDonald RA, Baker AH.
19 Smooth Muscle Enriched Long Noncoding RNA (SMILR) Regulates Cell Proliferation. *Circulation*
20 2016;**133**:2050-2065.
- 21 17. Kirk JM, Kim SO, Inoue K, Smola MJ, Lee DM, Schertzer MD, Wooten JS, Baker AR, Sprague D,
22 Collins DW, Horning CR, Wang S, Chen Q, Weeks KM, Mucha PJ, Calabrese JM. Functional
23 classification of long non-coding RNAs by k-mer content. *Nat Genet* 2018;**50**:1474-1482.
- 24 18. Carlevaro-Fita J, Polidori T, Das M, Navarro C, Zoller TI, Johnson R. Ancient exapted
25 transposable elements promote nuclear enrichment of human long noncoding RNAs. *Genome*
26 *Res* 2019;**29**:208-222.

- 1 19. Kapusta A, Kronenberg Z, Lynch VJ, Zhuo X, Ramsay L, Bourque G, Yandell M, Feschotte C.
2 Transposable elements are major contributors to the origin, diversification, and regulation of
3 vertebrate long noncoding RNAs. *PLoS Genet* 2013;**9**:e1003470.
- 4 20. Ziegler C, Kretz M. The More the Merrier-Complexity in Long Non-Coding RNA Loci. *Front*
5 *Endocrinol (Lausanne)* 2017;**8**:90.
- 6 21. Ounzain S, Micheletti R, Arnan C, Plaisance I, Cecchi D, Schroen B, Reverter F, Alexanian M,
7 Gonzales C, Ng SY, Bussotti G, Pezzuto I, Notredame C, Heymans S, Guigo R, Johnson R,
8 Pedrazzini T. CARMEN, a human super enhancer-associated long noncoding RNA controlling
9 cardiac specification, differentiation and homeostasis. *J Mol Cell Cardiol* 2015;**89**:98-112.
- 10 22. Vacante F, Denby L, Sluimer JC, Baker AH. The function of miR-143, miR-145 and the MiR-143
11 host gene in cardiovascular development and disease. *Vascul Pharmacol* 2019;**112**:24-30.
- 12 23. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N,
13 Gootenberg JS, Nishimasu H, Nureki O, Zhang F. Genome-scale transcriptional activation by an
14 engineered CRISPR-Cas9 complex. *Nature* 2015;**517**:583-588.
- 15 24. Shim WJ, Sinniah E, Xu J, Vitrinel B, Alexanian M, Andreoletti G, Shen S, Sun Y, Balderson B,
16 Boix C, Peng G, Jing N, Wang Y, Kellis M, Tam PPL, Smith A, Piper M, Christiaen L, Nguyen Q,
17 Boden M, Palpant NJ. Conserved Epigenetic Regulatory Logic Infers Genes Governing Cell
18 Identity. *Cell Syst* 2020;**11**:625-639 e613.
- 19 25. Lagarde J, Johnson R. Capturing a Long Look at Our Genetic Library. *Cell Syst* 2018;**6**:153-155.
- 20 26. Christoffels VM, Keijser AG, Houweling AC, Clout DE, Moorman AF. Patterning the embryonic
21 heart: identification of five mouse Iroquois homeobox genes in the developing heart. *Dev Biol*
22 2000;**224**:263-274.
- 23 27. Lachmann A, Torre D, Keenan AB, Jagodnik KM, Lee HJ, Wang L, Silverstein MC, Ma'ayan A.
24 Massive mining of publicly available RNA-seq data from human and mouse. *Nat Commun*
25 2018;**9**:1366.
- 26 28. Asp M, Giacomello S, Larsson L, Wu C, Furth D, Qian X, Wardell E, Custodio J, Reimegard J,
27 Salmen F, Osterholm C, Stahl PL, Sundstrom E, Akesson E, Bergmann O, Bienko M, Mansson-

- 1 Broberg A, Nilsson M, Sylven C, Lundeberg J. A Spatiotemporal Organ-Wide Gene Expression
2 and Cell Atlas of the Developing Human Heart. *Cell* 2019;**179**:1647-1660 e1619.
- 3 29. Mazina OM, Somarowthu S, Kadyrova LY, Baranovskiy AG, Tahirov TH, Kadyrov FA, Mazin AV.
4 Replication protein A binds RNA and promotes R-loop formation. *J Biol Chem* 2020;**295**:14203-
5 14213.
- 6 30. Kuo CC, Hanzelmann S, Senturk Cetin N, Frank S, Zajzon B, Derks JP, Akhade VS, Ahuja G,
7 Kanduri C, Grummt I, Kurian L, Costa IG. Detection of RNA-DNA binding sites in long noncoding
8 RNAs. *Nucleic Acids Res* 2019;**47**:e32.
- 9 31. Rockowitz S, Lien WH, Pedrosa E, Wei G, Lin M, Zhao K, Lachman HM, Fuchs E, Zheng D.
10 Comparison of REST cistromes across human cell types reveals common and context-specific
11 functions. *PLoS Comput Biol* 2014;**10**:e1003671.
- 12 32. Gao R, Liang X, Cheedipudi S, Cordero J, Jiang X, Zhang Q, Caputo L, Gunther S, Kuenne C,
13 Ren Y, Bhattacharya S, Yuan X, Barreto G, Chen Y, Braun T, Evans SM, Sun Y, Dobrova G.
14 Pioneering function of Isl1 in the epigenetic control of cardiomyocyte cell fate. *Cell Res*
15 2019;**29**:486-501.
- 16 33. Nosedá M, Harada M, McSweeney S, Leja T, Belian E, Stuckey DJ, Abreu Paiva MS, Habib J,
17 Macaulay I, de Smith AJ, al-Beidh F, Sampson R, Lumbers RT, Rao P, Harding SE, Blakemore
18 AI, Jacobsen SE, Barahona M, Schneider MD. PDGFR α demarcates the cardiogenic
19 clonogenic Sca1⁺ stem/progenitor cell in adult murine myocardium. *Nat Commun* 2015;**6**:6930.
- 20 34. Vacante F, Rodor J, Lalwani MK, Mahmoud AD, Bennett M, De Pace AL, Miller E, Van Kuijk K, de
21 Bruijn J, Gijbels M, Williams TC, Clark MB, Scanlon JP, Doran AC, Montgomery R, Newby DE,
22 Giacca M, O'Carroll D, Hadoke PWF, Denby L, Sluimer JC, Baker AH. CARMN Loss Regulates
23 Smooth Muscle Cells and Accelerates Atherosclerosis in Mice. *Circ Res* 2021;**128**:1258-1275.
- 24 35. Quijada P, Trembley MA, Small EM. The Role of the Epicardium During Heart Development and
25 Repair. *Circ Res* 2020;**126**:377-394.
- 26 36. Cai CL, Martin JC, Sun Y, Cui L, Wang L, Ouyang K, Yang L, Bu L, Liang X, Zhang X, Stallcup
27 WB, Denton CP, McCulloch A, Chen J, Evans SM. A myocardial lineage derives from Tbx18
28 epicardial cells. *Nature* 2008;**454**:104-108.

- 1 37. Zhou B, Ma Q, Rajagopal S, Wu SM, Domian I, Rivera-Feliciano J, Jiang D, von Gise A, Ikeda S,
2 Chien KR, Pu WT. Epicardial progenitors contribute to the cardiomyocyte lineage in the
3 developing heart. *Nature* 2008;**454**:109-113.
- 4 38. Lescroart F, Chabab S, Lin X, Rulands S, Paulissen C, Rodolosse A, Auer H, Achouri Y, Dubois
5 C, Bondue A, Simons BD, Blanpain C. Early lineage restriction in temporally distinct populations
6 of Mesp1 progenitors during mammalian heart development. *Nat Cell Biol* 2014;**16**:829-840.
- 7 39. Tyser RCV, Ibarra-Soria X, McDole K, Arcot Jayaram S, Godwin J, van den Brand TAH, Miranda
8 AMA, Scialdone A, Keller PJ, Marioni JC, Srinivas S. Characterization of a common progenitor
9 pool of the epicardium and myocardium. *Science* 2021;**371**.
- 10 40. Kim KH, Rosen A, Bruneau BG, Hui CC, Backx PH. Iroquois homeodomain transcription factors in
11 heart development and function. *Circ Res* 2012;**110**:1513-1524.
- 12 41. Protze SI, Lee JH, Keller GM. Human Pluripotent Stem Cell-Derived Cardiovascular Cells: From
13 Developmental Biology to Therapeutic Applications. *Cell Stem Cell* 2019;**25**:311-327.
- 14 42. Kuwahara K, Saito Y, Takano M, Arai Y, Yasuno S, Nakagawa Y, Takahashi N, Adachi Y,
15 Takemura G, Horie M, Miyamoto Y, Morisaki T, Kuratomi S, Noma A, Fujiwara H, Yoshimasa Y,
16 Kinoshita H, Kawakami R, Kishimoto I, Nakanishi M, Usami S, Saito Y, Harada M, Nakao K.
17 NRSF regulates the fetal cardiac gene program and maintains normal cardiac structure and
18 function. *EMBO J* 2003;**22**:6310-6321.
- 19 43. Zhang D, Wu B, Wang P, Wang Y, Lu P, Nechiporuk T, Floss T, Grealley JM, Zheng D, Zhou B.
20 Non-CpG methylation by DNMT3B facilitates REST binding and gene silencing in developing
21 mouse hearts. *Nucleic Acids Res* 2017;**45**:3102-3115.
- 22 44. Bingham AJ, Ooi L, Kozera L, White E, Wood IC. The repressor element 1-silencing transcription
23 factor regulates heart-specific gene expression using multiple chromatin-modifying complexes.
24 *Mol Cell Biol* 2007;**27**:4082-4092.
- 25 45. Kuwahara K. Role of NRSF/REST in the regulation of cardiac gene expression and function. *Circ*
26 *J* 2013;**77**:2682-2686.

- 1 46. Selmi T, Hussain S, Dietmann S, Heiss M, Borland K, Flad S, Carter JM, Dennison R, Huang YL,
2 Kellner S, Bornelov S, Frye M. Sequence- and structure-specific cytosine-5 mRNA methylation by
3 NSUN6. *Nucleic Acids Res* 2021;**49**:1006-1022.
- 4 47. Huang T, Chen W, Liu J, Gu N, Zhang R. Genome-wide identification of mRNA 5-methylcytosine
5 in mammals. *Nat Struct Mol Biol* 2019;**26**:380-388.
- 6 48. Aguilera A, Garcia-Muse T. R loops: from transcription byproducts to threats to genome stability.
7 *Mol Cell* 2012;**46**:115-124.
- 8 49. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. *Nat*
9 *Rev Genet* 2015;**16**:583-597.
- 10 50. Papatheodorou I, Moreno P, Manning J, Fuentes AM, George N, Fexova S, Fonseca NA,
11 Fullgrabe A, Green M, Huang N, Huerta L, Iqbal H, Jianu M, Mohammed S, Zhao L, Jarnuczak
12 AF, Jupp S, Marioni J, Meyer K, Petryszak R, Prada Medina CA, Talavera-Lopez C, Teichmann
13 S, Vizcaino JA, Brazma A. Expression Atlas update: from tissues to single cells. *Nucleic Acids*
14 *Res* 2020;**48**:D77-D83.

15

1 **FIGURE LEGENDS**

2

3 **Figure 1.** *CARMEN-201* controls SMC specification via its second exon

4 (a-b) Representative images and quantification of ACTN2-positive TNNI-positive CMs and SMMHC-
 5 positive SMCs in differentiating fetal and adult CPC cultures. Scale bar: 50 μ m. (c) Annotated *CARMEN*
 6 isoforms. The *CARMEN-201* second exon is highlighted in red. (d) Absolute quantification of *CARMEN-*
 7 *201 (C-201)*, *CARMEN-205 (C-205)* and *CARMEN-217 (C-217)* in differentiating fetal and adult CPCs. (e)
 8 Nuclear and cytoplasmic levels of *CARMEN-201*, *CARMEN-205*, *CARMEN-217*, *ACTB*, and *NEAT*. (f)
 9 Expression of *CARMEN-201*, *CARMEN-205*, *CARMEN-217*, SMC markers (*MYH11*; *CNN1*; *TAGLN*), and
 10 CM markers (*MYH6*; *MYH7*) in adult WT and Δ 201Ex2 CPC clones lacking *CARMEN-201* Exon 2. (g-h)
 11 Representative images and quantification of SMMHC-positive CNN1-positive TAGLN-positive SMCs in
 12 cultures of differentiating adult WT or Δ 201Ex2 CPC clones. Scale bar: 50 μ m. Data represent means \pm
 13 SEM; *p < 0.05 as compared to fetal CPCs in expansion; §p < 0.05 compared to the indicated conditions
 14 (n=3-6). ANOVA with post-hoc Tukey. See also Supplemental Fig. S1, S2 and S3.

15

16 **Figure 2.** The *CARMEN-201* Exon 2 contains a functional transposable element implicated in SMC
 17 specification.

18 (a) Expression of *CARMEN* isoforms and (b) SMC markers (*MYH11*; *CNN1*; *TAGLN*; *CALD1*) in
 19 differentiating fetal cells either untransfected (None), transfected with the SAM system in the absence of
 20 gRNA (SAM) or with the SAM system with a gRNA targeting sequences upstream the *CARMEN* TSS
 21 (SAM/gRNA). (c) Representative images and quantification of SMCs in cultures of differentiating fetal
 22 CPCs transfected as in (a). Scale bar: 50 μ m. (d) Position of the MIRc transposable element in the
 23 *CARMEN-201* second exon, and sequence conservation. (e) Expression of *CARMEN-201* using either a
 24 primer pair specific for the endogenous transcript (P1) or the exogenous exon 2 (P3), and (f) SMC
 25 markers (*MYH11*; *CNN1*; *TAGLN*; *CALD1*) in differentiating fetal CPCs either not transduced (None),
 26 transduced with a lentiviral vector encoding *CARMEN(C)-201* Ex2 or transduced with a lentiviral vector
 27 encoding a mutated *C-201* Ex2 (*C-201* mutEx2). (g) Representative images and quantification of
 28 SMMHC-positive CNN1-positive SMCs in cultures of differentiating fetal CPCs transfected as in (f). Scale

1 bar: 50 μ m. Data represent means \pm SEM; *p < 0.05 as compared to fetal CPCs in expansion; § p < 0.05
 2 compared to the indicated conditions (n=3-6). ANOVA with post-hoc Tukey. See also Supplemental Fig.
 3 S3 and S4.

4
 5 **Figure 3.** Transcriptomic analysis and identification of upstream regulators of CM and SMC specification
 6 in fetal CPCs with or without C-201 Exon 2 overexpression.

7 (a) Principal Component Analysis (PCA) visualizing transcriptomic data in a two-dimensional space. (b)
 8 Expression heatmap of regulators and markers of the CM and the SMC fate. The heatmaps show scaled
 9 TPM values. (c-d) TRIAGE transformation of input RNA-Seq data predicts regulatory genes controlling cell
 10 differentiation based on TRIAGE rank order (left) compared to ranking observed using simple gene
 11 expression (right). Control differentiation (None) vs. differentiation following C-201 Exon 2 overexpression
 12 (C-201 Ex2) 1 (d1) or 7 days (d7) after transduction. Regulators implicated in cardiovascular differentiation
 13 are highlighted in red. (e) Human heart single-cell data analysis reveals genes positively and negatively
 14 correlated with *CARMEN* expression during development. (f) UMAP plots showing epicardium-, pericyte-
 15 and SMC-specific expression of *CARMEN* in the human heart 6.5 weeks post-conception. See also
 16 Supplemental Fig. S4 and S5.

17
 18 **Figure 4.** Identification of *CARMEN-201* protein partners.

19 (a) Quantification of REST, (b) NSUN6 and (c) RPA1 by Western blotting in a protein pulldown assay
 20 using biotinylated sense or antisense *CARMEN-201* transcript in adult CPC lysates. Graphs show means
 21 \pm SEM *p < 0.05 as compared to input; § p < 0.05 comparing sense and antisense probe (n = 3). (d-f)
 22 Quantification of *CARMEN-201* enrichment after RNA immunoprecipitation using a control
 23 immunoglobulin G (IgG) or IgG directed against REST (anti-REST IgG), NSUN6 (anti-NSUN6 IgG) and
 24 RPA1 (anti-RPA1 IgG). Graphs show means \pm SEM; *p < 0.05 as compared to control (n=3). (g)
 25 Expression of *REST*, *CARMEN-201* and SMC markers (*MYH11*; *CNN1*; *TAGLN*) in differentiating fetal
 26 CPCs either not transduced (None), transduced with a lentiviral vector encoding C-201 Exon 2 (C-201
 27 Ex2), and treated with either a scrambled siRNA (Scr siRNA) or a siRNA directed against REST (Anti-
 28 *REST* siRNA). (h) Representative images and quantification of SMMHC-positive CNN1-positive SMCs

1 and ACTN2-positive CMs in cultures of differentiating fetal CPCs transfected as in (g). Scale bars: 50 μ m.
 2 Data represent means \pm SEM; * $p < 0.05$ as compared to fetal CPCs in expansion; $^{\S}p < 0.05$ compared to
 3 indicated conditions (n=3-6). ANOVA with post-hoc Tukey. See also Supplemental Fig. S6.

4
 5 **Figure 5.** Identification of *IRX1*, *IRX5*, *SFRP1* and *ISL1* as target genes of *C-201* action

6 (a) Significant DNA-binding domains (DBD) identified in the mature sequence of *C-201* and *C-205* when
 7 analyzed against the differentially downregulated genes on day 1 and day 7 following induction of
 8 differentiation. Graph shows the number of DNA-binding sites (DBS) for each DBD. (b) Venn diagram
 9 illustrating the overlap of promoters predicted to form triple helices with *C-201* and *C-205*. Functional
 10 enrichment analysis of the isoform-specific bound promoters. Graph shows the negative logarithm of the P
 11 value. (c) Venn diagram illustrating the identification of *IRX1*, *IRX5*, *ISL1* and *SFRP1* as common to the
 12 indicated lists of genes. Hypergeometric tests were performed to explore the significance of the overlap
 13 (d) REST occupancy at the promoters of *IRX1*, *IRX5*, *SFRP1* and *ISL1* in adult CPCs with or without
 14 GapmeR-mediated *C-201* silencing as determined by ChIP-qPCR. Occupancy at the *GAPDH* and the
 15 *SYN1* promoters was used as negative and positive controls respectively. (e) Expression of *IRX1*, *IRX5*,
 16 *ISL1* and *SFRP1* in differentiating fetal cells either untransfected (None), transfected with the SAM system
 17 in the absence of gRNA (SAM) or with the SAM system with a gRNA targeting sequences upstream the
 18 *CARMEN* TSS (SAM/gRNA). (f) Expression of *IRX1*, *IRX5*, *ISL1* and *SFRP1* in differentiating adult CPCs
 19 treated with Scrambled siRNA (Scr siRNA) or Anti-REST siRNA. (g) Expression of *IRX1*, *IRX5*, *ISL1* and
 20 *SFRP1* in differentiating fetal CPCs either untransfected, transfected with a lentiviral vector encoding *C-*
 21 *201* Ex2 or a mutated *C-201* Ex2 (*C-201* mutEx2), treated with either a scrambled siRNA (Scr siRNA) or a
 22 siRNA directed against REST (Anti-REST siRNA). Data represent means \pm SEM; * $p < 0.05$ as compared
 23 to CPCs in expansion; $^{\S}p < 0.05$ compared to indicated conditions (n=3-6). ANOVA with post-hoc Tukey.
 24 See also Supplemental Fig. S7.

25
 26 **Figure 6.** Validation of candidate cardiogenic factor downregulation

27 (a-d) Expression of epicardial genes (*WT1*, *TCF21*, *TBX18*) in differentiating fetal CPCs following *IRX1*,
 28 *IRX5*, *ISL1* and *SFRP1* silencing. (e) Epicardial (*WT1*, *TCF21*, *TBX18*) SMC (*MYH11*, *TAGLN*, *CNN1*)

1 and CM marker (*MYH6*) expression in differentiating fetal CPCs overexpressing *C-201* Ex2, and in fetal
2 CPCs treated with siRNAs directed against the indicated factors in combination. (f) Representative images
3 and quantification of ACTN2-positive CMs and SMMHC-positive CNN1-positive SMCs in cultures of
4 differentiating fetal CPCs transfected as in (e). Scale bars: 50 μ m. Expression of *CARMEN-201* (*C-201*),
5 *CARMEN-205* (*C-205*) and *CARMEN-217* (*C-217*) in differentiating fetal CPCs treated as in (e). Data
6 represent means \pm SEM; * $p < 0.05$ as compared to fetal CPCs in expansion; $\S p < 0.05$ as compared to
7 fetal CPCs in differentiation (n=6-12). ANOVA with post-hoc Tukey. See also Supplemental Fig. S8.

8
9 **Figure 7.** *CARMEN-201* expression is increased in response to myocardial infarction in humans
10 (a) Association of *CARMEN* with cardiovascular traits using CTG-VIEW. (b-e) Detection of *C-201* in the
11 failing human heart by RNA in situ hybridization assay (BaseScope). (b and c) Representative images of
12 sections of an explanted heart obtained from a heart failure patient. (b). Top: Hematoxylin/Eosin; scale
13 bar: 500 μ m; Bottom: Masson Trichrome staining; scale bar: 500 μ m. (c) Left: Masson Trichrome staining;
14 scale bar: 100 μ m; Right: Hematoxylin/Eosin; scale bar: 20 μ m. Red dots: Positive *C-201* BaseScope
15 signals. (d-e) Quantification of *C-201* expression in CMs and mural cells. Data represent means \pm SEM; * p
16 < 0.05 ; ANOVA with post-hoc Tukey. Two patients; 5 sections per patient; 5 to 10 different areas per
17 section. LV: Left ventricle; RV: Right ventricle; LA: Left atria; RA: Right atria. (f) Immunostaining detection
18 of VIMENTIN-positive cells (endothelial cells and fibroblasts) and SMMHC-positive cells (SMCs) in
19 adjacent sections of that used in b and c; scale bar: 25 μ m. (g) Time course of blood sample collection,
20 and expression of *C-201* in plasma of STEMI and NSTEMI patients (Data represent means \pm SEM; n=11;
21 * $p < 0.01$; ANOVA with post-hoc Tukey), and Table presenting patient characteristics. See also
22 Supplemental Fig. S8.

23
24

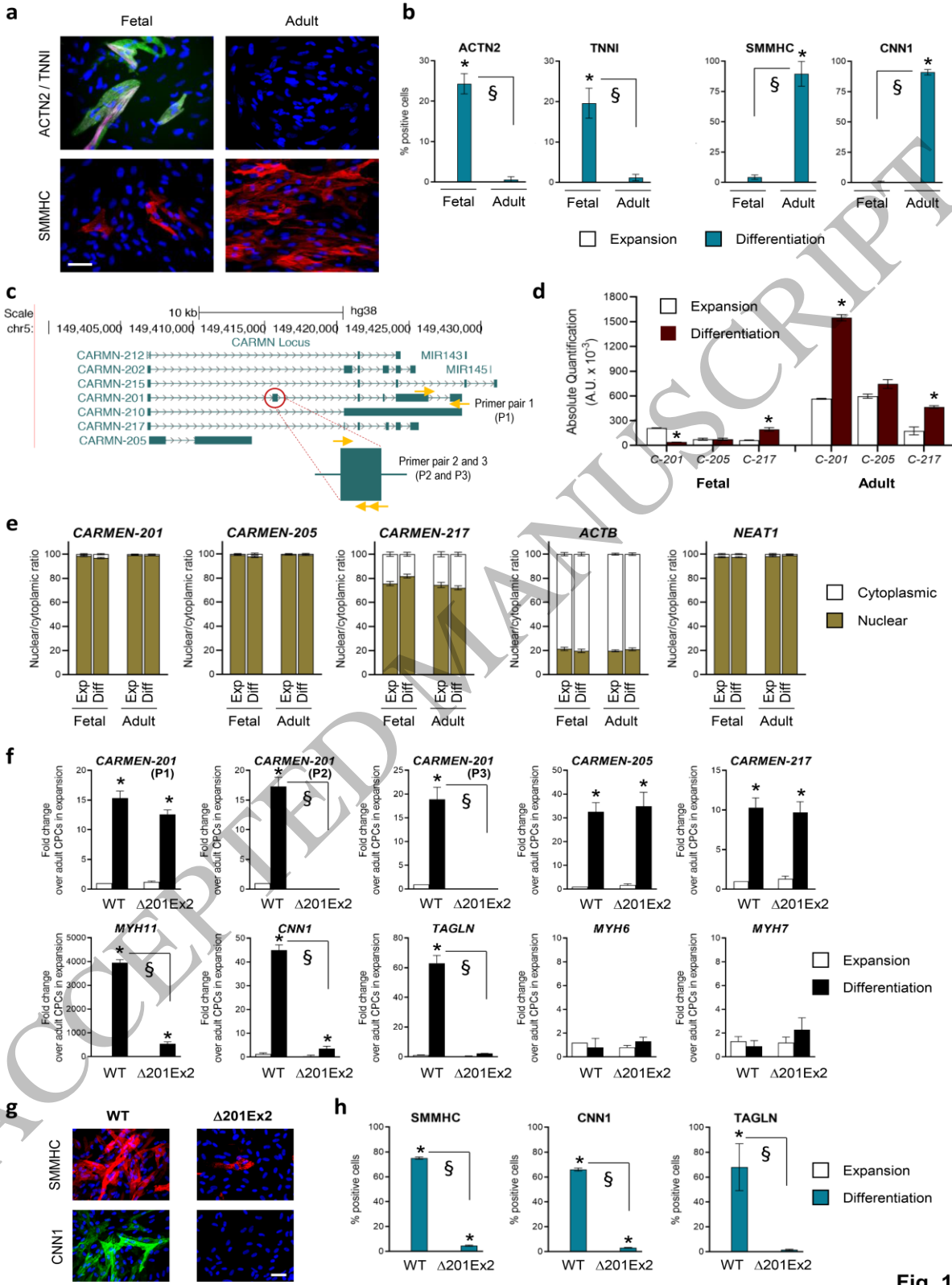


Fig. 1

Figure 1
160x231 mm (x DPI)

1
2
3

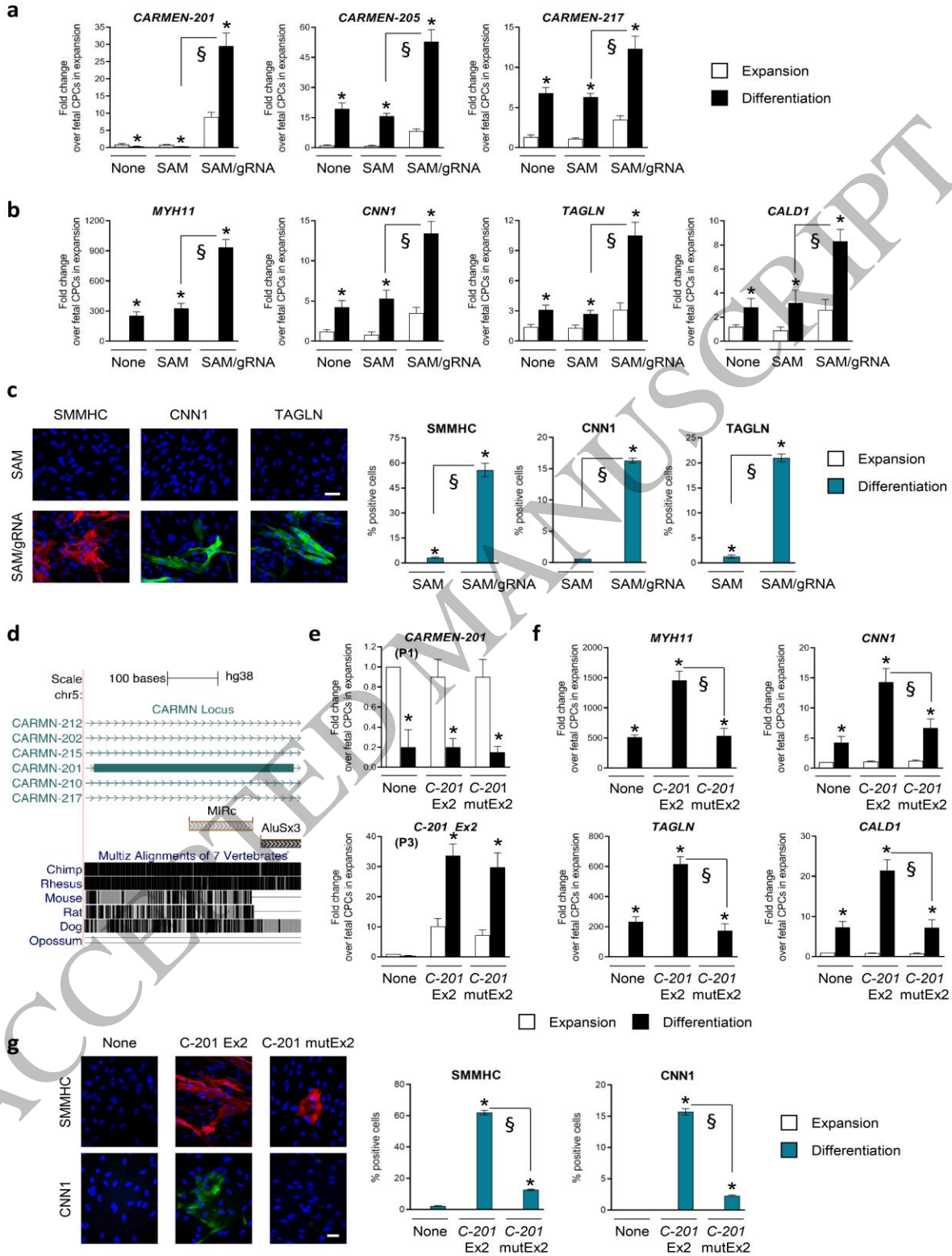


Fig. 2

Figure 2
160x231 mm (x DPI)

1
2
3

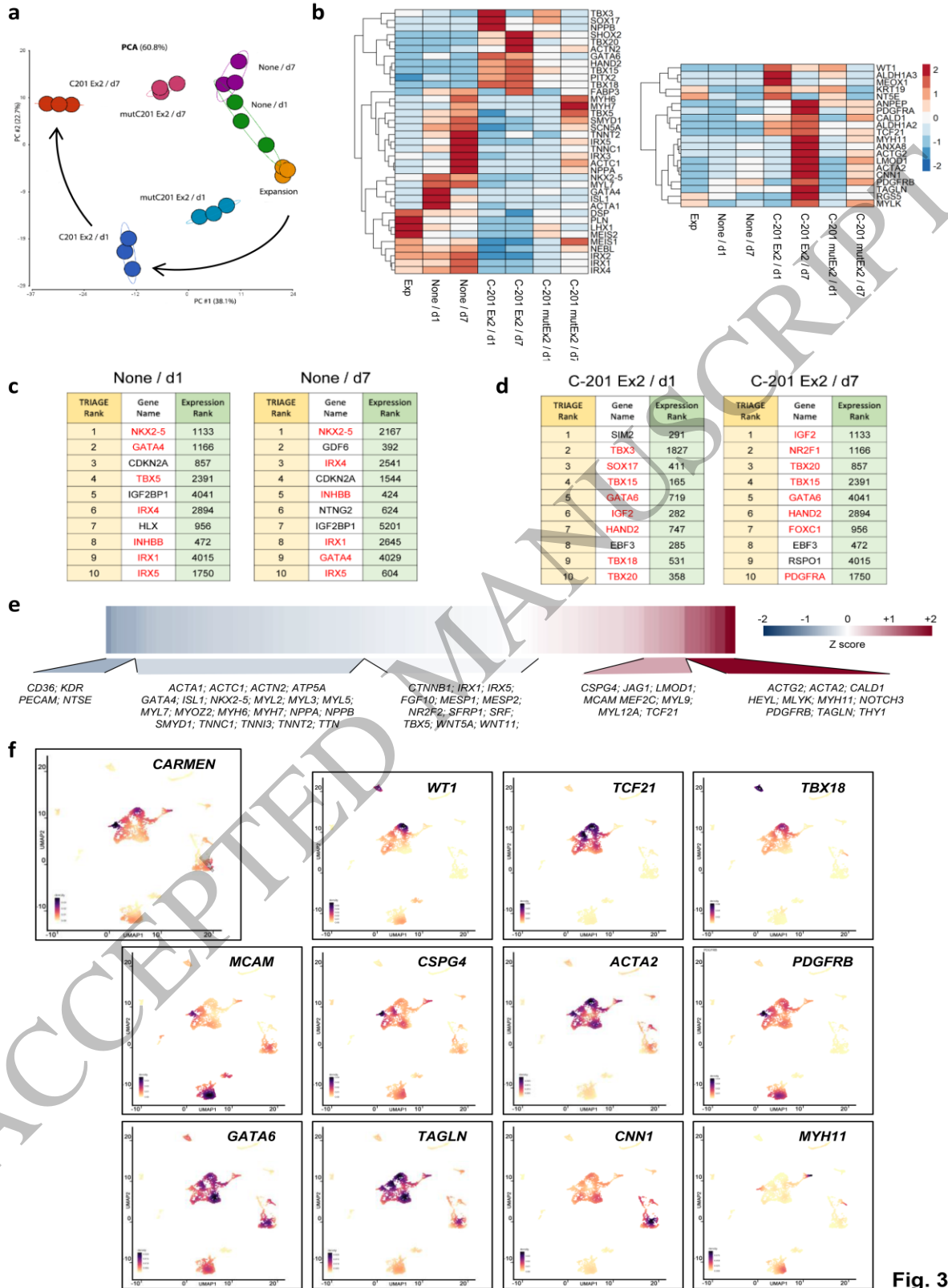


Fig. 3

Figure 3
160x231 mm (x DPI)

1
2
3

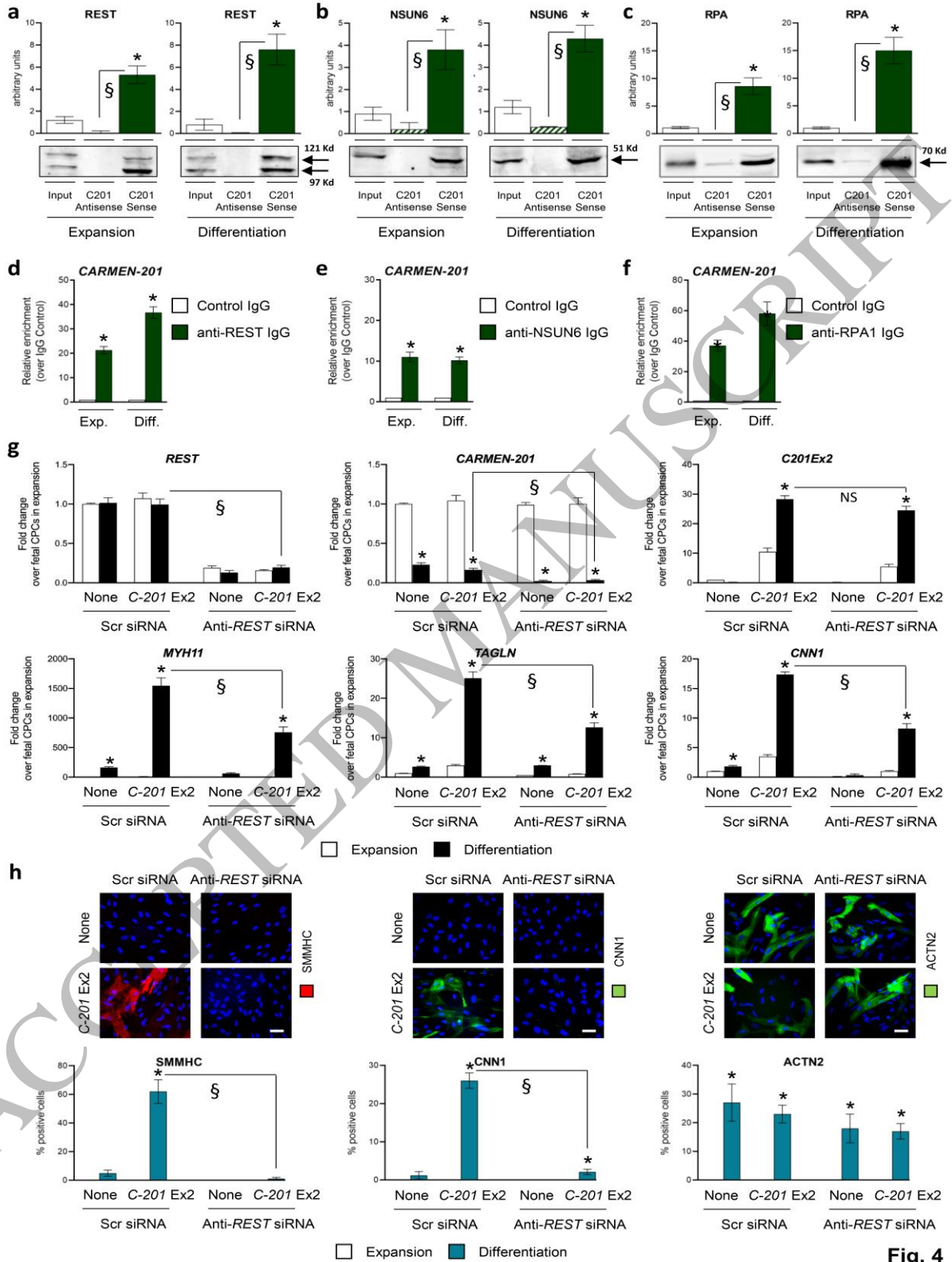


Fig. 4

Figure 4
160x231 mm (x DPI)

1
2
3

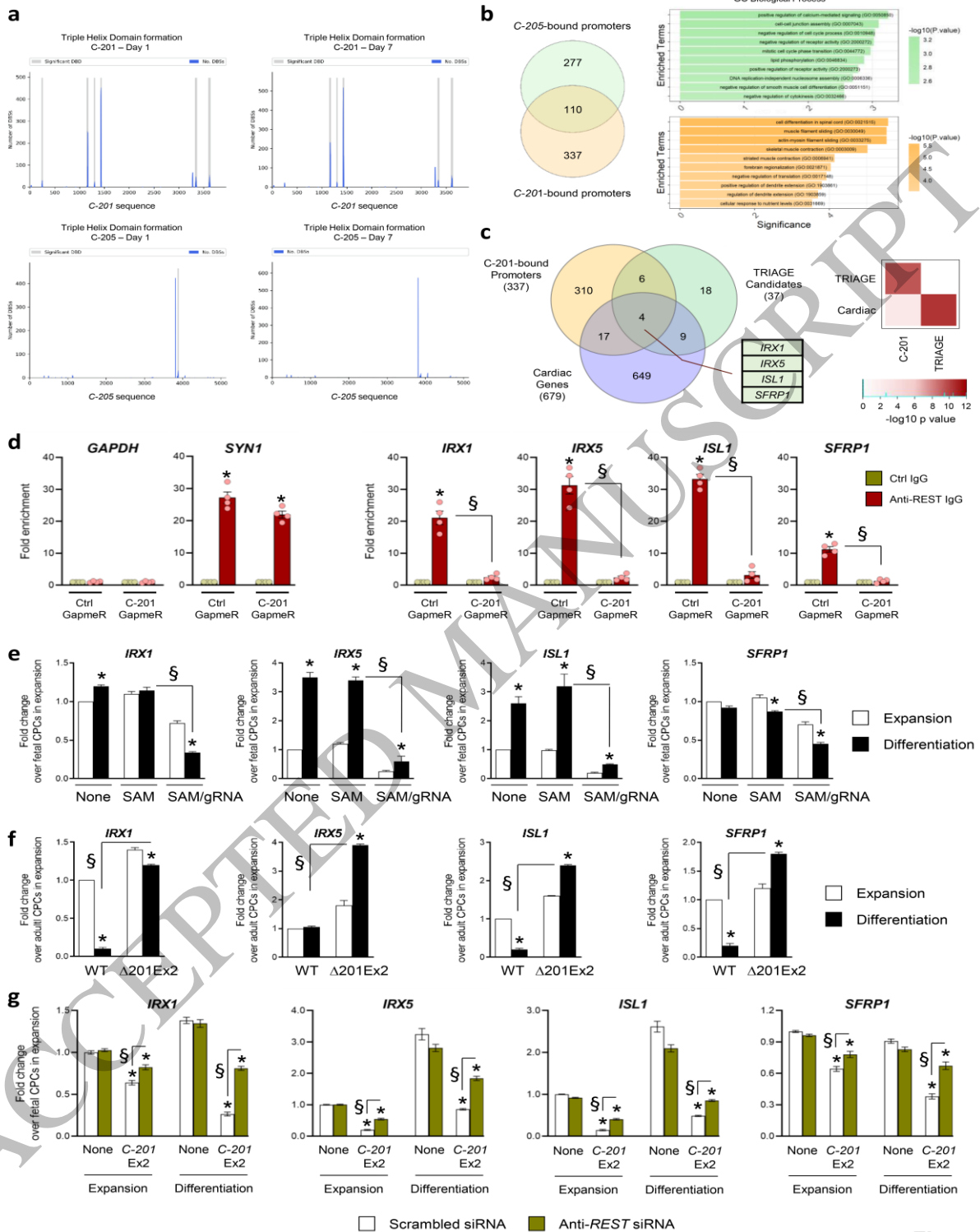


Fig. 5

Figure 5
160x231 mm (x DPI)

1
2
3
4

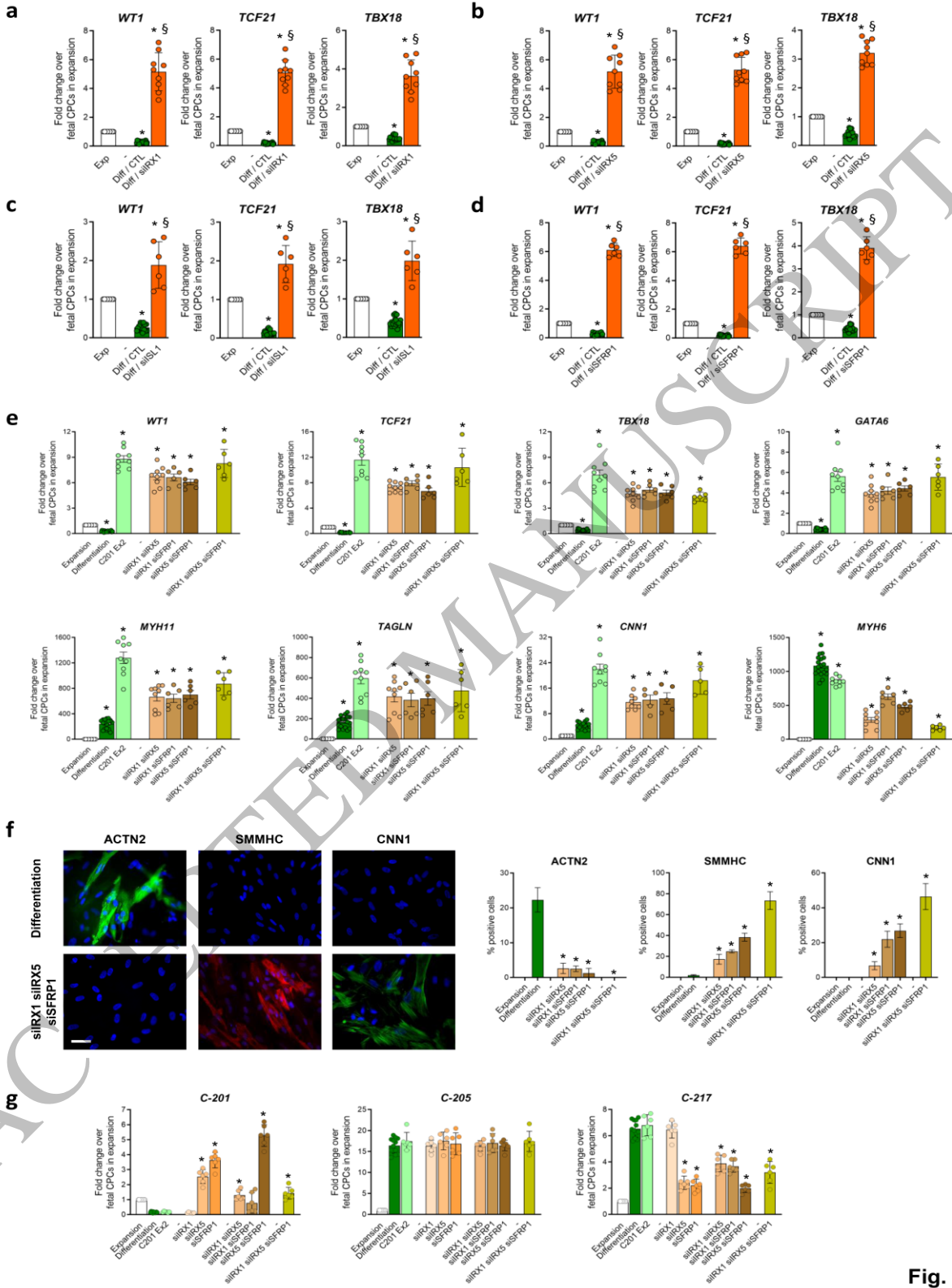
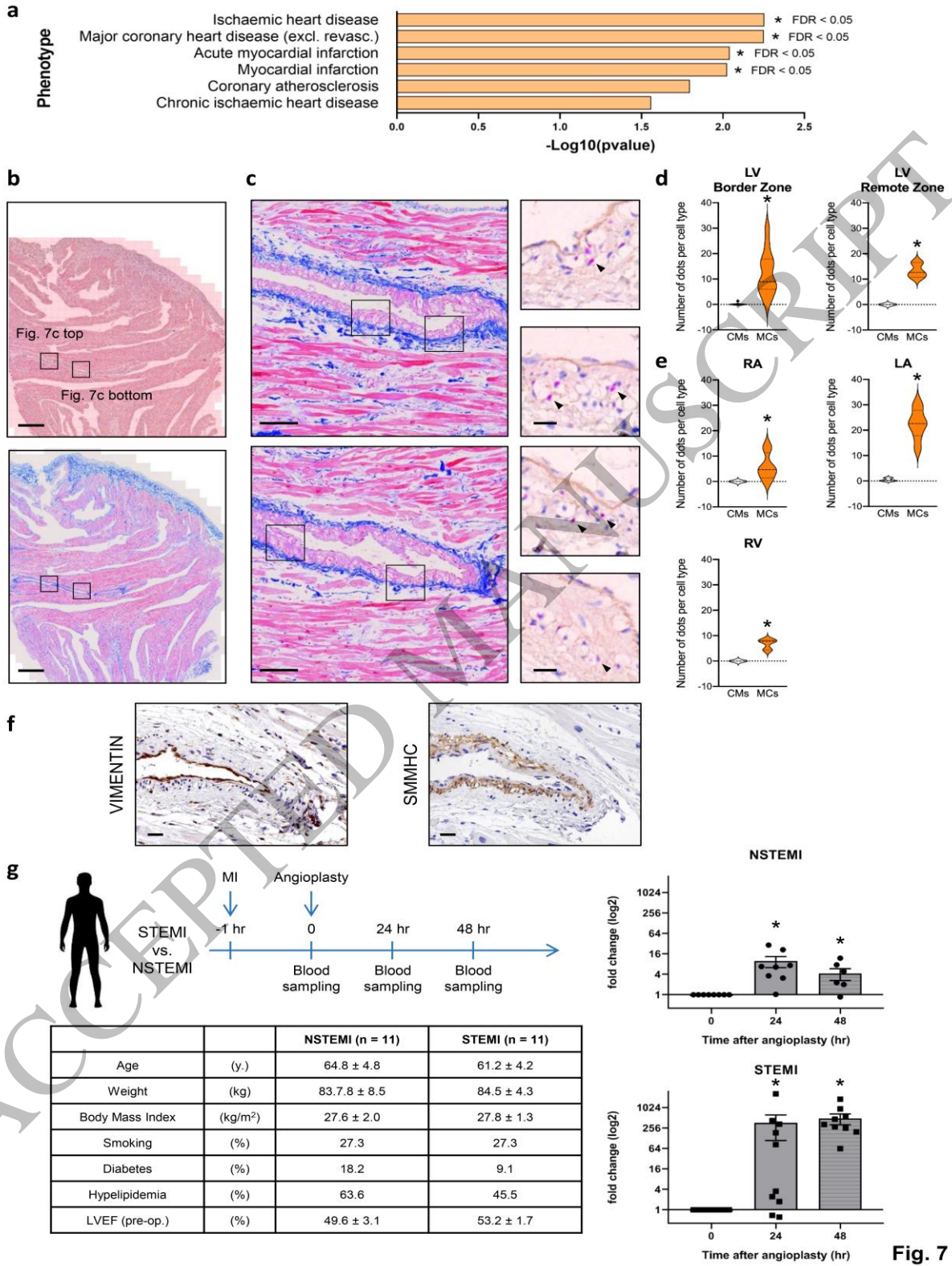


Fig. 6

Figure 6
160x231 mm (x DPI)

1
2
3

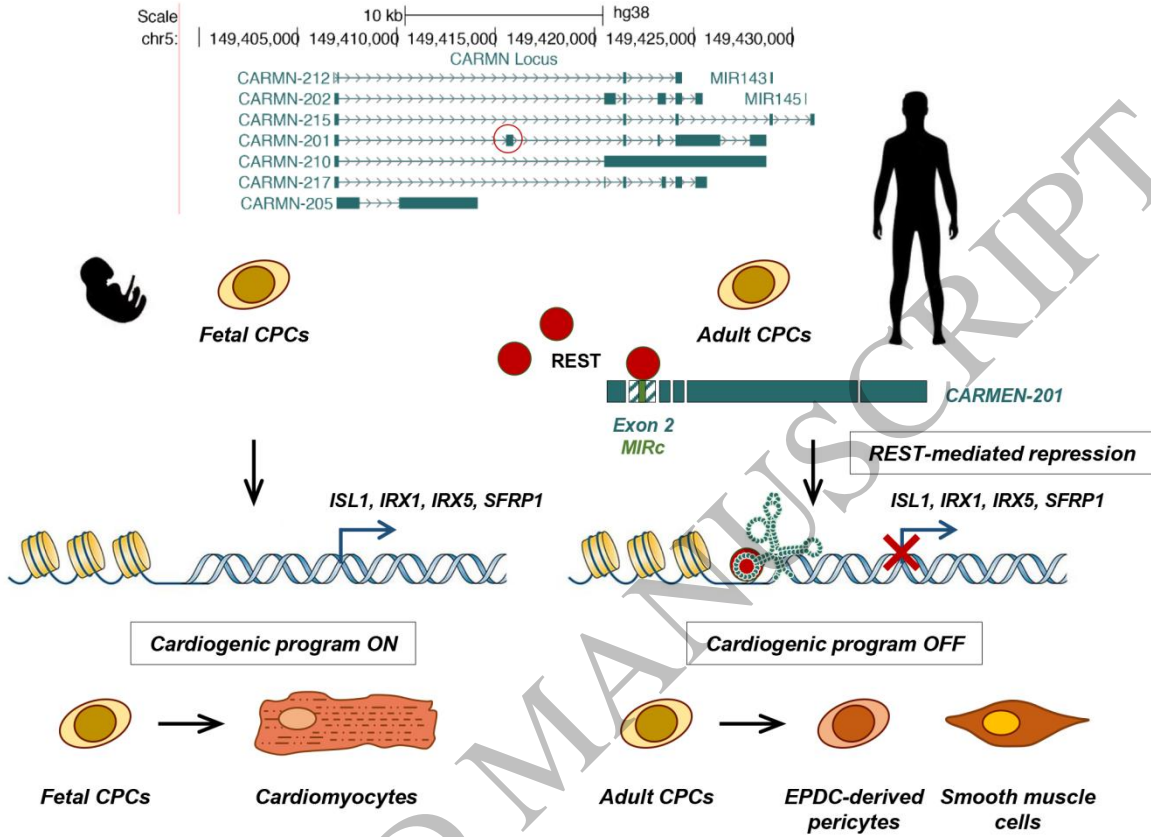


1
2
3

Figure 7
160x231 mm (x DPI)

Fig. 7

1



2

3

Graphical Abstract