Special series article

CARMEN, a human super enhancer-associated long noncoding RNA controlling cardiac specification, differentiation and homeostasis

Samir Ounzain a,⁎, Rudi Micheletti a, Carme Arnan b, Isabelle Plaisance a, Dario Cecchi b, Blanche Schroen c, Ferran Reverter b, Michael Alexanian a, Christine Gonzales a, Shi Yan Ng d,e, Giovanni Bussotti f, Iole Pezzuto a, Cedric Notredame f, Stéphane Heymans c, Roderic Guigó b, Rory Johnson b,⁎⁎, Thierry Pedrazzini a,⁎

a Experimental Cardiology Unit, Department of Medicine, University of Lausanne Medical School, Lausanne, Switzerland
b Bioinformatics and Genomics Group, Centre for Genomic Regulation, Barcelona, Spain
c Centre for Heart Failure Research, Cardiovascular Research Institute, Maastricht University, The Netherlands
d Bioinformatics and Genomics Group, Centre for Genomic Regulation, Barcelona, Spain
e NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore
f Comparative Bioinformatics Group, Centre for Genomic Regulation, Barcelona, Spain

⁎ Corresponding author at: Experimental Cardiology Unit, Department of Medicine, University of Lausanne Medical School, CH-1011 Lausanne, Switzerland.
E-mail addresses: samir.ounzain@chuv.ch (S. Ounzain), rory.johnson@crg.es (R. Johnson), thierry.pedrazzini@chuv.ch (T. Pedrazzini).

1. Introduction

The postnatal heart has minimal regenerative potential, and heart failure typically develops following injury [1]. Currently, heart transplantation remains the only viable therapeutic option for end-stage heart failure. Within this context, cell replacement therapy via injection of precursor cells into the damaged myocardium to promote cardiac regeneration and prevent pathological remodeling represents an attractive therapeutic approach [2]. However, the main challenge for utilizing cell therapies for heart disease in a clinical setting is the identification of a suitable source of human cardiac precursor cells (CPCs) that give rise to functionally integrated cardiomyocytes (CMs). The existence of resident CPCs in the adult mammalian heart, including the human heart, capable of differentiating into mature CMs has been demonstrated [2]. However, the number of CPCs in the adult heart is low and expansion of human CPCs in culture is a mandatory step to produce sufficient numbers of precursors for cell therapies. Furthermore, our relative inefficiency to induce robust differentiation of CPCs into functional CMs in vivo is a severe limitation to clinical application. Therefore, to promote cardiac regeneration, it is imperative to improve our understanding of the molecular pathways and regulatory circuitry that controls recruitment and differentiation of CPCs toward the cardiac lineage.

At the molecular level, cardiac specification and differentiation is a complex biological process dictated by the activity of gene regulatory networks (GRNs). Cardiac GRNs are under the control of the core transcription factors (TFs), including NKX2.5, GATA4, MESP1, MEF2C and TBOX proteins [3]. These factors interact in a combinatorial manner with chromatin remodeling complexes to target cis-regulatory sequences and elicit specific temporal and spatial gene expression...
programs. However, the notion that these networks are primarily protein-based regulatory systems has been somewhat premature [4]. Several recent studies have demonstrated that gene regulatory network activity is under the control of diverse noncoding RNAs (ncRNAs) [5,6]. These ncRNAs control numerous aspects of GRN activity including transcriptional control, post-transcriptional processing, and chromatin remodeling [6]. Currently, the best characterized ncRNAs in the heart are the microRNAs (miRNAs), which fine tune mRNA expression through post-transcriptional silencing [5]. However, the recent advent of unbiased high-throughput RNA-sequencing (RNA-Seq) has led to the discovery of the numerically greater and more diverse class of long noncoding RNAs (lncRNAs) [7,8]. LncRNAs are transcripts that are operationally defined as being larger than 200 nucleotides in size and that lack discernible protein coding potential. Latest estimates have more than 15,000 such transcripts in human gene collections (GENCODEv22), a number which continues to rise with accumulating data from cell-specific RNA-sequencing (RNA-Seq) [4,7]. Less than 1% of human lncRNAs have so far been characterized experimentally [9]. However, an increasingly significant number have been shown to play diverse regulatory roles in the nucleus and cytoplasm. These include chromosome X inactivation, genomic imprinting, splicing, transcriptional regulation and targeting of epigenetic chromatin modifying complexes to specific genomic loci [4].

LncRNAs were originally annotated based on their genomic location. However, functional annotations based on underlying chromatin states have recently emerged. A class of lncRNAs has been associated with active enhancer states (H3K27Ac/H3K4me1/p300), and these are classified as enhancer-associated lncRNAs [10,11]. Interestingly, these transcripts are able to modulate gene expression both in cis and trans, expanding the canonical roles of enhancer sequences. Enhancers are an important but enigmatic class of regulatory sequences, which are the key information processing units that integrate temporal, spatial and environmental cues within the genome [12]. Importantly, chromatin looping and promoter pause/release, both important processes that mediate enhancer-promoter communication are dependent on the production of enhancer-derived ncRNAs [13]. Recently, a novel class of enhancers has been described, the class of super enhancers (SE) [14,15]. SEs are much larger than classical enhancers (typically 1–10 kbp in size) and are master regulators of cell identity genes, including developmental TFs and other components determining cell-specific biology. Their function is associated with increased production of enhancer-associated ncRNAs as well as enrichment of specific histone modifications and chromatin remodeling complexes [14,15]. Not surprisingly, a number of important lineage-determining lncRNAs are derived from tissue-specific SEs [16,17].

Within this context, lncRNAs are emerging as an important class of regulatory mediators of cardiac lineage-specific commitment during development and of specialized cellular functions involved in maintaining cardiac integrity [18–24]. Two recently described lncRNAs, Braveheart (Bvht) and Fendrr, were shown to be key regulators of cardiac mesoderm specification and subsequent differentiation via their ability to epigenetically modulate the expression of important cardiac transcription factors [18,19]. Recent studies have also identified thousands of novel human and mouse heart enriched lncRNAs modulated in the adult heart post injury, implicating these in the pathological response [21–23,25], and highlighting their potential utility as biomarkers [21,22,24,26]. Moreover, considering the unique characteristics of SEs, cardiac lncRNAs associated with SE sequences are likely to be functionally important and relevant modulators of cardiac specification and differentiation. However, the roles of lncRNAs, and specifically SE-associated lncRNAs, in human CPCs have not been investigated. Therefore, there is a clear need to identify cardiogenic lncRNAs important in human development and CPC differentiation. In this study, we have identified and characterized CARMEN, a SE-associated lncRNA important for cardiac specification and differentiation in human CPCs.

2. Results

2.1. Profiling of the transcriptome in differentiating CPCs identifies differentially expressed lncRNAs

To address the role of lncRNAs in human CPC specification, differentiation and maturation, we first identified lncRNAs that were differentially expressed in human CPCs upon differentiation. CPCs were isolated from the fetal human heart as previously described [27]. These cells expressed early cardiac marker such as NKX2.5 and GATA4 but not proteins expressed by fully differentiated CMs, smooth muscle cells and endothelial cells (Suppl. Fig. 1A). CPCs were then induced to differentiate into CMs in vitro using an established method. Cardiac differentiation was associated with increased expression of mature CM marker genes including MYH6, CAV1.2 and RYR2 (Fig. 1A). This protocol typically resulted in efficient cardiac differentiation, with up to 16% of NKX2.5- and α-actinin-positive CMs presenting organized sarcomeres following seven days of differentiation (Fig. 1B). We therefore extracted RNA from undifferentiated (proliferating) and differentiated cell cultures and hybridized RNA samples to custom microarrays targeting essentially every transcript in the recently published GENCODEv7 lncRNA catalogue [7]. The majority of the lncRNAs that are in the GENCODEv7 catalogue are spliced, polyadenylated and primarily annotated based on ESTs and cDNAs from oligo-dT reverse transcribed cDNA libraries [7]. The array comprised four distinct, strand-specific probes against 14,880 transcripts, in addition to a set of 3231 randomly selected probes targeting human mRNAs.

We used a linear model approach to identify differentially expressed transcripts between proliferating and differentiated CPCs, taking advantage of the limma package for differential gene expression analysis [28]. This identified 570 (266 down- and 304 upregulated) differentially expressed lncRNAs during CPC differentiation at a false discovery rate (FDR) of <0.05 (Fig. 1C, Suppl. Table S2). Unsupervised hierarchical clustering of lncRNA and mRNA expression segregates the undifferentiated CPCs from differentiated CPCs (Fig. 1D), confirming that the differentiation process in vitro led to a robust change in the cardiac noncoding transcriptome. Consistent with previous reports in other tissues [7], lncRNAs in CPCs were detected at lower levels than mRNA (Suppl. Fig. 1B) and displayed a comparable lack of protein coding potential when compared to GENCODE mRNAs and lncRNAs (Suppl. Fig. 1C). Differentially expressed lncRNAs encompassed the various subtypes of lncRNAs (Fig. 1E) based on genomic geography with respect to proximal coding genes (i.e., intragenic vs. intergenic; exonic vs. intronic; sense vs. antisense; overlapping vs. non-overlapping). There were no significant differences in the conservation status of modulated lncRNAs with most being either highly or poorly conserved across 21 mammalian species (Suppl. Fig. 1D). However, as previously observed lncRNA exons were significantly less conserved than coding exons from mRNAs, although promoter sequences were equally well conserved between these two classes of transcripts (Suppl. Fig. 1E). In order to determine the expression patterns and tissue specificity of the modulated lncRNAs, we examined their expression in a wide variety of human tissues using available custom GENCODE lncRNA array datasets [7]. Tissue-specific enrichment profiles were obtained for all modulated lncRNAs across tissues. Globally, their expression levels and breadth of expression across tissues was comparable to all GENCODE lncRNAs (Suppl. Fig. 1F). Interestingly, many upregulated lncRNAs were more expressed in adult heart tissue as compared with downregulated and all GENCODE lncRNAs. This supported the notion that upregulated lncRNAs could promote a cardiac differentiated state, which was comparable to that seen in the adult heart.

2.2. Identification of high priority candidate lncRNAs

We manually selected eleven lncRNAs as high priority candidates based on specific criteria, including the presence of an active
enhancer-associated epigenetic signature (p300 enrichment) in fetal or adult human hearts, proximity to important coding and noncoding genes, significant modulation during CPC differentiation (log2FC > 0.5, \( < - 1.0 \)) (Fig. 2A) and classification subtype (Suppl. Fig. 2A). First of all, we utilized a previously executed genome-wide ChIP-Seq screen [29], and determined the occupancy profiles of enhancer-associated co-activator proteins at candidate lncRNA loci in fetal and adult human hearts. A number of candidates were associated with significant p300/ CBP enrichment, and specifically within the fetal heart (Lnc-869, Lnc-866, Lnc-518, Suppl. Fig. 2B), consistent with their expression in fetal CPCs. These data also suggested that some of these lncRNAs were representative of the previously described enhancer-associated lncRNAs [10, 11]. Furthermore, using previously published data sets, we assessed the nuclear and cytoplasmic enrichment profiles of high priority candidates in ENCODE validated subcellular fractions (Fig. 2B) [7]. The majority of our candidates exhibited nuclear enrichment (e.g. Lnc-866). Interestingly, several lncRNAs displayed different subcellular localization depending on the cell type in which they were expressed (e.g. S-KANK). Finally, we validated differential expression of all high priority candidates in fetal CPCs one-week post differentiation using quantitative RT-PCR (qRT-PCR) (Fig. 2C). Of note, three candidate lncRNAs that were proximal to or overlapping coding genes (i.e. S-KANK1, AS-TNIK, AS-LHX1) exhibited highly correlated expression patterns with their corresponding coding genes, indicative of potential cis-regulation (Fig. 2D).

2.3. CARMEN is a human super enhancer-associated lncRNA associated with cardiac differentiation

Amongst the modulated lncRNAs, three transcripts, namely ENST00000519898, ENST00000518014 and ENST00000509909, represent distinct annotated isoforms of a single lncRNA gene (Fig. 3A and B). Importantly, CARMEN isoforms were predicted to have no apparent coding potential using a variety of computational tools for (Suppl. Fig. 3A). In particular, we tested CARMEN using PhyloCSF (Suppl. Fig. 3A) [30]. The resulting score of \( -44.5 \) (using default settings), below the accepted threshold of 28.34 for coding transcripts, indicates no evidence for translation of any possible ORF. To further experimentally validate this observation, we searched two experimental datasets investigating the human proteome, neither of which record a peptide originating from CARMEN [31,32]. This locus was of particular interest for several reasons. Firstly, it was directly upstream of MIR-143 and -145, two important developmental miRNAs implicated in CPC specification and differentiation (Fig. 3A). Specifically, these two miRNAs were previously shown to promote differentiation toward the smooth muscle cell lineage [33,34]. Secondly, ChIP-analysis demonstrates that these transcripts mapped to a highly active cardiac enhancer in both fetal and adult human hearts (Fig. 3B). Notably, this region was shown to harbor a cardiac SRF/NKX2.5-bound enhancer required for cardiac-specific expression of miR-143 and -145 during cardiac development [34]. Furthermore, this enhancer has been previously demonstrated to be a Notch-responsive enhancer, capable of modulating miR-143 and -145 expressions in various developmental contexts [35]. Accordingly, the enhancer was able to drive robust cardiac-specific activity in a mouse transgenic reporter assays at E11.5 p.c. (Fig. 3C; images available using enhancer ID: Hs1752 at the VISTA enhancer browser, http://enhancer.lbl.gov/). Finally, using super enhancer (SE) tracks generated from data obtained in the adult human left ventricle [14], we demonstrated that this locus encompassed a cardiac SE. These findings indicated that this genomic locus and its associated lncRNAs were potentially important regulators of cardiac identity genes and pathways. We named this gene CARMEN for (CARDiac (M)esoderm (E)nhancer-associated (N)oncoding RNA, and its individual isoforms CARMEN1, CARMEN2 and CARMEN3 (Fig. 3B). We proceeded to quantify the absolute expression of CARMEN isoforms using primers designed to distinguish individual transcripts (Suppl. Fig. 3B). All three isoforms were expressed at relatively high levels and significantly induced in differentiating fetal CPCs (Fig. 3D), supporting a role during cardiac specification and differentiation for these SE-associated lncRNAs. Interestingly, MIR-143 and -145 were modulated in an opposing manner during differentiation (Fig. 3E).

Evolutionary conservation supports the importance of CARMEN in cardiac precursor cell differentiation. A long noncoding RNA, (AK087736), is transcribed from the orthologous region of the mouse genome, which we designate as the mouse ortholog (Fig. 4A). Interestingly, the promoter region is highly conserved, supporting an important role in the regulation of transcription at this locus. In the mouse, another annotated convergent lncRNA is present on the opposite strand. This lncRNA, AK143260, has recently been demonstrated to be a cardiac lncRNA named Braveheart (Bvht), critically required for progression of nascent mesoderm toward the cardiac fate during cardiogenesis [18]. It is important to note that no ortholog of Bvht has been identified in human.

We proceeded to quantify the expression of both Carmen and Bvht during cardiogenesis in two independent murine models of cardiac specification and differentiation in vitro. We first induced mouse embryonic stem (ES) cells to differentiate using the hanging drop model [36]. This model recapitulates embryonic cardiac development in vitro, generating all appropriate cardiac lineages. We first examined the temporal gene expression patterns associated with pluripotency (ES), cardiac mesoderm (MES), cardiac precursors (CPC) and differentiating cardiomyocytes (CM) (Suppl. Fig. 4A). Upon differentiation, the pluripotency markers Oct4 and Nanog were rapidly downregulated (Suppl. Fig. 4B). This occurred concomitantly with the transient induction of cardiac mesoderm specifying TFs, Mesp1 and Eomes (Fig. 4B; Suppl. Fig. 4B). Three core cardiac TFs, Nkx2.5, Gata4 and Isl1, which specify CPCs and initiate the cardiac gene program, were significantly upregulated at the CPC stage (Fig. 4B; Suppl. Fig. 4B). Finally, this was followed by the robust expression of cardiac differentiation and structural markers, Myh6, Myh7 and Myl2 (Suppl. Fig. 4B). We then assessed Carmen and Bvht expression. Carmen was induced between the MES and CPC stages with maximal expression occurring in CPCs (Fig. 4C). Bvht was initially downregulated between ES and MES stages, before subsequently being upregulated in CPCs and CMs (Fig. 4C), in agreement with previously published observations [18]. Altogether, these different kinetics of expression supported a role for Carmen during cardiac specification and differentiation of mouse ES cells. Importantly, Carmen expression in the adult mouse heart was comparable to expression in differentiated ES cell-derived CMs (Suppl. Fig. 4D).

We sought further support for these findings in another model of cardiac differentiation using P19CL6 cells. These can readily be induced to differentiate into beating cardiomyocytes upon induction by dimethyl sulfoxide (DMSO) [37]. Indeed, differentiating P19CL6 cells recapitulated cardiac differentiation, and demonstrated the expected change in gene expression in pluripotency (Oct4, Nanog), MES (Mesp1, Eomes), CPC (Gata4, Isl1) and CM (Myh6, Myl2) marker genes (Fig. 5). Mesodermal specification occurred specifically between day 0 and day 2 while CPCs and CMs emerged between day 2 and day 4. Comparable to their expression kinetics during ES differentiation, both Carmen and Bvht were maximally expressed between the MES and CPC stages.
Considering that human CARMEN maps to a cardiac super enhancer, we determined stage-specific activation of chromatin states at the locus in the mouse. We took advantage of publicly available chromatin state maps, generated using chromatin immunoprecipitation followed by sequencing (ChIP-Seq) in differentiating ES cells [38]. Analyses were executed in pluripotent mouse ES cells (ES, i.e. Oct4-positive cells), at the cardiac mesoderm stage (MES, i.e. Mesp1-positive cells), the cardiac precursor stage (CPC, i.e. Nkx2.5-positive cells) and in differentiated cardiomyocytes (CM, i.e. Myl2-positive cells). Assessment of H3K27me3 and H3K4me3 (associated with inactive and active canonical promoters respectively), and H3K4me1 and H3K27Ac (associated with poised and active enhancers respectively) allowed us to analyze chromatin state transitions at the CARMEN locus during cardiogenic differentiation (Suppl. Fig. 4C). First of all, Carmen was associated with a super enhancer-like signature as judged by the broad and high H3K27Ac occupancy levels during cardiac differentiation. Maximal enhancer activity, as indicated by the peak of H3K27Ac enrichment, occurred between the MES and CPC stages. Carmen expression therefore correlated with enhancer activity, as recently observed for other fetal cardiac enhancers active during cardiac differentiation [11]. Interestingly, Bvht was associated with a canonical promoter signature at its transcriptional start site (i.e. H3K4me3), which was absent at the Carmen locus. This suggests that Bvht should be classed as a promoter-associated intergenic lncRNA. Altogether these data strongly support Carmen being a bona fide SE-associated lncRNA, whose expression correlates with both the chromatin state activation of its associated enhancer in CPCs as well as with specification and subsequent differentiation of CPCs into cardiomyocytes.

2.4. CARMEN controls cardiac specification and differentiation in mouse CPCs

To determine if Carmen actively regulated cardiac specification and differentiation, we stably transected P19CL6 cells with a shRNAi targeting Carmen. Cardiogenesis was induced in transected cells by DMSO addition. Four days following induction, loss-of-function resulted in a complete block of Carmen expression (Fig. 5). Interestingly, Bvht induction was also completely blocked upon Carmen silencing, suggesting that Bvht was controlled by Carmen. In accordance with the known roles of Bvht in regulating the cardiogenic program, the expression of Mespl, the downstream cardiac TFs Gata4 and Isl1, and the CM-specific genes Myl2, Myh6 were severely inhibited by Carmen loss-of-function. In addition, Eomes, a key TF upstream of Mesp1 in the cardiac mesoderm specification pathway and the pluripotency regulators Oct4 and Nanog were also affected by Carmen depletion. Previous work demonstrated that Eomes and the pluripotency genes were not affected by Bvht knockdown. These data therefore suggest the existence of feedback mechanisms able to affect pluripotency (Oct4, Nanog) and early mesodermal specification (Eomes) that are dependent on Carmen expression.

To extend these findings in human cells, fetal CPCs were transected with modified antisense oligonucleotides (GapmeRs) targeting the three human CARMEN isoforms. After 72 h, transected proliferating CPCs were cultured under differentiation conditions. Efficient depletion was observed in GapmeR-treated cells 10 days post induction of cardiac differentiation (Fig. 6A). This depletion was associated with a significant reduced expression of cardiac transcription factors and differentiation makers, including GATA4, NKX2.5, TBX5, MYH6, MYH7, and TNNI (Fig. 6B). Furthermore CARMEN depletion was associated with decreased expression of the smooth muscle cell marker, MYH11. Importantly, Carmen depletion was not associated with decreased expression of mir-143 and -145, further supporting the notion that CARMEN isoforms are not the precursor transcripts for these miRNAs (Fig. 6C). As a consequence, Carmen silencing impaired the capacity of human CPCs to produce differentiated cardiomyocytes (Fig. 6D). To further dissect the roles of CARMEN and the proximal miRNAs in this differentiation process, Carmen-depleted cells were also co-transfected with MIR-143 and MIR-145 mimics. This led to increased expression of both miRNAs in control and Carmen-depleted cells. The predicted downregulation of the prototypic miRNA targets ELK1 and KLF4 demonstrated efficient MIR mimic transfection (Fig. 6C). Importantly, and confirming a miRNA-independent role for CARMEN in cardiogenic differentiation, MIR mimics were not able to restore cardiogenic differentiation induced by Carmen knockdown. In contrast, expression of the smooth muscle cell marker MYH11 was induced by MIR-143 and MIR-145 mimics, in agreement with the known roles of the two miRNAs as promoters of smooth muscle cell differentiation. Altogether, this indicates that Carmen is able to control cardiac specification and differentiation independently of MIR-143/145.

2.5. CARMEN exhibits cis-acting function, and interacts with SUZ12 and EZH2

Based on the genomic overlap with bona fide cis-acting enhancer sequences and the lack of a canonical promoter signature at the Carmen transcriptional start site, we suspected that Carmen enacted regulatory activity akin to other recently described enhancer associated lncRNAs [16–19]. In order to test this hypothesis, we constructed a plasmid, in which Carmen expression was under the control of a Doxycycline (Dox)-inducible promoter (Fig. 7A). This promoter is completely inactive in the absence of Dox. A firefly luciferase reporter gene under the control of a basal SV40 promoter was then inserted directly downstream of the Carmen expression cassette. A control vector was also constructed containing a GFP sequence instead of Carmen. The presence of Carmen DNA in the construct resulted in significant luciferase activity, even in the absence of Carmen transcription (Dox neg) indicating the intrinsic enhancer activity of the locus. Crucially, induction of Carmen transcription by addition of Doxycycline caused additional activation of luciferase activity. Thus, RNA from the Carmen enhancer contributed to stimulate luciferase expression (Fig. 7B). These results support the idea that cis-enhancing activity of this enhancer locus is dependent on the production of an associated lncRNA, i.e. Carmen. Importantly, a number of recent reports have indicated that enhancer cis-acting lncRNAs can also enact trans functions through interaction with ubiquitously expressed chromatin-modifying complexes [19]. In particular, many lncRNAs have been shown to associate with le polycomb repressive complex 2 (PRC2). We therefore investigated if Carmen could physically interact with components of this complex in mouse embryonic stem cells. We performed RNA immunoprecipitation (RIP) assays, in which RNA-protein complexes were immunoprecipitated with antibodies specific to EZH2 and SUZ12 (Fig. 7C and D). We found that Carmen was associated with both proteins, suggesting that Carmen may have trans-repressive roles during cardiac differentiation.

2.6. CARMEN is modulated in cardiac disease and required for maintaining a differentiated cardiac fate

Our data indicated that Carmen was required for the specification and cardiac differentiation of pluripotent and cardiac precursor cells. Furthermore, utilizing data from previous RNA-Seq based studies demonstrated that this transcript was highly expressed in the adult mouse heart.
and human heart (not shown) [22]. We therefore investigated whether CARMEN expression was modulated in response to cardiac stress in murine models of heart disease and in the heart of human patients suffering from two different cardiac pathologies. We first determined expression in a mouse model of myocardial infarction. Fourteen days after infarction, the myocardium was characterized by intense remodeling and decreased cardiac function (Suppl. Fig. 5A), and induction of cardiac markers of stress (Fig. 8A). Carmen was significantly upregulated post myocardial infarction (Fig. 8A), supporting a role for Carmen in adult heart homeostasis and cardiac remodeling.

To determine the potential roles of human CARMEN in cardiac disease, we examined its expression in two human heart pathologies, i.e. idiopathic dilated cardiomyopathy (DCM) and aortic stenosis (AOS). DCM patients presented reduced cardiac function whereas AOS patients demonstrated preserved function. In addition, cardiac dilation was evident in DCM patients and AOS patients were characterized by cardiac hypertrophy (Suppl. Fig. 5B and C; Suppl. Table S3). The stereotypical cardiac stress marker NPPA was also significantly upregulated in both pathologies (Fig. 8B and C). We then assessed the expression of all three human CARMEN isoforms. CARMEN3 was significantly upregulated in both DCM and AOS whereas CARMEN1 and 2 were only induced in AOS patients (Fig. 8B and C). We further searched for independent validation of the relationship between CARMEN3 expression and cardiac disease state. The Gtex project has produced RNA-Seq transcriptome data from post mortem organs across a panel of human individuals with medical history [39]. We examined therefore two relevant heart tissues within Gtex, namely atrium and ventricle, for which 25 and 66 individual samples were respectively available. Of these two, CARMEN3 expression was markedly higher in atria (P = 1.9e−09, Wilcoxon test) (Fig. 8D). Next, we used medical data to separate patients based on a diagnosis of cardiovascular disease. Interestingly, hypertensive patients demonstrated increased expression in atria (P = 0.008, Wilcoxon test).
these processes can facilitate the targeted derivation of important cardiovascular lineages required for cell replacement therapies. Cardiovascular lineage specification and differentiation is a multistep process regulated by a network of transcription factors [3]. Though many of the genetic factors that control cardiac development and CPC lineage determination and differentiation are known, we suspect that IncRNAs represent an additional layer of regulation [40]. Here, we identify hundreds of differentially modulated human IncRNAs during cardiac differentiation of isolated human fetal CPCs. Interestingly, many of these validated candidates are mapping to bona fide human fetal and adult cardiac enhancers, suggesting they represent the newly identified enhancer-associated IncRNA class [10,11]. We have recently demonstrated that a significant fraction of novel murine cardiac polyadenylated and multi-exonic IncRNAs are derived from developmental cardiac enhancers [11,22]. The human IncRNAs described here are comparable in their chromatin state and likely cis-acting function. In support of this, we find that several human IncRNA candidates are highly correlated in their expression with proximal coding genes. We further characterized one of these candidate IncRNAs, which we named CARMEN. CARMEN is highly conserved in mouse, and appears to be an important regulator of cardiovascular cell specification and differentiation. Furthermore, CARMEN is derived from a human SE, which has been shown to be active in the adult heart. Considering the master regulatory roles of SEs in controlling key cell identity genes, we suspect that CARMEN represents a novel SE-associated IncRNA critical for cell fate determination and differentiation, akin to other recently described SE-derived IncRNAs such as MyoD enhancer RNA and CCATL-1 [16,17].

Interestingly, CARMEN is proximal to an important developmental miRNA locus, encoding the MIR-143 and –145. These miRNAs have been previously shown to be crucial post-transcriptional regulators of key pathways implicated in pluripotency and cardiovascular lineage commitment [33,34]. It is important to emphasize that the CARMEN isoforms assessed here do not represent precursor transcripts for MIR-143/-145. Indeed, previous analysis of promoter-specific histone modifications and poly(A) signal frequencies demonstrate that CARMEN transcript boundaries are well defined [7]. The majority of CDNA and EST fragments overlapping CARMEN terminate at the position annotated by GENCODE, i.e. upstream of miR-143. In addition, a non-canonical polyadenylation sequence is located within 20 nucleotides of the termination site. The end of CARMEN is also marked by an SVM-predicted polyadenylation site [41]. Furthermore, MIR-143 and -145 expression does not correlate with CARMEN expression during cardiac differentiation, and CARMEN depletion does not lead to reduced miRNA expression. Most importantly, miRNA overexpression in CARMEN-depleted fetal CPCs is not able to rescue the differentiation defect induced by CARMEN knockdown, suggesting that CARMEN-mediated induction of cardiogenesis is independent of the proximal miRNAs. Altogether, these different pieces of evidence refute that CARMEN represents a precursor of the two miRNAs.

Mouse ES cells with deletions of both miR-143 and miR-145 have been shown in previous studies not to phenocopy Bvht-depleted cells [18], supporting a miRNA-independent role for Bvht. In the mouse, Bvht is necessary for the activation of a core cardiac gene network and functions upstream of Mesp1 [18]. It regulates the cardiac gene network during CM differentiation through interacting with SUZ12, a component of PRC2, suggesting that Bvht mediates epigenetic regulation of cardiac commitment. Bvht is also important for maintaining a cardiac identity in neonatal CMs, as shown for CARMEN. Our luciferase reporter assay provides support that the cis-acting regulatory activity of the enhancer locus is potentiated by the active transcription of CARMEN, a finding comparable with recently published findings for other IncRNAs [42]. We also find that CARMEN directly interacts with PRC2 through both
A

![Graphs showing relative expression of CARMEN 1, 2, and 3 under Scrambled and CARMEN GapmeR conditions with MIR-143 and MIR-146 mimics.](image)

B

![Graphs showing relative expression of GATA4, NKK 2.5, and TBX5 under Scrambled and CARMEN GapmeR conditions with MIR-143 and MIR-146 mimics.](image)

C

![Graphs showing relative expression of MYH6, MYH7, TNNI, and MYH11 under Scrambled and CARMEN GapmeR conditions with MIR-143 and MIR-146 mimics.](image)

D

![Images showing α-Actinin-positive cells under Scrambled and CARMEN GapmeR conditions.](image)
EZH2 and SUZ12. How Carmen targets PRC2 within the genome either in cis or trans is, however, currently not known. Potentially, Carmen may, like other cardiogenic lncRNAs such as Bvht and Fendrr, either compete for PRC2 binding at specific loci during cardiac differentiation or actively target PRC2 [18,19]. Our work suggests therefore that Carmen mediates cardiac specification and differentiation, in part through cis- and trans-dependent epigenetic regulation, which appears to represent an emerging paradigm for lncRNAs.

The loss-of-function data suggest that Carmen is directly implicated in the earliest steps of lineage commitment, able to control the expression of both pluripotency and cardiac mesodermal specifying genes. Since the human CARMEN locus does not contain an ortholog of mouse Bvht, it suggests that Bvht is dispensable for cardiac specification in human CPCs. The loss of Bvht in the human lineage represents an interesting evolutionary event in species-specific heart developmental pathways, which requires further investigation. In contrast, the fact that CARMEN is conserved and expressed in human supports an evolutionary conserved role in this process. Our work suggests that CARMEN functions upstream of and in the same regulatory pathway as EOMES and MESP1 in a miRNA-independent fashion, as a reprogramming factor necessary for cardiac commitment and differentiation of CPCs. Importantly, EOMES and MESP1 can specify all cell types of the cardiovascular lineage [43,44]. MESP1 is able to induce transdifferentiation of dermal fibroblasts into CPCs [45]. This indicates that CARMEN could also be a critical factor for inducing somatic cell reprogramming toward a cardiovascular state. A number of recent studies have demonstrated that specific combinations of key cardiac TFs and miRNAs are able to reprogram cardiac fibroblasts into mature CMs in vitro and in vivo [46–49]. This represents a very attractive therapeutic strategy for the promotion of cardiac regeneration. Our data suggest that CARMEN, in addition to other lncRNAs identified in the present study, could represent interesting cardiogenic reprogramming factors.

Pathological cardiac remodeling and the subsequent transition to heart failure are typically associated with the re-activation of a fetal gene program [1]. We have previously demonstrated that many developmental cardiac enhancer-associated lncRNAs are transcriptionally induced in the remodeling and failing heart [11,22]. Interestingly, Carmen is induced in mouse models of cardiac disease, and human CARMEN isoforms are induced in AOS and DCM patients. Importantly, Carmen-depleted cardiomyocytes demonstrate reduced expression of cell identity genes. Carmen expression appears therefore critical for maintaining a differentiated state in cardiomyocytes. An increasing number of studies are beginning to illustrate the roles of lncRNAs in the adult remodeling heart [20–23], supporting the notion that cardiac lncRNAs, including CARMEN, could represent unique targets for modulating the cardiac pathological response, and possibly promoting regeneration [40]. Indeed, it is conceivable that increased CARMEN expression in the stressed heart reflects upregulation in a subset of cardiac cells other than cardiomyocytes, in particular CPCs. If proved to be true, CARMEN could be involved in initiating a cardiogenic program in CPCs that are mobilized in the adult heart upon damage. In a broader context, the manipulation of super enhancer-associated lncRNAs represents a new avenue for controlling the gene regulatory networks that neocardio genesis in the adult heart. Carmen could potentially represent therefore an attractive therapeutic target for future regenerative and cell-based therapies.

4. Methods

Detailed methods can be found in Supplementary information.

4.1. Human p300 and super enhancer ChIP-Seq data

For ChIP-Seq analysis of human fetal and adult hearts, we utilized previously published data sets [29,30]. Super enhancer annotations in the adult left ventricle were obtained from [14].

4.2. Cardiac injury models

Ligation of the left anterior descending artery — Myocardial infarction in mice was induced as previously described [22].

4.3. Echocardiography

Transthoracic echocardiographies were performed using a 30-MHz probe and the Vevo 770 Ultrasound machine (VisualSonics, Toronto, ON, Canada).

4.4. Primary cell cultures

Human fetal heart chambers and cardiac progenitor cells were isolated as previously described [27].

4.5. Immunohistochemistry on fetal CPCs

Please refer to Supplementary material online.

4.6. RNA isolation, reverse transcription, end-point PCR and quantitative PCR

Primer sequences for qRT-PCR are provided in Suppl. Table S1. For TaqMan probe based qRT-PCR expression was analyzed using fluorescent-labeled TaqMan Probes (ABI), which are described in Suppl. Table S1. Analysis was carried out using an ABI Prism 7500 cycler and relative expression quantified using the ΔΔCt method. For endpoint PCR aliquots of PCR mixtures were taken during different cycles for agarose gel analysis to determine linear range of amplification. All reactions were run on a 1.5% agarose gel stained with Ethidium Bromide.

4.7. Cell culture and transfection

Please refer to Supplemental material online.

4.8. Differential expression analysis of lncRNAs

Total RNA (100 ng) was labeled using Low Input Quick Amp Labeling kit (Agilent 5190-2305) following manufacturer instructions. mRNA was reverse transcribed in the presence of T7-oligo-dT primer to produce cDNA. cDNA was then in vitro transcribed with T7 RNA polymerase in the presence of Cy3-CTP to produce labeled cRNA. The labeled cRNA was hybridized to the Agilent SurePrint G3 gene expression 8x60K microarray according to the manufacturer’s protocol (Agilent SurePrint 8 × 60 nt technology). The arrays were washed, and scanned on an Agilent G2565CA microarray scanner at 100% PMT and 3um resolution. Intensity data was extracted using the Feature Extraction software (Agilent). Differentially expressed microarray probes were identified using the limma package in R comparing microarray data from 3 proliferating and 3 differentiating cell cultures. Data sets have been deposited in the Gene Expression Omnibus Database under accession number (pending).

4.9. LncRNA analysis

Phastcons data: Base-level mean PhastCons scores were calculated across the indicated genomic features [7].

4.10. Subcellular localization analysis

Nuclear, Cytoplasmic and Chromatin RNASeq data from human cell lines were obtained from ENCODE [51]. Localization of transcripts was estimated by the ratio of RPKM (reads per kilobase per million mapped reads) values, with the condition that both values must have a non-zero
4.11. Protein-coding potential

We used a series of widely used tools to evaluate protein-coding potential of CARMEN isoforms (ENST00000505254, ENST00000518014) using default settings, with the indicated cutoffs. CNCI [52]: >0.43, CPAT (PMID:23335781): >0.54, CPC [53]: >0.71, iSeeRNA [54]: >0.78, PhyloCSF [30]: >28.34 and RNACode [55]: >37.55. All methods indicate non-protein coding status for both transcript isoforms.

4.12. Cross-species screening pipeline

Please refer to Supplemental material online.

4.13. CARMEN expression in Gtex samples

Tissue RNAseq quantifications were obtained from the Broad Institute in the form of gene-level RPKM [39].


To analyze the presence of chromatin marker peaks at promoters during cardiac differentiation of mouse embryonic stem cells, we used publicly available data published by Wamstad et al. [38].

4.15. Embryonic stem cell culture and differentiation

Mouse embryonic stem (ES) cells were differentiated into CMs as previously described [36].

4.16. P19CL6 cell culture and transfection

P19CL6 cells (RCB2318, RIKEN Cell Bank, Japan) were cultured in DMEM with 10% FCS and antibiotics. Transfection of P19CL6 cells with pLKO.1-puro-UbC-Tag635™ (containing shRNAi, Sigma Aldrich) was performed with Lipofectamine 2000 (Invitrogen) according to manufacturer’s instructions. P19CL6 was a kind gift of Dr. Elizabeth Robertson, University of Oxford, UK.
4.17. Methods related to data obtained in humans

All human material was obtained during routine sampling used for clinical purposes or after obtaining informed consent, stored in a coded way and available for research purposes in accordance with the Declaration of Helsinki and the ethical committee at Maastricht University Medical Center. For further details please refer to Supplemental material online.

4.18. CARMEN LncRNA cloning

Human CARMEN1 was amplified from the FBI-1 cDNA template using primers C222F (TACCGAGCTCGATCCGACGCTCGACGCGC) and C222R (CTGGACTAGTGGATCCaggcacagtgttagagtttgct) with Expand polymerase (Roche) (PCR conditions: annealing 60 °C, extension 3 min, 1.5 mM Mg, 40 cycles amplification) and cloned into BamHI digested plasmid pcDNA as described [56]. Human CARMEN1 insert was subcloned into inducible plasmid pTretight-SV40-luciferase (pTretight vector from Clontech in which we added a SV40-luciferase ORF). All the clones were checked by Sanger sequencing. (*pTre-tight vector was gift from Susanna de la Luna lab).

4.19. Luciferase assay

We performed luciferase assay using dual-luciferase reporter kit (Promega) following standard protocols.

4.20. RNA immunoprecipitation (RIP)

Cells were detached with Accutase (Millipore), crosslinked in 1% formaldehyde for 15 min and quenched with 2.5 M glycine for 5 min. The cell pellet was resuspended in modified RIPA buffer (150 mM NaCl, 50 mM Tris, 0.5% sodium deoxycholate, 0.1% SDS, 1% NP-40) supplemented with RNase inhibitor Superase.In (Ambion) and Complete protease inhibitor (Roche). The cell suspension was briefly sonicated at low amplitude for 5 × 30 s cycles using a Bioruptor sonicator to lyse nuclei. Cell debris was removed by centrifugation at 4 °C, precleared with Protein G dynal beads (Invitrogen) before adding the respective antibodies pre-bound with Protein G dynal beads for 3 h at RT. 25 µg of rabbit anti-SUZ12 antibody (Abcam ab12073) or control rabbit IgG antibody was used for each RIP. Beads were then washed 3 times in modified RIP buffer, and twice in high salt RIPA buffer (1 M NaCl, 50 mM Tris, 0.5% sodium deoxycholate, 0.1% SDS, 1% NP-40). Crosslinks were reversed and proteins were digested with Proteinase K (Invitrogen) at 65 °C for 2 h. RNA was extracted in Trizol (Invitrogen) following the manufacturer’s instructions. RNA was then reverse transcribed using High Capacity cDNA kit (Ambion) and used for quantitative PCR (RIP-qPCR).

4.21. Statistical analysis

Data throughout the paper are expressed as mean ± SEM. One way ANOVA was used to test significance of data comparisons between experimental groups, with P values < 0.05 were considered significant.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.yjmcc.2015.09.016.

Sources of funding

This work is in part funded by a grant from the Swiss National Science Foundation within the frame of the National Research Program 63 on “Stem cells and Regenerative Medicine” (TP; grant no. 406340-128129).

Author contribution

SO, RJ and TP designed the research; SO, RJ, RM, CA, IP, DC, BS, FR, MA, CG, SYN, GB and IP performed the experiments; SO, GB, CN, SH, RG, RJ and TP performed the data analysis; SO and RJ wrote the paper; and SO, RJ and TP edited the manuscript.

Disclosure statement

None to disclose.

Acknowledgments

We thank Benoit Bruneau, Gladstone Institute, San Francisco, CA, for providing access to custom ChIP-Seq tracks. In the process of this analysis, we took advantage of CvDC data produced as part of the Bench-to-Bassinet program. We are grateful to the personnel of the Cardiovascular Assessment Facility, University of Lausanne, Lausanne, Switzerland, for microsurgery and evaluation of cardiac function in the mouse.

References


Fig. 8. Dysregulation of CARMEN in cardiac pathology. (A) qRT-PCR analysis of Carmen and classical stress marker genes, Nppa, Nppb, Myh6, Myh7, Tgfb1 and Col1a, in sham-operated and MI hearts. (B) qRT-PCR analysis of CARMEN1, 2 and 3, and NPPA in cardiac biopsies from patients suffering from DCM. (C) qRT-PCR analysis of CARMEN1, 2 and 3, and NPPA in cardiac biopsies from patients suffering from AOS. (D) Gtex-derived RNA-Seq-based CARMEN expression in human atrium and ventricle. (E) Gtex-derived RNA-Seq-based CARMEN expression in atria from patients suffering from hypertension. (F) Mouse neonatal CMs were transfected with GapmeRs targeting Carmen or random scrambled sequence. Cells were harvested 48 h post transfection and assayed for Carmen, Nkx2.5, Gata4, Myh6 and Myh7 expression. "p < 0.05, SEM (n = 3–12).


K.W. Vance, S.N. Sansom, S. Lee, V. Chalé, L. Kong, S.E. Cooper, et al., The long noncoding RNA Paupar regulates the expression of both local and distal genes, EMBO J 33 (4) (2014 Feb 18) 296–311.


