bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The long noncoding RNA *Charme* supervises cardiomyocytes maturation by controlling cell differentiation programs in the developing heart

Valeria Taliani[1‡†], Giulia Buonaiuto[1†], Fabio Desideri[2], Adriano Setti[1], Tiziana Santini[2], Silvia Galfrè[2], Carmine Nicoletti[3], Leonardo Schirone[4], Davide Mariani[5], Valentina Valenti[4], Sebastiano Sciarretta[4], Emerald Perlas[6], Antonio Musarò[3], Irene Bozzoni[1,2], Monica Ballarino[1]

[1]Dept. of Biology and Biotechnologies "*Charles Darwin*", Sapienza University of Rome, Rome, Italy;

[2]Center for Life Nano- & Neuro-Science of Istituto Italiano di Tecnologia (IIT), 00161, Rome, Italy;

[3]DAHFMO-Unit of Histology and Medical Embryology, Laboratory affiliated to Istituto Pasteur Italia-Fondazione Cenci Bolognetti, Sapienza University of Rome, Rome, Italy;

[4]Department of Medical Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina, Italy;

[5]Center for Human Technologies, Istituto Italiano di Tecnologia (IIT), Genova, Italy;

[6]EMBL-Rome, Epigenetics and Neurobiology Unit, Monterotondo, Italy.

Corresponding author: Monica Ballarino

e-mail address: monica.ballarino@uniroma1.it; *phone*: +39 06 49912201

[‡]Present address: EMBL, Genome Biology Unit, Heidelberg, Germany

[†]These authors contributed equally to this work

Running Head

Charme functions during development

Key Words

NcRNA; LncRNA; Development; Muscle; Myogenesis; Cardiomyogenesis; Heart

ABSTRACT

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Long noncoding RNAs (lncRNAs) are emerging as critical regulators of heart physiology and disease, although the studies unveiling their modes-of-action are still scarce and limited to few examples. We recently identified *pCharme*, a chromatin-associated lncRNA whose functional knockout in mice results in defective myogenesis and morphological remodelling of the cardiac muscle. By combining Cap-Analysis OF Gene Expression (CAGE) with single-cell (sc)RNA sequencing and whole-mount *in situ* hybridization analyses, here we have profiled *pCharme* levels and we found it expressed in cardiomyocytes since the first steps of cardiomyogenesis. At this early stage, *pCharme* is crucial to seed the formation of specific RNA-enriched nuclear condensates containing MATR3, its known protein interactor, as well as important regulators of cardiac development. Perturbation of this mechanism by *in vivo pCharme* ablation results in myocardium compaction and ventricular hypo-trabeculation due to a delayed maturation of cardiomyocytes. The identification of novel genes controlling cardiac morphology is crucially important, as in human's congenital defects in myocardium compaction are clinically relevant and predispose patients to major complications. Our studies *in vivo* offer unique insights into cardiac muscle maturation and bear relevance to *Charme* locus for future theranostic applications.

INTRODUCTION

In all vertebrates, heart development occurs through a cascade of events in which a subtle equilibrium between proliferation, migration, and differentiation ultimately leads cardiac precursor cells to mature into all the major cell types (Bruneau 2013; Moorman and Christoffels 2003). At the molecular level, the execution of the developmental program is governed by the dynamic interplay of several cardiac regulators, whose expression and functions are coordinated in time and space. Alteration of this process results in abnormal cardiac morphogenesis and other congenital heart defects that in humans represent the most common types of birth defects and cause of infant death (Zimmerman et al., 2020; Srivastava, 2006; Center for Disease Control and Prevention, 2020). Mutations of transcription

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

factors and protein co-factors have emerged as causative of a broad spectrum of heart malformations, although only 15-20% of all congenital heart defects are associated to known genetic conditions (Morton et al., 2022; Kodo et al., 2009; Bouveret et al., 2015; Ang et al., 2016). The need to find more targets for diagnosis and therapy has therefore evoked interest in new categories of disease-causing genes, whose discovery has been accelerated by impressive advancements in genomics. In this direction, improvements in next-generation sequencing have revolutionized many areas of biological research, including cardiology, by giving greater awareness of the noncoding (nc)RNA variety and potential (Mattick et al., 2004; Cipriano et al., 2018). Our understanding of the stagging abundance of ncRNAs culminated with the discovery of long ncRNAs (lncRNAs) (Rinn and Chang 2012; Ulitsky and Bartel 2013; Yao et al., 2019). These RNAs form a heterogeneous class of non-protein coding transcripts, longer than 200 nucleotides, participating in many physiological (i.e. cell stemness, differentiation or tissue development) and pathological (i.e. cancer, inflammation, cardiovascular or neurodegeneration diseases) processes (Azad et al., 2021; Hu et al., 2018; Riva et al., 2016). In the heart, several biologically relevant lncRNAs have been identified, with functions related to aging (Zhao et al., 2020; Ghafouri-Fard et al., 2021), regeneration (Pagano et al., 2020; Wang et al., 2021) and development (Sabour et al., 2018; Han et al., 2019; García-Padilla et al., 2018; Pinheiro et al., 2021). As of now, there are several examples of lncRNAs that have been proven to exert important functions for cardiac cells development *in vitro* (Klattenhoff et al., 2013; Ounzain et al., 2015; Kim et al., 2021); however, when it comes to understanding their role in *in vivo* heart development, the knowledge currently available is fragmentary and comprised of very few instances (Grote et al. 2013; Anderson et al., 2016; Ritter et al., 2019). Nevertheless, the importance of these class of transcripts for adult heart homeostasis has been described in both mouse and human systems by linking the aberrant expression of specific lncRNAs to heart anomalies (Han et al., 2014; Wang et al., 2016; Ballarino et al., 2018; Scheuermann et al., 2013; Ponnusamy et al., 2019; Anderson et al., 2021). Along this line, in 2015 we identified a new collection of muscle-specific lncRNAs (Ballarino

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

et al., 2015) and focused our attention on *Charme* (*Ch*romatin *ar*chitect of *m*uscle *e*xpression), a mammalian conserved lncRNA gene, whose loss-of-function in mice causes progressive myopathy and congenital heart remodelling (Ballarino et al., 2018). Here, we extend the functional characterization of *Charme* to a developmental window and assign an embryonic origin to the cardiac defects produced by its knockout in mouse. By using whole-tissue imaging approaches we successfully visualize the early activation of *Charme* gene in the heart, with the functional *pCharme* isoform being highly expressed in foetal tissues and progressively dropping down after birth. To gain a deeper understanding of the *pCharme* cell type-specific expression, we mined the available single-cell (sc)RNA sequencing datasets of mouse embryonal hearts and found its expression to be restricted to cardiomyocytes (CM). In line with this specific localization, high-throughput transcriptomic analyses combined with the phenotypic characterization of WT and *Charme*^{KO} developing hearts highlight a critical role of the lncRNA for CM maturation and for the formation of trabeculated myocardium during development. Furthermore, high-throughput sequencing of RNA isolated from embryonal cardiac biopsies upon cross-linking and immunoprecipitation (CLIP) of MATR3, a nuclear *pCharme* interactor (Desideri et al., 2020), revealed the existence of a specific RNA-rich condensate containing *pCharme* as well as important regulators of embryo development, cardiac function, and morphogenesis. In line with the functional importance of the *pCharme*/MATR3 interaction in developing cardiomyocytes, MATR3 depletion in mouse-derived cardiac primary cells leads to the down-regulation of this class of transcripts, suggesting an active crosstalk between MATR3 and *pCharme* regulatory pathways.

Understanding the basics of cardiac development from the lncRNA point-of-view is of interest not only for unravelling novel RNA-mediated circuitries but also for improving treatment options aimed to enhance cardiomyogenic differentiation. Indeed, despite the recent advancements in generating cardiomyocytes from pluripotent stem cells through tissue engineering-based methods, most protocols produce immature cells, which lack many attributes of adult cardiomyocytes (Uosaki et al.,

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

2015). Consequently, the cells generated cannot be used for efficient drug screening, modelling of adult-onset disease, or as a source for cell replacement therapy. Here, we discovered a new non-coding regulator of cardiac maturation and characterized the molecular interactome participating in this process *in vivo*. This research will both advance our understanding of heart physiology and disease and will also serve as a foundation for future studies in human diagnostics and therapeutics.

RESULTS

Charme locus expression in developing mouse embryos and cardiomyocytes.

Our earlier studies revealed the occurrence of muscle hyperplasia and cardiac remodelling upon the CRISPR/Cas9-mediated *Charme* loss-of-function in mice (Ballarino et al., 2018; Desideri et al., 2020). Intriguingly, the morphological malformations were clearly displayed in both adult and neonatal mice, which strongly suggests possible roles for the lncRNA during embryogenesis. With the purpose to trace back the developmental origins of *Charme* functions, we started by analysing the whole collection of FANTOM5 mouse promoterome data to quantify transcription initiation events captured by CAGE-seq across the lncRNA locus (Noguchi et al., 2017). In addition to its cardiac-restricted expression (**Figure 1A**, left and middle panels; **Table S1**), this profiling revealed a marked transcription of the locus at the early stages of cardiac specification, with the highest levels of expression reached at E14-E15 days (**Figure 1A**, right; **Table S1**). We then evaluated the expression of the two splice variants produced by the locus, *pCharme* and *mCharme,* by using primers for their specific amplification in embryonal (E15.5) and postnatal (day 2, PN) hearts. RT-qPCR analysis confirmed the CAGE output, as both the isoforms are more expressed at the embryonic than postnatal stage (**Figure 1B**). More importantly, we found *pCharme* to be 50% more abundant than *mCharme*, which is particularly intriguing considering the prominent role of this nuclear isoform in skeletal myogenesis (Desideri et al., 2020). We then applied an *in-situ* hybridization (ISH) approach for studying the spatio-temporal profile of *pCharme* expression *in vivo*. These analyses revealed the

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

emergence of *pCharme* already in E8.5 cardiac sections when the heart tube formation initiates (**Figure 1C**, E8.5). Of note, all the regions of the tubular heart were positive for *pCharme* staining, including the territories giving rise to the future atria and ventricles (V), though the inflow (IFT) and outflow tracts (OFT) displayed the highest signals. At later stages of development (**Figure 1C**, E13.5; **Figure 1-figure supplement 1A**), ISH confirmed the presence of *pCharme* in both cardiac tissues and somites (S), but also revealed the highest levels of expression in ventricles, while signals from atria were less prominent. A similar expression pattern was found over intact embryos and prenatal hearts by whole-mount *in-situ* hybridization (**Figure 1D**, left). It is worthy of mention that specific *pCharme* signals were detected only in wild-type heart muscles, whereas no signal was appreciated in *Charme* knockout (*Charme*^{KO}) mice (**Figure 1D**, left**; Figure 1-figure supplement 1B**).

Publicly available single-cell RNA sequencing (scRNA-seq) datasets from embryonal hearts (E12.5, Jackson-Weaver et al., 2020) were processed for studying the expression of *pCharme* across different cell sub-types. This analysis was necessary because distinct cell subpopulations contribute to the heart, each one carrying out a specialized function in cardiac development and physiology (de Soysa et al., 2019). We clustered the cardiac subpopulations based on the expression of representative marker genes (**Figure 1-figure supplement 1C-1E**) as previously described (Li et al., 2016; Jackson-Weaver et al., 2020; Franco et al., 2006; Meilhac et al., 2018). Interestingly, we found that, during embryonic heart development, *pCharme* RNA was restricted to cardiomyocytes, with an almost identical expression between atrial (CM-A), ventricular (CM-V) and the other CM clusters (CM-VP, CM-IVS, CM-OFT) (**Figure 1E**). These results provided an accurate resolution of *pCharme* expression at the level of individual cell types and offered a valuable input for exploring the possible integration of the lncRNA into pathways controlling the physiology of cardiomyocytes.

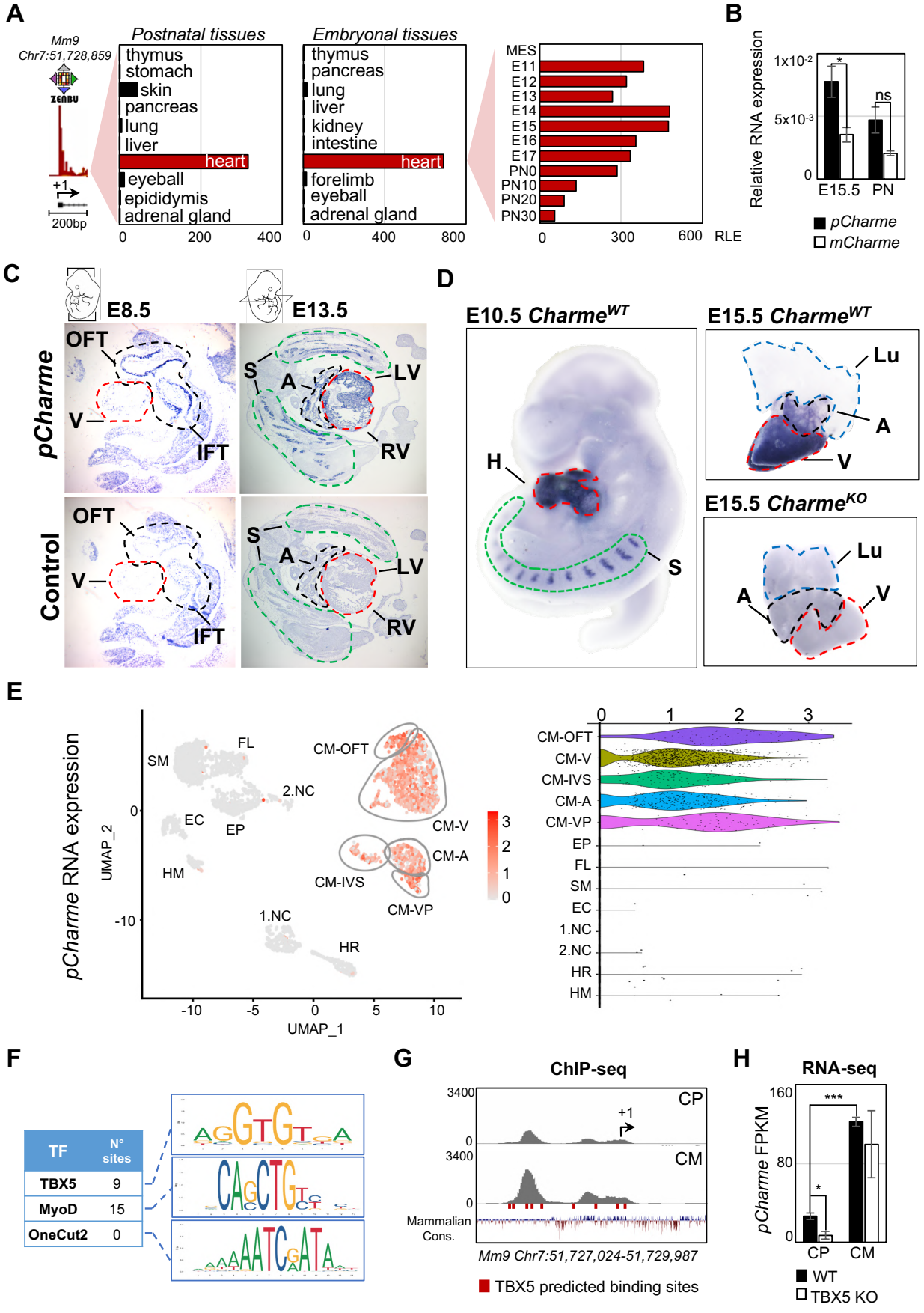
In an effort to identify the upstream regulators of *pCharme* expression in cardiomyocytes, we then performed an *in silico* prediction of *cis*-regulatory elements for transcription factors (TFs) playing a role in heart development. In accordance with our previous findings (Ballarino et al., 2015), Jaspar

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

153 database (https://jaspar.genereg.net; Castro-Mondragon et al., 2022) identified several MyoD
154 *consensus* sites in the 1kb region upstream of *pCharme* transcriptional start site (TSS, **Figure 1F**).
155 Although in skeletal muscle cells *pCharme* was found to be regulated by MyoD, this myogenic
156 regulator is not functional in the heart (Olson, 1993; Buckingham et al., 2017). Searching for other
157 cardiac regulators, we found canonical TBX5 binding motifs localized in *pCharme* regulatory regions
158 (**Figure 1F**). In contrast, no *consensus* site was found for OneCut2, a TF involved in neurogenesis
159 (Aydin et al., 2019) and chosen as a negative control. T-Box Transcription Factor 5 (TBX5) is a
160 member of the T-box family of TFs, known to activate genes associated with cardiomyocyte
161 maturation in early development, and with cardiac morphogenesis at later stages (Nadadur et al.,
162 2016; Steimle et al., 2017). Recently, TBX5 was found to be crucial for the expression of several
163 lncRNAs (Yang et al., 2017), a subset of them being enriched in the chromatin fraction of
164 cardiomyocytes (Hall et al., 2021). Inspection of Chromatin Immunoprecipitation (ChIP)-sequencing
165 (**Figure 1G)** and RNA-seq (**Figure 1H**) profiles from progenitor (CP) and differentiated (CM) murine
166 cardiomyocytes (Luna-Zurita et al., 2016) revealed the specific occupancy of TBX5 to highly
167 conserved spots in *pCharme* promoter, which increases with differentiation in concomitance to
168 *pCharme* expression. Remarkably, the lncRNA expression consistently decreases upon TBX5
169 knockout, which points to the relevance of this TF for the transcriptional regulation of *Charme* locus.
170 Nevertheless, the cooperation with additional factors cannot be excluded, also considering the attitude
171 of TBX5 to interact, physically and functionally, with several other cardiac regulators (Akerberg et
172 al., 2019).

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 1.



173

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 1. *Charme* locus expression in developing mouse embryos and cardiomyocytes

A) Transcriptional start site (TSS) usage analysis from FANTOM5 CAGE Phase1 and 2 datasets (skeletal muscle not included) on the last update of Zenbu browser (https://fantom.gsc.riken.jp/zenbu/; FANTOM5 Mouse mm9 promoterome) showing *Charme* locus expression in postnatal and embryonic body districts (left and middle panels) and during different stages of cardiac development (right panel, E11-PN30). MES=Mesoderm. Bars represent the Relative Logaritmic Expression (RLE) of the Tag Per Million values of *Charme* TSS usage in a specific sample**. B)** Quantitative Reverse Transcription PCR (RT-qPCR) amplification of *pCharme* and *mCharme* isoforms in RNA extract from *Charme*^{WT} E15.5 and postnatal (PN) hearts. Data were normalized to GAPDH mRNA and represent means ± SEM of 3 pools. **C)** *In-situ* hybridization (ISH) performed on embryonic cryo-sections using digoxigenin-labelled RNA antisense (*pCharme*, upper panel) or sense (control, lower panel) probes against *pCharme*. Representative images from two stages of embryonal development (E8.5 and E13.5) are shown. OFT: Outflow Tract; IFT: Inflow Tract; V: Ventricle; LV/RV: Left/Right Ventricle; A: Atria; S: Somites. **D)** Whole-mount *in-situ* hybridization (WISH) performed on *Charme*^{WT} intact embryos (E10.5, left panel) and *Charme*^{WT} and *Charme*^{KO} hearts at their definitive morphologies (E15.5, right panels). Signal is specifically detected in heart (H, red line) and somites (S, green line). Lungs (Lu, blue line) show no signal for *pCharme*. *Charme*^{KO} hearts were used as negative control. A: Atria (black line); LV and RV: Left and Right ventricle. **E)** Left panel: UMAP plot showing *pCharme* expression in single-cell transcriptomes of embryonal (E12.5) hearts (Jackson-Weaver et al., 2020). Right panel: Violin plot of *pCharme* expression in the different clusters (see **Material and Methods** for details). CM: Cardiomyocytes, CM-A: Atrial-CM, CM-V: Ventricular-CM, ISV: Interventricular Septum, VP: Venous Pole, OFT: Outflow Tract, NC: Neural Crest cell, EP: Epicardial cells, FL: Fibroblasts like cells, ED: Endothelial cells, SM: Smooth Muscle cells, HM: Hemopoietic Myeloid cells, HR: Hemopoietic Red blood cells. **F)** *In silico* analysis of TBX5, MyoD1 and Onecut2 transcription factors (TF) binding sites using Jaspar 2022 (relative profile score threshold=80%). MyoD1 and Onecut2 were used as positive and negative controls, respectively. N° sites =number of consensus motives. **G**) TBX5 ChIP-seq analysis across *Charme* promoter in murine cardiac-precursors (CP) and mature cardiomyocytes (CM) (GSE72223, Luna-Zurita et al., 2016). The genomic coordinates of the promoter, the *Charme* TSS (+1, black arrow), the TBX5 consensus sites (red lines) and the mammalian conservation track (Mammalian Cons.) from UCSC genome browser are reported. **H**) Quantification of *pCharme* expression levels from RNA-seq analysis performed in wild type (WT) and TBX5 Knock Out (KO) murine CP and CM (SRP062699, Luna-Zurita et al., 2016). FPKM: Fragments Per Kilobase of transcript per Million mapped reads. Data information: *$p < 0.05$; ***$p < 0.001$, unpaired Student's t test.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Genome-wide profiling of cardiac *Charme*^{WT} and *Charme*^{KO} transcriptomes

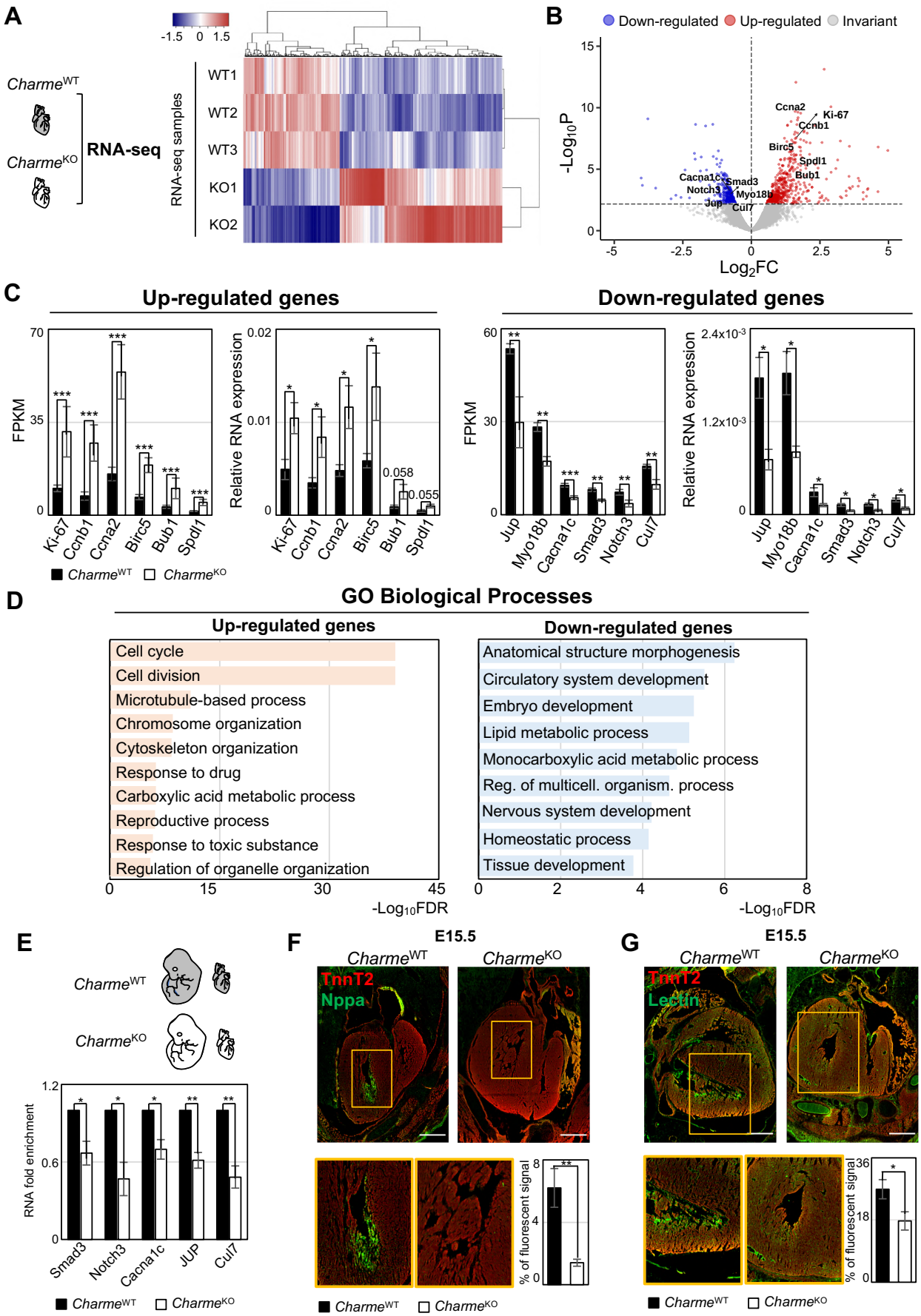
Over the years, we have accumulated strong evidence on the role played by *pCharme* in skeletal myogenesis (Ballarino et al., 2018; Desideri et al., 2020). However, important questions still pend on the role of *pCharme* in cardiomyogenesis, especially concerning the effect of its ablation on the heart (Ballarino et al., 2018). As a first step into the identification of the molecular signature underlying *Charme*-dependent cardiac anomalies, we performed a differential gene expression analysis on transcriptome profiles from developing WT and *Charme*^{KO} hearts (**Figure 2A**; **Figure 2-figure supplement 1A**). This analysis led to the identification of 913 differentially expressed genes (DEGs) (FDR<0.1, WT *vs* KO, **Table S2**), 573 of which were up-regulated and 340 down-regulated in *Charme*^{KO} hearts (**Figure 2B**; **Table S2**). Results were confirmed by RT-qPCR analyses performed from independent biological replicates on gene subsets (**Figure 2C**). A Gene Ontology (GO) term enrichment study was then applied separately to the up-regulated and down-regulated genes. The analyses revealed that the up-regulated DEGs were significantly enriched (FDR-values < 1,0E-10) in cell cycle and cell division GO categories (**Figure 2D**, left panel; **Figure 2-figure supplement 1B**, upper panel). This result correlates with the increased number of Ki-67^{+} mitotic nuclei as quantified in *Charme*^{KO} cardiac tissues as compared to WT (**Figure 2-figure supplement 1C**) and it is also highly consistent with the phenotype of cardiac hyperplasia previously observed in postnatal and young *Charme*^{KO} hearts (Ballarino et al 2018). Even more interestingly, the down-regulated genes clustered to more functional and morphogenetic categories, such as anatomical structure morphogenesis (FDR 5.77E-07) and circulatory system development (FDR 3.1338E-06) (**Figure 2D,** right panel; **Figure 2-figure supplement 1B,** lower panel). These categories include TFs involved in pivotal steps of embryo development such as Smad3 (Dunn et al., 2004) and Notch3 (MacGrogan et al., 2018), and functional components of cardiomyocytes such as Cacna1c, the alpha-1 subunit of a well-known voltage-dependent calcium channel (Wang et al., 2018), and the Myosin-18B (Myo18B) gene, known to regulate cardiac sarcomere organization (Latham et al., 2020). In line with

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

transcriptomic data, RT-qPCR analyses revealed a similar trend of down-regulation when the expression of these genes was compared between *Charme*^{KO} and wild-type cardiac RNA extracts (**Figure 2E**). Based on these results, we then narrowed our focus on the impact of *pCharme* ablation on cardiac development and morphogenesis and studied the occurrence of any histological phenotype associated with these biological processes at embryonal stages. In agreement with the gene expression analysis, haematoxylin and eosin staining of *Charme*^{KO} and WT cardiac cryo-sections from E13.5 embryos showed a pronounced alteration of the myocardium morphology, which results in an overgrowth of the compact layer (**Figure 2-figure supplement 1D**). To better understand the histological differences and further examine whether compaction occurs at the expense of ventricular trabeculae, we performed immunofluorescence staining and western blot analysis for the natriuretic peptide A (NPPA) factor, a known marker of the embryonic trabecular cardiomyocytes (Choquet et al., 2019; Horsthuis et al., 2008). In parallel, we also analysed the heart sections for cardiac vasculature, since the process of trabecular compaction is known to be temporally coupled with the formation of blood vessels during development (Samsa et al., 2013). To note, the circulatory system development category was among the main GO enriched for *pCharme* down-regulated targets. We found that mutant hearts (E13.5 and E15.5) display a significant reduction of Nppa^{+} trabeculae (**Figure 2F**; **Figure 2-figure supplement 1E**,**2F**), which parallels the diminished density of the capillary endothelium, as imaged by lectin staining (**Figure 2G**). Therefore, these data suggest that the early expression of *pCharme* ensures the achievement of cell-cycle and morphogenetic programs important for CM maturation, myocardial geometry and vascular network formation. The developmental need for *pCharme* is also important for the preservation of cardiac function and structure in adulthood, as a progressive deterioration of the systolic function, which becomes significant at 9 months of age (**Figure 2-figure supplement 1G**; **Table S3**), was observed in *Charme*^{KO} mice.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 2



263

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 2. Genome-wide profiling of cardiac *Charme*^{WT} and *Charme*^{KO} transcriptomes

A) Heatmap visualization from RNA-seq analysis of *Charme*^{WT} and *Charme*^{KO} PN hearts. Plot was produced by heatmap3 (https://cran.r-project.org/web/packages/heatmap3/vignettes/vignette.pdf) using the FPKM values for each RNA-seq sample. FPKM: Fragments Per Kilobase of transcript per Million mapped reads. Schematic representation of RNA samples collected for sequencing analysis from *Charme*^{WT} and *Charme*^{KO} hearts is shown. **B)** Volcano plots showing differential gene expression analysis from the RNA-seq analysis of *Charme*^{WT} *vs Charme*^{KO} PN hearts. Differentially expressed genes (DEGs) validated through RT-qPCR are in evidence. **C)** Average expression from RNA-seq (FPKM) and RT-qPCR (Relative RNA expression) quantification of up-regulated (left panel) and down-regulated (right panel) DEGs in *Charme*^{KO} *vs Charme*^{WT} PN hearts. Data were normalized to GAPDH mRNA and represent means ± SEM of WT (n=5) *vs* KO (n=4) pools. **D)** Gene ontology (GO) enrichment analysis performed by WEBGESTALT (http://www.webgestalt.org) on up-regulated (left panel) and down-regulated (right panel) DEGs in *Charme*^{WT} *vs* Charme^{KO} pools of PN hearts. Bars indicate the top categories of biological processes in decreasing order of $-Log_{10}$FDR. All the represented categories show a FDR value <0.05. **E)** RT-qPCR quantification of *pCharme* targets in embryonal (E15.5) hearts. Schematic representation of RNA samples collection is shown. DEGs belonging to the GO:anatomical structure morphogenesis were considered for the analysis. Data were normalized to GAPDH mRNA and represent means ± SEM of WT and KO (n=3) pools. **F)** Representative images from immunostaining for Nppa (green) and TnnT2 (red) in *Charme*^{WT} and *Charme*^{KO} (E15.5) cardiac sections. Scale bar: 500 µm. Quantification of the area covered by Nppa fluorescent signal is shown aside. Data represent the mean (%) ± SEM of at least 6 images from biological replicates for each genotype. **G)** Representative image from immunostaining for Lectin (green) and TnnT2 (red) in *Charme*^{WT} and *Charme*^{KO} (E15.5) cardiac section. Scale bars: 500 µm. Quantification of the area of labelled vascular tissue is shown aside. Data represent the mean (%) ± SEM of at least 6 images from biological replicates for each genotype. Data information: *$p < 0.05$; **$p < 0.01$, ***$p < 0.001$ unpaired Student's t test.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In foetal cardiomyocytes *pCharme* nucleates the formation of RNA-rich condensates by interacting with MATR3.

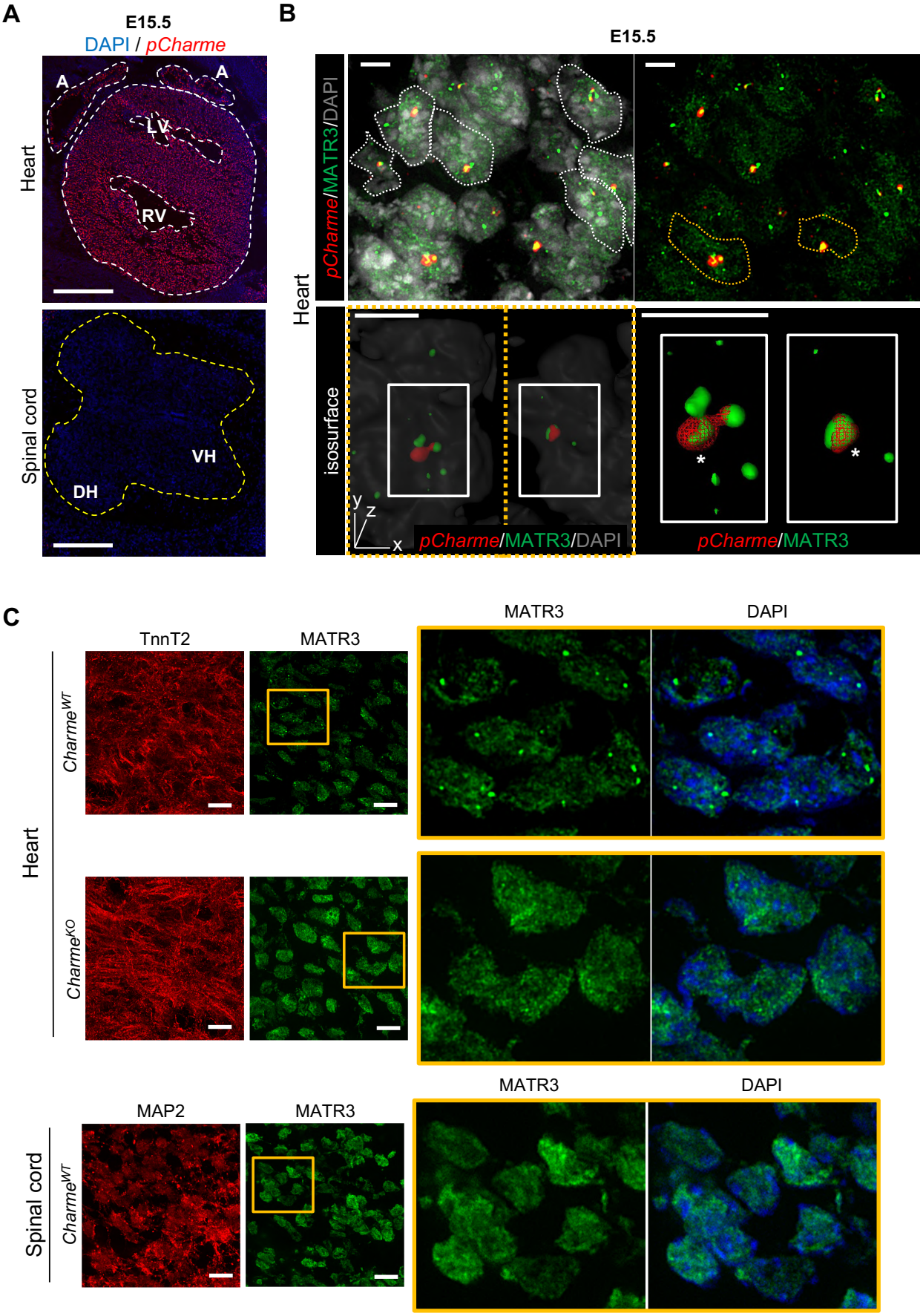
In differentiated myocytes, we previously demonstrated that CU-rich binding motifs inside *pCharme* coordinate the co-transcriptional recruitment and the subcellular localization of Matrin 3 (MATR3) (Desideri et al., 2020), a nuclear matrix RNA/DNA binding protein involved in multiple RNA biosynthetic processes (Coelho et al., 2016; Banerjee et al., 2017) and recently shown to play a role in chromatin repositioning during development (Cha et al., 2021). Interestingly, MATR3 protein was found very highly expressed in newborn mouse cardiomyocytes and heterozygous mice with mutations in MATR3 show congenital heart defects (Quintero-Rivera et al., 2015). Therefore, an intriguing possibility may be that *pCharme* plays a role in the localization of MATR3 in embryonal cardiac nuclei with possible implications in their physiology. To examine this hypothesis more directly, we first assessed the sub-cellular localization of *pCharme* in embryonal (E15.5) hearts by biochemical fractionation. RT-qPCR analyses revealed that, in line with what was previously observed in skeletal muscle, also in cardiomyocytes *pCharme* mainly localizes within the nucleus, while the fully spliced *mCharme* was enriched into the cytoplasm (**Figure 3-figure supplement 1A**). We next applied high-resolution RNA-fluorescence *in situ* hybridization (FISH) to visualize *pCharme* in embryonal hearts (**Figure 3A**; **Figure 3-figure supplement 1B**) and combined RNA-FISH with immunofluorescence to study its relative localization with MATR3 (**Figure 3B**). In agreement with subcellular fractionation, RNA-FISH experiments confirm the nuclear localization of *pCharme* and show its typical punctuate distribution overall the heart. Moreover, the analysis of the three-dimensional distribution of *pCharme* and MATR3 revealed a clear colocalization between their respective signals. Quantitative analysis of the over-lapping signals by 3D Pearson's correlation coefficient highlighted the formation of MATR3/*pCharme* nuclear condensates (**Figure 3-figure supplement 1C**). Based on these results, we then tested if the presence of *pCharme* can influence the nuclear localization of MATR3 *in vivo*. To this end, we performed MATR3 IF assays in wild-type

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

and *Charme*^{KO} embryonal (E15.5) muscle biopsies and spinal cord nuclei, as a control (**Figure 3C**; **Figure 3-figure supplement 1D**,**3E**). We observed intense and discrete MATR3 positive signals in wild-type skeletal (**Figure 3-figure supplement 1D**) and cardiac (**Figure 3C**, upper panel) muscles expressing the lncRNA. Coherently with a *pCharme*-dependent MATR3 condensation, MATR3 signals appeared more diffuse when *pCharme* was not expressed, as in *Charme*^{KO} muscles and wild-type spinal cord nuclei (**Figure 3C,** middle and lower panels; **Figure 3-figure supplement 1D**).

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 3



330

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 3. In foetal cardiomyocytes *pCharme* nucleates the nuclear localization of MATR3

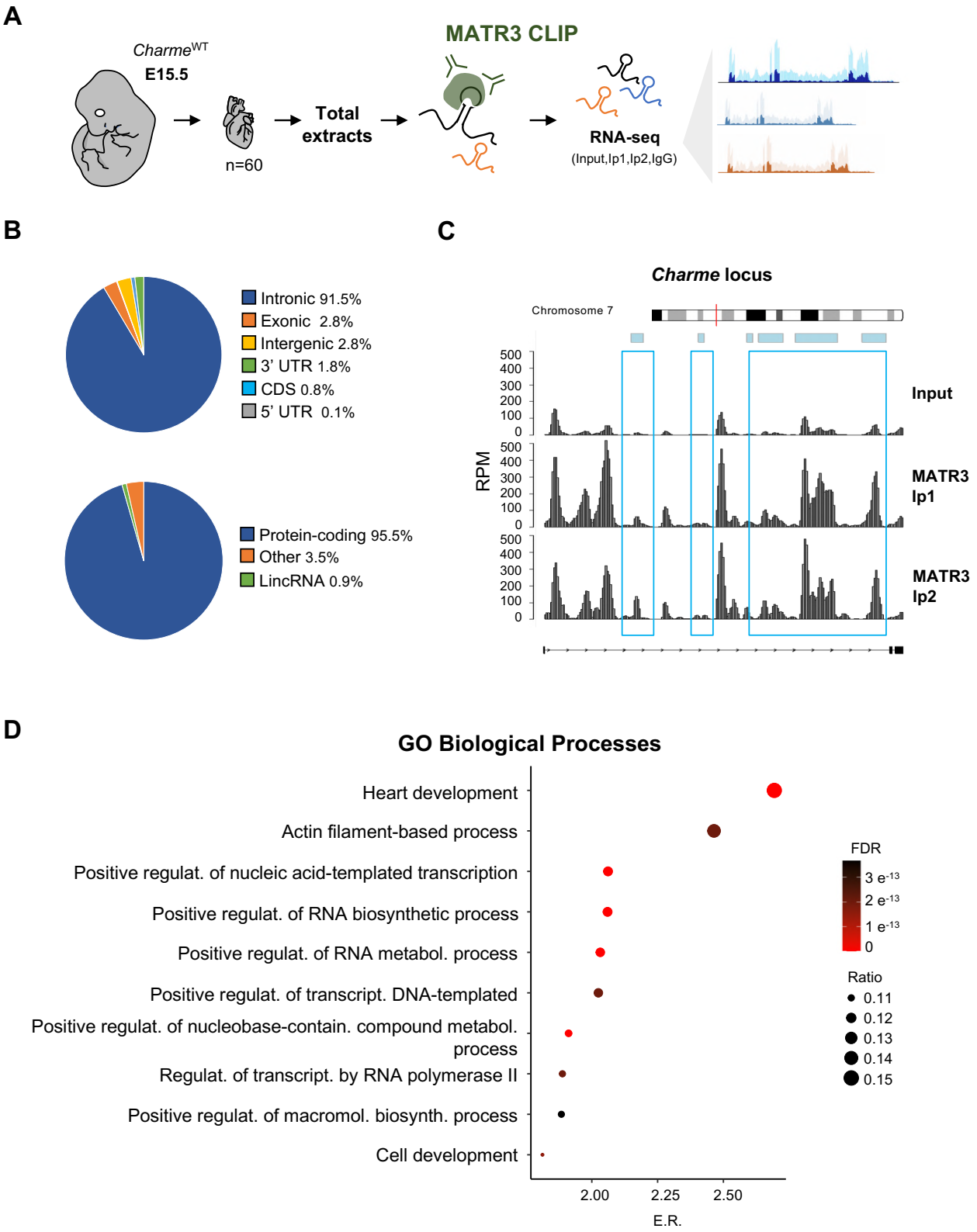
A) RNA-FISH for *pCharme* (red) and DAPI staining (blue) in *Charme*^{WT} hearts and spinal cord from E15.5 embryos sections. Signal is specifically detected in whole heart (white dashed lines) while no signal is detected in spinal cord (yellow dashed line). A: Atria; LV and RV: Left and Right ventricle; DH and VH: Dorsal and Ventral horn. Scale bars, 500 µm. **B)** Upper panel: RNA-FISH for *pCharme* (red) combined with immunofluorescence for MATR3 (green) and DAPI staining (gray) in *Charme*^{WT} hearts from E15.5 embryos cryosections. Dashed lines show the edge of the nuclei. Lower panel: selected nuclei of interest (yellow dashed lines in upper panel) were enlarged and processed for isosurface reconstruction (left panel) and digital magnification (right panel) where overlapped signals are shown (asterisks). Scale bars 5 µm. **C)** Upper panel: Immunostaining for MATR3 (green), TNNT2 (red) and DAPI (blue) in *Charme*^{WT} and *Charme*^{KO} hearts from E15.5 embryos. Lower panel: immunostaining for MATR3 (green), Map2 (red) and DAPI (blue) in *Charme*^{WT} and *Charme*^{KO} spinal cord from E15.5 embryos. The selected regions of interest (orange squares) were digitally enlarged on the right panels. Each image is a representative of three individual samples. Scale bars, 10 µm.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

We then applied a straightforward and RNA-centric approach to directly identify the MATR3-bound transcripts in the context of cardiac development. To this aim, we performed a MATR3 cross-linking immunoprecipitation (CLIP) assay in E15.5 hearts and sequenced the total RNA collected from two independent experiments (**Figure 4A**). Western blot analysis of protein extracts was carried out with anti-MATR3 antibodies to verify the enrichment and proper size of MATR3 in the immunoprecipitated samples and to quantify the efficiency of protein recovery (**Figure 4-figure supplement 1A**, left panel). Moreover, RT-qPCR analysis of the retrieved RNAs evidenced the specific enrichment of *pCharme*, but not *mCharme*, in the MATR3-immunoprecipitated samples, thus supporting the distinctive binding of the protein to the nuclear isoform (**Figure 4-figure supplement 1A**, right panel). Analysis of MATR3-associated RNAs across the sequenced transcriptome (**Figure 4-figure supplement 1B**) led to the identification of 951 cardiac expressed transcripts which were significantly bound by MATR3 (log2 Fold enrichment > 2 and FDR value <0.05; **Table S4**), a subset of which were also validated by RT-qPCR analyses (**Figure 4-figure supplement 1C**,**4D**). In line with the binding propensities exhibited in other cell systems (Uemura et al., 2017), most of the MATR3 enrichments were located within the introns of mRNA precursors (**Figure 4B**). Strikingly, while the class of lncRNAs was poorly represented, we found *pCharme* at the top of the MATR3 interactors (**Figure 4C**, **Table S4**). Together with the data provided by the imaging approaches, this result extends our previous evidence in skeletal muscle cells by showing *pCharme* as one of the most abundant lncRNA components of RNA-rich condensates formed together with MATR3 in embryonal cardiomyocytes and concentrating pre-mRNA encoding important cardiac regulators.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 4



bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 4. MATR3/*pCharme* nuclear condensates contain key regulators of heart development
A) Schematic representation of MATR3 cross-linked RNA immunoprecipitation (CLIP) and sequencing workflow on total extracts obtained from n=60 embryonal (E15.5) hearts. See **Materials and Methods** for details. **B)** Upper panel: pie-plot representing the genomic occupancy of MATR3 enriched peaks (log2 Fold enrichment > 2 and FDR < 0.05). Peaks overlapping multiple regions were assigned with the following priority: CDS, 3'UTR, 5'UTR, exons, introns and intergenic. Lower panel: pie plot representing transcript biotypes of 951 identified MATR3 interacting genes. Peaks overlapping multiple transcripts were assigned with the following priority: protein coding, lncRNA and other. Percentages relative to each group are shown. **C)** Genomic visualization of Input, Ip1 and Ip2 normalized read coverage tracks (RPM) from MATR3 CLIP-seq on *Charme* locus. MATR3 peaks displaying log2 Fold enrichment > 2 in both Ip1 and Ip2 samples compared to Input are depicted as light blue squares and are demarcated by light blue boxes. Plot obtained using Gviz R package. **D)** GO enriched categories obtained with WebGestalt (http://www.webgestalt.org) on protein-coding genes overlapping MATR3 peaks. Dots indicate the top categories of biological processes (description in y-axis) in decreasing order of Enrichment Ratio (E.R.= overlapped genes/expected genes, x-axis). Dot size (Ratio) represents the ratio between overlapped genes and GO categories size while dot color (FDR) represents significance. All the represented categories show an FDR<0.05.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The *pCharme*/MATR3 interaction in cardiomyocytes sustains developmental gene expression.

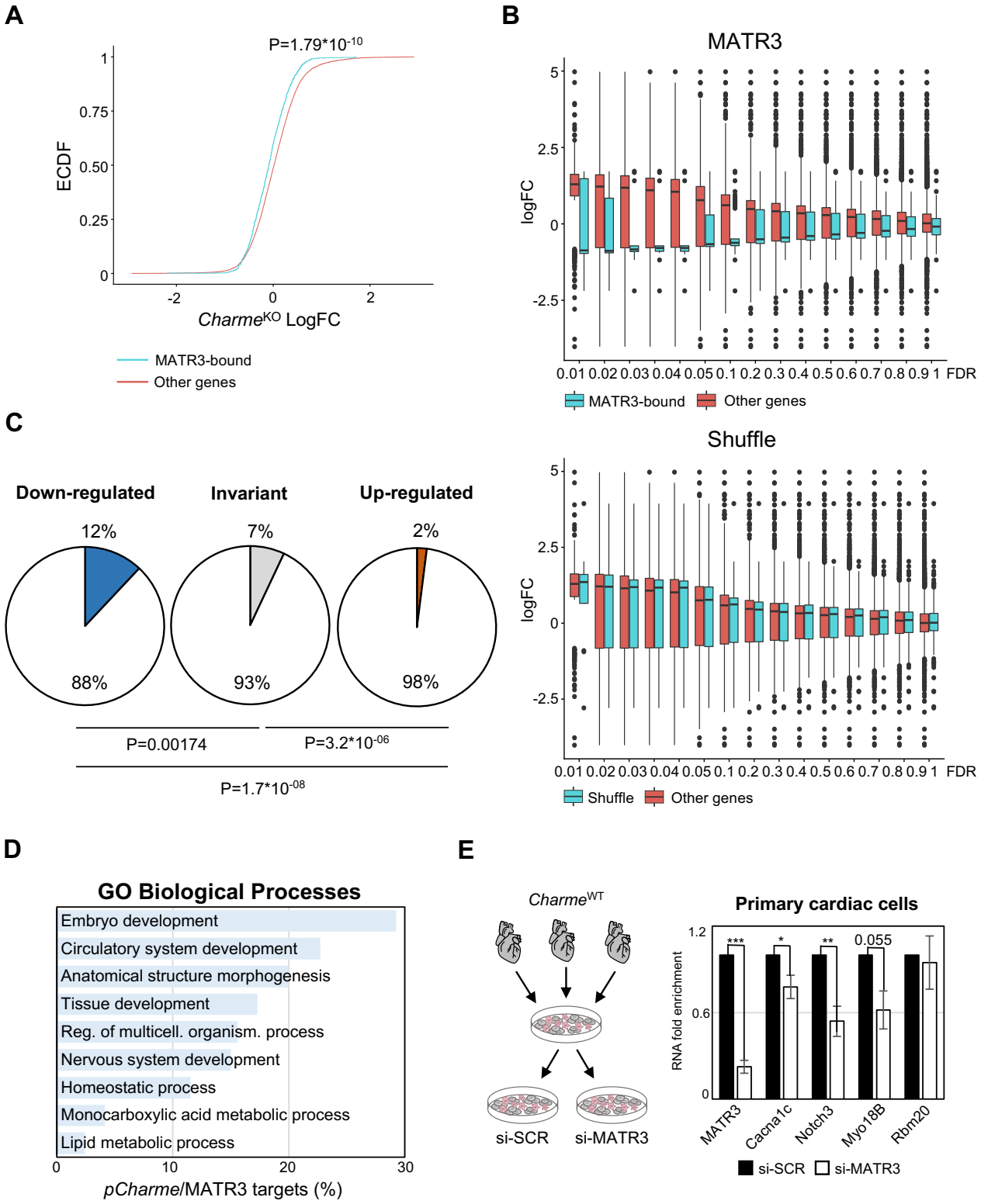
A GO enrichment analysis of the 882 protein-coding set of transcripts directly bound by MATR3, revealed heart development and RNA transcription as the most significantly enriched functional categories (**Figure 4D**). Together with the developmental requirement of *pCharme* for both cardiac gene expression and MATR3 nuclear localization, this result encouraged further studies on possible targets shared by *pCharme* and MATR3 in the heart. To this aim, we intersected the list of MATR3 RNA interactors with DEGs from cardiac *Charme*^{WT} *vs Charme*^{KO} transcriptomes. We found that the abundance of MATR3-bound RNAs shows a global tendency to be reduced in the *Charme*^{KO} condition (**Figure 5A**). In line with this, the enrichment of MATR3 targets was progressively increased as the significance (FDR) of down-regulation increased (**Figure 5B**, upper panel), in respect to a random group of genes with the same numerosity of MATR3 targets (**Figure 5B**, lower panel). Specifically, we found that 41 (12%) out of the 340 genes that were down-regulated in *Charme*^{KO} hearts(Figure were also bound by MATR3 (**Figure 5C**, left panel), while a reduced binding was found to transcripts which levels do not change (invariant) or result more abundant upon *pCharme* ablation (**Figure 5C**, middle and right panels). By checking the distribution of these common *pCharme*/MATR3 targets across the *Charme*^{KO} down-regulated GO terms (**Figure 2D**; **Figure 2-figure supplement 1B**), we found embryo development, anatomical structure morphogenesis and circulatory system development as the top three represented categories (**Figure 5D**). Intriguingly, inside these categories, we found genes such as Cacna1c, Notch3, Myo18B and Rbm20, whose role in cardiac physiopathology was extensively studied (Goonasekera et al., 2012; Tao et al., 2017; Ajima et al., 2008; van den Hoogenhof et al., 2018). Collectively, these data offer an unprecedented knowledge on the RNA-binding propensities of MATR3 in the embryonal heart and identify a subset of MATR3-bound RNAs whose levels are influenced by *pCharme*. To test whether the expression of these *pCharme* target genes was also influenced by MATR3, we then purified cardiac primary cells from wild-type hearts for transfection with control (si-SCR) or MATR3 (si-MATR3) siRNAs (**Figure**

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

5E) focusing our analysis on a subset of well-known cardiac effectors taken from the top three GO categories. RT-qPCR analysis performed on RNA extracted from these samples revealed that 3 out of 4 tested genes exhibited a significant expression decrease upon MATR3 depletion (**Figure 5E**).

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 5



bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 5. The *pCharme*/MATR3 interaction in cardiomyocytes sustains developmental genes expression

A) Empirical cumulative distribution functions (ECDF) showing that MATR3-bound RNAs have a reduced abundance in the *Charme*^{KO} condition compared to the other expressed genes. Significance was determined using a two-sided Kolmogorov–Smirnov (KS) test. **B)** *Charme*^{KO} logFC of MATR3 targets (red boxplots) and other genes (cyan boxplots) were compared selecting deregulated genes with decreasing FDR threshold. The difference between medians of groups were exacerbated when more significant deregulated genes were evaluated. This effect is specific for MATR3 targets, since it is lost when a random group of genes was evaluated (Shuffle). **C)** Pie chart depicting the percentage of MATR3 targets in *Charme*^{KO} down-regulated, invariant or up-regulated genes. Significance of enrichment or depletion was assessed with two-sided Fisher's Exact test. **D)** Percentage (%) of MATR3-bound transcripts in the GO categories identified for *Charme*^{KO} down-regulated genes (**Figure 2D**). **E)** Left panel: Schematic representation of primary cells extraction from *Charme*^{WT} hearts. Once isolated, cells were plated and transfected with the specific siRNA (si-MATR3) or control siRNA (si-SCR). See **Materials and Methods** for details. Right panel: RT-qPCR quantification of MATR3, Cacna1c, Notch3, Myo18B and Rbm20 levels in primary cardiac cells treated with si-SCR or si-MATR3. Data were normalized to GAPDH mRNA and represent mean ± SEM of 4 independent experiments. Data information: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$, unpaired Student's t test.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Our results demonstrate that in the developing heart RNA-rich MATR3-condensates form at the sites of *pCharme* transcription and control the expression of important regulators of embryo development, cardiac function, and morphogenesis. To the best of our knowledge, no previous research has given such an insight into the importance of specific lncRNA/RNA binding protein interactions occurring at the embryonal stages of mammalian heart development. As MATR3 is involved in multiple nuclear processes, further studies will be necessary to address how the early interaction of the protein with *pCharme* impacts on the expression of specific RNAs, either at transcriptional or post-transcriptional stages.

DISCUSSION

In living organisms, the dynamic assembly and disassembly of distinct RNA-rich molecular condensates influences several aspects of gene expression and disease (Roden et al., 2021). The engagement of specific lncRNAs can enhance the biochemical versatility of these condensates because of the extraordinary tissue-specificity, the structural flexibility and the propensity of this class of RNAs to gather macromolecules (Buonaiuto et al., 2021). Furthermore, the maintenance of specific lncRNAs on the chromatin combined with their scaffolding activity for RNA and proteins can cunningly seed high-local concentrations of molecules to specific loci (Bhat et al., 2021; Ribeiro et al., 2018). The occurrence of alternative RNA processing events eventually leads to the formation of diverse lncRNA isoforms, thus refining the biochemical properties and the binding affinities of these ncRNAs. This suggestive model perfectly fits with *pCharme* and provides mechanistic insights into the physiologic importance of this lncRNA in muscle. In fact, of the two different splicing isoforms produced by the *Charme* locus, only the unspliced and nuclear *pCharme* isoform was found to play an epigenetic, architectural function in skeletal myogenesis.

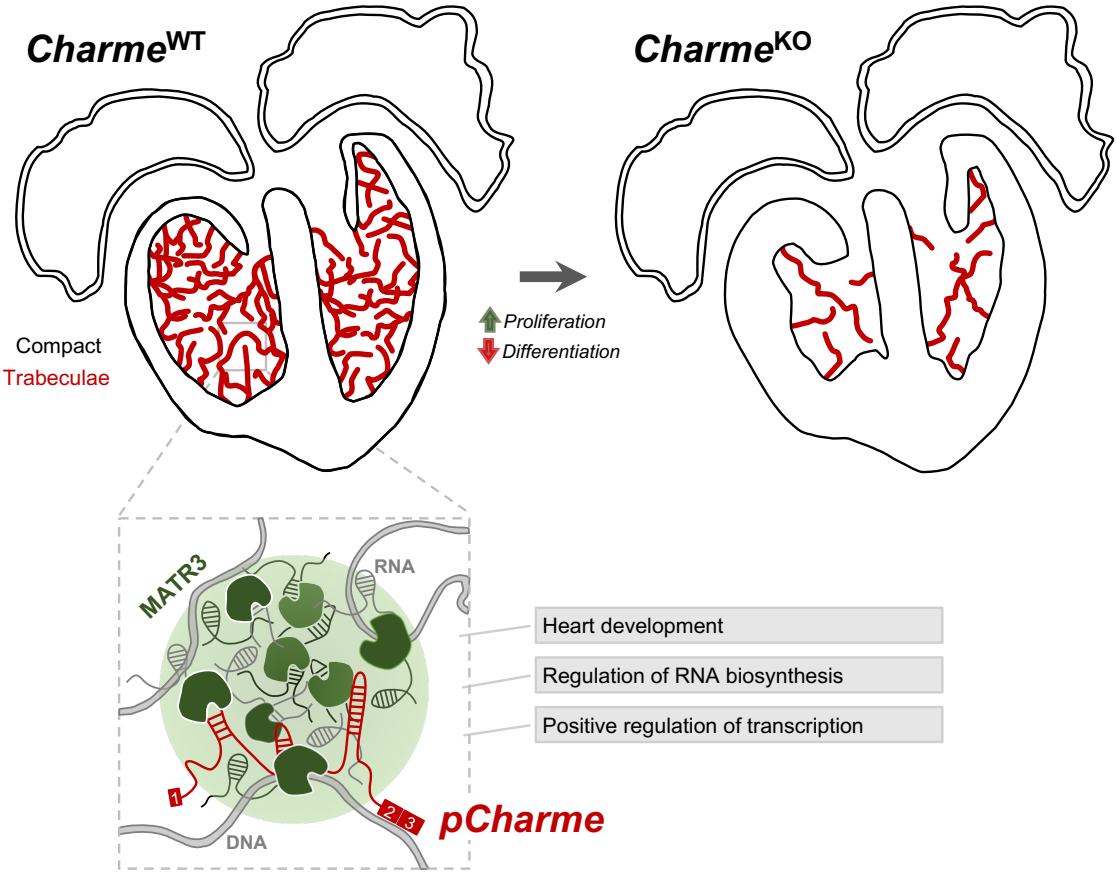
In this study, we extend our earlier findings to embryogenesis, taking advantage of our previously generated *Charme*^{KO} mouse model. We show that the temporal and cell-type specific expression of

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

506 *pCharme* in embryos supervises heart development and morphogenesis through the regulation of pro-
507 differentiation and cell cycle-related genes. Mechanistically, we provide evidence on the engagement
508 of *pCharme* in RNA-rich condensates acting as chromatin hubs for MATR3, a RNA-binding protein
509 highly abundant in the foetal heart and causing, upon mutation, congenital heart defects in both human
510 and mouse (Quintero-Rivera et al., 2015). In the absence of *pCharme*, these condensates do not form
511 *in vivo* with consequent impairment of the developmental gene expression, which ultimately leads to
512 cardiac malformations. Therefore, these findings add an important dowel to the characterization of
513 *pCharme*, by valorising the significance of its functional interaction with the nuclear matrix protein
514 MATR3 in the embryonal heart (**Figure 6**).

515

Figure 6



516

517

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 6. Proposed model for *pCharme* functions during heart development.
In physiological conditions (*Charme*^{WT}), *pCharme* is required for the expression of several genes that are involved in cardiomyocytes proliferation, differentiation and maturation at developmental stages. The lncRNA-dependent nucleation of MATR3 protein contributes to the expression of genes involved in heart development. *pCharme* absence (*Charme*^{KO}) leads to a substantial remodelling of the murine heart morphology.

Besides the clear implication in cardiac pathologies, these results also underlie the importance of studying cardiomyogenesis in animal models. Specifically, the proliferative response of cardiomyocytes to *pCharme* knockout was not expected and it has given a novel and interesting facet to its characterization. Indeed, while our previous transcriptomic studies, carried out in myoblast cell lines, minimized the role of this lncRNA by limiting its impact to the terminal differentiation of skeletal myocytes (Ballarino et al., 2018), the herein analyses revealed a much stronger connection between the regulation of proliferation and maturation of cardiomyocyte and enabled unprecedented advancement into the understanding of the *pCharme* function *in vivo*. Furthermore, this system was also crucial for demonstrating the importance of *pCharme* expression in cardiac cells for the preservation of the cardiac functions in adulthood, since adult *Charme*^{KO} mice develop a significant reduction of systolic function with initial signs of cardiac dilation, denoting an early phase of cardiomyopathy. These results are supported by recently published data suggesting that the transition of cardiomyocytes toward an immature phenotype *in vivo* is associated with the development of dilated cardiomyopathy (Ikeda et al., 2019). Finally, even more important is the observation that several regulators of embryonic development (i.e. Smad3, Notch3) and myogenic components (i.e. Cacna1c, Myosin-18B) are dysregulated upon *pCharme* knockout. Notably, mutations of these genes have been linked to the onset of several cardiomyopathies and heart failure (Goonasekera et al., 2012; Tao et al., 2017; Ajima et al., 2008; van den Hoogenhof et al., 2018). Therefore, the expression of *pCharme* in early development ensures the proper expression of genes that control the balance between proliferation and differentiation and, as consequence, the morphology of the heart.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Recent cardiovascular studies have uncovered essential roles for lncRNAs in cardiac development and disease (Scheuermann et al., 2013; Anderson et al., 2016; Ritter et al., 2019). However, a still unmet need is to disentangle non-canonical lncRNA-mediated mechanisms of action to gain insight into more successful diagnosis and therapeutics for enhancing cardiomyogenesis. In this direction, many efforts are aimed at the study of ncRNAs, not only to gain a deeper comprehension of the molecular mechanisms underneath disease onset and to facilitate more accurate classifications of patient subpopulations, but also to use them as possible diagnostic biomarkers or therapeutic targets (Buonaiuto et al., 2021). If, on the one hand, studying *pCharme* can provide new knowledge on cardiomyopathies, on the other, it might provide a new perspective on cardiac regeneration, as its depletion increases cell cycle rates. In fact, the induction of cardiomyocyte proliferation has represented a main challenge in the last decades as a possible approach to facing myocardial infarction (Chu et al., 2016), and it is still considered an important intervention strategy. In this view, we found in humans a newly discovered locus that produces a syntenic *pCharme* transcript with equivalent structure and tissue-specificities as the murine lncRNA (Ballarino et al., 2018). Future efforts will be devoted to clarifying the implication of this transcript in those human cardiomyopathies where pathological remodelling of the cardiac muscle like those observed in the murine model occurs.

MATERIALS AND METHODS

Ethics Statement and animal procedures

All mice used in this work were C57BL6/J mice and all procedures involving laboratory animals were performed according to the institutional and national guidelines and legislations of Italy and according to the guidelines of Good Laboratory Practice (GLP). All experiments were approved by the Institutional Animal Use and Care Committee and carried out in accordance with the law (Protocol number 82945.56). All animals were kept in a temperature of 22°C ± 3°C with a humidity between 50% and 60%, in animal cages with at least 5 animals.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

570

571 **Isolation, transfection, and subcellular fractionation of mouse primary heart cells**

572 For primary heart cells isolation and transfection, 5 to 10 postnatal (PN3) hearts for each replicate
573 (n=4) were pooled, harvested and kept at 37°C in culture medium (FBS 10%, 1X Non-Essential
574 Aminoacids, 1x PenStrep and DMEM high glucose). Hearts were mashed with pestles for 2 min and
575 cell isolation performed according to manufacturer's instructions (Neonatal Heart Dissociation Kit,
576 Miltenyi Biotec). Cell suspension was centrifuged for 5 min at 600 x g and cells were resuspended in
577 cell culture medium and plated in 22.1 mm plates. 1,5 million cells were transfected 48 hr later with
578 75 mM si-SCR or si-MATR3 in 3 µl/ml of Lipofectamine RNAiMAX (Thermo Fisher Scientific) and
579 100 µl/ml of Opti-MEM (Thermo Fisher Scientific), according to manufacturer's specifications. Total
580 RNA was collected 48 hr after transfection. See **Table S5** for siRNA sequences. Subcellular
581 fractionation of primary embryonal (E15.5) cardiac cells was performed using the Paris Kit (Thermo
582 Fisher Scientific, cat#AM1921), according to the manufacturer's instructions.

583

584 **Whole-mount *in situ* hybridization**

585 Embryos were fixed overnight in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS)
586 plus 0.1% Tween® 20 (PBT) at 4°C, dehydrated through a series of methanol/PBT solutions (25%,
587 50%, 75% and 100% methanol), and stored at -20°C until hybridization. Fixed embryos were
588 rehydrated and rinsed twice in PBT. At this point, embryos were either digested with DNase and/or
589 RNase or kept in PBT. All embryos were bleached in 6% hydrogen peroxide in PBT for 1h. Embryos
590 were then rinsed 3 times in PBT for 5 min, digested with proteinase K (10 µg/ml in PBT) for 5 min
591 at room temperature, washed once in 2 mg/ml glycine in PBT and twice in PBT for 5 min each, and
592 post-fixed in 4% PFA/0.2% glutaraldehyde in PBT for 20 min at room temperature. Embryos were
593 subsequently rinsed twice in PBT for 5 min and pre-hybridized at 70°C in hybridization solution
594 (50% Formamide, 5x SSC, pH 5, 0.1% Tween 20, 50 µg/ml heparin, 50 µg/ml Torula RNA, 50 µg/ml

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

salmon sperm DNA) for 2h. Embryos were then incubated overnight at 70°C in hybridization solution containing 500 ng/ml of denatured riboprobe. Riboprobes were generated by *in vitro* transcription in the presence of Digoxigenin-UTP (Roche Diagnostics). The antisense and sense probes were synthesized from linearized pBluescript-*Charme*_Ex2/3 plasmid. On the second day, embryos were washed twice in 50% formamide/4x SSC, pH 5/1% SDS and twice in 50% formamide/2x SSC, pH 5 for 30 min each at 55°C. Embryos were then rinsed three times for 5 min in MABT (100 mM maleic acid, 150 mM NaCl, pH 7.5, 0.1% Tween), blocked for 2 h at room temperature in 10% goat serum in MABT, and incubated overnight at 4°C in 1% goat serum in MABT with 1:5000 alkaline phosphatase-coupled anti-Digoxigenin antibody (Roche Diagnostics). On the third day, embryos were washed in MABT twice for 5 min and 5 more times for 1 h each. Embryos were then rinsed twice in NTMT (100 mM NaCl, 100 mM Tris-HCl, pH 9.5, 50 mM MgCl2, 0.1% Tween) for 15 min each, followed by the staining reaction in BM Purple (Roche Diagnostics) in the dark for 30 min to 12 h. Stained embryos were fixed overnight in 4% PFA in PBT, stored in PBT and photographed under a stereomicroscope.

Cryo-section *in situ* hybridization

Embryos were dissected in cold PBS (pH 7.4) and fixed in 4% w/v PFA for 24 h at 4°C. Following fixation, the embryos were cryoprotected either in 30% w/v sucrose in PBS (for PFA-fixed embryos) or in 30% w/v sucrose in 0.1 M Tris pH 7.5 (for Z7-fixed embryos), embedded in Tissue freezing medium (Leica Microsystems), sectioned at 12 µm using a cryostat (Leica 1900UV) and transferred to superfrost plus (ROTH) slides. The sections were air-dried for at least 30 min and stored at −80°C until later use. For chromogenic detection, sections were post-fixed in 4% w/v PFA in PBS for 10 min or in Z7, washed three times in PBS (or twice in 0.1 M Tris–HCl pH:7, 0.05 M NaCl and once in PBS for Z7) and incubated in acetylating solution (1.3% v/v triethanolamine, 0.03 N HCl, 0.25% v/v acetic anhydrite) for 10 min. Sections were then washed in PBS, incubated in 1% v/v Triton-X-

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

100 in PBS for 30 min and washed three times in PBS. Prehybridization was performed for 4-6 h in buffer H (50% v/v formamide, 5× SSC (0.75 M NaCl, 0.075 M sodium citrate), 5× Denhardt's (0.1% bovine serum albumin, 0.1% and 0.1% Polyvinylpyrrolidone), 250 µg/ml yeast RNA and 500 µg/ml salmon sperm DNA). Hybridization was performed in a humidified chamber for 16 h at 65°C in H buffer with DIG-labeled probe added (400 µg/ml). The probes were generated by *in vitro* transcription in the presence of Digoxigenin-UTP (Roche Diagnostics). Following hybridization, sections were sequentially washed in 5× SSC (5 min, 65°C), 0.2× SSC (1 h, 65°C), 0.2× SSC (5 min, RT). Then they were incubated in AB buffer (0.1 M Tris pH 7.5, 0.15 M NaCl) for 5 min, and in blocking solution (10% v/v Fetal Calf Serum in AB) for 1–2 h at RT. Antibody incubation was performed for 16 h at 4°C in AB buffer supplemented with 1% v/v Fetal Calf Serum and anti-DIG antibody coupled to alkaline phosphatase (1:5000 dilution; Roche). Sections were then washed thoroughly in AB and equilibrated in alkaline phosphatase buffer (AP - 0.1 M Tris–HCl pH: 9.5, 0.1 M NaCl, 0.05 M MgCl2) for 5 min. Alkaline phosphatase activity was detected in the dark in AP buffer supplemented with 45 mg/ml 4-nitrobluetetrazolium chloride (NBT, Roche) and 35 mg/ml 5-bromo-4-chloro-3-indolyl-phosphate (BCIP, Roche). The reaction was stopped with PBS and the sections were mounted in Glycergel (Dako). Sections were analysed and photographed under a stereomicroscope. Fluorescent detection was performed via Basescope™ assay (Advanced Cell Diagnostics, Bio-Techne) as previously described in D'Ambra et al., 2021, with little modifications according to the manufacturer's instructions for tissue processing. Probes used to detect pCharme RNA (ref. 1136321-C1) were custom produced by Advanced Cell Diagnostics and designed to specifically target the intronic sequence in order to detect the unspliced transcripts.

Preparation of probe templates for *in-situ* hybridization experiments

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

pCharme exon 2 and exon 3 were PCR-amplified from cDNA extracted from myotubes using Charme_Up-BamHI and Charme_Down-EcoRI primers (**Table S5**). PCR products were cloned into pBluescript ks(-) upon BamHI and EcoRI (Thermo Fisher Scientific) enzymatic restriction.

Histology

All hearts were fixed in 4% formaldehyde, embedded in OCT, and cut into 7 µm sections. After washing with PBS 3 times for 5 min, the sections were stained for 7 min with eosin (Merk, cat#109844). Subsequently, slides were washed 3 times with PBS and then incubated with haematoxylin (Merk, 105175) for 90 s.

Immunohistochemistry

Prior to staining, embryonal E13.5 sections (4 µm of thickness) were dewaxed in absolute xylene, rehydrated in decreasing concentrations of ethanol and subjected to antigen retrieval in 10 mM citrate buffer (pH 6.0) for 10 min. After PBS washing, sections were permeabilized in PBS-T buffer (PBS1X/0.1-0.3% Triton X-100) for 10 min at room temperature and blocked in 5% BSA. After overnight incubation with the primary antibody, slides were washed 3 times with TBS-T buffer (TBS 1X/ 0.1% Tween-20), and then incubated with the secondary antibody for 1h at room temperature. Coverslips were mounted using ProLong Diamond Antifade Mountant (Thermo Fischer Scientific, P-36961). Fluorescence signals are pseudo colored in green (Nppa) or red (TnnT2) protein (**Figure 4B**). Nuclei were counterstained with DAPI. Fresh E15.5 and PN tissues were embedded in OCT and then frozen in isopentane pre-chilled in liquid nitrogen. Cryo-sections (10 µm of thickness) were fixed in PFA 4% at 4°C for 20 min prior staining with primary antibodies, as previously described (Cazzella et al., 2012). Antibodies and dilutions are reported in **Table S5**. DAPI, Ki-67, Nppa, TnnT2 and G. simplicifolia lectin immunofluorescence signals (**Figure 2E; Figure 3B** and **3D**) were acquired with Carl Zeiss Microscopy GmbH Imager A2 equipped with Axiocam503 color camera. MATR3, TnnT2

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

and Map2 signals (**Figure 4D** and **4E; Figure 4-figure supplement 1E**) were acquired as Z stacks (200 nm path) by inverted confocal Olympus IX73 microscope equipped with a Crestoptics X-LIGHT V3 spinning disk system and a Prime BSI Express Scientific CMOS camera. Images were acquired as 16bit 2048x2048 pixel file by using 100X NA 1.40 and 60X NA 1.35 oil (UPLANSApo) objectives and were collected with the MetaMorph software (Molecular Devices). The average number of Ki-67 positive nuclei from n=4 *Charme*^{WT} and *Charme*^{KO} post-natal cardiac sections was determined by dividing the number of immunolabeled nuclei over the total number of nuclei in each microscope field. For each replicate, from 4 to 18 fields were analysed with ImageJ. Nppa and *G. simplicifolia* lectins sub-tissutal localization (**Figure 4B, 4C** and **4D**) was analysed by quantifying the ratio between the area occupied by the immunofluorescence signals and the total area of the left ventricle. Statistical analysis was performed using t-test, and the differences between means were considered significant at $P \leq 0.05$.

RNA extraction and RT-qPCR analysis

Total RNA from cultured cells and tissues was isolated using TRI Reagent (Zymo Research), extracted with Direct-zol™ RNA MiniPrep (Zymo Research), treated with DNase (Zymo Research), retrotranscribed using PrimeScript Reagent Kit (Takara) and amplified by RT-qPCR using PowerUp SYBR-Green MasterMix (Thermo Fisher Scientific), as described in Desideri et al., 2020. See **Table S5** for oligos details.

RNA-Seq Analysis

To reduce biological variability, *Charme*^{WT} and *Charme*^{KO} 2-days littermates were sacrificed and hearts from the corresponding genotypes pooled together before RNA extraction (3 *Charme*^{WT} pools, 9 hearts each, 2 *Charme*^{KO} pools, 3 hearts each). Validation analyses were performed on 2 additional *Charme*^{WT} pools (6 hearts each) and 2 *Charme*^{KO} pools (4 hearts each). Principal component analysis

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(PCA) conducted on the RNA-seq data, revealed that the two groups were evidently distinguished for the first principal component (**Figure 2-figure supplement 1A**). Illumina Stranded mRNA Prep was used to prepare cDNA libraries for RNA-Seq that was performed on an Illumina Novaseq 6000 Sequencing system at IIT-Istituto Italiano di Tecnologia (Genova, Italy). RNA-seq experiment produced an average of 26 million 150 nucleotide long paired-end reads per sample. Dark cycles in sequencing from Novaseq 6000 machines can lead to a high quality stretches of Guaninines artifacts; in order to remove these artifacts, low quality bases and N stretches from reads were removed by Cutadapt software using "-u -U","--trim-n" and "—nextseq-trim=20" parameters (Martin et al., 2011). Illumina adapter remotion were performed using Trimmomatic software (Bolger et al., 2014). Reads whose length after trimming was <35 nt were discarded. Reads were aligned to GRCm38 assembly using STAR aligner software (Dobin et al., 2013). Gene loci fragment quantification were performed on Ensemble (release 87) gene annotation gtf using STAR –quantMode GeneCounts parameter. Read counts of "reverse" configuration files were combined into a count matrix file, which was given as input to edgeR (Robinson et al., 2010) R package for differential expression analysis, after removing genes with less than 10 counts in at least two samples. Samples were normalized using TMM. Model fitting and testing were performed using glmFIT and glmLRT functions. Gene-level FPKM values were calculated using rpkm function from the edgeR package. FDR cutoff for selecting significant differentially expressed genes was set to 0.1. Genes with less than 1 average FPKM in both conditions were filtered out. Heatmap of differentially expressed genes was generated using heatmap3 R package (Zhao et al., 2014) from log2 transformed FPKM values. Volcano plot were generated using "Enhanced Volcano" R package (bioconductor.org/packages/release/bioc/vignettes/EnhancedVolcano/inst/doc/EnhancedVolcano). Gene Ontology analyses were performed on up-regulated and down-regulated protein coding genes using WebGestalt R package (Liao Y et al., 2019) applying Weighted Set Cover dimensionality reduction.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Cross-linking immunoprecipitation (CLIP) assay

A total of 60 *Charme*^{WT} E15.5 embryonal hearts were collected and pestled for 2 min in PBS, 1x PIC and 1x PMSF. The solution was filtered in 70 µm strainer and the isolated cells were plated and UV-crosslinked (4,000 mJ) in a Spectrolinker UV Crosslinker (Spectronics corporation). Upon harvesting, cells were centrifuged 5 min at 600 x g and pellets resuspended in NP40 lysis buffer (50 mM HEPES pH 7.5, 150 mM KCl, 2 mM EDTA, 1 mM NaF, 0.5% (v/v) NP40, 0.5 mM DTT, 1x PIC,) and incubated on ice for 15 min. Lysate was diluted to a final concentration of 1 mg/ml. 30 µl of Dynabeads Protein G magnetic particles (Thermo Fisher Scientific) per ml of total lysate were washed twice with 1 mL of PBS-Tween (0.02%), resuspended with 5 µg of MATR3 (**Table S5**) or IgG antibodies (Immunoreagents Inc.) and incubated for 2 h at room temperature. Beads were then washed twice with 1 mL of PBS-T and incubated with total extract overnight at 4°C. Beads were washed three times with 1 mL of HighSalt NP40 wash buffer (50 mM HEPES-KOH, pH 7.5, 500 mM KCl, 0.05% (v/v) NP40, 0.5 mM DTT, 1x PIC) and resuspended in 100 µl of NP40 lysis buffer. For RNA sequencing, 75 µl of the sample were treated for 30 min at 50°C with 1.2 mg/ml Proteinase K (Roche) in Proteinase K Buffer (100 mM Tris-HCl, pH 7.5, 150mM NaCl, 12.5 mM EDTA, 2% (w/v) SDS). T For Western Blot analysis, 25 µl of the sample were heated at 95°C for 5 min and resuspended in 4x Laemmli sample buffer (BioRad)/50 mM DTT before SDS-PAGE.

MATR3 CLIP-seq analysis

Trio RNA-Seq (Tecan Genomics, Redwood City, CA) has been used for library preparation following the manufacturer's instructions. The sequencing reactions were performed on an Illumina Novaseq 6000 Sequencing system at IGA Technology services. CLIP-sequencing reactions produced an average of 25 million 150 nucleotide long paired-end reads per sample Adaptor sequences and poor quality bases were removed from raw reads using a combination of Trimmomatic version 0.39

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

743 (Bolger et al., 2014) and Cutadapt version 3.2 (Martin et al., 2011) softwares. Reads whose length
744 after trimming was <35 nt were discarded. Alignment to mouse GRCm38 genome and Ensembl
745 87 transcriptome was performed using STAR aligner version 2.7.7a (Dobin et al., 2013). Alignment
746 files were further processed by collapsing PCR duplicates using the MarkDuplicates tool included in
747 the Picard suite version 2.24.1 (http://broadinstitute.github.io/picard/) and discarding the multi-
748 mapped reads using BamTools version 2.5.1 (Barnett et al., 2011). Properly paired reads were
749 extracted using SAMtools version 1.7 (Li et al., 2009). GRCm38 genome was divided into 200 bp
750 long non-overlapping bins using the BEDtools makewindows tool (Quinlan et al., 2010). Properly
751 paired fragments falling in each bin were counted using the BEDtools intersect tool filtering out reads
752 mapping to rRNAs, tRNAs or mitochondrial genome in order to create sample-specific count files.
753 These files were given as input to Piranha version 1.2.1 (Uren et al., 2012) using *–x -s –u 0* parameters
754 to call significant bins for MATR3 Ip and IgG samples. BEDtools intersect was used to assign each
755 genomic bin to genes using Ensembl 87 annotation. For each gene the bin signal distribution in the
756 input sample was calculated after normalization of fragment counts by the total number of mapping
757 fragments. Ip significant bins presenting normalized signal lower than the upper-quartile value of the
758 related gene distribution were filtered out. After this filter, significant bins belonging to Ip samples
759 were merged using BEDTools merge tool. The number of fragments overlapping identified peaks and
760 the number overlapping the same genomic region in the Input sample were counted and used to
761 calculate fold enrichment (normalized by total mapping fragments counts in each data set), with
762 enrichment P-value calculated by Yates' Chi-Square test or Fisher Exact Test where the observed or
763 expected fragments number was below 5. Benjamini-Hochberg FDR procedure was applied for
764 multiple testing corrections. Peaks presenting log2 fold enrichment over Input >2 and FDR < 0.05 in
765 both Ip samples were selected as enriched regions (**Table S3).** BEDtools intersect tool was used to
766 annotate such regions based on their overlap with Ensembl 87 gene annotation and to filter out
767 transcripts hosting regions enriched in the IgG sample. Furthermore, htseq-count software (Anders et

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

al., 2015) with *-s no -m union -t gene* parameters was used to count reads from deduplicated BAM files. Peaks overlapping transcripts with Input CPM (counts per million) > Ip CPM in both Ip samples were filtered out. Gene Ontology analysis was performed on protein coding genes overlapping enriched regions using WebGestalt R package (Liao et al., 2019). Bigwig of normalized coverage (RPM, reads per million) files were produced using bamCoverage 3.5.1 from deepTools tool set (http://deeptools.readthedocs.io/en/develop) on BAM files of uniquely mapping and deduplicated reads using *--normalizeUsing CPM* parameter (Ramirez et al., 2014). Normalized coverage tracks were visualized with IGV software (https://software.broadinstitute.org/software/igv/) and Gviz R package (Hahne et al., 2016).

Protein analyses

Protein extracts were prepared and analysed by western blot as in Desideri et al., 2020. See **Table S5** for antibodies details.

Single cell transcriptomics

scRNAseq analysis was performed on publicly available datasets of E12.5 mouse hearts (SRR10969391, Jackson-Weaver et al. 2020). FASTQ reads were aligned, filtered and counted through the Cell Ranger pipeline (v4.0) using standard parameters. GRCm38 genome version was used in alignment step and annotation refers to Ensembl Release87**.** The dataset was cleaned (nFeature_RNA > 200 and < 6000, percent.mt > 0 and < 5, nCount_RNA > 500 and < 40000) and cells were clustered using Seurat 4.0.5 (Stuart et al., 2019). Cluster uniformity was then checked using COTAN (Galfrè et al., 2020) by evaluating if less than 1% of genes were over the threshold of 1.5 of GDI. If a cluster resulted not uniform, with more than 1% of genes above 1.5, a new round of clustering was performed. After this iterative procedure, the few remaining cells not fitting any cluster, were discarded. A dataset of 4014 cells and 34 clusters was obtained. COTAN function

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

cotan_on_cluster_DE (github devel branch) was used to obtain a correlation coefficient and a *p*-value for each gene and each cluster. From the correlation matrix, a cosine dissimilarity between clusters was estimated and used to plot a dendrogram (with the ward.D2 algorithm, **Figure 1-figure supplement 1D**). The tree was used to decide which clusters could be merged. Cell type for each final cluster was assigned based on a list of markers (Jackson-Weaver et al., 2020; Li et al., 2016) as follows. Cardiomyocytes (1834 cells): Myh6+, Nppa+, Atrial CM (333 cells): are also Myl1+, Myl4+ , Ventricular CM (1289): are also Myl2+, Myl3+, interventricular septum CM (117 cells): are also TBX20+, Gja5+ (Franco et al., 2006), Venous Pole CM (95 cells): also Osr1+ (Meilhac et al., 2018), Outflow Tract CM (72 cells): also Isl1+, Sema3c+, Neural crest cells (there are two clusters expressing their markers: 1.NC with 48 cells and 2.NC with 68 cells): Msx1+, Twist1+, Sox9+, Epicardial cells (359 cells): Cebpb+, Krt18+, Fibroblasts like cells (278 cells): Tcf21+, Fn1+, Endothelial cells (168 cells): Klf2+, Pecam1+, Cdh5+, Smooth muscle cells (710 cells): Cnn1+, Acta2+, Tagln2+, Tagln+, Hemopoietic myeloid cells (79 cells): Fcer1g+, Hemopoietic red blood cells (397 cells): Hba-a1+. The final UMAP plot with the cell assignment is shown in **Figure 1-figure supplement 1E**. Heatmap in **Figure 1-figure supplement 1F** shows a coherent assignment between final clusters and cell type with additional marker genes.

Echocardiography

The echocardiographer was blinded to the phenotypes. Mice were anaesthetised with 2.5% Avertin (Sigma T48402) to perform echocardiographic structural (measurement of left ventricular diameters and wall thickness) and functional (fractional shortening) analyses with a VEVO 3100 (Visualsonics) using a mx550d probe. We used avertin since it does not induce significant cardiodepressant effects, potentially affecting our echocardiographic experiments compared to ketamine combinations, such as ketamine+xylazine. The fractional shortening (FS) of the left ventricle was calculated as FS% = (left ventricular end-diastolic diameter (LVEDD)-left ventricular end-systolic diameter

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(LVESD)/LVEDD) x 100, representing the relative change of the left ventricular diameters during the cardiac cycle. The mean FS of the left ventricle was determined by the average of FS measurements of the left ventricular contraction over 3 beats.

Data accessibility

The data presented in this study will be openly available in NCBI Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/), reference number GSE200878 for RNA-seq data and GSE200877 for MATR3 CLIP-seq data.

Statistical methods and rigor

Statistical tests, p-values, and n for each analysis are reported in the corresponding figure legend. For each experiment, at least three individual animals or pools of littermates were used (see each figure legend for detail). No sex bias was introduced by randomly choosing among male and female mice. All analyses were performed by 1 or more blinded investigators.

Competing interest

The authors declare no competing interests.

Acknowledgments

The authors acknowledge Pietro Laneve, Francesca Pagano and Andrea Cipriano for helpful discussion, Alessandro Calicchio for cloning the probe templates for *in-situ* hybridization analyses and Marcella Marchioni for technical help. This work was supported by grants from Sapienza University (prot. RM11916B7A39DCE5 and RM12117A5DE7A45B) and POR FESR Lazio 2020-T0002E0001 to MB and ERC-2019-SyG (855923-ASTRA), AIRC (IG 2019 Id. 23053), EC-H2020

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

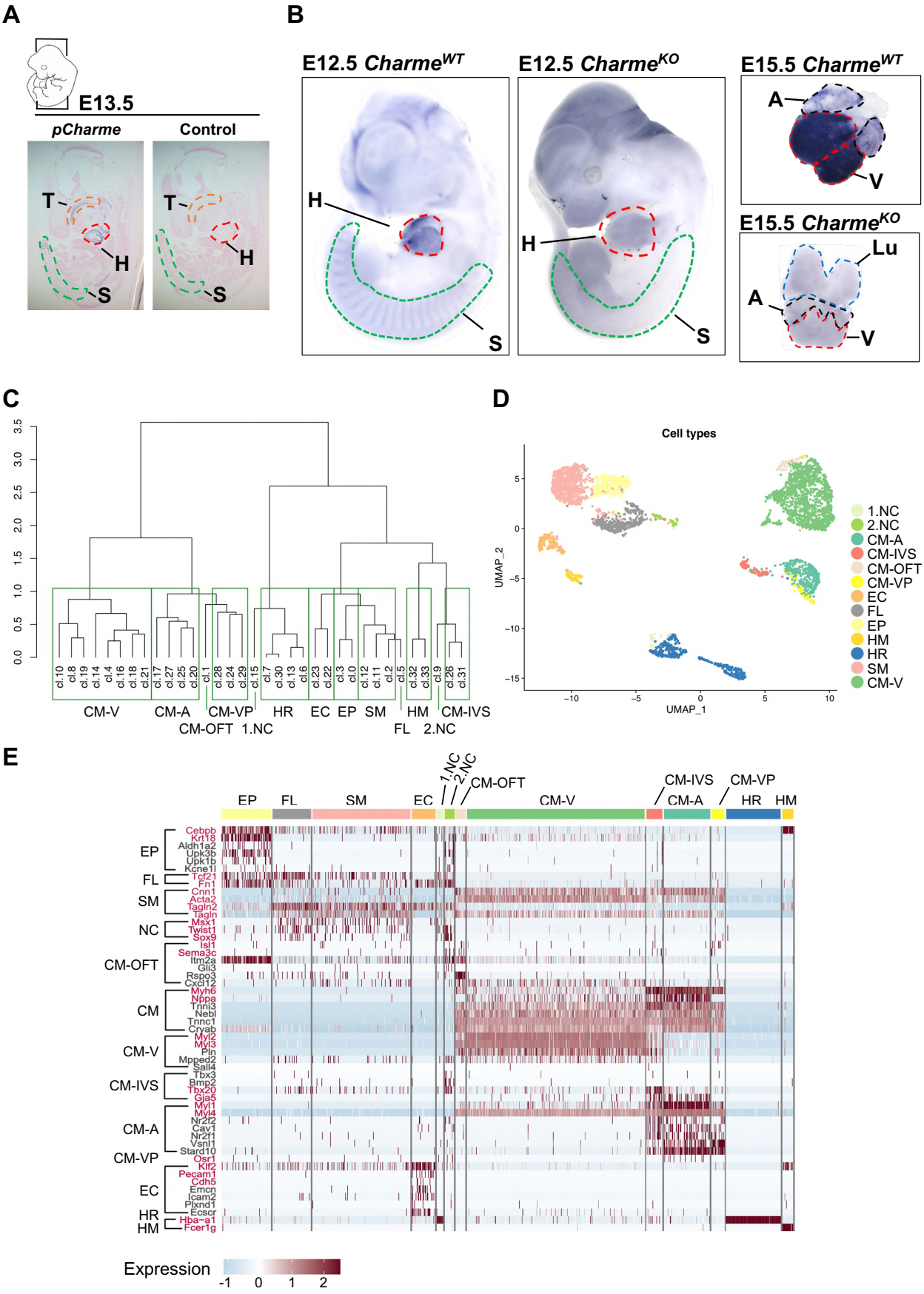
Marie Sklodowska-Curie Action ITN 2016 PN 721890; PRIN 2017 (2017P352Z4) and H2020 Program "Sapienza Progetti Collaborativi" to IB.

Author contribution

M.B. designed, conceived and supervised the study; V.T. performed immunofluorescence assays on mouse sections, WB and subcellular fractionation analysis on heart extracts, MATR3-CLIP assay and cardiac primary cells isolation and transfection; G.B. performed TSS usage, TF binding motif and gene expression analysis on cardiac extracts; F.D. helped to perform MATR3-CLIP assay and carried out its validation and helped the retrieval of cardiac primary cells and their transfection; A.S. performed the bioinformatic analysis on RNA-seq and CLIP-seq data; T. S. performed RNA-FISH and RNA-FISH/IF assays; S.G. performed bioinformatic analysis on the cardiac single-cell dataset; D.M. prepared the libraries for RNA-seq experiments; L.S. and V.V. performed echocardiography experiments that were supervised by S.S.; E.P. performed *in situ* hybridization and whole-mount *in situ* hybridization assays; C.N. and A.M. provided expertise and supervised the experiments in mice. The paper was written by M.B with major contributions from I.B, V.T., G.B. and F.D. and suggestions from all the other authors.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 1-figure supplement 1



bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

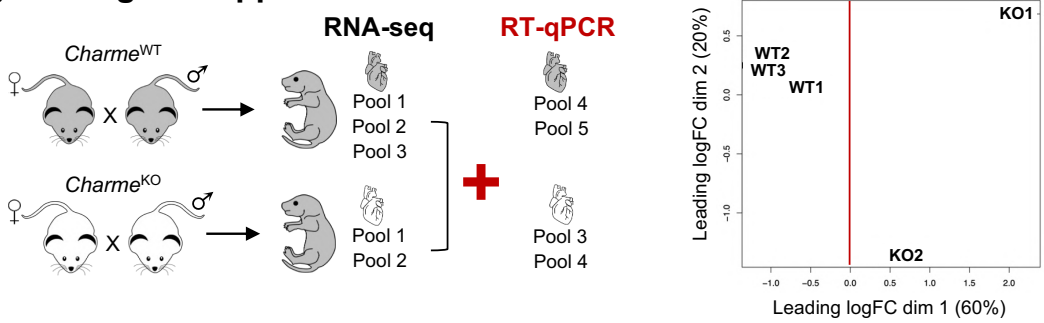
Figure 1-figure supplement 1

A) *In-situ* hybridization (ISH) performed on E13.5 embryonic cryo-sections using digoxigenin-labelled RNA antisense (*pCharme*, left panel) or sense (control, right panel) probes against *pCharme*. T: Tongue; H: Heart; S: Somites. **B)** Whole-mount *in-situ* hybridization (WISH) performed on *Charme*^{WT} and *Charme*^{KO} intact embryos (E12.5, left panels) and hearts (E15.5, right panels). Signal is specifically detected in heart (H, red line) and somites (S, green line). The specificity of the staining can be appreciated by the complete absence of signal in *Charme*^{KO} samples. **C)** Dendrogram showing the relationships between homogeneous clusters. It was obtained using 1 - cosine similarity between clusters computed on the COTAN correlation values for each gene in each cluster. All the informative transcriptome was used to create a hierarchical clustering between homogeneous cell clusters. Green boxes mark which clusters were merged for the final clustering (**Figure 1-figure supplement 1D** and **S1E).** CM: Cardiomyocytes, A-CM: Atrial-CM, V-CM: Ventricular-CM, ISV: Interventricular Septum, VP: Venous Pole, OFT: Outflow Tract, NC: Neural Crest cell, EP: Epicardial cells, FL: Fibroblasts like cells, ED: Endothelial Cells, SM: Smooth Muscle cells, HM: Hemopoietic Myeloid cells, HR: Hemopoietic Red blood cells. **D)** Seurat (Stuart et al., 2019) UMAP plot coloured by final cell assignments. See **Materials and Methods** for details. **E)** Heatmap was generated by the Seurat DoHeatmap function (Stuart et al., 2019) and represents the log normalized expression of cell identity markers (listed on the left) for each cell of the identified sub-populations. Genes used for cell cluster assignment are marked in red. In the heatmap red represents the maximum expression value, light blue the minimum. Correspondence between gene markers and cell types is indicated on the left.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 2-figure supplement 1

A



B

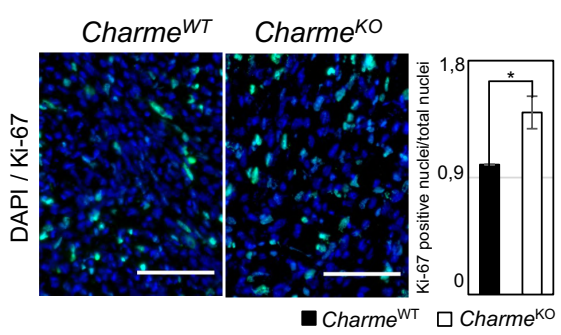
Up-regulated genes

Gene Seta	Description	Size	Expect	Ratio	P Value	FDR	-Log$_{10}$FDR
GO:0007049	Cell cycle	1204	53.926	2.6703	8.65E-40	0	11.70
GO:0051301	Cell division	454	20.334	3.9343	8.65E-40	0	11.70
GO:0007017	Microtubule-based process	521	23.335	2.5712	8.65E-12	3.30E-09	8.48
GO:0051276	Chromosome organization	816	36.548	2.0247	2.44E-09	5.32E-07	6.27
GO:0007010	Cytoskeleton organization	937	41.967	1.9301	3.61E-09	7.43E-07	6.13
GO:0042493	Response to drug	629	28.172	1.9878	5.03E-07	6.0792E-05	4.22
GO:0019752	Carboxylic acid metabolic process	582	26.067	2.0332	5.19E-07	6.1793E-05	4.21
GO:0022414	Reproductive process	782	35.025	1.8558	6.74E-07	7.6659E-05	4.12
GO:0009636	Response to toxic substance	340	15.228	2.364	1.2898E-06	0.00013101	3.88
GO:0033043	Regulation of organelle organization	919	41.161	1.7249	2.9369E-06	0.00025717	3.59

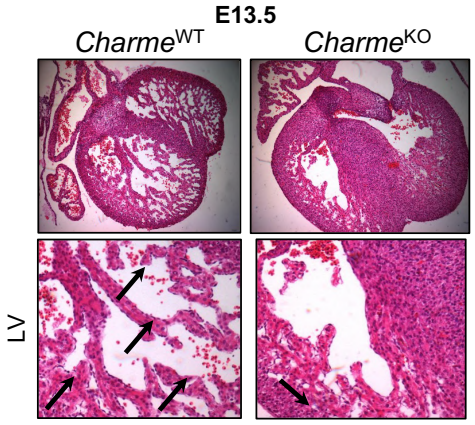
Down-regulated genes

Gene Set	Description	Size	Expect	Ratio	P Value	FDR	-Log$_{10}$FDR
GO:0009653	Anatomical structure morphogenesis	1624	43.33	1.7309	5.77E-07	0.0043983	2.36
GO:0072359	Circulatory system development	802	21.398	2.0563	3.1338E-06	0.0079579	2.10
GO:0009790	Embryo development	740	19.744	2.0766	5.6073E-06	0.010679	1.97
GO:0006629	Lipid metabolic process	748	19.957	2.0544	7.2954E-06	0.011115	1.95
GO:0032787	Monocarboxylic acid metabolic process	342	9.1248	2.6302	0.000014439	0.014156	1.85
GO:0051239	Regulation of multicellular organismal process	1852	49.413	1.5583	0.000022223	0.016748	1.78
GO:0007399	Nervous system development	1374	36.659	1.6367	0.000060188	0.026971	1.57
GO:0042592	Homeostatic process	1142	30.469	1.7066	0.000071747	0.027428	1.56
GO:0009888	Tissue development	1180	31.483	1.6517	0.00016629	0.045853	1.34

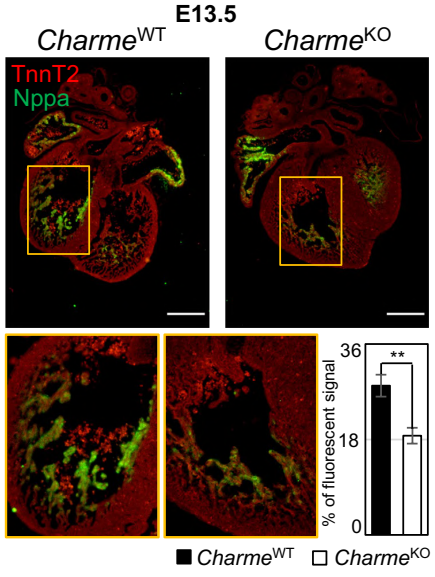
C



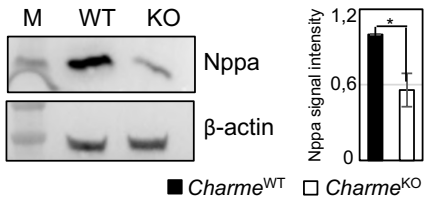
D



E



F



880

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

G

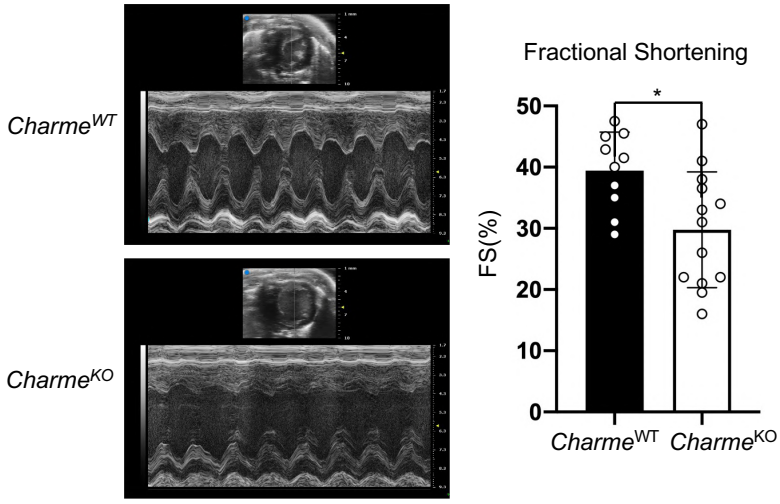
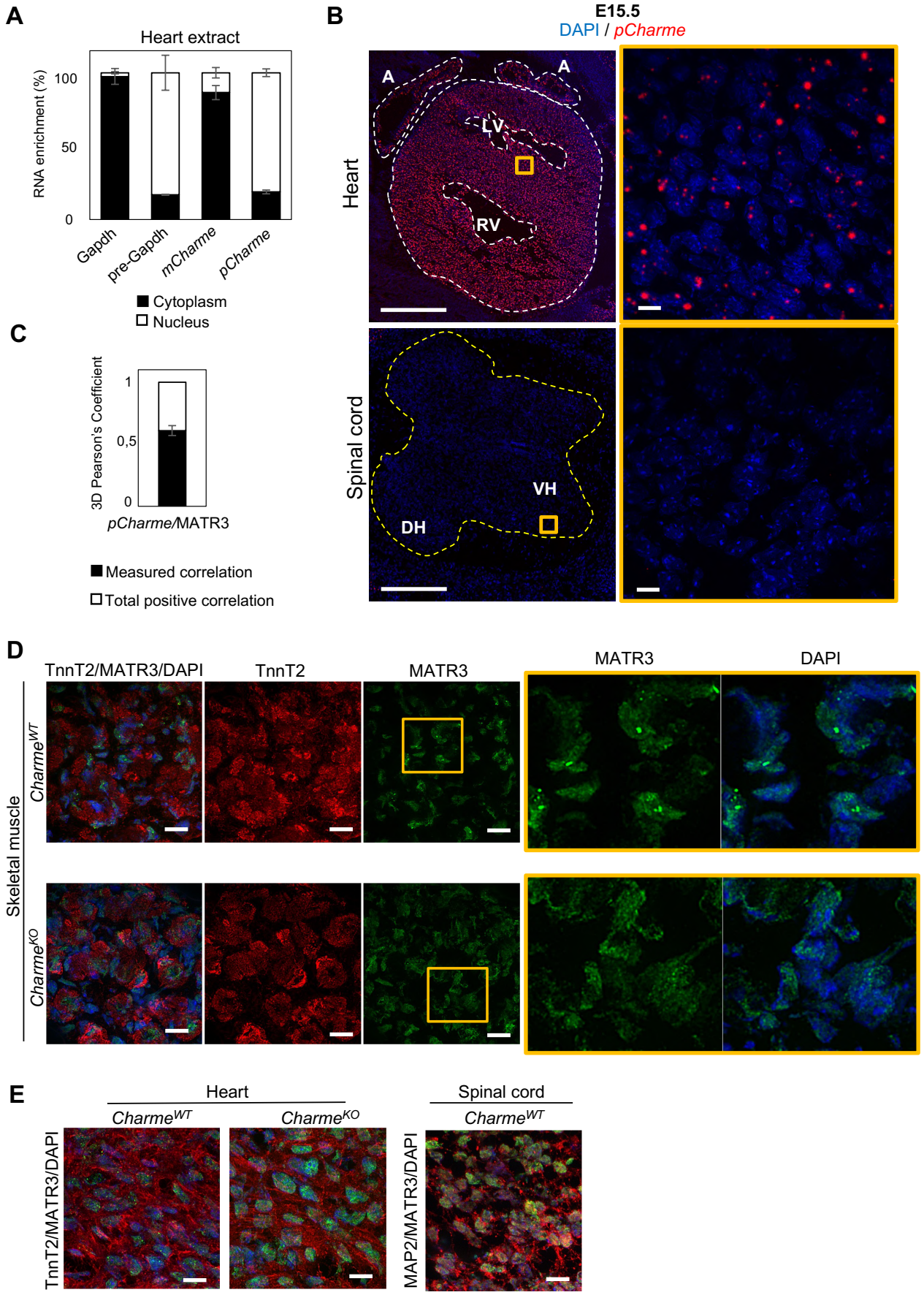


Figure 2-figure supplement 1

A) Left panel: schematic overview of the workflow used to identify DEGs from *Charme*^{WT} and *Charme*^{KO} cardiac transcriptomes. Right panel: multi-dimensional scaling plot of leading fold-change between each pair of *Charme*^{WT} and *Charme*^{KO} RNA-seq samples. Plot was obtained by using the plotMDS function from edger package (https://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf). **B)** Table reports the over-represented GO categories of biological processes in decreasing order of $-\text{Log}_{10}$FDR as identified by WebGestalt (http://www.webgestalt.org). All the represented categories show a FDR value <0.05. **C)** Left panel: Representative image from immunostaining for Ki-67 (green) and DAPI (blue) on C*harme*^{WT} and *Charme*^{KO} (PN) cardiac sections. Data are mean ± SEM.; n = 4. Each image is representative of three individual samples. Scale bars: 70 µm; Right panel: quantification of Ki-67 positive nuclei/total nuclei on C*harme*^{WT} and *Charme*^{KO} heart sections from postnatal mice. Data are mean ± SEM.; n = 4. Each image is representative of four individual samples. **D)** Haematoxylin eosin staining from *Charme*^{WT} and *Charme*^{KO} E13.5 hearts transverse sections (upper panels) and their magnifications (lower panels). Black arrows indicate trabeculated myocardium. Scale bars, 50 µm. **E)** Representative images from immunostaining for Nppa (green) and TnnT2 (red) in *Charme*^{WT} and *Charme*^{KO} (E13.5) cardiac sections. Scale bar: 300 µm. Quantification of the area covered by Nppa fluorescent signal is shown aside. Data represent the mean (%) ± SEM of at least 6 images from biological replicates for each genotype. **F)** Left panel: Representative image from western blot analysis for Nppa in *Charme*^{WT} and *Charme*^{KO} E15.5 hearts extract. β-actin was used as a loading control. Right panel: Quantification of Nppa signal intensity relative to β-actin. Data are mean ± SEM (n = 3). **G)** Representative M-mode echocardiographic track of *Charme*^{WT} and *Charme*^{KO} 9-12 months aged mice. Quantification of heart function was evaluated by FS (Fractional Shortening) =(EDD-ESD)/EDD); EDD= end-diastolic diameter and ESD= end-systolic diameter. Data are mean ± SEM (n = 10-13). Data information: *p < 0.05, **p < 0.01, ***p < 0.001, unpaired Student's t test.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 3-figure supplement 1



909

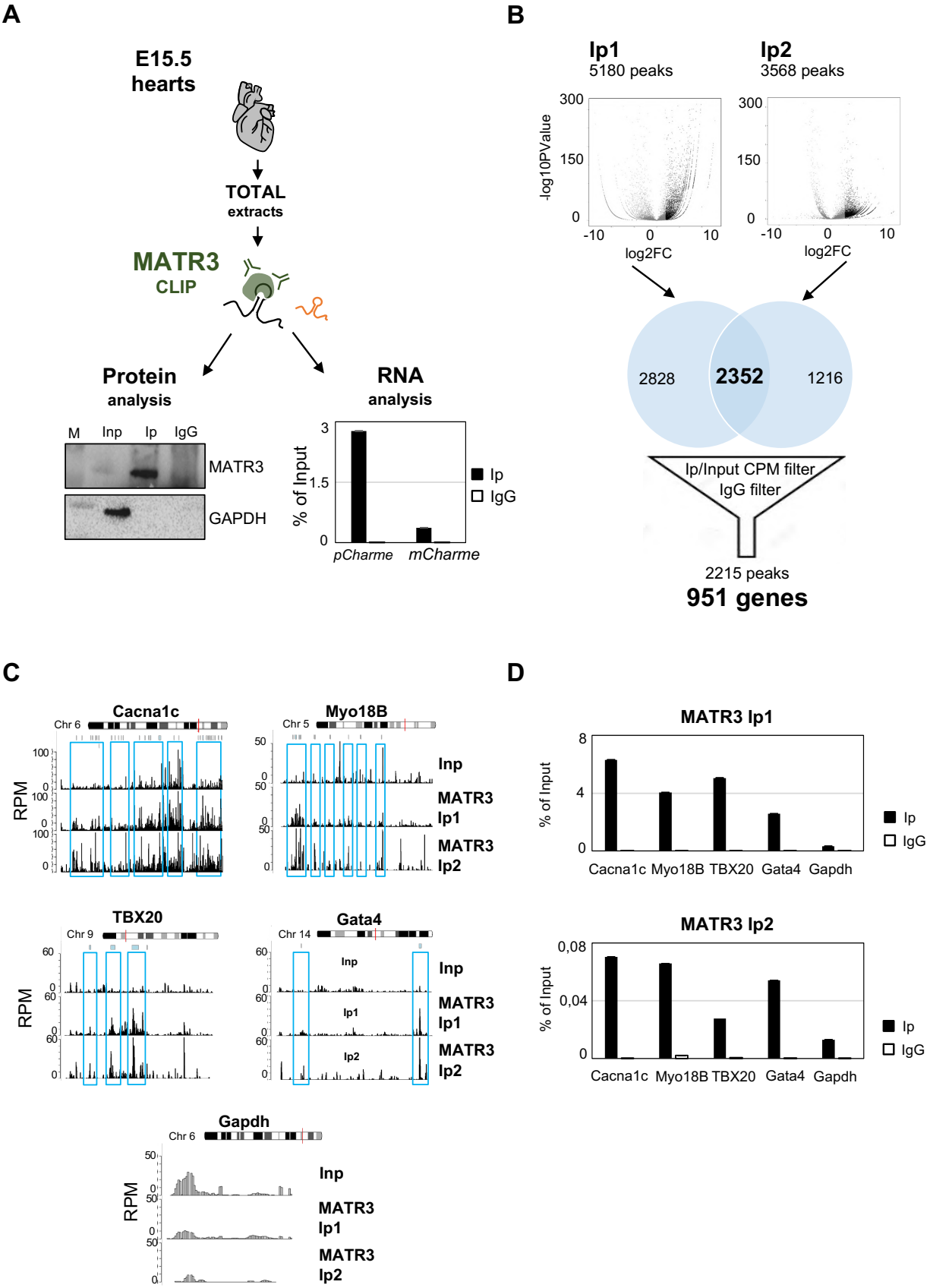
bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 3-figure supplement 1

A) Left panel: Quantification of the subcellular distribution of *pCharme* and *mCharme* in cardiac tissues from PN mice. Histogram shows the RT-qPCR quantification of the relative % of RNA abundance in cytoplasmic *versus* nuclear compartments. GAPDH and pre-GAPDH RNAs were used, respectively, as cytoplasmic and nuclear controls. **B)** RNA-FISH for *pCharme* (red) and DAPI staining (blue) in *Charme*^{WT} hearts and spinal cord from E15.5 embryos sections (left panels) and their magnification (right panels). Signal is specifically detected in whole heart (white dashed line) while no signal is detected in spinal cord (yellow dashed line). A: Atria; LV and RV: Left and Right ventricle; DH and VH: Dorsal and Ventral horn. Scale bars, 500 µm; 10 µm for magnifications. **C)** 3D Pearson's correlation coefficient of *pCharme*/MATR3 overlapping signals. Histogram indicates the mean ±SEM calculated over 237 colocalized signals inside the nuclei from 3 independent experiments. **D)** Immunostaining for MATR3 (green), TnnT2 (red) and DAPI (blue) in *Charme*^{WT} and *Charme*^{KO} skeletal muscles from E15.5 embryos cryosections. The selected regions of interest (orange squares) were digitally enlarged on the right panels. Each image is a representative of three individual samples. Scale bars, 10 µm. **E)** Left panel: Immunostaining for MATR3 (green), TnnT2 (red) and DAPI (blue) in *Charme*^{WT} and *Charme*^{KO} heart from E15.5 embryos. Right panel: Immunostaining for MATR3 (green), Map2 (red) and DAPI (blue) in *Charme*^{WT} and *Charme*^{KO} spinal cord from E15.5 embryos. Each image is representative of three individual samples. Scale bars, 10 µm.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 4-figure supplement 1



bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 4-figure supplement 1

A) Schematic representation of MATR3-CLIP assay performed on total extracts from embryonal hearts. Left panel: representative image of western blot analysis performed to test MATR3 recovery in the Ip and IgG samples. GAPDH protein serves as a loading control. Right panel: RT-qPCR quantification of *pCharme and mCharme* transcripts recovery in the Ip and IgG samples. Values are expressed as percentage (%) of Input. **B)** Schematic representation of the workflow of MATR3 CLIP-seq analysis: the Volcano plots represent the fold-enrichment over Input (log2 fold-enrichment, x-axis) and significance (-log10 PValue, y-axis) of MATR3 peaks in Ip1 sample (left panel) and Ip2 sample (right panel). Each dot represents a peak and the black dots represent significantly enriched peaks (log2 Fold enrichment > 2 and FDR < 0.05). The Venn diagrams depict the intersection of significantly enriched peaks identified in Ip1 and Ip2 that, after filters applying, correspond to 951 common MATR3-bound transcripts. See **Materials and Methods** for details. **C)** Genomic visualization of Input (Inp), Ip1 and Ip2 normalized read coverage tracks (RPM) from MATR3 CLIP-seq on Cacna1c, Myo18B, TBX20, Gata4 and GAPDH loci. MATR3 peaks displaying log2 Fold enrichment > 2 in both Ip1 and Ip2 samples compared to Input are depicted as light blue squares and are demarcated by light blue boxes. Plot obtained using Gviz R package. **D)** RT-qPCR quantification of Cacna1c, Myo18B, TBX20 and Gata4 transcripts recovery in Ip1 (upper panel), Ip2 (lower panel) and IgG samples. GAPDH transcript serves as negative control. Values are expressed as percentage (%) of Input.

REFERENCES

Akerberg BN, Gu F, VanDusen NJ, Zhang X, Dong R, Li K, Zhang B, Zhou B, Sethi I, Ma Q, et al. 2019. A reference map of murine cardiac transcription factor chromatin occupancy identifies dynamic and conserved enhancers. *Nat Commun* **10**: 4907.

Ajima R, Akazawa H, Kodama M, Takeshita F, Otsuka A, Kohno T, Komuro I, Ochiya T, Yokota J. 2008. Deficiency of Myo18B in mice results in embryonic lethality with cardiac myofibrillar aberrations. *Genes Cells* **13**:987-99.

Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**:166-169.

Anderson K, Anderson D, McAnally J, Shelton JM, Bassel-Duby R, Olson EN. 2016. Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. *Nature* **539**: 433–436.

Anderson DM, Anderson KM, Nelson BR, McAnally JR, Bezprozvannaya S, Shelton JM, Bassel-Duby R, & Olson EN. 2021. A myocardin-adjacent lncRNA balances SRF-dependent gene transcription in the heart. *Genes & development* **35**: 835–840.

Ang YS, Rivas RN, Ribeiro A, Srivas R, Rivera J, Stone NR, Pratt K, Mohamed T, Fu JD, Spencer CI, et al. 2016. Disease Model of GATA4 Mutation Reveals Transcription Factor Cooperativity in Human Cardiogenesis. *Cell* **167**: 1734–1749.

Aydin B, Kakumanu A, Rossillo M, Moreno-Estellés M, Garipler G, Ringstad N, Flames N, Mahony S, & Mazzoni EO. 2019. Proneural factors Ascl1 and Neurog2 contribute to neuronal subtype identities by establishing distinct chromatin landscapes. *Nature Neuroscience* **22**: 897.

Azad FM, Polignano LI, Proserpio V, Oliviero S. 2021. Long Noncoding RNAs in Human Stemness and Differentiation. *Trends in Cell Biology* **31**: 542-545.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Ballarino M, Cazzella V, D'Andrea D, Grassi L, Bisceglie L, Cipriano A, Santini T, Pinnarò C, Morlando M, Tramontano A, Bozzoni I. 2015. Novel Long Noncoding RNAs (lncRNAs) in Myogenesis: a miR-31 Overlapping lncRNA Transcript Controls Myoblast Differentiation. *Molecular and Cellular Biology*, **35**: 728–736.

Ballarino M, Cipriano A, Tita R, Santini T, Desideri F, Morlando M, Colantoni A, Carrieri C, Nicoletti C, Musarò A, Carroll DO, Bozzoni I. 2018. Deficiency in the nuclear long noncoding RNA Charme causes myogenic defects and heart remodeling in mice. *The EMBO Journal* **37**: e99697.

Banerjee A, Vest KE, Pavlath GK, Corbett AH. 2017. Nuclear poly(A) binding protein 1 (PABPN1) and Matrin3 interact in muscle cells and regulate RNA processing. *Nucleic Acids Research* **45**: 10706–10725.

Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT. 2011. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* **27**:1691-2.

Bhat P, Honson D, Guttman M. 2021. Nuclear compartmentalization as a mechanism of quantitative control of gene expression. *Nat Rev Mol Cell Biol* **22**:653-670.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bouveret R, Waardenberg AJ, Schonrock N, Ramialison M, Doan T, de Jong D, Bondue A, Kaur G, Mohamed S, Fonoudi H, et al. 2015. NKX2-5 mutations causative for congenital heart disease retain functionality and are directed to hundreds of targets. *eLife* **4***:* e06942.

Bruneau BG. 2013. Signaling and transcriptional networks in heart development and regeneration. *Cold Spring Harbor perspectives in biology* **5**: a008292.

Buckingham M. 2017. Gene regulatory networks and cell lineages that underlie the formation of skeletal muscle. *PNAS* **114**: 5830-5837.

Buonaiuto G, Desideri F, Taliani V, Ballarino M. 2021. Muscle Regeneration and RNA: New Perspectives for Ancient Molecules. *Cells* 10: 2512.

Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, Lemma RB, Turchi L, Blanc-Mathieu R, Lucas J, Boddie P, Khan A, Manosalva Pérez N et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 2022 Jan 7;50(D1):D165-D173. doi: 10.1093/nar/gkab1113. PMID: 34850907; PMCID: PMC8728201.

Cazzella V, Martone J, Pinnarò C, Santini T, Twayana SS, Sthandier O, D'Amico A, Ricotti V, Bertini E, Muntoni F, Bozzoni I. 2012. Exon 45 skipping through U1-snRNA antisense molecules recovers the Dys-nNOS pathway and muscle differentiation in human DMD myoblasts. *Mol Ther.* **20**: 2134-42.

Center for Disease Control and Prevention, (2020, December 9). Data and Statistics on Congenital Heart Defects. https://www.cdc.gov/ncbddd/heartdefects/data.html

Cha HJ, Uyan Ö, Kai Y, Liu T, Zhu Q, Tothova Z, Botten GA, Xu J, Yuan GC, Dekker J, Orkin SH. 2021. Inner nuclear protein Matrin-3 coordinates cell differentiation by stabilizing chromatin architecture. *Nat Commun*. **12**: 6241.

Choquet C, Kelly RG, & Miquerol L. 2019. Defects in Trabecular Development Contribute to Left Ventricular Noncompaction. *Pediatric Cardiology* **40**: 1331–1338.

Chu H, Kohane DS, & Langer R. 2016. RNA therapeutics - The potential treatment for myocardial infarction. *Regenerative Therapy* **4**: 83–91.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Cipriano A, Ballarino M. 2018. The Ever-Evolving Concept of the Gene: The Use of RNA/Protein Experimental Techniques to Understand Genome Functions. *Front Mol Biosci* **6**: 5-20.

Coelho MB, Attig J, Ule J, Smith CW. 2016. Matrin3: connecting gene expression with the nuclear matrix. *Wiley Interdiscip Rev RNA* **7**: 303-15.

D'Ambra E, Santini T, Vitiello E, D'Uva S, Silenzi V, Morlando M, Bozzoni I. 2021. Circ-Hdgfrp3 shuttles along neurites and is trapped in aggregates formed by ALS-associated mutant FUS. *iScience* **24**: 103504.

DeLaughter DM, Bick AG, Wakimoto H, McKean D, Gorham JM, Kathiriya IS, Hinson JT, Homsy J, Gray J, Pu W, et al. 2016. Single-Cell Resolution of Temporal Gene Expression during Heart Development. *Dev Cell* **39**:480-490.

Desideri F, Cipriano A, Petrezselyova S, Buonaiuto G, Santini T, Kasparek P, Prochazka J, Janson, G, Paiardini A, et al. 2020. Intronic Determinants Coordinate Charme lncRNA Nuclear Activity through the Interaction with MATR3 and PTBP1. *Cell Reports* **33**: 108548.

de Soysa TY, Ranade SS, Okawa S, et al. 2019 Single-cell analysis of cardiogenesis reveals basis for organ-level developmental defects. *Nature* **572**: 120-124.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15-21.

Dunn NR, Vincent SD, Oxburgh L, Robertson EJ, Bikoff EK. 2004. Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo. *Development* **131**: 1717-28.

Franco D, Meilhac SM, Christoffels VM, Kispert A, Buckingham M, Kelly RG. 2006. Left and right ventricular contributions to the formation of the interventricular septum in the mouse heart. *Dev Biol*. **294**: 366-375.

Galfrè SG, Morandin F, Pietrosanto M, Cremisi F, Helmer-Citterich M. 2020. COTAN: scRNA-seq data analysis based on gene co-expression, *NAR Genomics and Bioinformatics* 3: lqab072.

Ghafouri-Fard S, Abak A., Talebi SF, Shoorei H, Branicki W, Taheri M, Akbari Dilmaghani N. 2021. Role of miRNA and lncRNAs in organ fibrosis and aging. *Biomedicine & pharmacotherapy* **143**: 112132.

García-Padilla C, Aránega A, Franco D. 2018. The role of long non-coding RNAs in cardiac development and disease. *AIMS Genet* **5**: 124-140.

Goonasekera SA, Hammer K, Auger-Messier M, Bodi I, Chen X, Zhang H, Reiken S, Elrod JW, Correll RN, York AJ, et al. 2012. Decreased cardiac L-type Ca^{2+} channel activity induces hypertrophy and heart failure in mice. *J Clin Invest* **122**: 280-290.

Grote P, Wittler L, Hendrix D, Koch F, Währisch S, Beisaw A, Macura K, Bläss G, Kellis M, Werber M, et al. 2013. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev Cell* **24**:206-214.

Hahne F, Ivanek R. 2016. Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol Biol* **18**: 335-51.

Hall JM, Kainth AS, Rowton MJ, Lu E, Haddad G, Keplinger AJ, Šćepanović J, Perez-Cervantes S, Chan SSK, Piccirilli JA, et al. 2021.The cardiac lncRNA Chantico directly regulates *Cxcl1* chemokine transcription. *bioRxiv* doi: 10.1101/2021.12.23.474040.

Han P, Li W, Lin CH, Yang J, Shang C, Nurnberg ST, Jin KK, Xu W, Lin CY, Lin CJ, et al. 2014. A long noncoding RNA protects the heart from pathological hypertrophy. *Nature* **514**: 102–106.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Han X, Zhang J, Liu X, Fan X, Ai S, Luo Y, Li X, Jin H, Luo S, Zheng H, et al. 2019. The lncRNA Hand2os1/Uph locus orchestrates heart development through regulation of precise expression of Hand2. *Development* **146**: dev176198.

Horsthuis T, Houweling AC, Habets PE. MH, De Lange FJ, El Azzouzi H, Clout DEW, Moorman AFM, Christoffels VM. 2008. Distinct regulation of developmental and heart disease-induced atrial natriuretic factor expression by two separate distal sequences. *Circulation Research* **102**: 849–859.

Hu G, Niu F, Humburg BA, Liao K, Bendi S, Callen S, Fox HS, Buch S. 2018. Molecular mechanisms of long noncoding RNAs and their role in disease pathogenesis. *Oncotarget* 9: 18648–18663.

Ikeda S, Mizushima W, Sciarretta S, Abdellatif M, Zhai P, Mukai R, Fefelova N, Oka SI, Nakamura M, Del Re DP, et al. 2019. Hippo Deficiency Leads to Cardiac Dysfunction Accompanied by Cardiomyocyte Dedifferentiation During Pressure Overload. *Circ Res.* **124**: 292-305.

Jackson-Weaver O, Ungvijanpunya N, Yuan Y, Qian J, Gou Y, Wu J, Shen H, Chen Y, Li M, et al. 2020. PRMT1-p53 Pathway Controls Epicardial EMT and Invasion. *Cell Rep* **31**:107739.

Kim NJ, Lee KH, Son Y, Nam AR, Moon EH, Pyun JH, Park J, Kang JS, Lee YJ, Cho JY. 2021. Spatiotemporal expression of long noncoding RNA Moshe modulates heart cell lineage commitment. *RNA Biol*. **18**: 640-654.

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

Kodo K, Nishizawa T, Furutani M, Arai S, Yamamura E, Joo K, Takahashi T, Matsuoka R, Yamagishi H. 2009. GATA6 mutations cause human cardiac outflow tract defects by disrupting semaphorin-plexin signaling. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 13933–13938.

Latham SL, Weiß N, Schwanke K, Thiel C, Croucher DR, Zweigerdt R, Manstein DJ, Taft MH. 2020. Myosin-18B Regulates Higher-Order Organization of the Cardiac Sarcomere through Thin Filament Cross-Linking and Thick Filament Dynamics. *Cell Rep* **32**: 108090.

Li G, Xu A, Sim S, Priest JR, Tian X, Khan T, Quertermous T, Zhou B, Tsao PS, Quake SR, et al. 2016. Transcriptomic Profiling Maps Anatomically Patterned Subpopulations among Single Embryonic Cardiac Cells. *Dev Cell* **39**:491-507.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009.1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.

Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. 2019. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res*. **47**:W199-W205.

Luna-Zurita L, Stirnimann CU, Glatt S, Kaynak BL, Thomas S, Baudin F, Samee MAH, He D, Small EM, Mileikovsky M, et al. 2016. Complex Interdependence Regulates Heterotypic Transcription Factor Distribution and Coordinates Cardiogenesis. *Cell* **164**: 999–1014.

MacGrogan D, Münch J, de la Pompa JL. 2018. Notch and interacting signalling pathways in cardiac development, disease, and regeneration. *Nat Rev Cardiol.* **15**: 685-704.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**: 10-12.

Mattick J. 2004. RNA regulation: a new genetics?. *Nat Rev Genet* **5**: 316–323.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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

Morton SU, Quiat D, Seidman, JG, Seidman C E. 2022. Genomic frontiers in congenital heart disease. *Nature reviews. Cardiology* **19**: 26–42.

Moorman AF, Christoffels VM. 2003.Cardiac chamber formation: development, genes, and evolution. *Physiol Rev* **83**: 1223-67.

Nadadur RD, Broman MT, Boukens B, Mazurek SR, Yang X, van den Boogaard M, Bekeny J, Gadek M, Ward T, Zhang M, et al. 2016. Pitx2 modulates a Tbx5-dependent gene regulatory network to maintain atrial rhythm. *Science translational medicine* **8**: 354ra115.

Noguchi S, Arakawa T, Fukuda S, Furuno M, Hasegawa A, Hori F, Ishikawa-Kato S, Kaida K, Kaiho A, Kanamori-Katayama M, et al. 2017. FANTOM5 CAGE profiles of human and mouse samples. *Scientific Data* **4**: 1–10.

Olson E. 1993. Regulation of muscle transcription by the MyoD family. The heart of the matter. *Circulation research* **72**: 1-6.

Ounzain S, Micheletti R, Arnan C, Plaisance I, Cecchi D, Schroen B, Reverter F, Alexanian M, Gonzales C, Ng, SY, Bussotti G et al., 2015. CARMEN, a human super enhancer-associated long noncoding RNA controlling cardiac specification, differentiation and homeostasis. *Journal of molecular and cellular cardiology* **89**: 98–112.

Pagano F, Calicchio A, Picchio V, Ballarino M. 2020. The Noncoding Side of Cardiac Differentiation and Regeneration. *Current Stem Cell Research & Therapy* **15**: 723–738.

Pinheiro A, Naya FJ. 2021. The Key Lnc (RNA)s in Cardiac and Skeletal Muscle Development, Regeneration, and Disease. *Journal of Cardiovascular Development and Disease* **8**: 84.

Ponnusamy M, Liu F, Zhang YH, Li RB, Zhai M, Liu F, Zhou LY, Liu CY, Yan KW, Dong YH, et al. 2019. Long Noncoding RNA CPR (Cardiomyocyte Proliferation Regulator) Regulates Cardiomyocyte Proliferation and Cardiac Repair. *Circulation* **139**: 2668–2684.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841-842.

Quintero-Rivera F, Xi QJ, Keppler-Noreuil KM, Lee JH, Higgins AW, Anchan RM, Roberts AE, Seong IS, Fan X, Lage K, et al. 2015. MATR3 disruption in human and mouse associated with bicuspid aortic valve, aortic coarctation and patent ductus arteriosus. *Human molecular genetics* **24**: 2375–2389.

Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**: W187-91.

Ribeiro DM, Zanzoni A, Cipriano A, Delli Ponti R, Spinelli L, Ballarino M, Bozzoni I, Tartaglia GG, Brun C. 2018. Protein complex scaffolding predicted as a prevalent function of long non-coding RNAs. *Nucleic Acids Research* **46**: 917–928.

Rinn JL, Chang HY. 2012. Genome regulation by long noncoding RNAs. *Annual review of biochemistry* **81**: 145–166.

Ritter N, Ali T, Kopitchinski N, Schuster P, Beisaw A, Hendrix DA, Schulz MH, Müller-McNicoll M, Dimmeler S, Grote P. 2019. The lncRNA Locus Handsdown Regulates Cardiac Gene Programs and Is Essential for Early Mouse Development. *Dev Cell*. **50**:644-657.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Riva P, Ratti A, Venturin M. 2016. The Long Non-Coding RNAs in Neurodegenerative Diseases: Novel Mechanisms of Pathogenesis. *Current Alzheimer research* **13**: 1219–1231.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**:139-40.

Roden C, Gladfelter AS. 2021. RNA contributions to the form and function of biomolecular condensates. *Nat Rev Mol Cell Biol* **22**:183-195.

Sabour D, Machado R, Pinto JP, Rohani S, Sahito R, Hescheler J, Futschik ME, Sachinidis A. 2018. Parallel Genome-wide Profiling of Coding and Non-coding RNAs to Identify Novel Regulatory Elements in Embryonic and Maturated Heart. *Molecular therapy. Nucleic acids* **12**: 158–173.

Samsa LA, Yang B, & Liu J. 2013. Embryonic cardiac chamber maturation: Trabeculation, conduction, and cardiomyocyte proliferation. *American Journal of Medical Genetics, Part C: Seminars in Medical Genetics* **163**: 157–168.

Scheuermann JC, Boyer LA. 2013. Getting to the heart of the matter: long non-coding RNAs in cardiac development and disease. *EMBO J* **32**: 1805-16.

Srivastava D. 2006. Genetic regulation of cardiogenesis and congenital heart disease. *Annual review of pathology* **1**: 199–213.

Steimle JD, Moskowitz IP. 2017. TBX5: A Key Regulator of Heart Development. *Current topics in developmental biology*, **122:** 195–221.

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, Satija R. 2019. Comprehensive Integration of Single-Cell Data. *Cell* **177**:1888-1902.

Tao YK, Zeng H, Zhang GQ, Chen ST, Xie XJ, He X, Wang S, Wen H, Chen JX. 2017. Notch3 deficiency impairs coronary microvascular maturation and reduces cardiac recovery after myocardial ischemia. *Int J Cardiol* **236**:413-422.

Uemura Y, Oshima T, Yamamoto M, Reyes CJ, Costa Cruz PH, Shibuya T, & Kawahara Y. 2017. Matrin3 binds directly to intronic pyrimidine-rich sequences and controls alternative splicing. *Genes to cells : devoted to molecular & cellular mechanisms* **22**: 785–798.

Ulitsky I, Bartel DP. 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell* **154**: 26–46.

Uosaki H, Cahan P, Lee DI, Wang S, Miyamoto M, Fernandez L, Kass DA, Kwon C. 2015. Transcriptional landscape of cardiomyocyte maturation. *Cell Rep* **13**:1705–1716.

Uren PJ, Bahrami-Samani E, Burns SC, Qiao M, Karginov FV, Hodges E, Hannon GJ, Sanford JR, Penalva LOF, Smith AD. 2012. Site identification in high-throughput RNA–protein interaction data. *Bioinformatics* **28**: 3013–3020.

van den Hoogenhof MMG, Beqqali A, Amin AS, van der Made I, Aufiero S, Khan MAF, Schumacher CA, Jansweijer JA, van Spaendonck-Zwarts KY, Remme CA, et al. 2018. RBM20 Mutations Induce an Arrhythmogenic Dilated Cardiomyopathy Related to Disturbed Calcium Handling. *Circulation* **138**: 1330-1342.

Wang K, Holt C, Lu J, Brohus M, Larsen KT, Overgaard MT, Wimmer R, Van Petegem F. 2018. Arrhythmia mutations in calmodulin cause conformational changes that affect interactions with the cardiac voltage-gated calcium channel. *Proc Natl Acad Sci U S A* **115**: E10556-E10565.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.05.498800; this version posted July 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Wang Y, Chen J, Cowan DB, Wang DZ. 2021. Non-coding RNAs in cardiac regeneration: Mechanism of action and therapeutic potential. *Seminars in cell & developmental biology* **118**: 150–162.

Wang Z, Zhang XJ, Ji YX, Zhang P, Deng KQ, Gong J, Ren S, Wang X, Chen I, Wang I, et al. 2016. The long noncoding RNA Chaer defines an epigenetic checkpoint in cardiac hypertrophy. *Nat. Med.* **22**: 1131–1139.

Yang XH, Nadadur RD, Hilvering CR, Bianchi V, Werner M, Mazurek SR, Gadek M, Shen KM, Goldman JA, Tyan L et al. 2017. Transcription-factor-dependent enhancer transcription defines a gene regulatory network for cardiac rhythm. *Elife* **6**: e31683.

Yao RW, Wang Y, Chen LL. 2019. Cellular functions of long noncoding RNAs. *Nature cell biology* **21**: 542–551.

Zhao C, Li G, Li J. 2020. Non-coding RNAs and Cardiac Aging. *Advances in experimental medicine and biology* **1229**: 247–258.

Zhao S, Guo Y, Sheng Q, Shyr Y. 2014. Advanced heat map and clustering analysis using heatmap3. *Biomed. Res. Int* e986048.

Zimmerman MS, Carswell Smith AG, Sable CA, Echko MM, Wilner LB, Olsen HE, Tasew Atalay H, Awasthi A, Bhutta ZA, Boucher JL, et al. 2020. Global, regional, and national burden of congenital heart disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet. Child & adolescent health* **4**: 185–200.