1   **The long noncoding RNA Charme supervises cardiomyocyte maturation by**

2   **controlling cell differentiation programs in the developing heart**

3

4   **Valeria Taliani[1‡†], Giulia Buonaiuto[1†], Fabio Desideri[2], Adriano Setti[1], Tiziana Santini[1], Silvia Galfrè[2],**

5   **Leonardo Schirone[3], Davide Mariani[4], Giacomo Frati[3], Valentina Valenti[3], Sebastiano Sciarretta[3],**

6   **Emerald Perlas[5], Carmine Nicoletti[6], Antonio Musarò[6], Monica Ballarino[1]**

7   [1]Department of Biology and Biotechnologies "*Charles Darwin*", Sapienza University of Rome, Rome, Italy;

8   [2]Center for Life Nano- & Neuro-Science of Istituto Italiano di Tecnologia (IIT), 00161, Rome, Italy;

9   [3]Department of Medical Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina, Italy;

10   [4]Center for Human Technologies, Istituto Italiano di Tecnologia (IIT), Genova, Italy;

11   [5]EMBL-Rome, Epigenetics and Neurobiology Unit, Monterotondo, Italy;

12   [6]DAHFMO-Unit of Histology and Medical Embryology, Laboratory affiliated to Istituto Pasteur Italia-

13   Fondazione Cenci Bolognetti, Sapienza University of Rome, Rome, Italy.

14

15   *Corresponding author:* Monica Ballarino

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

17

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

19   [†]These authors contributed equally to this work

20   **Running Head**

21   Charme functions during development

22   **Key Words**

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

24

25

26

ABSTRACT

Long noncoding RNAs (lncRNAs) are emerging as critical regulators of heart physiology and disease, although the studies unveiling their modes-of-action are still 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. Here, we combined Cap-Analysis of Gene Expression (CAGE), single-cell (sc)RNA sequencing and whole-mount *in situ* hybridization analyses to study pCharme cardiac expression. Since the early steps of cardiomyogenesis, we found the lncRNA being specifically restricted to cardiomyocytes, where it assists the formation of specific nuclear condensates containing MATR3, as well as important RNAs for cardiac development. In line with the functional significance of these activities, pCharme ablation in mice results in a delayed maturation of cardiomyocytes, which ultimately leads to morphological alterations of the myocardium and ventricular hypo-trabeculation. Since congenital anomalies in myocardium are clinically relevant in humans and predispose patients to major complications, the identification of novel genes controlling cardiac morphology becomes crucial. Our study offers unique insights into a novel lncRNA-mediated regulatory mechanism promoting cardiomyocyte maturation and bears 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 precursor cells to mature into all the major cardiac 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

2

al., 2020; Srivastava, 2006; Center for Disease Control and Prevention, 2020). Mutations of transcription factors and their protein cofactors have emerged as causative of a broad spectrum of heart malformations, although only 15-20% of all congenital heart defects have been 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 the next-generation sequencing technologies have revolutionised many areas of the biomedical research, including cardiology, by the discovery of long noncoding RNA (lncRNAs) (Mattick et al., 2004; Cipriano et al., 2018, Rinn and Chang 2012; Ulitsky and Bartel 2013; Yao et al., 2019). LncRNAs 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) 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 (Jusic et al., 2022) and regeneration (Pagano et al., 2020; Wang et al., 2021). Specifically for cardiac development (Pinheiro et al., 2021), most of the lncRNA-mediated mechanisms were dissected by cell culture studies (Klattenhoff et al., 2013; Ounzain et al., 2015; Kim et al., 2021), while *in vivo* characterizations remained limited to few instances (Grote et al. 2013; Anderson et al., 2016; Ritter et al., 2019; Hazra,R,et al., 2022).

The importance of lncRNA for adult heart homeostasis has been often described by linking their aberrant expression to cardiac anomalies (Han et al., 2014; Wang et al., 2016; Ballarino et al., 2018; Han et al., 2019; Ponnusamy et al., 2019; Anderson et al., 2021). Along this line, in 2015 we identified a new collection of muscle-specific lncRNAs (Ballarino et al., 2015) and focused our attention on Charme (Chromatin architect of muscle expression) (mm9 chr7:51,728,798-51,740,770), a mammalian conserved lncRNA gene, whose loss-of-function in mice causes progressive myopathy

3

77   and, intriguingly, congenital heart remodelling (Ballarino et al., 2018). Here, we look at the function

78   of this lncRNA in a developmental window and assign an embryonal origin to the cardiac defects

79   produced by its knockout. By using molecular and whole-tissue imaging approaches, we give an

80   integrated view of Charme locus activation in the developing heart and show that the functional

81   pCharme isoform, retaining intron-1, is expressed already in the fetal tissues, progressively dropping

82   down after birth. Single-cell (sc)RNA sequencing analyses from available datasets also reveal a

83   strong cell type-specific expression of pCharme in the embryonal heart, where it is restricted to

84   cardiomyocytes (CM). In line with this specific location, high-throughput transcriptomic analyses

85   combined with the phenotypic characterization of WT and Charme^{KO} developing hearts, highlight a

86   functional role of the lncRNA in CM maturation and in the developmental formation of trabeculated

87   myocardium. Furthermore, high-throughput sequencing of RNA isolated from fetal cardiac biopsies

88   upon cross-linking and immunoprecipitation (CLIP) of MATR3, a nuclear pCharme interactor

89   (Desideri et al., 2020), reveals the existence of a specific RNA-rich condensate containing pCharme

90   as well as RNAs for important regulators of embryo development, cardiac function, and

91   morphogenesis. In accordance with the functional importance of the pCharme/MATR3 interaction in

92   developing cardiomyocytes, the binding of these transcripts with MATR3 was altered in Charme^{KO}

93   hearts. Moreover, MATR3 depletion in mouse-derived cardiac primary cells leads to the down-

94   regulation of such RNAs, which suggests an active crosstalk between MATR3, pCharme and cardiac

95   regulatory pathways.

96   Understanding the basics of cardiac development from a lncRNA *point-of-view* is of interest not only

97   for unravelling novel RNA-mediated circuitries but also for improving treatment options aimed at

98   enhancing cardiomyogenic differentiation. Indeed, despite the recent advancements in generating

99   cardiomyocytes from pluripotent stem cells through tissue engineering-based methods, most

100 protocols produce immature cells, which lack many attributes of adult cardiomyocytes (Uosaki et al.,

101 2015). Consequently, the cells generated cannot be used for efficient drug screening, modelling of

4

102   adult-onset disease, or as a source for cell replacement therapy. We identify here pCharme as a new

103   non-coding regulator of cardiac maturation and characterise the molecular interactome acting with

104   the lncRNA *in vivo*. This research not only advances our understanding of heart physiology but in the

105   next few years may serve as a foundation for new human diagnostic and therapeutic approaches.

106

107   **RESULTS**

108   **Charme locus expression in developing mouse embryos and cardiomyocytes.**

109   Our earlier studies revealed the occurrence of muscle hyperplasia and cardiac remodelling upon the

110   CRISPR/Cas9-mediated Charme loss-of-function in mice (Ballarino et al., 2018; Desideri et al.,

111   2020). Intriguingly, the morphological malformations were clearly displayed in both adult and

112   neonatal mice, strongly suggesting possible roles for the lncRNA during embryogenesis. With the

113   purpose to trace back the developmental origins of Charme functions, we started by analysing the

114   whole collection of FANTOM5 mouse promoterome data to quantify transcription initiation events

115   captured by CAGE-seq across the lncRNA locus (Noguchi et al., 2017). In addition to its cardiac

116   specificity, this profiling revealed the highest expression of the locus during development (E11-E15)

117   followed by a gradual decrease at postnatal stages (**Figure 1A** and **Supplementary File 1**). We then

118   evaluated the expression of the two splice variants, pCharme and mCharme, produced by the locus at

119   E15.5 and postnatally (day 2, after birth). RT-qPCR analyses with specific primers confirmed the

120   CAGE output and revealed that both the isoforms are more expressed at the fetal than the postnatal

121   stage (**Figure 1B**). More importantly, we found pCharme to be 50% more abundant than mCharme,

122   which is particularly intriguing in consideration of the prominent role of the nuclear isoform

123   previously shown in skeletal myogenesis (Desideri et al., 2020). Using *in-situ* hybridization (ISH)

124   approaches, we then profiled by imaging the spatio-temporal expression of Charme during

125   development. These analyses revealed an initial expression of Charme locus already in the tubular

126   heart (**Figure 1C**, E8.5), within territories which will give rise to the future atria and ventricles (V),

5

127 with the inflow (IFT) and outflow (OFT) tracts displaying the highest signals. ISH also confirmed the

128 expression of the locus at later stages of development, both in cardiac tissues and somites (S) (**Figure**

129 **1C**, E13.5; **Figure 1-figure supplement 1A**). A similar expression pattern was found within intact

130 embryos and fetal hearts by whole-mount *in-situ* hybridization (WISH) (**Figure 1D**). To note, specific

131 signals were detected only in wild-type (Charme^{WT}) heart muscles, whereas no signal was found in

132 Charme knockout organs (Charme^{KO}) (**Figure 1D**, left; **Figure 1-figure supplement 1B**), taken from

133 our previously generated mice (Ballarino et al., 2018). As distinct cell subpopulations are known to

134 form the heart, each one carrying out a specialised function in cardiac development and physiology

135 (de Soysa et al., 2019), we then processed publicly available single-cell RNA sequencing (scRNA-

136 seq) datasets from embryonal hearts (E12.5, Jackson-Weaver et al., 2020) for studying at a deeper

137 resolution the locus expression across individual cell types. Upon clustering (**Figure 1E**, upper) the

138 different cardiac cell-types based on the expression of representative marker genes (Li et al., 2016;

139 Jackson-Weaver et al., 2020; Franco et al., 2006; Meilhac et al., 2018) (**Figure 1-figure supplement**

140 **1C-1E**), we found pCharme expression restricted to cardiomyocytes, with an almost identical

141 distribution between the atrial (CM-A), ventricular (CM-V) and other (CM-VP, CM-IVS, CM-OFT)
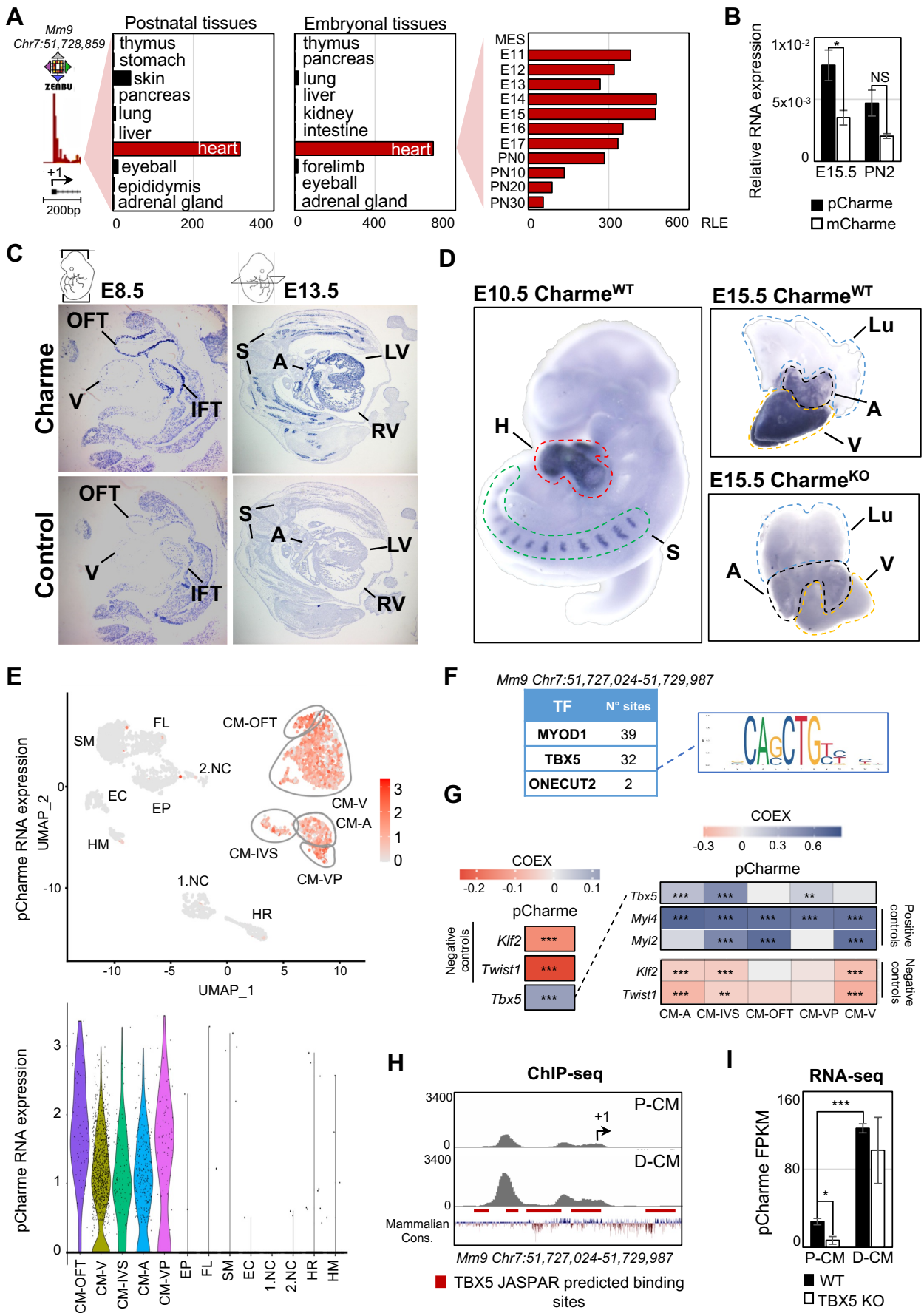
142 cell clusters (**Figure 1E**, lower).

143 These findings offered valuable input for exploring the possible integration of pCharme functions

144 into pathways controlling the maturation of cardiomyocytes. In this direction, we analysed *in silico*

145 the lncRNA promoter to search for transcription factors (TFs) acting as upstream regulators of

146 pCharme expression in cardiomyocytes. In accordance with our previous findings (Ballarino et al.,

147 2015), Jaspar database (https://jaspar.genereg.net; Castro-Mondragon et al., 2022) identified several

148 MYOD1 *consensus* sites in the 1kb region upstream of pCharme transcriptional start site (TSS,

149 **Figure 1F)**, although it is known that this myogenic regulator is not expressed in the heart (Olson,

150 1993; Buckingham et al., 2017). To note, a very low number of *consensus* sites was found for

151 ONECUT2, a TF involved in neurogenesis (Aydin et al., 2019). Searching for cardiac regulators, we

6

152 identified canonical motifs (**Figure 1-figure supplement 1F**) and enriched sites (**Figure 1F**) for the

153 T-box transcription factor-5 (TBX5). In addition to being a key regulator of heart development,

154 known to activate genes associated to CM maturation in early development and morphogenesis at

155 later stages (Nadadur et al., 2016; Steimle et al., 2017), TBX5 was recently found to control the

156 expression of several lncRNAs (Yang et al., 2017), a subset of them intriguingly enriched, as for

157 pCharme, in the chromatin fraction of cardiomyocytes (Hall et al., 2021). By further analysing the

158 scRNA-seq dataset (Jackson-Weaver et al., 2020), we observed a positive and highly significant

159 correlation between *Tbx5* and pCharme (**Figure 1G**, left) in CM, with some differences occurring

160 across the CM cluster subtypes (see for instance CM-IVS/CM-OFT, **Figure 1G**, right). To note,

161 pCharme positively correlates with markers of cardiomyocyte identity, Myosin light chain 2 (*Myl2*)

162 and 4 (*Myl4*), while significant anticorrelation was found for the Kruppel like factor 2 (*Klf2*) and the

163 Twist family BHLH transcription factor 1 (*Twist1*) transcripts, that respectively mark endothelial and

164 neural crest cells (Jackson-Weaver et al., 2020; Li et al., 2016).

165 To provide functional support to a possible role of TBX5 for pCharme transcription in CM, we

166 examined available Chromatin Immunoprecipitation (ChIP)-sequencing (**Figure 1H**) and RNA-seq

167 (**Figure 1I**) datasets from murine cardiomyocytes at progenitor (P-CM) and differentiated (D-CM)

168 stages (Luna-Zurita et al., 2016). In both conditions, ChIP-seq analyses revealed the specific

169 occupancy of TBX5 to highly conserved spots within the pCharme promoter, which increases with

170 differentiation. However, while in differentiated cardiomyocytes pCharme expression was insensitive

171 to TBX5 ablation, the lncRNA abundance was consistently decreased in TBX5 knockout progenitor

172 cells (**Figure 1I**). As progenitor CM represent a cell resource for recapitulating *ex-vivo* the early

173 establishment of the cardiac lineage and for mimicking cardiomyocyte development (Kattman et al.,

174 2011), these results are consistent with the role of TBX5 for Charme transcription at fetal stages. The

175 presence of *consensus* sites for other TFs (**Figure 1-figure supplement 1F**) and the known attitude

176 of TBX5 to interact, physically and functionally, with several cardiac regulators (Akerberg et al.,

177   2019), also predict the possible contribution of other regulators to pCharme transcription which may

178   act, in cooperation or not, with TBX5 during CM maturation.

8

Figure 1.

A

Mm9
Chr7:51,728,859

ZENBU

+1

200bp

Postnatal tissues
thymus
stomach
skin
pancreas
lung
liver
heart
eyeball
epididymis
adrenal gland

0   200   400

Embryonal tissues
thymus
pancreas
lung
liver
kidney
intestine
heart
forelimb
eyeball
adrenal gland

0   400   800

MES
E11
E12
E13
E14
E15
E16
E17
PN0
PN10
PN20
PN30

0   300   600   RLE

B

Relative RNA expression

1x10⁻²: 1×10^{-2}
5x10⁻³: 5×10^{-3}

E15.5   PN2

*   NS

■ pCharme
□ mCharme

C

E8.5   E13.5

Charme
OFT
V   IFT
S   A   LV
RV

Control
OFT
V   IFT
S   A   LV
RV

D

E10.5 Charme^WT
H
S

E15.5 Charme^WT
Lu
A
V

E15.5 Charme^KO
Lu
A   V

E

pCharme RNA expression
UMAP_2

SM   FL   CM-OFT
EC   2.NC
EP   CM-V
HM   CM-A
CM-IVS
1.NC   CM-VP
HR

UMAP_1

3
2
1
0

pCharme RNA expression

3

2

1

0

CM-OFT CM-V CM-IVS CM-A CM-VP EP FL SM EC 1.NC 2.NC HR HM

F

Mm9 Chr7:51,727,024-51,729,987

TF	N° sites
MYOD1	39
TBX5	32
ONECUT2	2

CAGCTG

G

COEX
-0.2 -0.1 0 0.1

pCharme

Negative controls	
Klf2	***
Twist1	***
Tbx5	***

COEX
-0.3 0 0.3 0.6

pCharme

	CM-A	CM-IVS	CM-OFT	CM-VP	CM-V
Tbx5	***	***		**	
Myl4	***	***	***	***	***
Myl2		***	***		***
Klf2	***	***			***
Twist1	***	**			***

Positive controls
Negative controls

H

ChIP-seq

3400   P-CM
0   +1

3400   D-CM
0

Mammalian Cons.

Mm9 Chr7:51,727,024-51,729,987

■ TBX5 JASPAR predicted binding sites

I

RNA-seq

pCharme FPKM

160

80

0

P-CM   D-CM

*   ***

■ WT
□ TBX5 KO

179

9

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 is 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 embryonal 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 the specific samples.

B) Quantitative Reverse Transcription PCR (RT-qPCR) amplification of pCharme and mCharme isoforms in RNA extracts from Charme^{WT} E15.5 and neonatal (PN2) hearts. Data were normalized to *Gapdh* mRNA and represent means ± SEM of 3 pools.

C) *In-situ* hybridization (ISH) performed on embryonal cryo-sections using digoxigenin-labelled RNA antisense (Charme, upper panel) or sense (control, lower panel) probes against Charme. 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 Charme. The specificity of the staining can be appreciated by the complete absence of signal in explanted hearts from Charme^{KO} mice (Ballarino et al., 2018). A: Atria (black line); V: Ventricles (yellow line).

E) Upper panel: UMAP plot showing pCharme expression in single-cell transcriptomes of embryonal (E12.5) hearts (Jackson-Weaver et al., 2020). Lower panel: Violin plot of pCharme expression in the different clusters (see Materials 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 cells, EP: Epicardial cells, FL: Fibroblasts like cells, EC: Endothelial Cells, SM: Smooth Muscle cells, HM: Hemopoietic Myeloid cells, HR: Hemopoietic Red blood cells.

F) *In silico* analysis of MYOD1, TBX5 and ONECUT2 transcription factors (TF) binding sites using Jaspar 2022 (relative profile score threshold=80%) (Castro-Mondragon et al., 2022). MyoD1 and Onecut2 were used as positive and negative controls, respectively. N° sites =number of consensus motifs.

G) COTAN heatmap obtained using the whole scRNA-seq (left) and contrasted subsetted dataset (right) showing pCharme and *Tbx5* expression correlation (COEX). *Myl4* and *Myl2* were used as positive controls for cardiomyocytes while *Klf2* and *Twist1* as negative controls (markers of EC and NC, respectively). See Materials and Methods for details.

H) TBX5 ChIP-seq analysis across Charme promoter in murine precursors (P-CM) and differentiated cardiomyocytes (D-CM) (GSE72223, Luna-Zurita et al., 2016). The genomic coordinates of the promoter, the Charme TSS (+1, black arrow), the TBX5 JASPAR predicted binding sites (red lines) and the mammalian conservation track (Mammalian Cons.) from UCSC genome browser are reported.

I) Quantification of pCharme expression from RNA-seq analysis performed in wild type (WT) and TBX5 knockout (KO) murine P-CM and D-CM (SRP062699, Luna-Zurita et al., 2016). FPKM: Fragments Per Kilobase of transcript per Million mapped reads.

Data information: $*p < 0.05$; $***p < 0.001$; $NS > 0.05$, unpaired Student's t-test.

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 are still pending on the role of pCharme in cardiomyogenesis. As a first step towards the identification of the molecular signature underlying Charme-dependent cardiac anomalies, we performed a differential

gene expression analysis on transcriptome profiles from Charme[WT] and Charme[KO] neonatal (PN2)

hearts (**Figure 2A**; **Figure 2-figure supplement 1A**). At this stage, the heart continues to change

undergoing maturation, a process in which its morphology and cell composition keep evolving

(Padula et al., 2021; Tian et al., 2017).

RNA-seq analysis of whole hearts led to the identification of 913 differentially expressed genes

(DEGs) (FDR<0.1, WT *vs* KO, **Supplementary File 2**), 573 of which were Up-regulated and 340

were Down-regulated in Charme[KO] hearts (**Figure 2B**; **Supplementary File 2**). These results were

confirmed by RT-qPCR analyses performed on gene subsets from independent biological replicates

(**Figure 2C**; **Figure 2-figure supplement 1B**). Similar to pCharme activity in skeletal myocytes

(Ballarino et al., 2018), we did not find DEGs within the neighbouring chromatin environment (370

Kb around Charme) (**Figure 2-figure supplement 1C**).

A Gene Ontology (GO) term enrichment study was then applied separately to the Up- and Down-

regulated DEGs. These analyses revealed that the Up-regulated DEGs were primarily enriched (FDR-

values < 1,0E-10) in cell cycle and cell division categories (**Figure 2D**, left panel), which parallels

with a slight increase in the number of KI-67[+] mitotic Charme[KO] nuclei, as quantified from neonatal

cardiac sections (**Figure 2-figure supplement 1D**). The same study applied to the Down-regulated

DEGs revealed their enrichment to more functional and morphogenetic GO categories and primarily

referred to anatomical structure morphogenesis (FDR 5.77E-07) and circulatory system development

(FDR 3.1338E-06) (**Figure 2D,** right panel). Interestingly, these top-ranked categories include TFs

involved in pivotal steps of embryo development, such as *Smad3* (Dunn et al., 2004) and *Nocth3*

(MacGrogan et al., 2018), and functional components of cardiomyocytes, such as the voltage-

dependent calcium channel subunit *Cacna1c* (Wang et al., 2018), and Myosin-18B (*Myo18B*), known

to regulate cardiac sarcomere organisation (Latham et al., 2020). By expanding the analysis of these

genes from neonatal to embryonal timepoints (E12.5, E15.5 and E18.5), we found that their

expression significantly decreases upon pCharme ablation from the fetal E15.5 stage onward, as

compared to WT (**Figure 2E**). Similarly, the *Myh6*/*Myh7* ratio, which measures CM maturation (England and Loughna, 2013; Scheuermann and Boyer, 2013) and that increases over the wild-type timepoints, displays a gradual decrease in the Charme[KO] hearts from E15.5 onward (**Figure 2F**). To note, the expression of cell-cycle and proliferation genes was not affected by pCharme ablation over the same developmental window (**Figure 2-figure supplement 1E**). Hence, their upregulation in neonatal KO hearts could be due to the impairment of cardiac maturation, which leads to a shift to a more fetal and proliferative state of cardiomyocytes upon birth.

Overall, our findings suggest that during cardiac development, pCharme regulation is achieved through the regulation of morphogenetic pathways whose dysregulation causes a pathological remodelling of the heart. Accordingly, histological analyses performed by haematoxylin and eosin staining of Charme[WT] and Charme[KO] cardiac cryo-sections (E15.5) showed a pronounced alteration of the myocardium, with a clear decrease of the area of the ventricular cavities (**Figure 2G**). The histological investigation was then deepened to the trabeculated myocardium, the tissue directly surrounding ventricle cavities. To this purpose, we performed immunofluorescence staining and western blot analysis for the natriuretic peptide A (NPPA) factor, a known marker of the embryonal trabecular cardiomyocytes (Choquet et al., 2019; Horsthuis et al., 2008). In parallel, heart sections were analysed for cardiac vasculature, since i) the process of trabeculae formation is known to be temporally coupled with the formation of blood vessels during development (Samsa et al., 2013) and, importantly, ii) the "circulatory system development" category was among the main GO enriched for pCharme down-regulated targets. We found that, with respect to WT, mutant hearts display a significant reduction of NPPA[+] trabeculae (**Figure 2H; Figure 2- figure supplement 1F**), which parallels the decreased expression of the *Irx3* and *Sema3a* trabecular markers (Choquet et al., 2019) (**Figure 2-figure supplement 1G**) as well as the diminished density of the capillary endothelium, as imaged by lectin staining (**Figure 2-figure supplement 1H**). Therefore, the fetal expression of pCharme is necessary for the achievement of morphogenetic programs important for CM maturation,

280   myocardial geometry and vascular network formation. Overall, these ensure the preservation of the

281   cardiac function and structure in adulthood, as a progressive deterioration of the systolic function,

282   which becomes significant at 9 months of age (**Figure 2-figure supplement 1I**; **Supplementary File**

283   **3**), was observed in Charme^{KO} mice. To note, a similar alteration in heart efficiency was observed in

284   the murine model with a specific mutation in the pCharme intron-1 (Charme$^{\Delta Int}$) (Desideri et al.,

285   2020), further confirming the distinct activity of this isoform in cardiac processes.

Figure 2

A

CharmeWT
CharmeKO — RNA-seq

B
- Down-regulated
- Up-regulated
- Invariant

C

Up-regulated genes: Ki-67, Ccnb1, Ccna2, Birc5, Bub1 (0.058), Spdl1 (0.055)

Down-regulated genes: Myo18B, Smad3, Notch3, Cacna1c, JUP, Cul7

■ CharmeWT □ CharmeKO

D

GO Biological Processes

Up-regulated genes
- Cell cycle
- Cell division
- Microtubule-based process
- Chromosome organization
- Cytoskeleton organization
- Response to drug
- Carbox. acid metabolic process
- Reproductive process
- Response to toxic substance
- Regulation of organelle organization

Down-regulated genes
- Anatom. structure morphogenesis
- Circulatory system development
- Embryo development
- Lipid metabolic process
- Monocarbox. acid metabolic process
- Reg. of multicell. organism. process
- Nervous system development
- Homeostatic process
- Tissue development

E

E12.5 E15.5 E18.5

Smad3, Notch3, Cacna1c, JUP, Cul7

F

Fetal — Myh7 / Myh6 — Adult

E12.5, E15.5, E18.5, PN2

■ CharmeWT □ CharmeKO

G

E15.5 CharmeWT CharmeKO

Total area, LV, RV, IVS

■ CharmeWT □ CharmeKO

H

E15.5 CharmeWT CharmeKO

TNNT2 NPPA

% of fluorescent signal

■ CharmeWT □ CharmeKO

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

288 **A)** Heatmap visualization from RNA-seq analysis of Charme^{WT} and Charme^{KO} neonatal (PN2) hearts.

289 Plot was produced by heatmap3 (https://cran.r-

290 project.org/web/packages/heatmap3/vignettes/vignette.pdf). Expression values were calculated as

291 FPKM, were log$_2$-transformed and mean-centered. FPKM: Fragments Per Kilobase of transcript per

292 Million mapped reads.

293 **B)** Volcano plots showing differential gene expression from transcriptome analysis of Charme^{WT} *vs*

294 Charme^{KO} PN2 hearts. Differentially expressed genes (DEGs) validated through RT-qPCR (**Figure**

295 **2C**) are in evidence. FC: Fold Change

296 **C)** RT-qPCR quantification of Up-regulated (left panel) and Down-regulated (right panel) DEGs in

297 Charme^{WT} *vs* Charme^{KO} neonatal hearts. Data were normalized to *Gapdh* mRNA and represent means

298 ± SEM of WT (n=5) *vs* KO (n=4) pools

299 **D)** Gene Ontology (GO) enrichment analysis performed by WEBGESTALT

300 (http://www.webgestalt.org) on Up-regulated (left panel) and Down-regulated (right panel) DEGs in

301 Charme^{WT} *vs* Charme^{KO} pools of neonatal hearts. Bars indicate the top categories of Biological

302 processes in decreasing order of – log$_{10}$FDR. All the represented categories show a False Discovery

303 Rate (FDR) value <0.05.

304 **E)** RT-qPCR quantification of pCharme targets in Charme^{WT} and Charme^{KO} extracts from E12.5,

305 E15.5 and E18.5 hearts. DEGs belonging to the GO category "anatomical structure morphogenesis"

306 were considered for the analysis. Data were normalized to *Gapdh* mRNA and represent means ± SEM

307 of WT and KO (n=3) pools.

308 **F)** RT-qPCR quantification of the Myh6/Myh7 ratio in Charme^{WT} and Charme^{KO} extracts from E12.5,

309 E15.5 and E18.5 and neonatal hearts. Data were normalized to *Gapdh* mRNA and represent means ±

310 SEM of WT and KO (n=3) pools. Schematic representation of the physiological *Myh6/Myh7*

311 expression trend is shown.

G) Upper panel: Haematoxylin-eosin staining from Charme[WT] and Charme[KO] E15.5 cardiac transverse sections. Scale bars: 500 μm. Lower panel: Quantification of the total area, the left and right ventricle cavities and the thickness of the interventricular septum (IVS) in Charme[WT] and Charme[KO] E15.5 hearts. For each genotype, data represent the mean ± SEM (n=3).

H) Representative images of Nppa (green) and TnnT2 (red) immunostaining in Charme[WT] and Charme[KO] (E15.5) cardiac sections. Regions of interest (ROI, orange squares) were digitally enlarged on the lower panels. Scale bar: 500 μm. Quantification of the area covered by the Nppa fluorescent signal is shown aside. Data represent the mean (%) ± SEM of WT (n=4) and KO (n=3).

Data information: *p < 0.05; **p < 0.01, NS > 0.05, unpaired Student's t-test

pCharme nucleates the formation of RNA-rich condensates by interacting with MATR3 in fetal cardiomyocytes.

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 was found highly expressed in cardiomyocytes from newborn mice and heterozygous mutations in MATR3 resulted in congenital heart defects (Quintero-Rivera et al., 2015). This raises the intriguing possibility that pCharme may play a role in cardiomyocyte maturation through MATR3. To examine this hypothesis, we first assessed the subcellular localization of pCharme in fetal (E15.5) hearts by biochemical fractionation. RT-qPCR analyses revealed that, in line with what was previously observed in skeletal muscle, in cardiomyocytes pCharme mainly localises in the nucleus, while the fully spliced mCharme is enriched in the cytoplasm (**Figure 3-figure supplement 1A**). We next applied high-resolution RNA-fluorescence *in situ* hybridization (RNA-FISH) to visualise pCharme

337 alone (**Figure 3A**; **Figure 3-figure supplement 1B**) or relatively to MATR3 through a combined

338 immunofluorescence IF/RNA-FISH approach (**Figure 3B**). In agreement with the subcellular

339 fractionation, the imaging experiments confirmed the nuclear localization of pCharme, which exhibits

340 its typical punctuate pattern. Further analysis of the three-dimensional distribution of pCharme and

341 MATR3 revealed a clear colocalization of their signals and the formation of nuclear condensates, as

342 quantified by 3D Pearson's correlation coefficient applied on the overlapping signals (**Figure 3-**

343 **figure supplement 1C**). Based on these results, we then tested if the presence of pCharme could

344 influence the nuclear localization of MATR3. To this end, we performed MATR3 IF assays in wild-

345 type and Charme[KO] fetal (E15.5) muscle biopsies and in spinal cord nuclei (**Figure 3C**; **Figure 3-**

346 **figure supplement 1D-F**). A striking heterogeneous distribution of MATR3 positive signals was

347 observed within the nucleus of wild-type skeletal (**Figure 3-figure supplement 1D**) and cardiac

348 (**Figure 3C**, upper panel) muscles, both expressing the lncRNA. Coherently with their pCharme-

349 dependent formation, these condensates appeared more diffuse in tissues where the lncRNA is not

350 expressed, as Charme[KO] muscles and WT spinal cords (**Figure 3C**, middle and lower panels; **Figure**

351 **3-figure supplement 1D**). More accurate quantification of the MATR3 fluorescence distribution in

352 the nucleus of Charme[KO] cardiomyocytes (Coefficient of variation (CV), **Figure 3-figure**

353 **supplement 1F**), revealed a MATR3 IF pattern which was less discrete and more homogeneous in

354 respect to WT.

Figure 3

A

E15.5

DAPI / pCharme

Heart

A   A

LV

RV

Spinal cord

DH   VH

B

E15.5

Heart

pCharme/MATR3/DAPI

isosurface

y z x

pCharme/MATR3/DAPI

pCharme/MATR3

C

Heart

TNNT2   MATR3   MATR3   DAPI

Charme^WT

Charme^KO

Spinal cord

MAP2   MATR3   MATR3   DAPI
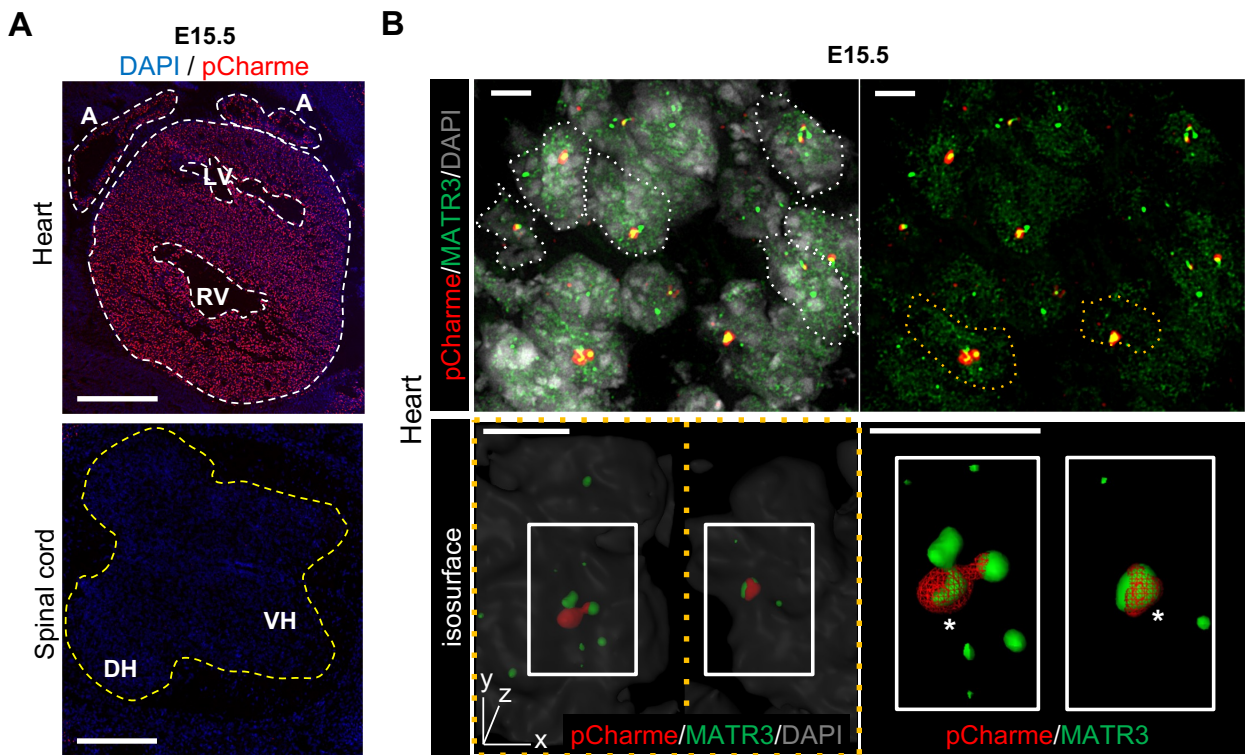
Charme^WT

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

A) RNA-FISH for pCharme (red) and DAPI staining (blue) in Charme^{WT} cardiac and spinal cord from E15.5 tissue sections. Whole heart (white dashed lines), 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} from E15.5 cardiac sections. Dashed lines show the edge of nuclei. Lower panel: selected nuclei (yellow dashed lines in the upper panel) were enlarged and processed for isosurface reconstruction (left panel) and digital magnification (right panel). Overlapped signals are shown (asterisks). Scale bars, 5 μm.

C) Upper panel: Representative images for MATR3 (green), TnnT2 (red) and DAPI (blue) stainings on Charme^{WT} and Charme^{KO} E15.5 cardiac sections. Lower panel: Representative images for for MATR3 (green), Map2 (red) and DAPI (blue) stainings on Charme^{WT} and Charme^{KO} E15.5 spinal cord sections. ROI (orange squares) were digitally enlarged on the right panels. Each image is representative of three individual biological replicates. Scale bars, 10 μm.

The known ability of MATR3 to interact with both DNA and RNA and the high retention of pCharme on the chromatin may predict the presence of chromatin and/or specific transcripts within these MATR3-enriched condensates. In skeletal muscle cells, we have previously observed on a genome-wide scale, a global reduction of MATR3 chromatin binding in the absence of pCharme (Desideri et al., 2020). Nevertheless, the broad distribution of the protein over the genome made the identification of specific targets through MATR3-ChIP challenging. With the purpose to functionally characterise MATR3 interactome, here we applied a more straightforward MATR3 cross-linking immunoprecipitation (CLIP)-seq approach to fetal Charme^{WT} and Charme^{KO} hearts for the identification of MATR3 bound RNA targets (**Figure 4A**). Western blot analyses with antibodies

against MATR3 allowed to test the efficiency of protein recovery after the immunoprecipitation, in both the WT and KO conditions (**Figure 4-figure supplement 1A**). Subsequent analysis of the RNA, led to the identification of 951 cardiac-expressed transcripts significantly bound by MATR3 in the WT heart (log2 Fold enrichment >2 and FDR value <0.05, **Figure 4-figure supplement 1B** and **Supplementary File 4**). Four candidates (*Cacna1c*, *Myo18B*, *Tbx20*, *Gata4*) were selected based on their FDR values (**Supplementary File 4**) for further validation by RT-qPCR, together with *Gapdh* as a negative control (**Figure 4-figure supplement 1C**).

In line with the binding propensities exhibited by MATR3 in other cell systems (Uemura et al., 2017), most of its enrichment was located within the introns of protein-coding mRNA precursors (**Figure 4B**). Strikingly, even though the class of lncRNAs was poorly represented (**Figure 4B**, upper panel), we found pCharme at the top of the MATR3 interactors in Charme[WT] hearts (**Supplementary File 4**). Quantification of the retrieved RNAs by RT-qPCR evidenced the specific enrichment of pCharme but not mCharme in the wild-type samples, which supports the distinctive binding of MATR3 to the nuclear isoform (**Figure 4-figure supplement 1D**). As expected, we observed a strong reduction of MATR3-CLIP signal on Charme RNA in Charme[KO] samples (**Figure 4C**, **Supplementary File 4**).

To gain further insight into the specificity of MATR3 binding, we proceeded with the analysis of the CLIP-identified peaks for the identification of RNA-binding proteins consensus sites, including MATR3. Motif enrichment analysis using 93 RNA-binding motifs catalogued by CISBP-RNA database (Ray D, et al., 2013) pinpointed the MATR3 consensus sequence (UUCUU, Desideri et al., 2020) among the most over-represented pyrimidine-enriched motifs (Evalue < 0.05) (**Figure 4D**). When analysed in respect to the intensity of the CLIP signals, the MATR3 motif was positioned at the peak summit (+/- 50 nt) (**Figure 4E**), close to the strongest enrichment of MATR3, further confirming a direct and highly specific binding of the protein to these sites. Upon this quality assessment, we finally performed a GO enrichment analysis on the 882 protein-coding transcripts directly bound by MATR3 in the WT heart and found "Heart development" as the most significantly

406   enriched functional category (**Figure 4F**). Therefore, our results indicate the formation of MATR3-

407   containing condensates inside the nucleus of cardiomyocytes which contain pCharme, as the most

408   abundant lncRNA, and pre-mRNA encoding for important cardiac regulators.

Figure 4

A

E15.5 hearts

Charme[WT]
Charme[KO]

↓

TOTAL extracts

↓

MATR3 CLIP

↓

RNA-seq

Charme[WT] (Input,Ip1,Ip2,IgG) Charme[KO] (Input,Ip1,Ip2,IgG)

B

- Protein-coding 95.5%
- Other 3.5%
- LncRNA 0.9%

- Intronic 91.5%
- Exonic 2.8%
- Intergenic 2.8%
- 3' UTR 1.8%
- CDS 0.8%
- 5' UTR 0.1%

C

Charme locus

Chromosome 7

Input

MATR3 Ip1 — Charme[WT]

MATR3 Ip2

Input

MATR3 Ip1 — Charme[KO]

MATR3 Ip2

FPM

D

Motif enrichment analysis

Motif	Motif ID	Protein name	E-Value
cUUucc	M207 0.6	PCBP1	$4.40e^{-26}$
CCuₙCC	M043 0.6	PCBP2	$2.17e^{-22}$
UUCUU	Desideri et al.,2020	MATR3	$1.48e^{-17}$
cUUuᵤ	M228 0.6	PTBP1	$1.18e^{-16}$
uuuᵤCCₙ	M188 0.6	PCBP3	$2.03e^{-16}$
uUUuᵤCu	M227 0.6	ROD1	$3.43e^{-14}$
GCUUᵤc	M037 0.6	MBNL1	$1.49e^{-5}$

E

UUCUU

Charme[WT] MATR3 IP peaks

Median FPM (Rep1 + Rep2)

-50 bp ← peak summit → +50 bp

log₂ (odds-ratio)
2 1 0 -1 -2

F

GO Biological Processes

- Heart development
- Actin filament-based process
- Positive regulat. of nucleic acid-templated transcription
- Positive regulat. of RNA biosynthetic process
- Positive regulat. of RNA metabol. process
- Positive regulat. of transcript. DNA-templated
- Positive regulat. of nucleob. contain. compound metabol. process
- Regulat. of transcript. by RNA PolII
- Positive regulat. of macromol. biosynth. process
- Cell development

E.R. 2.00 2.25 2.50

FDR $3e^{-13}$ $2e^{-13}$ $1e^{-13}$ 0

Ratio ● 0.11 ● 0.12 ● 0.13 ● 0.14 ● 0.15

409

23

410

Figure 4. MATR3/pCharme nuclear condensates contain key regulators of heart development

A) Schematic representation of MATR3 CLIP-seq workflow from fetal (E15.5) Charme^{WT} and Charme^{KO} hearts. See **Materials and Methods** for details.

B) MATR3 CLIP-seq from fetal hearts. Upper panel: a pie-plot projection representing transcript biotypes of 951 identified MATR3 interacting RNAs. Peaks overlapping multiple transcripts were assigned with the following priority: protein coding, lncRNA and others. Lower panel: a pie-plot projection representing the location 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. Percentages relative to each group are shown.

C) MATR3 CLIP-seq (Input, Ip1 and Ip2) normalized read coverage tracks (FPM) across pCharme from fetal hearts. Significant MATR3 peaks, displaying log_2 Fold enrichment > 2 in both Ip1 and Ip2 samples compared to Input, are demarcated by light-blue boxes.

Normalized read coverage tracks (FPM) from MATR3 CLIP-seq in Charme^{KO} fetal hearts on Charme locus are also shown. Plot obtained using Gviz R package.

D) Motif enrichment analysis perfomed on MATR3 CLIP-seq peaks (Charme^{WT}) with AME software using 93 RNA binding motifs from CISBP-RNA database. 7 consensus motifs resulted significantly over-represented (Evalue < 0.05) among the MATR3 peaks, as compared to control regions. See **Material and Methods** for details.

E) Positional enrichment analysis of MATR3 motif in MATR3 CLIP-seq peaks (Charme^{WT}). For each analyzed position close to peak summit, line plot displays the median CLIP-seq signal (FPM, IP1 + IP2), while heatmap displays the log2 odds-ratio of UUCUU motif enrichment. Significant enrichments (p-value < 0.05) are shown in red. See **Material and Methods** for details.

F) GO enriched categories obtained with WebGestalt (http://www.webgestalt.org) on protein-coding genes overlapping Charme^{WT} MATR3 peaks. Dots indicate the top categories of biological processes

435 (description in y-axis) in decreasing order of Enrichment Ratio (E.R.= overlapped genes/expected

436 genes, x-axis). Dot size (Ratio) represents the ratio between overlapped gens and GO categories size

437 while dot color (FDR) represents significance. All the represented categories show an FDR<0.05.

438

439 **The pCharme/MATR3 interaction sustains developmental gene expression in fetal**

440 **cardiomyocytes.**

441 On a transcriptome scale, we found that in the heart the expression of MATR3-bound RNAs was

442 globally reduced by pCharme knockout (**Figure 5-figure supplement 1A**). Indeed, while MATR3

443 targets were under-represented among transcripts whose expression was unaffected (invariant) or

444 increased (up-regulated) by pCharme ablation, 12% of the downregulated DEGs (41 out of 340) were

445 bound by MATR3 (**Figure 5A**). To note, these common targets were significantly enriched across

446 the top-three GO classes, previously defined for Charme^{KO} downregulated genes (**Figure 5B**). This

447 evidence suggests that pCharme is needed to sustain the cardiac expression of MATR3 targets. To

448 investigate the possible implication of the lncRNA for MATR3 binding to RNA, we proceeded with

449 a differential binding (DB) analysis of MATR3 CLIP-seq datasets between Charme^{WT} and Charme^{KO}

450 conditions (**Figure 5-figure supplement 1B,C**). Among the shared peaks, DB revealed that the main

451 consequence of pCharme depletion was a significant decrease of MATR3 enrichment on the RNA

452 (**Figure 5C**, Loss), while the increase of protein binding was observed on a lower fraction of peaks

453 (**Figure 5C**, Gain). In line with the specificity for MATR3 binding, also in these sets of differentially

454 bound regions, the MATR3 motif was positioned close to the summit of the peak (**Figure 5-figure**

455 **supplement 1D**). By mapping these differentially bound regions to the MATR3 targets which are

456 also responsive to pCharme ablation (DEGs), we found that while the gained peaks were equally

457 distributed, the loss peaks were significantly enriched in a subset (20 out of 40, 50%) of down-

458 regulated DEGs (**Figure 5D**, upper panel). Overall, these results suggest that cooperation between

459 pCharme and MATR3 activities is needed for the expression of specific RNAs, which majority (13

out of 20) is known to play important roles in embryo development, anatomical structure

morphogenesis and development of the circulatory system (**Figure 5D**, lower panel). Indeed, we

found interesting candidates 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).

Finally, to univocally distinguish the contribution of pCharme and MATR3 on their expression, we

tested the effect of MATR3 depletion on the abundance of these RNAs. We then purified cardiac

primary cells from wild-type hearts for transfection with control (si-SCR) or MATR3 (si-*Matr3*)

siRNAs (**Figure 5E**, left panel). RT-qPCR analysis revealed that 3 out of 4 tested genes exhibited a

significant expression decrease upon MATR3 depletion (**Figure 5E**, right panel).

In sum, these data offer unprecedented knowledge on the RNA-binding propensies of MATR3 in

the fetal heart and identify a subset of RNAs whose expression and MATR3 binding are influenced

by pCharme. 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.

Figure 5

A

B

C

D

E

480
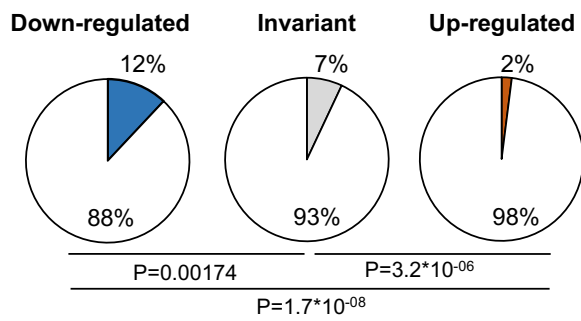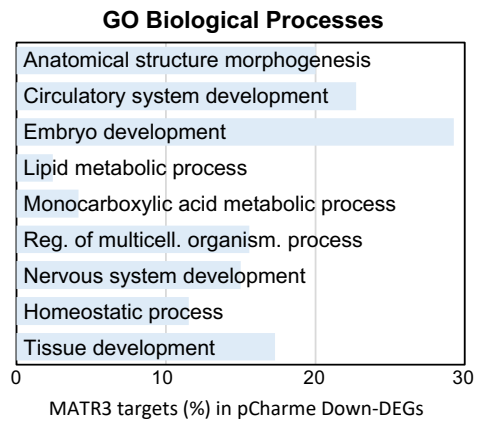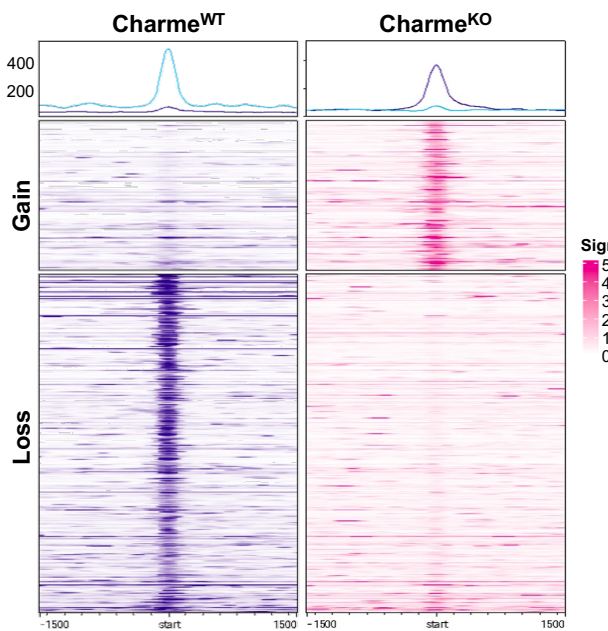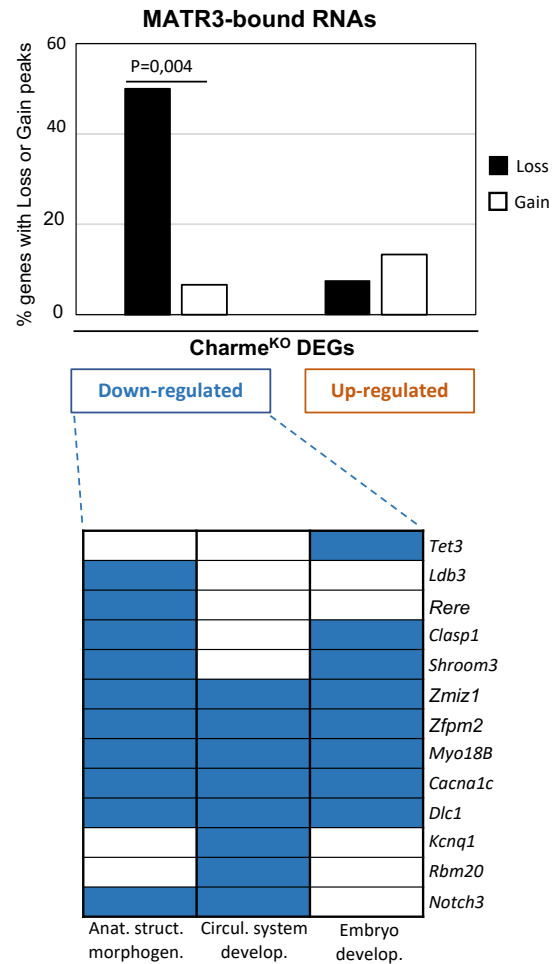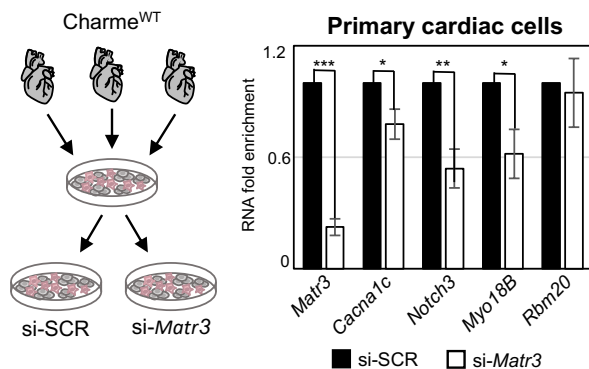
Figure 5. The pCharme/MATR3 interaction in cardiomyocytes sustains developmental genes expression

A) Pie charts showing the percentage of MATR3 targets in Charme^{KO} down-regulated, invariant or up-regulated DEGs. Significance of enrichment or depletion was assessed with two-sided Fisher's Exact test, shown below.

B) MATR3 targets (%) in the GO categories enriching Charme^{KO} down-regulated DEGs (Down-DEGs) (see also **Figure 2A**).

C) Profile Heatmaps of differential MATR3 CLIP-seq peaks (Charme^{WT} *vs* Charme^{KO}). Normalized mean read counts of both IP samples are shown only for significant (FDR < 0.05) «Gain» and «Loss» peaks.

D) Upper panel: histogram showing the distribution (%) of «Gain» and «Loss» MATR3 peaks in pCharme DEGs. Significance of enrichment was assessed with two-sided Fisher's Exact test. Lower panel: distribution of the subset (13 out of 20) of Down-DEGs with Loss peaks in the first three GO categories identified for down-regulated genes (see also Figure 2A).

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* RNA 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.

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, structural flexibility and the propensity of this class of RNAs to gather macromolecules (Buonaiuto et al., 2021). Furthermore, the maintenance of specific

28

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 under physiological circumstances.

Here, we took advantage of our Charme[KO] mouse model to extend the characterization of the lncRNA in embryogenesis. We show that the temporal and cell-type specific expression of pCharme is critical for the activation of pro-differentiation cardiac genes during development. Mechanistically, we provide evidence that this activity relies on the formation of pCharme-dependent RNA-rich condensates acting as chromatin hubs for MATR3, a nuclear matrix DNA/RNA-binding protein highly abundant in the fetal heart and source, upon mutation, of congenital defects (Quintero-Rivera et al., 2015). MATR3 ability to bind RNA can lead to the formation of dynamic shell-like nuclear condensates, whose composition depends on the integrity of the two RNA Recognition Motifs (RRM1 and 2) (Malik et al., 2018, Sprunger et al., 2022). Evidence supporting the physiological relevance of RNA-MATR3 interplay is emerging in both neural and muscle pathologies (Ramesh N et al., 2020; Senderek J, et al. 2009; Feit H, et al 1998.). In skeletal muscle, we previously showed that the presence of an intronic element bearing ~100 MATR3 binding sites inside pCharme, is needed for MATR3 nuclear localization (Desideri et al., 2020). Here, we add an important dowel into the functional characterization of pCharme and give a developmental meaning to its interaction with MATR3 (**Figure 6**). We show that the absence of the lncRNA in cardiomyocytes alters the MATR3 nuclear distribution and its binding to specific RNAs, with a consequent mis-regulation of their developmental expression. Among them, we found a subset of shared pCharme/MATR3 targets,

whose role in cardiac physiopathology was extensively studied, such as *Smad3*, *Notch3* and the myogenic components *Cacna1c* and *Myosin-18B* ([Goonasekera et al., 2012](#); [Tao et al., 2017](#); [Ajima et al., 2008](#); [van den Hoogenhof et al., 2018](#)).

Figure 6
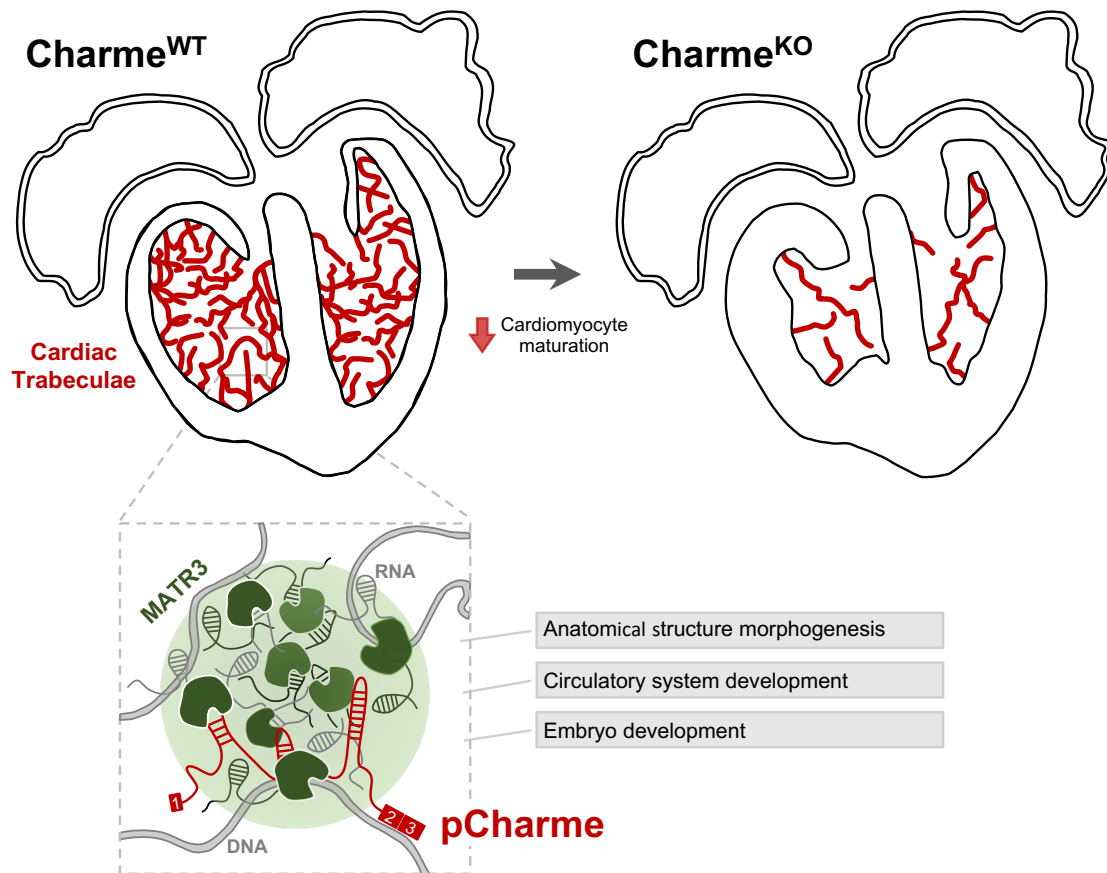
Figure 6. Proposed model for pCharme functions during heart development.

At developmental stages (Charme^WT), pCharme is required for the expression of genes involved in cardiomyocyte maturation. This activity is accompanied by the formation of nuclear condensates, through the interaction with the RNA-binding protein MATR3, which enrich transcripts involved in cardiac development. pCharme absence (Charme^KO) leads to the reduction of cardiac trabeculae during development and to the remodelling of heart morphology.

It is interesting to note that, in the cardiac context, the reactivation of fetal-specific RNA-binding proteins, including MATR3, was recently found to drive transcriptome-wide switches through the

regulation of early steps of RNA transcription and processing (D'Antonio et al., 2022). Along this view, the pCharme-dependent regulation of MATR3 binding in developing cardiomyocytes emphasises the need of cell-type specific lncRNAs for ruling the protein activities to specific RNAs, in a tissue-distinctive and spatio-temporal manner. As the impairment in the nuclear distribution of MATR3 can exert adverse effects in diverse tissues, our study may provide a more inclusive viewpoint on myopathies as well as other neuromuscular disorders where alteration in muscle development involves not only the myogenic lineage but also the interaction of the myogenic cells with the surrounding tissues.

These results also underlie the importance of studying cardiomyogenesis in animal models, which allow a better appreciation of the balance between the proliferation and maturation processes in cardiomyocytes and, importantly, the study of the cardiac functions in adulthood. In our case, adult Charme^{KO} mice develop a significant reduction of systolic function with initial signs of cardiac dilation, which denotes an early phase of cardiomyopathy. The 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).

Recent cardiovascular studies have uncovered essential roles for lncRNAs in cardiac development and disease (Scheuermann and Boyer, 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 classification of patient subpopulation but also to use them as possible diagnostic biomarkers or therapeutic targets (Buonaiuto et al., 2021). Future efforts will be devoted to clarifying the implication of the syntenic pCharme transcript in those human cardiomyopathies where pathological remodelling of the cardiac muscle occurs.

MATERIALS AND METHODS

Ethics Statement and animal procedures

31

571 C57Bl/6J mice were used in this work. The Charme[KO] animals were previously derived through the

572 insertion of a PolyA signal in the Charme locus, as detailed in (Ballarino et al., 2018). All procedures

573 involving laboratory animals were performed according to the institutional and national guidelines

574 and legislations of Italy and according to the guidelines of Good Laboratory Practice (GLP). All

575 experiments were approved by the Institutional Animal Use and Care Committee and carried out in

576 accordance with the law (Protocol number 82945.56). All animals were kept in a temperature of 22°C

577 ± 3°C with a humidity between 50% and 60%, in animal cages with at least 5 animals.

578

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

580 For primary heart cells isolation and transfection, 5 to 10 postnatal (PN3) hearts for each replicate

581 (n=4) were pooled, harvested and kept at 37°C in culture medium (FBS 10%, 1X Non-Essential

582 aminoacids, 1x PenStrep and DMEM high glucose). Hearts were mashed with pestles for 2 min and

583 cell isolation performed according to manufacturer's instructions (Neonatal Heart Dissociation Kit,

584 Miltenyi Biotec). Cell suspension was centrifuged for 5 min at 600 x g and cells were resuspended in

585 cell culture medium and plated in 22.1 mm plates. 1,5 million cells were transfected 48 hr later with

586 75 mM si-SCR or si-MATR3 in 3 µl/ml of Lipofectamine RNAiMAX (Thermo Fisher Scientific) and

587 100 µl/ml of Opti-MEM (Thermo Fisher Scientific), according to manufacturer's specifications. Total

588 RNA was collected 48 hr after transfection. See **Supplementary File 5** for siRNA sequences.

589 Subcellular fractionation of primary embryonal (E15.5) cardiac cells was performed using the Paris

590 Kit (Thermo Fisher Scientific, cat#AM1921), according to the manufacturer's instructions.

591

592 **Whole-mount *in situ* hybridization**

593 Embryos were fixed overnight in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS)

594 plus 0.1% Tween® 20 (PBT) at 4°C, dehydrated through a series of methanol/PBT solutions (25%,

595 50%, 75% and 100% methanol), and stored at -20°C until hybridization. Fixed embryos were

32

rehydrated and rinsed twice in PBT. At this point, embryos were either digested with DNase and/or

RNase or kept in PBT. All embryos were bleached in 6% hydrogen peroxide in PBT for 1h. Embryos

were then rinsed 3 times in PBT for 5 min, digested with proteinase K (10 μg/ml in PBT) for 5 min

at room temperature, washed once in 2 mg/ml glycine in PBT and twice in PBT for 5 min each, and

post-fixed in 4% PFA/0.2% glutaraldehyde in PBT for 20 min at room temperature. Embryos were

subsequently rinsed twice in PBT for 5 min and pre-hybridized at 70°C in hybridization solution

(50% Formamide, 5x SSC, pH 5, 0.1% Tween 20, 50 μg/ml heparin, 50 μg/ml Torula RNA, 50 μg/ml

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). Antisense and sense Charme probes, used as

specificity controls, were synthesised 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-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 specifically 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

pCharme exon 2 and exon 3 were PCR-amplified from cDNA extracted from myotubes using Charme_Up-BamHI and Charme_Down-EcoRI primers (**Supplementary File 5**). 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

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. Antibodies and dilutions are reported in **Supplementary File 5**. DAPI, KI-67, NPPA, TNNT2 and G. simplicifolia lectin immunofluorescence signals (Figure 2H; Figure 2-figure supplement 1D,E,H) were acquired with Carl Zeiss Microscopy GmbH Imager A2 equipped with Axiocam503 color camera. MATR3, TNNT2 and MAP2 signals (Figure 3C; Figure 3-figure supplement 1D,E) were acquired as Z stacks (200 nm path) by inverted confocal Olympus IX73 microscope equipped with a Crestoptics X-LIGHT V3 spinning disk system

35

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} neonatal 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 Software (Schneider C.A. et al., 2012). NPPA and G.

simplicifolia lectins sub-tissutal staining (Figure 2H; Figure 2-figure supplement 1H) was analysed

using ImageJ software (Schneider C.A. et al., 2012) and according to Fukuda et al., 2013, with little

modifications. Briefly, we quantified the ratio between the area covered by the NPPA/lectins signal

immunofluorescence signals and the total area of the left ventricle. To calculate the area, the left

ventricle was selected with the line selection tool. A threshold was then applied to the selected region

to measure the area covered by the signal. Plot in Fig. 2F represents the percentage of area covered

by the NPPA/lectins signal divided by the area of the left ventricle.

Quantification of MATR3 puncta in Figure 3C was performed using ImageJ software (Schneider C.A.

et al., 2012) according to (Higaki T. et al., 2020), with little modifications. Briefly, a "mask" from

binarized images was created and used to measure the intensity of MATR3 fluorescence inside the

nuclei and the standard deviation (SD). MATR3 distribution was evaluated by Coefficient Variation

(CV) defined as CV= mean SD/mean MATR3 intensity. A higher CV value indicates a higher spatial

variability of fluorescence distribution (punctate staining); conversely, a lower CV value corresponds

to a higher uniformity of fluorescence distribution (diffuse staining). 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^{TM} RNA MiniPrep (Zymo Research), treated with DNase (Zymo Research),

696 retrotranscribed using PrimeScript Reagent Kit (Takara) and amplified by RT-qPCR using PowerUp

697 SYBR-Green MasterMix (Thermo Fisher Scientific), as described in Desideri et al., 2020. See

698 Supplementary File 5 for oligos details.

699

700 **RNA-Seq Analysis**

701 To reduce biological variability, Charme[WT] and Charme[KO] neonatal littermates (PN2) were sacrificed

702 and hearts from the corresponding genotypes pooled together before RNA extraction (3 Charme[WT]

703 pools, 9 hearts each, 2 Charme[KO] pools, 3 hearts each). Validation analyses were performed on 2

704 additional Charme[WT] pools (6 hearts each) and 2 Charme[KO] pools (4 hearts each). Principal

705 component analysis (PCA) conducted on the RNA-seq data, revealed that the two groups were

706 evidently distinguished for the first principal component (**Figure 2-figure supplement 1A**). Illumina

707 Stranded mRNA Prep was used to prepare cDNA libraries for RNA-Seq that was performed on an

708 Illumina Novaseq 6000 Sequencing system at IIT-Istituto Italiano di Tecnologia (Genova, Italy).

709 RNA-seq experiment produced an average of 26 million 150 nucleotide long paired-end reads per

710 sample. Dark cycles in sequencing from Novaseq 6000 machines can lead to a high quality stretches

711 of Guaninines artifacts; in order to remove these artifacts, low quality bases and N stretches from

712 reads were removed by Cutadapt software using "-u -U","--trim-n" and "—nextseq-trim=20"

713 parameters (Martin et al., 2011). Illumina adapter remotion were performed using Trimmomatic

714 software (Bolger et al., 2014). Reads whose length after trimming was <35 nt were discarded. Reads

715 were aligned to GRCm38 assembly using STAR aligner software (Dobin et al., 2013). Gene loci

716 fragment quantification was performed on Ensemble (release 87) gene annotation gtf using STAR –

717 quantMode GeneCounts parameter. Read counts of "reverse" configuration files were combined into

718 a count matrix file, which was given as input to edgeR (Robinson et al., 2010) R package for

719 differential expression analysis, after removing genes with less than 10 counts in at least two samples.

720 Samples were normalized using TMM. Model fitting and testing were performed using the 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.

Cross-linking immunoprecipitation (CLIP) assay

A total of 60 Charme^{WT} and 56 Charme^{KO} E15.5 embryonal hearts were collected, divided in two distinct biological replicates and pestled for 2 min in PBS, 1x PIC and 1x PMSF. For each replicate, the solution was filtered in a 70 μm strainer and the isolated cells were plated and UV-crosslinked (4,000 μJ) 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,), incubated on ice for 15 min and sonicated at low intensity six times with Bioruptor® Plus sonication device to ensure nuclear membrane lysis. 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 (Supplementary File 5) 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).

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

(Bolger et al., 2014) and Cutadapt version 3.2 (Martin et al., 2011) softwares. Reads whose length

after trimming was <35 nt were discarded. Alignment to mouse GRCm38 genome and Ensembl

87 transcriptome was performed using STAR aligner version 2.7.7a (Dobin et al., 2013). Alignment

files were further processed by collapsing PCR duplicates using the MarkDuplicates tool included in

the Picard suite version 2.24.1 (http://broadinstitute.github.io/picard/) and discarding the multi-

mapped reads using BamTools version 2.5.1 (Barnett et al., 2011). Properly paired reads were

extracted using SAMtools version 1.7 (Li et al., 2009). GRCm38 genome was divided into 200 bp

long non-overlapping bins using the BEDtools makewindows tool (Quinlan et al., 2010). Properly

paired fragments falling in each bin were counted using the BEDtools intersect tool filtering out reads

mapping to rRNAs, tRNAs or mitochondrial genome in order to create sample-specific count files.

These files were given as input to Piranha version 1.2.1 (Uren et al., 2012) using –x -s –u 0 parameters

to call significant bins for MATR3 Ip and IgG samples. BEDtools intersect was used to assign each

genomic bin to genes using Ensembl 87 annotation. For each gene the bin signal distribution in the

input sample was calculated after normalization of fragment counts by the total number of mapping

771 fragments. Ip significant bins presenting normalized signals lower than the upper-quartile value of

772 the related gene distribution were filtered out. After this filter, significant bins belonging to Ip samples

773 were merged using BEDTools merge tool. The number of fragments overlapping identified peaks and

774 the number overlapping the same genomic region in the Input sample were counted and used to

775 calculate fold enrichment (normalized by total mapping fragments counts in each data set), with

776 enrichment P-value calculated by Yates' Chi-Square test or Fisher Exact Test where the observed or

777 expected fragments number was below 5. Benjamini-Hochberg FDR procedure was applied for

778 multiple testing corrections. Peaks presenting log2 fold enrichment over Input >2 and FDR < 0.05 in

779 both Ip samples were selected as enriched regions (**Supplementary File 3).** BEDtools intersect tool

780 was used to annotate such regions based on their overlap with Ensembl 87 gene annotation and to

781 filter out transcripts hosting regions enriched in the IgG sample. Furthermore, htseq-count software

782 ([Anders et al., 2015](#)) with *-s no -m union -t gene* parameters was used to count reads from deduplicated

783 BAM files. Peaks overlapping transcripts with Input CPM (counts per million) > Ip CPM in both Ip

784 samples were filtered out. Gene Ontology analysis was performed on protein coding genes

785 overlapping enriched regions using WebGestalt R package ([Liao et al., 2019](#)). Bigwig of normalized

786 coverage (RPM, reads per million) files were produced using bamCoverage 3.5.1 from deepTools

787 tool set (http://deeptools.readthedocs.io/en/develop) on BAM files of uniquely mapping and

788 deduplicated reads using *--normalizeUsing CPM* parameter ([Ramirez et al., 2014](#)). Normalized

789 coverage tracks were visualized with IGV software (https://software.broadinstitute.org/software/igv/)

790 and Gviz R package ([Hahne et al., 2016](#)).

791 Local Motif Enrichment analysis was performed by using 10 nt-long sliding windows. For each

792 relative peak position/control set, the proportion of windows containing or not MATR3 motif was

793 calculated and tested for significant differences using two tails Fisher exact test. CLIP-seq IP signal

794 (FPM) relative to each position was retrieved from bigwig files using pyBigWig

795 (https://github.com/deeptools/pyBigWig). Motif Enrichment Analysis was performed using AME

software from MEME suite (Bailey et al., 2015) on a 100 nt window centred on the peak summit. For each identified peak, a control set was by selecting 10 random regions from expressed (CPM>0) RNA precursors. Differential analysis of MATR3 binding between Charme^{WT} and Charme^{KO} conditions was assessed using DiffBind software (Ross-Innes et al., 2012).

Protein analyses

Protein extracts were prepared and analysed by western blot as in Desideri et al., 2020. See **Supplementary File 5** 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 *clustersDeltaExpression* (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

821 follows. Cardiomyocytes (1834 cells): Myh6+, Nppa+, Atrial CM (333 cells): are also Myl1+, Myl4+

822 , Ventricular CM (1289): are also Myl2+, Myl3+, interventricular septum CM (117 cells): are also

823 TBX20+, Gja5+ (Franco et al., 2006), Venous Pole CM (95 cells): also Osr1+ (Meilhac et al., 2018),

824 Outflow Tract CM (72 cells): also Isl1+, Sema3c+, Neural crest cells (there are two clusters

825 expressing their markers: 1.NC with 48 cells and 2.NC with 68 cells): Msx1+, Twist1+, Sox9+,

826 Epicardial cells (359 cells): Cebpb+, Krt18+, Fibroblasts like cells (278 cells): Tcf21+, Fn1+,

827 Endothelial cells (168 cells): Klf2+, Pecam1+, Cdh5+, Smooth muscle cells (710 cells): Cnn1+,

828 Acta2+, Tagln2+, Tagln+, Hemopoietic myeloid cells (79 cells): Fcer1g+, Hemopoietic red blood

829 cells (397 cells): Hba-a1+. The final UMAP plot with the cell assignment is shown in **Figure 1-figure**

830 **supplement 1E**. Heatmap in **Figure 1-figure supplement 1F** shows a coherent assignment between

831 final clusters and cell type with additional marker genes. For correlation analyses (**Figure 1G**),

832 relevant genes in each subpopulation of interest (CM-VP, CM-OFT, A-CM, CM-IVS, V-CM) were

833 analysed separately by applying COTAN on a test dataset composed by the subpopulation in exam,

834 together with a fixed contrasting cell group with the SM and EC cells (see **Figure 1E**).

835

836 **Echocardiography**

837 The echocardiographer was blinded to the phenotypes. Mice were anaesthetised with 2.5% Avertin

838 (Sigma T48402) to perform echocardiographic structural (measurement of left ventricular diameters

839 and wall thickness) and functional (fractional shortening) analyses with a VEVO 3100 (Visualsonics)

840 using a mx550d probe. We used avertin since it does not induce significant cardiodepressant effects,

841 potentially affecting our echocardiographic experiments compared to ketamine combinations, such

842 as ketamine+xylazine. The fractional shortening (FS) of the left ventricle was calculated as FS% =

843 (left ventricular end-diastolic diameter (LVEDD)-left ventricular end-systolic diameter

844 (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 rigour

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 details). 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.

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.

Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME Suite. *Nucleic Acids Res*. **43**(W1):W39-49.

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.

44

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.

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.

D'Antonio M, Nguyen JP, Arthur TD, Matsui H, Donovan MKR, D'Antonio-Chronowska A, Frazer KA. 2022. In heart failure reactivation of RNA-binding proteins is associated with the expression of 1,523 fetal-specific isoforms. *PLoS computational biology* **18**: e1009918.

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.

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

955 England J, Loughna S. 2013. Heavy and light roles: myosin in the morphogenesis of the heart. *Cell Mol Life*
956 *Sci*. **70**:1221-39.

957 Feit H, Silbergleit A, Schneider LB, Gutierrez JA, Fitoussi RP, Réyès C, Rouleau GA, Brais B, Jackson CE,
958 Beckmann JS, Seboun E. 1998. Vocal cord and pharyngeal weakness with autosomal dominant distal
959 myopathy: clinical description and gene localization to 5q31. *Am J Hum Genet*. **63**:1732-42.

960 Franco D, Meilhac SM, Christoffels VM, Kispert A, Buckingham M, Kelly RG. 2006. Left and right
961 ventricular contributions to the formation of the interventricular septum in the mouse heart. *Dev Biol*. **294**:
962 366-375.

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

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

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

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

973 Hahne F, Ivanek R. 2016. Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol Biol* **18**:
974 335-51.

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

978 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
979 noncoding RNA protects the heart from pathological hypertrophy. *Nature* **514**: 102–106.

980 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
981 Hand2os1/Uph locus orchestrates heart development through regulation of precise expression of Hand2.
982 *Development* **146**: dev176198.

983 Hazra R, Brine L, Garcia L, Benz B, Chirathivat N, Shen MM, Wilkinson JE, Lyons SK, Spector DL. 2022.
984 Platr4 is an early embryonic lncRNA that exerts its function downstream on cardiogenic mesodermal lineage
985 commitment. *Developmental cell* **57**: 2450–2468.e7.

986 Higaki T, Akita K, Katoh K. 2020. Coefficient of variation as an image-intensity metric for cytoskeleton
987 bundling. *Scientific Reports* **10**:22187.
988
989 Horsthuis T, Houweling AC, Habets PE. MH, De Lange FJ, El Azzouzi H, Clout DEW, Moorman AFM,
990 Christoffels VM. 2008. Distinct regulation of developmental and heart disease-induced atrial natriuretic factor
991 expression by two separate distal sequences. *Circulation Research* **102**: 849–859.

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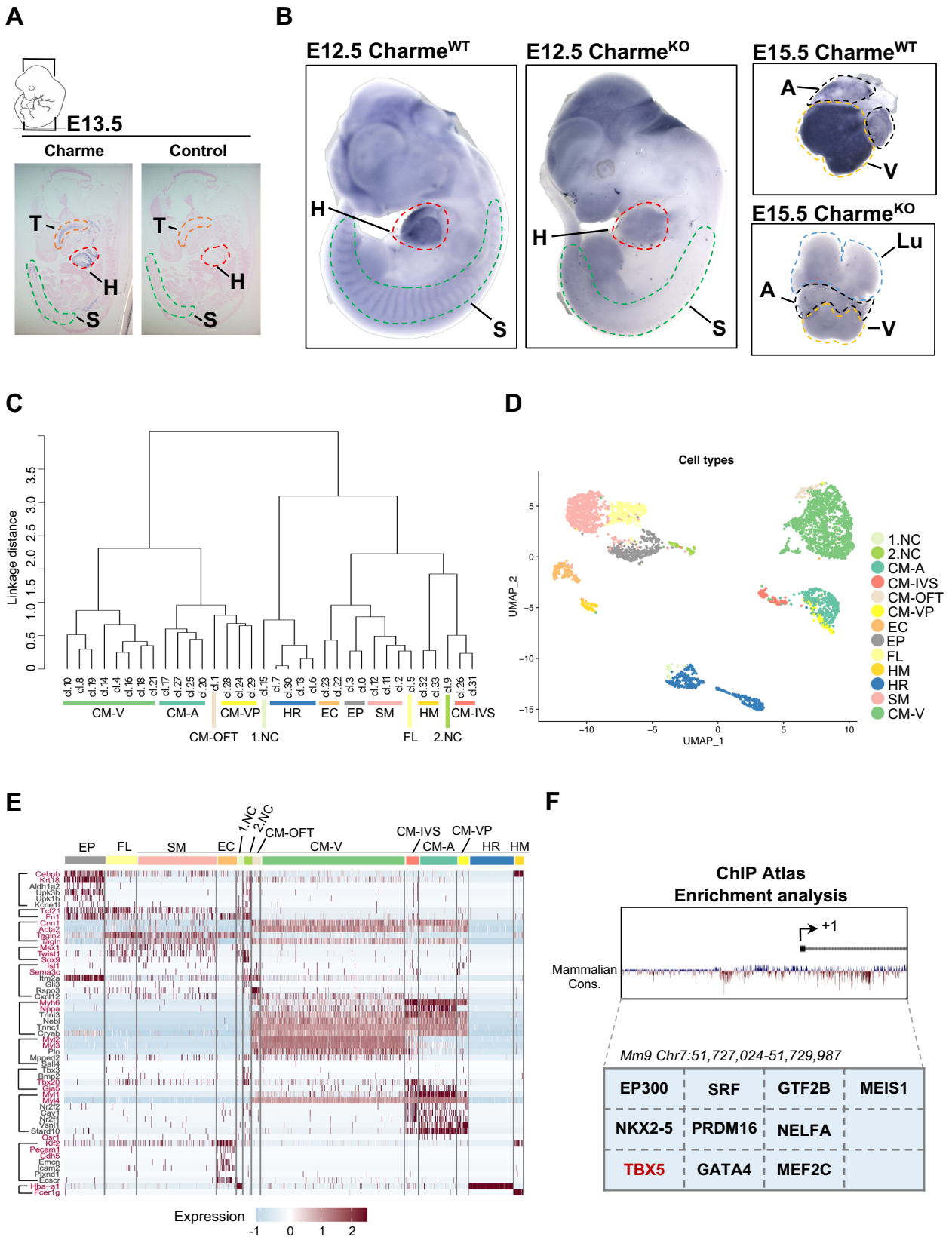993 noncoding RNAs and their role in disease pathogenesis. *Oncotarget* 9: 18648–18663.

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

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

999 Jusic A, Thomas PB, Wettinger SB, Dogan S, Farrugia R, Gaetano C, Tuna BG, Pinet F, Robinson EL, Tual-
1000 Chalot S, Stellos K, Devaux Y. 2022. EU-CardioRNA COST Action CA17129. Noncoding RNAs in age-
1001 related cardiovascular diseases. *Ageing Res Rev*. **77**:101610.

1002 Kattman SJ, Witty AD, Gagliardi M, Dubois NC, Niapour M, Hotta A, Ellis J, Keller G. 2011. Stage-specific
1003 optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human
1004 pluripotent stem cell lines. *Cell Stem Cell*. **8**:228-40.

1005 Kim NJ, Lee KH, Son Y, Nam AR, Moon EH, Pyun JH, Park J, Kang JS, Lee YJ, Cho JY. 2021.
1006 Spatiotemporal expression of long noncoding RNA Moshe modulates heart cell lineage commitment. *RNA*
1007 *Biol*. **18**: 640-654.

1008 Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, Ding H, Butty VL,
1009 Torrey L, Haas S et al., 2013. Braveheart, a long noncoding RNA required for cardiovascular lineage
1010 commitment. *Cell* **152**: 570–583.

1011 Kodo K, Nishizawa T, Furutani M, Arai S, Yamamura E, Joo K, Takahashi T, Matsuoka R, Yamagishi H.
1012 2009. GATA6 mutations cause human cardiac outflow tract defects by disrupting semaphorin-plexin signaling.
1013 *Proceedings of the National Academy of Sciences of the United States of America* **106**: 13933–13938.

1014 Latham SL, Weiß N, Schwanke K, Thiel C, Croucher DR, Zweigerdt R, Manstein DJ, Taft MH. 2020. Myosin-
1015 18B Regulates Higher-Order Organization of the Cardiac Sarcomere through Thin Filament Cross-Linking
1016 and Thick Filament Dynamics. *Cell Rep* **32**: 108090.

1017 Li G, Xu A, Sim S, Priest JR, Tian X, Khan T, Quertermous T, Zhou B, Tsao PS, Quake SR, et al. 2016.
1018 Transcriptomic Profiling Maps Anatomically Patterned Subpopulations among Single Embryonic Cardiac
1019 Cells. *Dev Cell* **39**:491-507.

1020 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009.1000
1021 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools.
1022 *Bioinformatics* **25**: 2078-2079.

1023 Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. 2019. WebGestalt 2019: gene set analysis toolkit with revamped
1024 UIs and APIs. *Nucleic Acids Res*. **47**:W199-W205.

1025 Luna-Zurita L, Stirnimann CU, Glatt S, Kaynak BL, Thomas S, Baudin F, Samee MAH, He D, Small EM,
1026 Mileikovsky M, et al. 2016. Complex Interdependence Regulates Heterotypic Transcription Factor
1027 Distribution and Coordinates Cardiogenesis. *Cell* **164**: 999–1014.

1028 MacGrogan D, Münch J, de la Pompa JL. 2018. Notch and interacting signalling pathways in cardiac
1029 development, disease, and regeneration. *Nat Rev Cardiol*. **15**: 685-704.

1030 Malik AM, Miguez RA, Li X, Ho YS, Feldman EL, Barmada SJ. 2018. Matrin 3-dependent neurotoxicity is
1031 modified by nucleic acid binding and nucleocytoplasmic localization. *Elife* **17**;7:e35977.

1032 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads.
1033 *EMBnet.journal* **17**: 10-12.

1034 Mattick J. 2004. RNA regulation: a new genetics?. *Nat Rev Genet* **5**: 316–323.

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.

Padula SL, Velayutham N, Yutzey KE. 2021. Transcriptional Regulation of Postnatal Cardiomyocyte Maturation and Regeneration. *Int J Mol Sci.* **22**(6):3288.

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.

Ramesh N, Kour S, Anderson EN, Rajasundaram D, Pandey UB. 2020. RNA-recognition motif in Matrin-3 mediates neurodegeneration through interaction with hnRNPM. *Acta neuropathol commun* **8**, 138.

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.

Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, Li X, Gueroussov S, Albu M, Zheng H, Yang A, et al. 2013, A compendium of RNA-binding motifs for decoding gene regulation. Nature 499: 172–177.

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.

48

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.

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.

Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, etal.,. 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**.

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.

Schneider CA, Rasband WS. Eliceiri KW, 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9** 671–675,

Senderek J, Garvey SM, Krieger M, Guergueltcheva V, Urtizberea A, Roos A, Elbracht M, Stendel C, Tournev I, Mihailova V, et al.,. 2009. Autosomal-dominant distal myopathy associated with a recurrent missense mutation in the gene encoding the nuclear matrix protein, matrin 3. *Am J Hum Genet*. **84**:511-8.

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

Sprunger ML, Lee K, Sohn BS, Jackrel ME. 2022. Molecular determinants and modifiers of Matrin-3 toxicity, condensate dynamics, and droplet morphology. *iScience* **25**:103900.

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.

Tian X, Li Y, He L, Zhang H, Huang X, Liu Q, Pu W, Zhang L, Li Y, Zhao H, et al., 2017. Identification of a hybrid myocardial zone in the mammalian heart after birth. *Nat Commun*. **20**:8(1):87

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.

1122   Ulitsky I, Bartel DP. 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell* **154**: 26–46.

1123   Uosaki H, Cahan P, Lee DI, Wang S, Miyamoto M, Fernandez L, Kass DA, Kwon C. 2015. Transcriptional
1124   landscape of cardiomyocyte maturation. *Cell Rep* **13**:1705–1716.

1125   Uren PJ, Bahrami-Samani E, Burns SC, Qiao M, Karginov FV, Hodges E, Hannon GJ, Sanford JR, Penalva
1126   LOF, Smith AD. 2012. Site identification in high-throughput RNA–protein interaction data. *Bioinformatics*
1127   **28**: 3013–3020.

1128   van den Hoogenhof MMG, Beqqali A, Amin AS, van der Made I, Aufiero S, Khan MAF, Schumacher CA,
1129   Jansweijer JA, van Spaendonck-Zwarts KY, Remme CA, et al. 2018. RBM20 Mutations Induce an
1130   Arrhythmogenic Dilated Cardiomyopathy Related to Disturbed Calcium Handling. *Circulation* **138**: 1330-
1131   1342.

1132   Wang K, Holt C, Lu J, Brohus M, Larsen KT, Overgaard MT, Wimmer R, Van Petegem F. 2018. Arrhythmia
1133   mutations in calmodulin cause conformational changes that affect interactions with the cardiac voltage-gated
1134   calcium channel. *Proc Natl Acad Sci U S A* **115**: E10556-E10565.

1135   Wang Y, Chen J, Cowan DB, Wang DZ. 2021. Non-coding RNAs in cardiac regeneration: Mechanism of
1136   action and therapeutic potential. *Seminars in cell & developmental biology* **118**: 150–162.

1137   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
1138   noncoding RNA Chaer defines an epigenetic checkpoint in cardiac hypertrophy. *Nat. Med.* **22**: 1131–1139.

1139   Yang XH, Nadadur RD, Hilvering CR, Bianchi V, Werner M, Mazurek SR, Gadek M, Shen KM, Goldman
1140   JA, Tyan L et al. 2017. Transcription-factor-dependent enhancer transcription defines a gene regulatory
1141   network for cardiac rhythm. *Elife* **6**: e31683.

1142   Yao RW, Wang Y, Chen LL. 2019. Cellular functions of long noncoding RNAs. *Nature cell biology* **21**: 542–
1143   551.

1144   Zhao S, Guo Y, Sheng Q, Shyr Y. 2014. Advanced heat map and clustering analysis using heatmap3. *Biomed.*
1145   *Res. Int* e986048.

1146   Zimmerman MS, Carswell Smith AG, Sable CA, Echko MM, Wilner LB, Olsen HE, Tasew Atalay H, Awasthi
1147   A, Bhutta ZA, Boucher JL, et al. 2020. Global, regional, and national burden of congenital heart disease, 1990–
1148   2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet. Child & adolescent*
1149   *health* **4**: 185–200.

Figure 1-figure supplement 1

A

B

E12.5 Charme^WT   E12.5 Charme^KO   E15.5 Charme^WT

E15.5 Charme^KO

C

D

Cell types

E

F

ChIP Atlas
Enrichment analysis

Mm9 Chr7:51,727,024-51,729,987

EP300	SRF	GTF2B	MEIS1
NKX2-5	PRDM16	NELFA	
TBX5	GATA4	MEF2C	

Expression
-1 0 1 2

1150

Figure 1-Figure supplement 1. Related to Figure 1

51

A) *In-situ* hybridization (ISH) performed on E13.5 embryonal cryo-sections using digoxigenin-labelled RNA antisense (Charme, left panel) or sense (control, right panel) probes against Charme. T: Tongue (orange); H: Heart (red); S: Somites (green).

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 signals in the Charme^{KO} samples. Heart (H, red line); Somites (S, green line); A: Atria (black line); V: Ventricles (yellow line).

C) Dendrogram showing the relationships between homogeneous clusters. All the informative transcriptome was used to create a hierarchical clustering between homogeneous cell clusters. Colored lines mark which clusters were merged for the final clustering (Figure 1-Figure supplement 1D-E). 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, EC: 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 Seurat DoHeatmap (Stuart et al., 2019) and represents, for each cell of the identified sub-populations, the log normalized expression of cell identity markers (listed on the left). Genes used for cell clustering are marked in red. Maximum expression value (red), minimum expression value (light blue). Correspondence between gene markers and cell types is indicated on the left.

F) *In silico* analysis of cardiovascular TF binding sites on Charme promoter using the enrichment analysis tool on the ChIP Atlas database (https://chip-atlas.org/enrichment_analysis) by setting the threshold at 200.

Figure 2-figure supplement 1
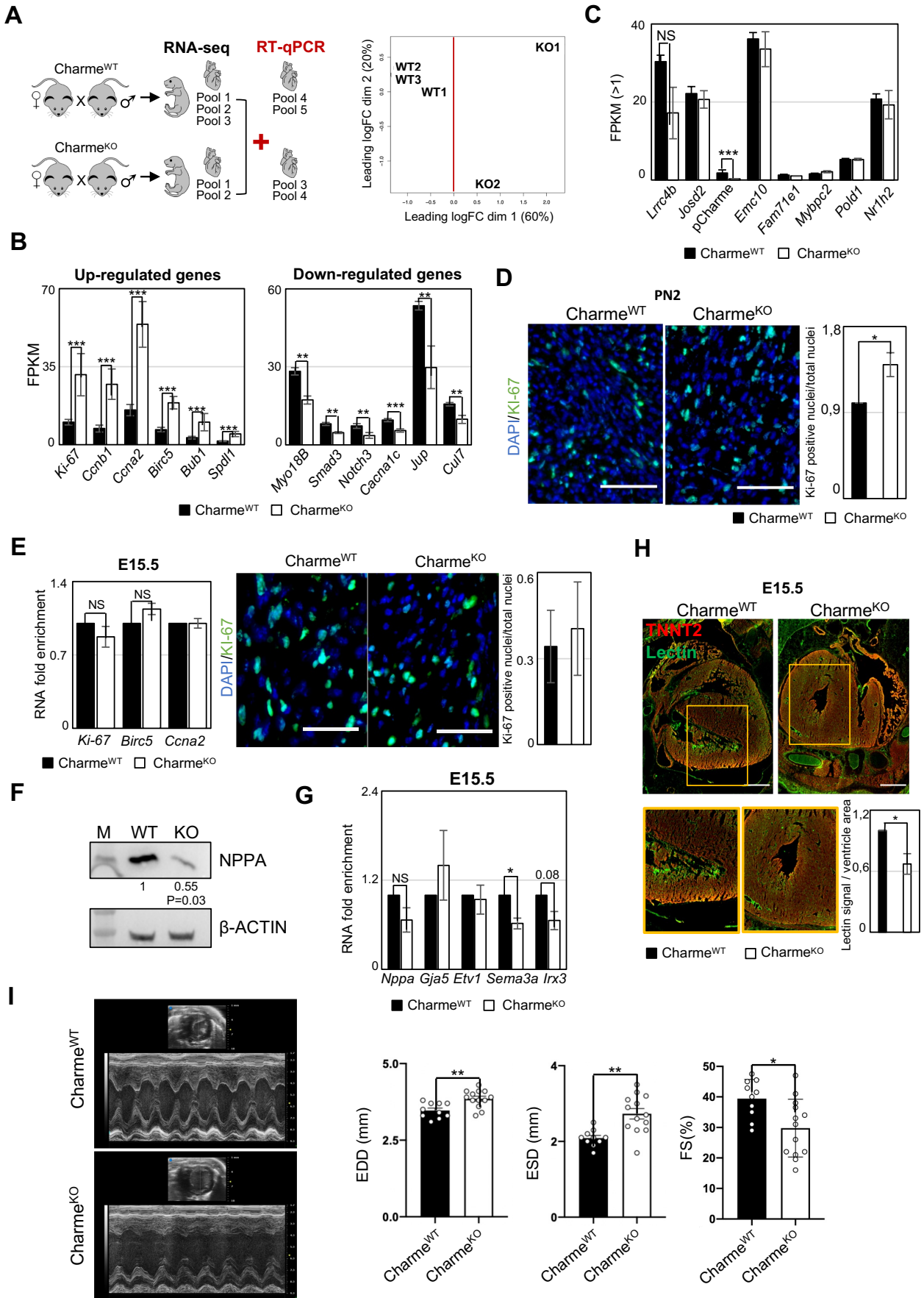
Figure 2-Figure supplement 1. Related to Figure 2

A) Schematic overview of the workflow to identifying DEGs from Charme^{WT} and Charme^{KO} transcriptomes (left panel). Multi-dimensional scaling plot of leading fold-change (FC) between each pair of Charme^{WT} and Charme^{KO} RNA-seq samples (right panel). 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) Quantification by RNA-seq (FPKM) of Up-regulated (left panel) and Down-regulated (right panel) DEGs in Charme^{KO} *vs* Charme^{WT} neonatal hearts.

C) Quantification by RNA-seq (FPKM) of pCharme and Charme-neighbouring genes expression in Charme^{WT} *vs* Charme^{KO} neonatal hearts.

D) Left panel: Representative images for KI-67 (green) and DAPI (blue) stainings on Charme^{WT} and Charme^{KO} neonatal cardiac sections. Scale bars: 70 μm. Right panel: Quantification of KI-67 positive nuclei/total nuclei on Charme^{WT} and Charme^{KO} cardiac sections from neonatal mice. Data are expressed as mean ± SEM, n = 4.

E) Left panel: RT-qPCR quantification of DEGs belonging to the "cell cycle" GO class from Charme^{WT} and Charme^{KO} E15.5 hearts. Data were normalized to *Gapdh* mRNA and represent means ± SEM of n=3 heart pools. Right panel: Representative images for KI-67 (green) and DAPI (blue) stainings on Charme^{WT} and Charme^{KO} E15.5 cardiac sections are shown. Quantification of KI-67 positive nuclei/total nuclei on Charme^{WT} and Charme^{KO} cardiac sections from E15.5 mice. Data are mean ± SEM.; n =3. Scale bars: 140 μm.

F) Western blot analysis for NPPA in Charme^{WT} and Charme^{KO} E15.5 cardiac extracts. β-ACTIN was used as a loading control. Quantification of NPPA signal intensity relative to β-ACTIN is shown below. Data are mean ± SEM (n = 3). *p= 0.03. A representative image is shown.

1203 **G)** RT-qPCR quantification of trabeculae markers expression in Charme^{WT} *vs* Charme^{KO} E15.5

1204 cardiac extracts. Data were normalized to *Gapdh* mRNA and represent means ± SEM of n=3 heart
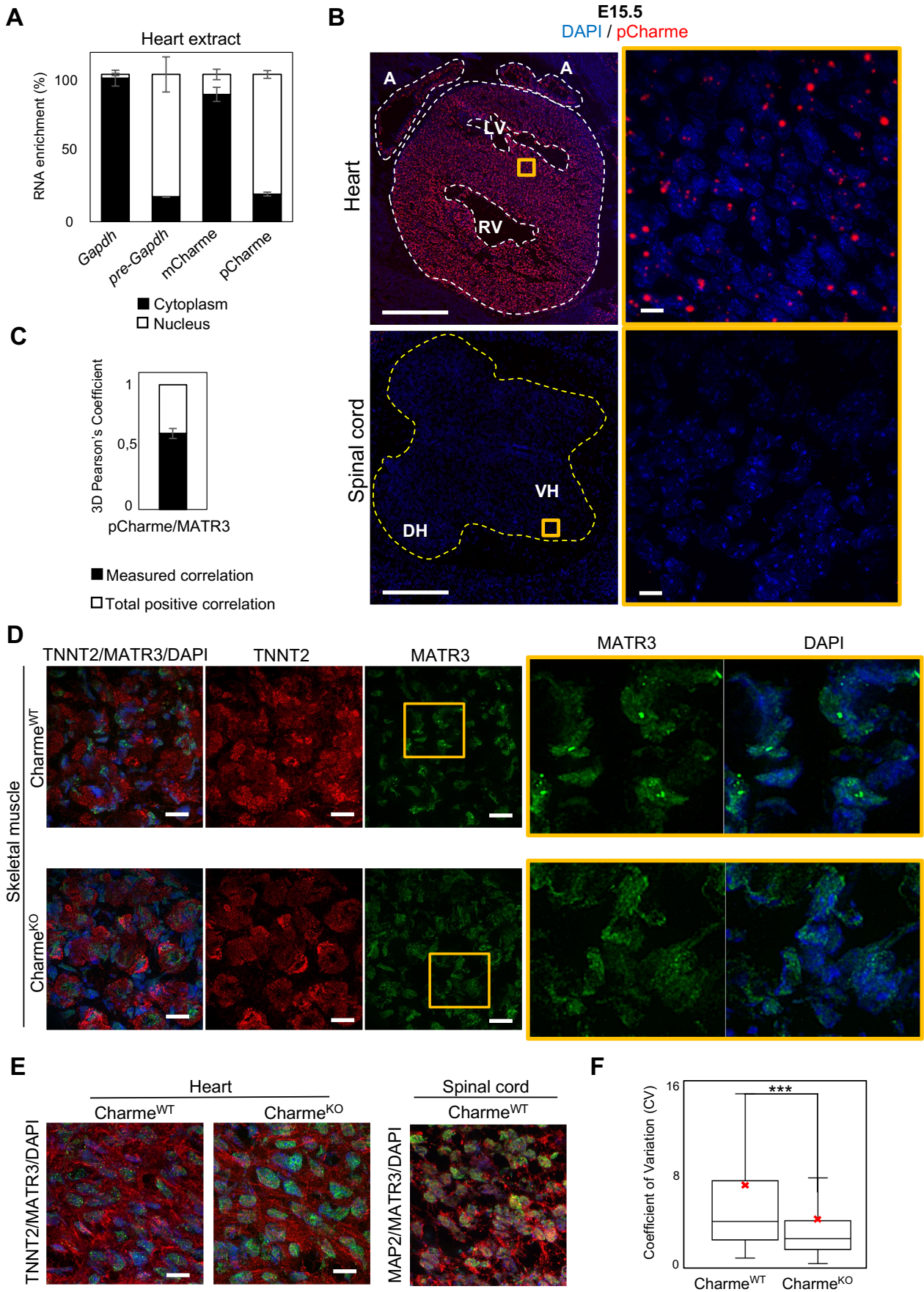
1205 pools.

1206 **H)** Representative image of Lectin (green) and TnnT2 (red) immunostainings in Charme^{WT} and

1207 Charme^{KO} E15.5 cardiac sections. ROI (orange squares) were digitally enlarged on the lower panels.

1208 Scale bars: 500 μm. Quantification of the area covered by the Lectin fluorescence signal is shown

1209 aside. For each genotype, data represent the mean ± SEM of (at least) 6 images. n=3.

1210 **I)** Representative M-mode echocardiographic track of Charme^{WT} and Charme^{KO} 9-12 months aged

1211 mice. Quantification of heart morphology (EDD= end-diastolic diameter and ESD= end-systolic

1212 diameter) and function (FS: Fractional Shortening =(EDD-ESD)/EDD) were evaluated. Data

1213 represent the mean ± SEM (n = 10-13)

1214 Data information: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, unpaired Student's t test.

1215

Figure 3-figure supplement 1

A Heart extract

RNA enrichment (%)

Gapdh   *pre-Gapdh*   mCharme   pCharme

■ Cytoplasm
☐ Nucleus

B   E15.5
DAPI / pCharme

Heart

A   A
LV
RV

Spinal cord

VH
DH

C

3D Pearson's Coefficient

pCharme/MATR3

■ Measured correlation
☐ Total positive correlation

D

TNNT2/MATR3/DAPI   TNNT2   MATR3   MATR3   DAPI

Skeletal muscle

Charme^WT

Charme^KO

E

Heart
Charme^WT   Charme^KO

Spinal cord
Charme^WT

TNNT2/MATR3/DAPI

MAP2/MATR3/DAPI

F

Coefficient of Variation (CV)

16

8

0

***

Charme^WT   Charme^KO

Figure 3-Figure supplement 1. Related to Figure 3

A) Quantification of the subcellular distribution of pCharme and mCharme in cardiac tissues from neonatal mice. Histogram shows the quantification by RT-qPCR of the 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 tissue sections (left panels) and their magnification (right panels). Whole heart (white dashed lines), 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 shows the mean ± SEM calculated over 237 colocalized nuclear signals from 3 independent experiments.

D) Representative images for for MATR3 (green), TnnT2 (red) and DAPI (blue) stainings on Charme[WT] and Charme[KO] skeletal muscles from E15.5 tissue sections. ROI (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: Representative images for for MATR3 (green), TnnT2 (red) and DAPI (blue) stainings on Charme[WT] and Charme[KO] heart from E15.5 tissue sections. Right panel: Representative images for MATR3 (green), Map2 (red) and DAPI (blue) stainings on Charme[WT] and Charme[KO] spinal cord from E15.5 tissue sections. Each image is a representative of three individual samples. Scale bars, 10 μm.

F) Quantification of MATR3 fluorescence intensity distribution (CV=Coefficient of variation) in Charme[WT] (N=1346 nuclei) and Charme[KO] (N=1404 nuclei). The red X indicate the mean value of CV distribution.

Figure 4-figure supplement 1
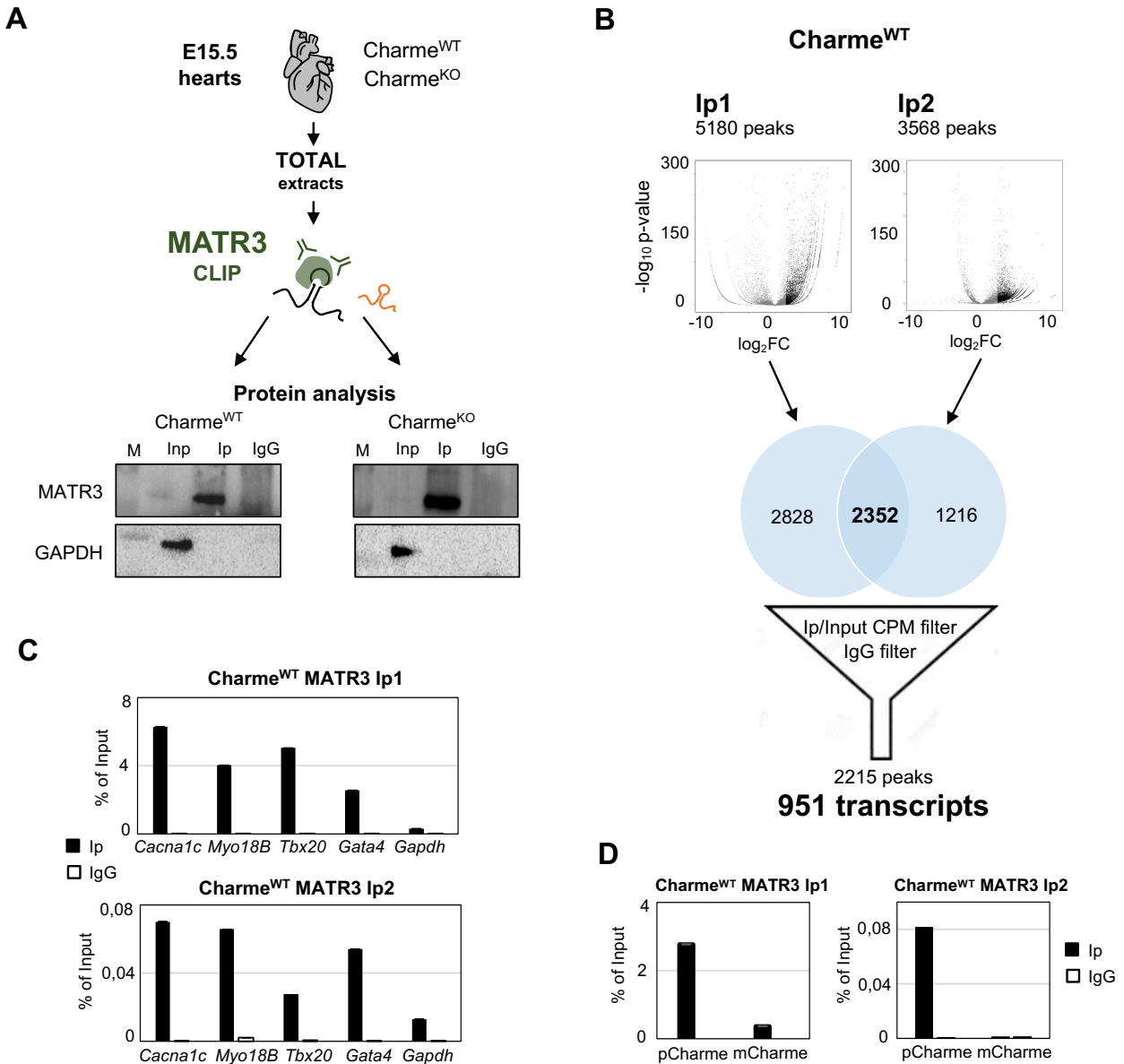
1240

1241 **Figure 4-Figure supplement 1. Related to Figure 4**

1242 **A)** Schematic representation of MATR3-CLIP assay as performed from fetal hearts (E15.5) in

1243 Charme^{WT} and Charme^{KO} conditions. MATR3 western blot analysis on the retrieved protein fractions

1244 is shown. GAPDH protein serves as a loading control. Input (Inp) samples represent 10% of the total

1245 protein extracts.

1246 **B)** Schematic representation of the workflow used for MATR3 CLIP-seq analysis. Volcano plots

1247 represent the fold-enrichment over Input (log$_2$ Fold enrichment, x-axis) and significance (-log$_{10}$ p-

1248 value, y-axis) of MATR3 peaks in the Ip1 (left panel) and Ip2 (right panel) samples. Black dots

1249 represent the significantly enriched peaks ($\log_2$ Fold enrichment > 2 and FDR < 0.05). Venn diagrams

1250 depict the intersection between Ip1 and Ip2 significantly enriched peaks. The 2215 filtered peaks

1251 correspond to the 951 MATR3-bound transcripts. CPM: Counts Per Million. See Materials and

1252 Methods for details.

1253 **C)** RT-qPCR quantification of *Cacna1c*, *Myo18B*, Tbx20 and *Gata4* RNA recovery in MATR3 Ip1

1254 (upper panel), Ip2 (lower panel) and IgG Charme[WT] samples. *Gapdh* transcript serves as negative

1255 control. Values are expressed as percentage (%) of Input.

1256 **D)** RT-qPCR quantification of pCharme and mCharme RNA recovery in MATR3 Ip1 (left panel),

1257 Ip2 (right panel) and IgG Charme[WT] samples. Values are expressed as percentage (%) of Input.

1258

Figure 5-figure supplement 1

A

B

C

D

Charme^WT vs Charme^KO MATR3 IP peaks
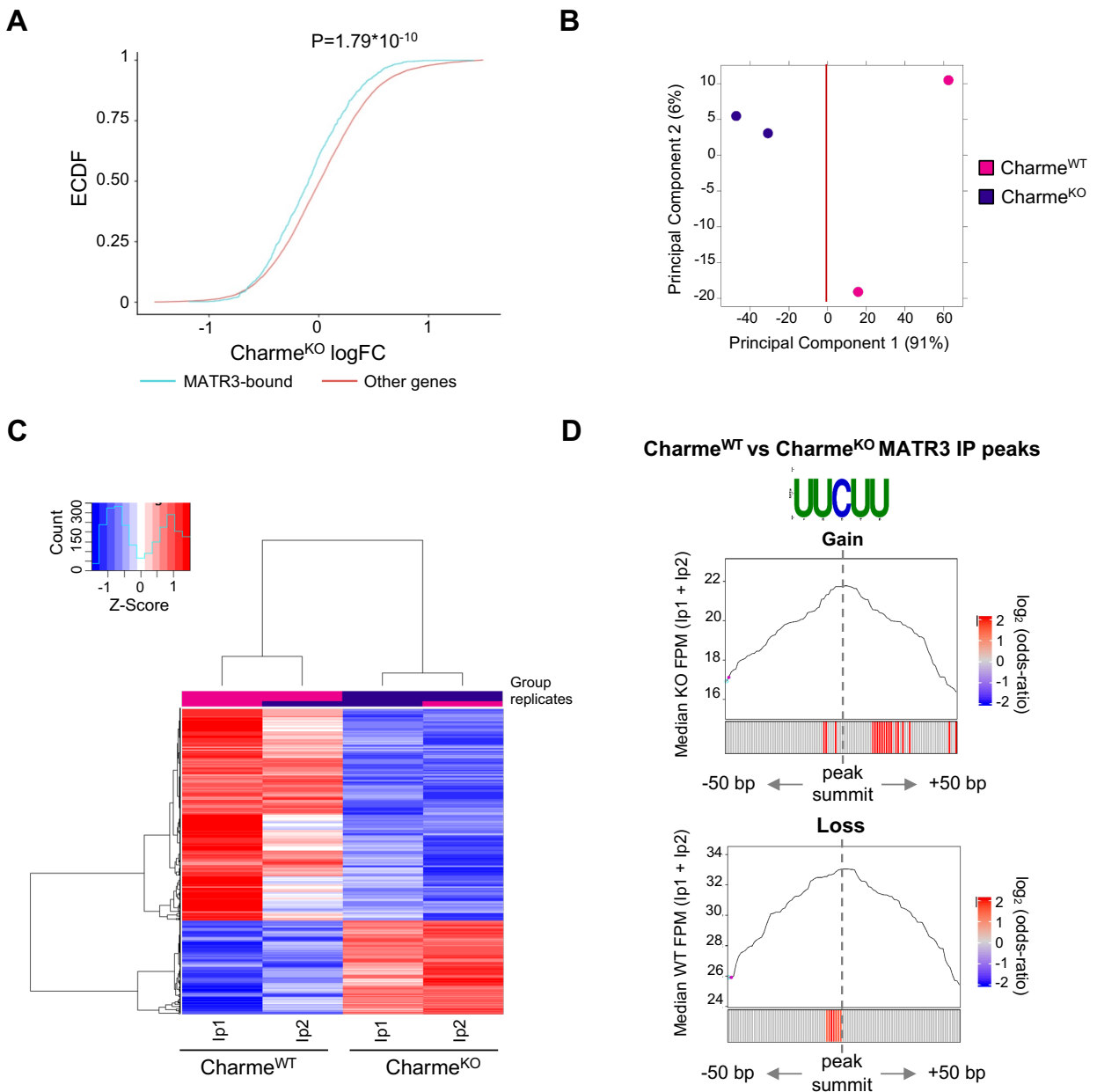
1259

Figure 5-Figure supplement 1. Related to Figure 5

A) Empirical Cumulative Distribution Functions (ECDF) showing the RNA abundance of MATR3

targets in Charme^KO condition compared to the other expressed genes. Significance was determined

using a two-sided Kolmogorov–Smirnov (KS) test.

B) Principal Component analysis (PCA) plot of Charme^WT and Charme^KO MATR3 CLIP-seq peaks.

Plot was obtained from DiffBind package. See Material and Methods for details.

C) Heatmap visualization of Charme^{WT} and Charme^{KO} MATR3 CLIP-seq peaks. Plot was produced by DiffBind package. Z-Score of normalized reads counts is shown. See Material and Methods for details.

D) Positional enrichment analysis of MATR3 motif in differential MATR3 CLIP-seq peaks (Charme^{WT} *vs* Charme^{KO}). For each analysed position close to peak summit, line plot displays the median CLIP-seq signal (FPM, IP1 + IP2) while heatmap displays the $\log_2$ odds-ratio of UUCUU motif enrichment. Significant enrichments (p-value < 0.05) are shown in red. See Material and Methods for details. CLIP-seq signal in Charme^{KO} and Charme^{WT} conditions is displayed for «Gain» and «Loss» peaks, respectively.

Supplementary File 1. Charme Tss usage data collected from Zenbu genome browser

Supplementary File 2. RNA-seq in Charme^{WT} and Charme^{KO} neonatal hearts

Supplementary File 3. Ecocardiography measurement for Charme^{WT} and Charme^{KO} animals.

Supplementary File 4. MATR3 CLIP-seq in Charme^{WT} and Charme^{KO} fetal hearts

Supplementary File 5. List and sequences of the oligonucleotides, siRNAs, antibodies and imaging probes used.