

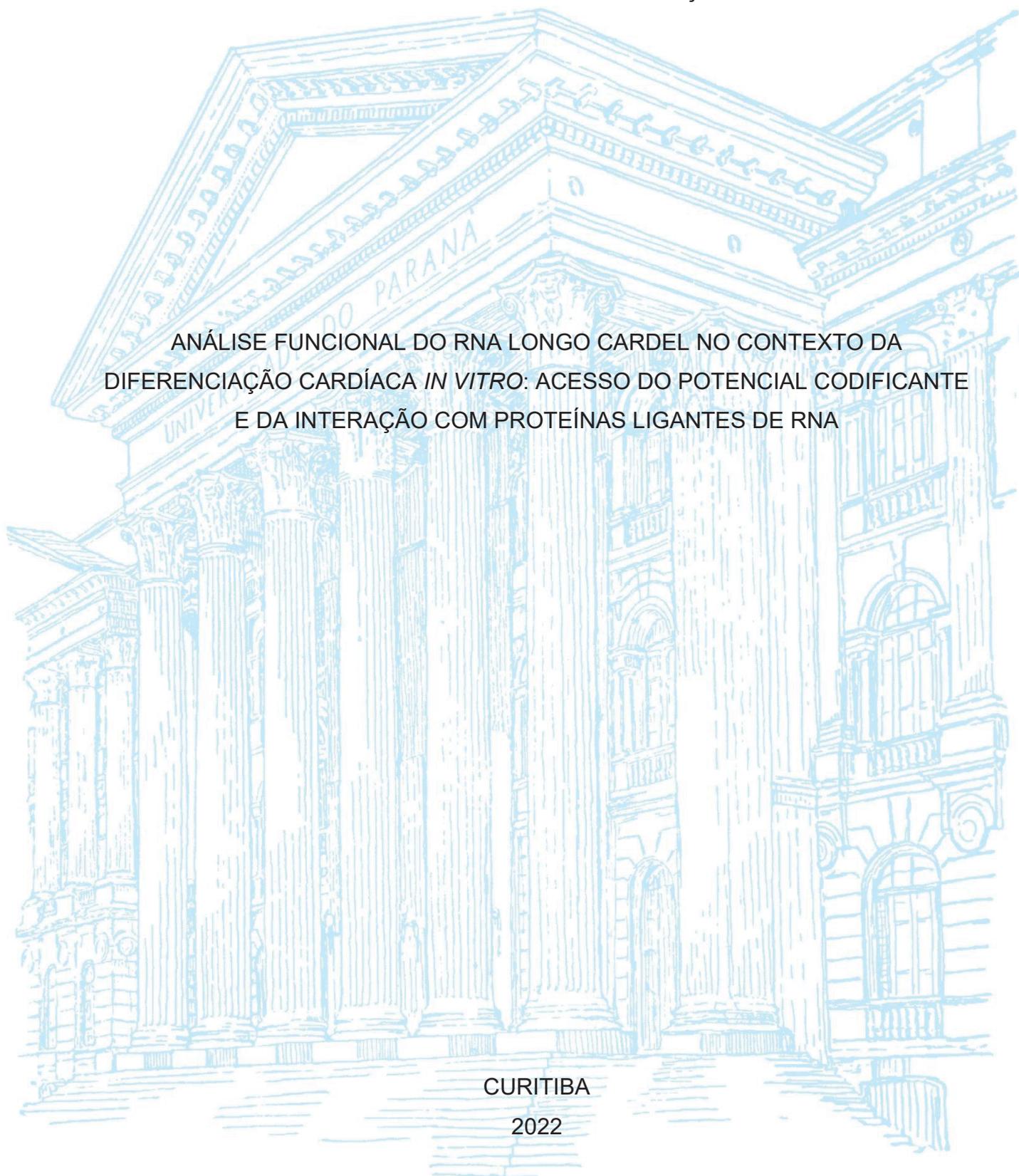
UNIVERSIDADE FEDERAL DO PARANÁ

RHAÍZA STELLA VIEIRA FRANÇA

ANÁLISE FUNCIONAL DO RNA LONGO CARDEL NO CONTEXTO DA  
DIFERENCIAÇÃO CARDÍACA *IN VITRO*: ACESSO DO POTENCIAL CODIFICANTE  
E DA INTERAÇÃO COM PROTEÍNAS LIGANTES DE RNA

CURITIBA

2022



RHAÍZA STELLA VIEIRA FRANÇA

ANÁLISE FUNCIONAL DO RNA LONGO CARDEL NO CONTEXTO DA  
DIFERENCIAÇÃO CARDÍACA *IN VITRO*: ACESSO DO POTENCIAL CODIFICANTE  
E DA INTERAÇÃO COM PROTEÍNAS LIGANTES DE RNA

Dissertação apresentada ao curso de Pós-Graduação em Biologia Celular e Molecular, Setor de Ciências Biológicas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Biologia Celular e Molecular.

Orientador: Prof. Dr. Bruno Dallagiovanna Muñiz

CURITIBA

2022

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)  
UNIVERSIDADE FEDERAL DO PARANÁ  
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIAS BIOLÓGICAS

França, Rhaíza Stella Vieira

Análise funcional do RNA longo *CARDEL* no contexto da diferenciação cardíaca *in vitro* : acesso do potencial codificante e da interação com proteínas ligantes de RNA / Rhaíza Stella Vieira França. – Curitiba, 2022.

1 recurso on-line : PDF.

Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Ciências Biológicas, Programa de Pós-Graduação em Biologia Celular e Molecular.

Orientador: Prof. Dr. Bruno Dallagiovanna Muñiz.

1. RNA longo não codificante. 2. Diferenciação celular. 3. Proteínas de ligação a RNA. I. Dallagiovanna Muñiz, Bruno. II. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Biologia Celular e Molecular. III. Título.

Bibliotecária: Giana Mara Seniski Silva CRB-9/1406



## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação BIOLOGIA CELULAR E MOLECULAR da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **RHAÍZA STELLA VIEIRA FRANÇA** intitulada: **Análise funcional do RNA longo CARDEL no contexto da diferenciação cardíaca in vitro: acesso do potencial codificante e da interação com proteínas ligantes de RNA**, sob orientação do Prof. Dr. BRUNO DALLAGIOVANNA MUNIZ, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestra está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 06 de Setembro de 2022.

Assinatura Eletrônica

14/09/2022 17:02:35.0

BRUNO DALLAGIOVANNA MUNIZ

Presidente da Banca Examinadora

Assinatura Eletrônica

14/09/2022 15:05:21.0

DANIELA FIORI GRADIA

Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

14/09/2022 15:25:26.0

DANIELA PARADA PAVONI

Avaliador Interno

Dedico essa dissertação a todas as pessoas que, com sua coragem e curiosidade, colorem a monotonia dos dias.

## **AGRADECIMENTOS**

Aos meus pais, Natalia e Marcio, e minha avó Otília, pelo amor e suporte incondicionais. Aos amigos novos e antigos, que acreditaram nos meus talentos e me apoiaram durante a vida. Ao Dr. Bruno Dallagiovanna, pela orientação, pelas discussões científicas e por acreditar no meu potencial. Aos amigos e (ex) membros do Laboratório de Biologia Básica de Células-tronco (LABCET), pelo ambiente de trabalho divertido e acolhedor. Um agradecimento especial a Isabela Tiemy Pereira, Rubens Gomes Júnior e Cíntia Horiouchi, pela paciência e gentileza em me guiarem nessa jornada. Aos membros da banca examinadora: Dra. Daniela Gradia, Dra. Daniela Parada Pavoni e Dr. Alejandro Correa Dominguez, pelo tempo dedicado à leitura deste documento e por contribuírem para a minha formação. À Pós-graduação em Biologia Celular e Molecular pela oportunidade de me especializar nessa área fascinante. Às agências financiadoras FIOCRUZ, CNPq, CAPES e FUNDAÇÃO ARAUCÁRIA pelo auxílio financeiro.

“Gilliat [...] contemplava a natureza de um modo singular. Tinha visto algumas vezes, na água do mar, completamente límpida, animais inesperados, de grandes dimensões, de formas diversas, os quais, fora da água, assemelhavam-se a cristal mole, e, tornados à água, confundiam-se com ela [...]; disto concluía ele que, se a água era habitada por transparências vivas, bem podia ser que o ar fosse habitado por transparências igualmente vivas. [...] Criaturas cor do ar podem escapar aos nossos olhos por causa da luz; quem nos prova que essas criaturas não existem?”. (Trabalhadores do Mar-Victor Hugo, 1866, p. 40)

## RESUMO

Além de uma complexa rede de fatores de transcrição, o processo de diferenciação cardiomiogênica em vertebrados envolve também mecanismos de regulação mediados por proteínas ligantes de RNA (RBPs) e RNAs não codificantes (ncRNAs). Dentre os ncRNAs, os microRNAs e os RNAs longos (lncRNAs) desempenham funções importantes na cardiomiogênese, atuando na regulação de mRNAs. Os miRNAs geralmente inibem mRNAs, enquanto lncRNAs podem regular mRNAs de maneira direta ou mediada por RBPs. Em um estudo de expressão gênica de diferentes estágios de diferenciação cardíaca *in vitro*, um lncRNA específico denominado *CARDEL* (*CARDiac DEvelopment LncRNA*) despertou interesse por apresentar expressão aumentada em células progenitoras cardíacas e estar associado aos polissomos. Duas hipóteses foram levantadas sobre os motivos desta associação e foram testadas ao longo do trabalho. Uma primeira hipótese considerou que *CARDEL* pode ter potencial codificante para pequenos peptídeos, como já descrito para outros lncRNAs. Bases de dados recentemente anotaram o *CARDEL* como codificante para o peptídeo SERTM2, a partir de um de seus pequenos quadros abertos de leitura (sORF-1). A avaliação do potencial de tradução do sORF-1 de *CARDEL* em células humanas envolveu a construção de um plasmídeo pcDNA 3.1 contendo o sORF-1 fusionado à sequência codificante (CDS) de *enhanced GFP* (eGFP), via clonagem com as endonucleases BamHI/EcoRI. A microscopia de fluorescência de HEK293 transfectadas com a construção resultante pcDNA 3.1-CARDEL1A-eGFP permitiu observar um sinal correspondente a eGFP. O *Western Blot* dos extratos indicou a presença do peptídeo SERTM2-eGFP, reforçando o potencial codificante de *CARDEL*. Para caracterizar o possível papel de SERTM2 na função do *CARDEL* durante a cardiomiogênese, uma nova linhagem de células-tronco pluripotentes PLZ foi gerada e usada na superexpressão induzida do *CARDEL* com uma mutação em seus sORFs (*CARDEL mut*). Células PLZ foram transfectadas com o vetor integrativo AAVS1-CARDEL mut e selecionadas por puromicina. Um dos clones PLZ *CARDEL mut* (C8.1) apresentou sistema de indução funcional, expressão de marcadores de pluripotência NANOG e OCT4A e cariótipo adequado, sendo escolhido para experimentos de diferenciação cardíaca já iniciados. A segunda hipótese do presente trabalho foi de que o *CARDEL* atua na regulação pós-transcricional de mRNAs durante a cardiomiogênese ao interagir com RBPs. Esse potencial foi explorado *in silico* por meio da anotação de sítios de ligação de RBPs no *CARDEL* e da estimativa da probabilidade de interação. Três proteínas candidatas foram selecionadas: LN28A, YBOX1 e RBFOX1. Interações *CARDEL*-RBPs também foram analisadas pela técnica de incPRINT, que consistiu na superexpressão conjunta de RBPs e do cDNA de *CARDEL* e estimou sua força de interação por unidades relativas de luminescência (RLU). Testes preliminares de interação do *CARDEL* com RBPs comuns resultaram em RLU abaixo do limite mínimo requerido. Como a maioria das proteínas testadas interage com sequências intrônicas ou 5'/3' – UTR, ausentes no cDNA de *CARDEL*, a inclusão dessas sequências pode viabilizar o uso do incPRINT. Apesar da natureza da associação do *CARDEL* a polissomos não ter sido definitivamente elucidada, o presente trabalho conseguiu fortalecer hipóteses e desenvolver ferramentas para estudos mais aprofundados sobre o papel desse transcrito na cardiomiogênese.

Palavras-chave: lncRNA, diferenciação cardíaca, *CARDEL*, RBP.

## ABSTRACT

In addition to a complex network of transcription factors, the process of cardiomyogenic differentiation in vertebrates also entails regulatory mechanisms mediated by RNA-binding proteins (RBPs) and non-coding RNAs (ncRNAs). Among ncRNAs, microRNAs and long non-coding RNAs (lncRNAs) perform important functions in cardiomyogenesis, such as regulation of mRNAs. MiRNAs generally inhibit mRNAs, while lncRNAs might regulate mRNA directly or siding with RBPs. In a study of gene expression of different stages of cardiac differentiation *in vitro*, a specific lncRNA named *CARDEL* (CARDiac DEVELOPMENT lncRNA) came to light, as it showed an increased expression in cardiac progenitors and was associated with polysomes. Two hypotheses were raised about these associations and were tested through this study. The first hypothesis considered that *CARDEL* has a coding potential to small peptides, as described to other lncRNAs. Recently, *CARDEL* was annotated as protein coding for the SERTM2 peptide, transcribed from one of its small open reading frames (sORF-1). Evaluation of *CARDEL* sORF-1 coding potential in human cells required the construction of a pcDNA 3.1 vector comprising a sORF-1 fused to enhanced GFP (eGFP) coding sequence (CDS), using BamHI/EcoRI endonucleases. Fluorescent microscopy of HEK293 cells transfected with pcDNA 3.1-*CARDEL*1A-eGFP plasmid enabled eGFP signal detection. Western Blot analysis of cell extracts pointed to SERTM2-eGFP peptide synthesis, strengthening the coding potential hypothesis for *CARDEL*. Aiming to characterize a possible role of SERTM2 in *CARDEL* function during cardiomyogenesis, a novel stem-cell PLZ lineage was made and used to promote induced overexpression of *CARDEL* containing a frameshift mutation in its sORFs (*CARDEL mut*). PLZ cells were transfected with the integrative vector AAVS1-*CARDEL mut* and selected by puromycin treatment. One of the PLZ *CARDEL mut* clones (C8.1) presented a functional expression cassette, expressed *NANOG* and *OCT4A* pluripotency markers and a correct karyotype, being chosen for cardiac differentiation experiments that have just started. The second hypothesis of the current work was that *CARDEL* acts in mRNAs post-transcriptional regulation during cardiomyogenesis, interacting with RBPs. This potential was explored *in silico* through annotation of *CARDEL*'s RBP binding sites and calculation of its interaction probability. Seven candidate proteins were selected: LN28A, DAZP1, HNRPL, MSI1, RBFOX1, SMAG1/SMAUG1 and YBOX1. *CARDEL*-RBPs interactions were also analysed by incPRINT, a technique consisting in co-overexpression of RBPs and *CARDEL* cDNA that estimated *CARDEL*-RBPs interaction strength in relative luminescent units (RLU). Preliminary interaction tests between *CARDEL* and common binding RBPs resulted in RLUs below the required minimum threshold. As most of the tested proteins interact with intronic or 5'/3' – UTR sequences lacking in *CARDEL* cDNA, the inclusion of these sequences could make the incPRINT technique viable. Although the mechanism underlying *CARDEL* post-transcriptional regulation remains elusive, the current work managed to strengthen hypotheses and to develop tools for more detailed studies about the role of this transcript to cardiomyogenesis

Keywords: lncRNA, cardiac differentiation, *CARDEL*, RBP.

## LISTA DE FIGURAS

FIGURA 1 - REDE DE REGULAÇÃO DA CARDIOMIOGÊNESE .....	20
FIGURA 2 - EXPRESSÃO E EFEITO DE CARDEL EM MARCADORES CARDÍACOS.....	28
FIGURA 3 – VARIANTES DE <i>SPLICING</i> DE <i>CARDEL</i> E ESTRATÉGIA PARA MUTAÇÃO DA CDS PREDITA .....	30
FIGURA 4– ESTRATÉGIA DE OBTENÇÃO DE <i>CARDEL1A.eGFP</i> .....	33
FIGURA 5 – ETAPAS PARA OBTENÇÃO DE <i>pcDNA 3.1-CARDEL1A.eGFP</i> .....	47
FIGURA 6 – AMPLIFICAÇÃO DE SEQUÊNCIA FUSIONADA <i>CARDEL1A.eGFP</i> ...	48
FIGURA 7 – EXPRESSÃO DE <i>pcDNA 3.1-CARDEL1A.eGFP</i> EM HEK293 .....	50
FIGURA 8 – <i>WESTERN BLOT</i> DE EXTRATOS DE HEK293 TRANSFECTADAS COM <i>pcDNA 3.1-CARDEL1A.eGFP</i> .....	51
FIGURA 9 – PREDIÇÃO <i>IN SILICO</i> DE MICROPEPTÍDEO SERTM2 .....	52
FIGURA 10 – TRANSFECCÃO COM <i>AAVS1--eGFP</i> E EXPRESSÃO DE <i>CARDEL</i> EM CÉLULAS HEK293 .....	54
FIGURA 11 – TESTES DE TRANSFECCÃO E CARACTERIZAÇÃO DE PLZ <i>CARDEL mut.</i> .....	56
FIGURA 12 – TESTES DE INTEGRAÇÃO DO SISTEMA DE SUPEREXPRESSÃO DE <i>CARDEL mut</i> EM CÉLULAS PLZ .....	58
FIGURA 13 – PERFIS DE EXPRESSÃO DE HEK293 TRANSFECTADAS COM <i>pcDNA3.1-CARDEL-10XMS2</i> .....	61
FIGURA 14 – INTERAÇÃO DE <i>CARDEL</i> COM LIGANTES COMUNS.....	63
FIGURA 15 – DIAGRAMA DE VENN DE RBPs COM POTENCIAL DE LIGAÇÃO PARA <i>CARDEL SERTM2-201</i> E <i>SERTM2-202</i> .....	64
FIGURA 16 – MAPA DO VETOR <i>AAVS1-TRE-GW-rtTA</i> COM <i>CARDEL</i> COM MUTAÇÃO <i>FRAMESHIFT</i> .....	87
FIGURA 17 – MAPA DO VETOR <i>pcDNA3.1-CARDEL10xMS2</i> .....	87
FIGURA 18 – MAPA DO VETOR <i>pcDNA 3.1/myc-His B CARDEL1A.eGFP</i> .....	91
FIGURA 19 – TESTE DE SUPEREXPRESSÃO DE TRANSCRITO DE <i>CARDEL</i> EM PLZ <i>CARDEL mut</i> .....	91

## LISTA DE TABELAS

TABELA 1 - <i>PRIMERS</i> PARA AMPLIFICAÇÃO DE SEQUÊNCIAS DE sORF <i>CARDEL</i> -eGFP .....	34
TABELA 2 – <i>PRIMERS</i> UTILIZADOS PARA qPCR .....	41
TABELA 3 – PROTEÍNAS LIGANTES DE RNA COM POTENCIAL DE INTERAÇÃO COM <i>CARDEL</i> .....	65
TABELA 4 – INFORMAÇÕES FUNCIONAIS DE PROTEÍNAS SELECIONADAS APÓS APLICAÇÃO DOS CRITÉRIOS DE EXCLUSÃO .....	70
TABELA 5 – MEIOS DE CULTIVO E KITS COMERCIAIS UTILIZADOS.....	89

## LISTA DE SIGLAS

5'/3'-UTR	-	5'/3' <i>Untranslated Region</i>
AGIL	-	<i>G-rich RNA motif</i>
ASRPS	-	<i>A Small Regulatory Peptide of STAT3</i>
bGH	-	fator de crescimento bovino
BMP	-	<i>Bone Morphogenetic Protein</i>
CACNA1C	-	<i>Calcium Voltage-Gated Channel Subunit Alpha1 C</i>
CAMK2D	-	<i>Ca<sup>2+</sup>/calmodulin-dependent protein kinase type II subunit delta</i>
CARDEL	-	<i>Cardiac Development RNA</i>
CARMEN	-	<i>(CAR)diac (M)esoderm (E)nhancer-associated (N)oncoding RNA</i>
CCR4-NOT	-	<i>Carbon Catabolite Repression Negative on TATA-less</i>
cDNA	-	DNA Complementar
CDS	-	Sequência(s) Codificante(s)
CE(s)	-	Corpos Embrioides
CMV	-	Citomegalovírus
C-MYC	-	<i>Cellular myelocytomatosis oncogene</i>
CNBP	-	<i>Cellular Nucleic Acid Binding Protein</i>
D0/5/9/10	-	Dia 0/5/9/10 de protocolo de diferenciação cardíaca
DAZP1	-	<i>DAZ-associated Protein 1</i>
DMEM	-	Meio <i>Dulbecco's Modified Eagle</i>
DNA	-	Ácido Desoxirribonucleico
DOX	-	Doxiciclina
EDTA	-	Ácido Etilenodiamino Tetra-Acético
eGFP	-	<i>enhanced Green Fluorescent Protein</i>
eIF4F	-	<i>Eukaryotic initiation factor 4F</i>
EOMES	-	Fator de transcrição T-box Eomesodermina
EPR	-	<i>Epithelial cell Program Regulator lncRNA</i>
ESC (s)	-	Células-Tronco Embrionárias
F-12	-	Mistura nutriente <i>Ham's F-12</i>
FENDRR	-	<i>Fetal-lethal Non-coding Developmental Regulatory RNA</i>
FGF	-	<i>Fibroblast Growth Factor</i>
FORCP	-	<i>FOXA1-Regulated Conserved Small Protein</i>
FOXF1	-	<i>Forkhead Box F1</i>
G3BP1	-	<i>Ras GTPase-activating Protein-Binding Protein 1</i>
GAPDH	-	<i>Glyceraldehyde-3-Phosphate Dehydrogenase</i>
GATA	-	<i>Gata binding protein</i>
HAND	-	<i>Heart and Neural Crest Derivatives Expressed</i>
HNRPC	-	<i>Heterogeneous Nuclear Ribonucleoprotein C gene</i>
IGF2	-	<i>Insulin Like Growth Factor 2</i>
incPRINT	-	<i>In-cell protein-RNA interaction</i>
iPSC(s)	-	Células-tronco Pluripotentes Induzidas
iROCK	-	<i>Rho-associated, coiled-coil containing protein kinase inhibitor</i>
IRX3	-	<i>Iroquois Homeobox 3</i>
ISL1	-	<i>ISL LIM Homeobox 1</i>
KDR	-	<i>Kinase Insert Domain Receptor</i>
KI4	-	Krüppel-like transcription factor 4
KLF4	-	<i>Krueppel-like factor 4</i>
KO	-	<i>Knockout; nocaute, silenciado</i>

LB	-	Meio Luria-Bertani
LEMP	-	lncRNA <i>Encoded Micropeptide</i>
L-Gln	-	L-glutamina
lincRNA	-	RNA Longo intergênico não-codificante
LincRNA	-	RNA longo não codificante intergênico
LN28A	-	<i>Lin28 Homolog A</i>
lncRNA	-	RNA Longo não-codificante
MADS	-	<i>Minichromosome Maintenance</i>
MEF2	-	<i>Myocyte Enhancer Factor 2</i>
Menhir	-	<i>MElaNoma HIIndrance RNA</i>
MESP1	-	<i>Mesoderm Posterior BHLH Transcription Factor 1</i>
MIAT	-	<i>Myocardial infarction associated transcript</i>
miRNA	-	micro RNA
MOV-10	-	<i>Mov10 RISC Complex RNA Helicase gene</i>
mRNA	-	RNA mensageiro
MS2	-	RNA MS2 <i>stem loop</i>
MS2CP	-	<i>MS2 Bacteriophage Coat Protein</i>
MSI1	-	Musashi1 Protein
MSI1H	-	<i>Musashi RNA Binding Protein 1</i>
ncRNA	-	RNA não-codificante
NKX2-5	-	<i>NK2 Homeobox 5</i>
Oct3/4	-	<i>Octamer-binding transcription fator 3/4</i>
PAINT	-	<i>Phylogenetic Annotation and INference Tool</i>
PAN2/3	-	Poly(A)-nuclease 2/3
PBS	-	Salina Tamponada com Fosfato
PCR	-	Reação em Cadeia da Polimerase
PCR	-	Reação em cadeia de polimerase
PEI	-	Poli-etilenamina
PITX2	-	<i>Paired Like Homeodomain 2</i>
pLDDT	-	pontuação de confiança por resíduo de aminoácido
PRC2	-	<i>Polycomb Repressive Complex 2</i>
qPCR	-	PCR quantitativa
RBFOX	-	<i>RNA-binding protein, fox-1 homolog</i>
RBM	-	<i>RNA-binding Motif</i>
RBP	-	Proteína Ligante de RNA
RF	-	<i>Random Forest Classifier</i>
RISC	-	<i>RNA-induced Silencing Complex</i>
RLU	-	<i>Relative Luminescence Unit</i>
RNA POL	-	RNA Polimerase II
RNA-seq	-	Sequenciamento de RNAs em larga escala
RNP	-	Complexo Ribonucleoproteico
RPKM	-	<i>Reads Per Kilobase per Million mapped reads</i>
RPMI	-	Meio Roswell Park Memorial Institute
rRNA	-	RNA ribossomal
rtTA	-	<i>reverse tetracycline-controlled TransActivator</i>
SDS	-	Dodecil Sulfato de Sódio
SDS-PAGE	-	Eletroforese em Gel de PoliAcrilamida com SDS
SERCA2	-	<i>Sarco(Endo)Plasmic Reticulum Calcium-ATPase 2 Protein</i>
SERTM2	-	<i>Serine rich and transmembrane domain containing 2</i>
SERTM2-201	-	Variante de <i>splicing</i> 1 de <i>CARDEL</i>

SERTM2-202	- Variante de <i>splicing</i> 2 de CARDEL
SFB	- Soro Fetal Bovino
SMAG1	- <i>Protein Smaug homolog 1 (SAMD4) gene</i>
SNRNP70	- <i>Small Nuclear Ribonucleoprotein U1 Subunit 70 gene</i>
sORF	- <i>small Open Reading Frame</i> /pequeno quadro aberto de leitura
Sox2	- <i>SRY-Box Transcription Factor 2</i>
SSEA-3/4	- Stage-specific embryonic antigen-3/4
SVM	- <i>Support Vector Machine Classifier</i>
TBX	- <i>T-box protein</i>
TGF- $\beta$	- <i>Transforming Growth Factor Beta</i>
TNFR2	- <i>Tumor Necrosis Factor Receptor 2</i>
TNNI	- Troponina I, isoforma fetal
TNNI3	- Troponina I, isoforma adulta
TNNT2	- <i>Cardiac Troponin T2</i>
TP53TG1	- <i>TP53 Target 1 lncRNA</i>
TRA-1-60	- T cell receptor alpha locus
TRA-1-81	- Podocalixina
TRE	- <i>Tetracyclin Responsive Element</i>
VEGFR2	- <i>Vascular Endothelial Growth Factor Receptor 2</i>
Wnt	- <i>Wingless/Integrated</i>
WT	- <i>Wild Type</i> ; linhagem selvagem parental
XIST	- <i>X-Inactive Specific Transcript</i>
XRN1	- <i>5'-3' Exoribonuclease 1</i>
YBOX1	- <i>Y-Box Binding Protein 1</i>
ZFN	- <i>Zinc Finger Nucleases</i>

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>16</b>
<b>2 REVISÃO DE LITERATURA</b> .....	<b>19</b>
2.1 ASPECTOS MORFOLÓGICOS E MOLECULARES DA CARDIOMIOGÊNESE	19
2.2 ATUAÇÃO DE RNAs NÃO CODIFICANTES E PROTEÍNAS LIGANTES DE RNA NA CARDIOMIOGÊNESE .....	21
2.3 LNCrNA CARDEL NA DIFERENCIAÇÃO CARDÍACA .....	27
2.4 CÉLULAS-TRONCO PLURIPOTENTES.....	31
<b>3 OBJETIVOS</b> .....	<b>32</b>
3.1 Objetivo geral .....	32
3.2 Objetivos específicos .....	32
<b>4 MATERIAL E MÉTODOS</b> .....	<b>32</b>
4.1 CONSTRUÇÃO DE PLASMÍDEO pCDNA 3.1-HIS B/MYC-CARDEL1A.eGFP	32
4.1.1 Obtenção de amplicons CARDEL1A, sCARDEL1.eGFP e CARDEL1A.eGFP	32
4.1.2 Clonagem de construção em <i>E. coli</i> DH5 $\alpha$ , seleção e <i>screening</i> de clones.....	34
4.2 AVALIAÇÃO DO POTENCIAL CODIFICANTE DO SORF-1 DE CARDEL EM CÉLULAS HEK293.....	35
4.2.1 Eletroforese em gel de poliacrilamida-SDS (SDS-PAGE) e <i>Western Blot</i> .....	36
4.3 CONSTRUÇÃO DE FERRAMENTAS PARA AVALIAÇÃO DO EFEITO DO CARDEL COM MUTAÇÃO NOS SORFS PUTATIVOS .....	38
4.3.1 Transfecção de células embrionárias de rim (HEK293) .....	39
4.3.2 Transfecção de iPSC PLZ com AAVS1-CARDEL mut e seleção da linhagem	39
4.3.3 Quantificação de expressão de RNA por PCR quantitativa (qPCR).....	41
4.3.4 Detecção da sequência integrativa de AAVS1-GW-rtTA e outros testes .....	42
4.3.5 Preparo de células PLZ para cariotipagem.....	43
4.4 DETECÇÃO DE INTERAÇÕES CARDEL-PROTEÍNA POR INCPRINT .....	43
4.4.1 Teste de funcionalidade de pCDNA 3.1 – CARDEL10x MS2.....	43
4.4.2 Técnica de incPRINT.....	44
4.5 ANÁLISES <i>IN SILICO</i> DO POTENCIAL DE INTERAÇÃO CARDEL-RBP .....	45
<b>5 RESULTADOS E DISCUSSÃO</b> .....	<b>46</b>

5.1	OBTENÇÃO DE CONSTRUÇÃO PCDNA 3.1-HIS B/MYC-CARDEL1A.EGFP ..	46
5.2	ANÁLISE DE EXPRESSÃO DE CÉLULAS HEK293 TRANSFECTADAS COM PCDNA 3.1-HIS B/MYC-CARDEL1A.EGFP.....	48
5.3	CARACTERIZAÇÃO DE LINHAGEM DE IPSC PLZ ICARDEL MUT .....	53
5.4	ESTUDO DA INTERAÇÃO DO CARDEL COM RBPS.....	59
5.4.1	Teste de funcionalidade de construção pcDNA 3.1 – CARDEL 10x MS2 .....	60
5.4.2	Teste de interação de CARDEL com proteínas ligantes comuns.....	62
5.4.3	Análise <i>in silico</i> de interações de <i>CARDEL</i> com RBPs .....	63
<b>6</b>	<b>CONSIDERAÇÕES FINAIS E PERSPECTIVAS .....</b>	<b>71</b>
	<b>REFERÊNCIAS.....</b>	<b>74</b>
	<b>ANEXO 1 – MAPA DE VETORES .....</b>	<b>87</b>
	<b>ANEXO 2 – LAUDO DO CARIÓTIPO DE PLZ CARDEL MUT C8.1 .....</b>	<b>88</b>
	<b>APÊNDICE 1 – REAGENTES.....</b>	<b>89</b>
	<b>APÊNDICE 2 – INFORMAÇÕES SUPLEMENTARES DE PLZ CARDEL MUT.....</b>	<b>91</b>

## 1 INTRODUÇÃO

O processo de diferenciação cardiomiogênica em vertebrados é extremamente complexo e depende de uma extensa rede de fatores de transcrição atuando em configurações tempo/espço específicas (MEILHAC; BUCKINGHAM, 2018). Além dos fatores de transcrição, a regulação da expressão gênica durante a cardiomiogênese demanda outras moléculas que atuam tanto em nível transcricional, interagindo com promotores, potenciadores ou fatores epigenéticos, quanto em nível pós-transcricional, participando do processamento e tradução de mRNAs.

Biomoléculas comumente envolvidas nesse processo são as proteínas ligantes de RNA (RBPs) (BLECHHERMONI; LADD, 2013) e os RNAs não-codificantes (ncRNAs) (CHAHAL et al., 2019). As RBPs estão envolvidas na modulação de mecanismos de *splicing*, transporte e tradução de mRNAs. Alguns exemplos de RBPs importantes para o desenvolvimento e funcionalidade de cardiomiócitos são enzimas envolvidas no ciclo do ácido tricarbóxico, proteínas mitocondriais, a ATPase Serca2 (LIAO et al., 2016), bem como os fatores de *splicing* RBM20 (*RNA Binding Motif 20*) e RBM24 (BERTERO et al., 2019 e YANG et al., 2014). Já os ncRNAs são RNAs que a princípio não são traduzidos em proteínas, sendo os microRNAs (miRNAs) e os RNAs longos não codificantes (lncRNAs) os tipos mais estudados.

Os microRNAs possuem como mecanismo canônico de ação a formação de complexos com RISC (*RNA-Induced Silencing Complex*) e são responsáveis pela hidrólise, degradação ou sequestro de RNAs mensageiros (mRNAs), inibindo sua tradução (IWAKAWA e TOMARI, 2022). Os lncRNAs, por sua vez, compreendem a maioria do transcriptoma não-codificante humano. Muitos desses transcritos são poliadenilados e sintetizados pela RNA polimerase II, de maneira semelhante a mRNAs (AMARAL et al., 2011). Em âmbito pós-transcricional, alguns lncRNAs estão envolvidos no *splicing* (GONZALEZ et al., 2015; ZHANG et al., 2017), tempo de meia-vida (CHEN et al., 2016; WU et al., 2019), localização celular (CAI et al., 2018)

e tradução (BOOY et al., 2021) de mRNAs. Ademais, esses RNAs comumente interagem com miRNAs de maneira tempo e sequência específicos, por meio de mecanismos variados. Nesse contexto, lncRNAs podem ter impacto negativo ou positivo na expressão de determinados miRNAs. Alguns lncRNAs possuem papel de “esponjas”, que consiste no sequestro de miRNAs mediado por complementaridade de sequência, capaz de impedir a interação desses pequenos nucleotídeos com seus mRNAs-alvo. De forma inversa, alguns lncRNAs são expressos de maneira co-dependente a determinados miRNAs e por vezes até servem como sequências precursoras para a biogênese desses transcritos curtos (revisado por HUANG, 2018).

Alguns lncRNAs e miRNAs já provaram estar associados à diferenciação cardíaca. Os lncRNAs *CARMEN* e *Moshe*, por exemplo, influenciam no comprometimento das células à mesoderme e a um tipo de progenitor cardíaco, respectivamente (OUNZAIN et al., 2015; KIM et al., 2021). Similarmente, muitos miRNAs expressos durante a diferenciação cardíaca possuem o papel de inibir rotas de diferenciação cujo ponto de partida é a mesoderme, como descrito por Garate et al. (2018), e podem trabalhar de maneira antagônica e coordenada, como os miRNAs miR-1 e miR-133 (WOJCIECHOWSKA et al., 2017).

Em busca de melhor compreender a regulação pós-transcricional envolvida na cardiomiogênese, Pereira (2019) utilizou protocolos de diferenciação cardíaca *in vitro* (KATTMAN et al., 2011; LIAN et al., 2012) para investigar padrões de expressão de células originalmente pluripotentes em diferentes estágios da diferenciação cardíaca. O sequenciamento do transcriptoma dessas células revelou um lncRNA de interesse, o *LINC00890*. Esse transcrito apresentou um pico de expressão no estágio de progenitor cardíaco, tanto em frações livres de ribossomos como em frações correspondentes a poliribossomos ou polissomos, definidos como complexos compostos por uma molécula de RNA associada a múltiplos ribossomos (ALBERTS et al., 2017, p. 349). Experimentos de ganho e perda de função associados à aplicação de protocolos de diferenciação cardíaca estabeleceram uma correlação entre a expressão do *LINC00890* e o comprometimento das células a

cardiomiócitos. Essas conclusões estimularam a investigação mais detalhada do transcrito de *LINC00890*, renomeado *CARDEL* (CARdiac DEvelopment LncRNA).

De maneira inesperada, a consulta a bases de dados GenBank/UniProt e Ensembl indicou que a anotação do transcrito *CARDEL* havia sido modificada de não codificante para codificante. Essa atualização se baseou em sistemas de curadoria de dados de sequenciamento e anotação de genomas, como o PAINT (*Phylogenetic Annotation and Inference Tool*) (GAUDET et al., 2011), no caso do GenBank, e o PhyloCSF (MUDGE et al., 2019), no caso do Ensembl. Em ambos, o transcrito de *CARDEL* foi anotado como codificante para um pequeno peptídeo nomeado SERTM2, cujo quadro aberto de leitura (ORF) está posicionado no éxon 3 de *CARDEL*. A anotação promovida por PhyloCSF ainda especifica a variante de *splicing* codificante a esse peptídeo, nomeada SERTM2-202.

Com base nessas recentes informações, formulou-se a hipótese de que o transcrito *CARDEL* SERTM2- 202 se associaria a polissomos para ser traduzido em no peptídeo SERTM2, de 90 aminoácidos. Por ser derivado de um ORF pequeno e ser composto por menos de 100 aminoácidos, SERTM2 pode ser referido como um micropeptídeo, como discutido por Sousa; Farkas (2018b). Dessa maneira, o presente trabalho empregou estratégias de superexpressão a fim de determinar se a sequência codificante (CDS) anotada pode ser traduzida em células humanas, e se o efeito observado nos experimentos iniciais de diferenciação cardíaca é provocado pelo transcrito ou pelo micropeptídeo putativo.

Paralelamente, explicações alternativas à associação de *CARDEL* a polissomos foram exploradas pela criação de uma segunda hipótese, na qual esse transcrito poderia estar regulando a tradução de mRNAs. Essa especulação encontra respaldo na literatura, uma vez que já são descritos mecanismos nos quais lncRNAs regulam mRNAs de maneira direta (BOOY et al., 2021) ou através de complexos ribonucleoproteicos (RNPs) com capacidade de sequestrar mRNAs (WANG et al., 2020) ou proteínas reguladoras da tradução (ELGUINDY; MENDELL, 2021). Nesse contexto, o presente trabalho também buscou detectar as proteínas

associadas ao *CARDEL*, visando identificar possíveis alvos e reguladores do transcrito.

## 2 REVISÃO DE LITERATURA

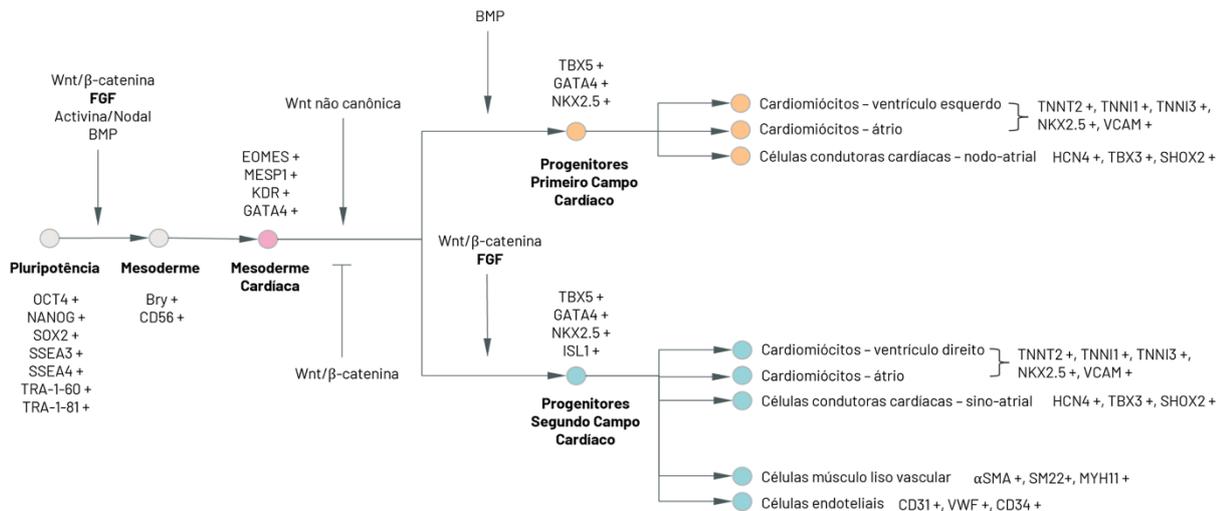
### 2.1 ASPECTOS MORFOLÓGICOS E MOLECULARES DA CARDIOMIOGÊNESE

O coração é constituído por camadas formadas majoritariamente a partir da mesoderme cardíaca, sendo o miocárdio a porção responsável pela contratilidade do coração. Embora essa camada seja composta por vários tipos celulares, são os cardiomiócitos que definem sua função, sendo capazes de contrair-se coordenadamente. Essas células diferenciam-se a partir do epiblasto, uma camada epitelial interna do embrião em estágio de blastocisto que contém células destinadas a se tornar células somáticas. Durante a gastrulação, algumas células do epiblasto diferenciam-se em células de mesoderme pela interação com ligantes das vias de sinalização WNT (*Wingless/Integrated*), Fgf (*Fibroblast Growth Factor*) e BMP (*Bone Morphogenetic Protein*) (revisado por PARIKH et al., 2015). Esse processo é marcado pela expressão sucessiva do fator de transcrição T-box Brachyury/T e do fator de crescimento vascular endotelial 2 (VEGFR2) (revisado por SPÄTER et al., 2014). Em seguida, as células começam a se especializar em mesoderme lateral com a inibição da expressão de Brachyury/T e ativação do Mesp1 (*Mesoderm Posterior BHLH Transcription Factor 1*), promovida pelo fator de transcrição T-box Eomesodermina (Eomes) (BRADE et al., 2013; COSTELLO et al., 2011).

Em camundongos, células Mesp1<sup>+</sup> migram pelo tecido epitelial, posicionando-se na porção anterior da dobra que dará origem à cabeça do embrião, em uma estrutura denominada crescente cardíaco. Nesse ponto, as células se especializam em progenitores do primeiro campo cardíaco e do segundo campo cardíaco (revisado por BRADE et al., 2013). A diferença principal entre esses progenitores é que células do primeiro campo cardíaco estão comprometidas ao miocárdio, sendo capazes de se diferenciar em cardiomiócitos ou células condutoras cardíacas, já progenitores do segundo campo cardíaco podem dar origem a

cardiomiócitos, células endoteliais ou células do músculo esquelético/liso. No quesito molecular, a distinção é notada pela maior expressão de *Isl1* por células do segundo campo cardíaco (revisado por SPÄTER et al., 2014 e MEILHAC; BUCKINGHAM, 2018), como evidente na Figura 1.

FIGURA 1 - REDE DE REGULAÇÃO DA CARDIOMIOGÊNESE



NOTA: Vias de sinalização envolvidas durante a diferenciação cardíaca, podendo ser ativadas (representadas por pontas de setas) ou inibidas (indicadas por ponta em barra) dependendo do estágio de diferenciação; abaixo de cada estágio de diferenciação, são indicados conjuntos genes marcadores, que teoricamente permitem discernir tipos celulares similares, como por exemplo cardiomiócitos atriais, do ventrículo esquerdo e do direito.

Fonte: Pereira (2019), adaptado de Später et al. (2014), com informações adicionais de Brade et al. (2013) e Skelton et al. (2017).

O mecanismo básico de diferenciação de progenitores a cardiomiócitos envolve membros da família GATA dedo de zinco (GATA4 e GATA6), as proteínas T-box 5 (como TBX5), do homeodomínio NK (NKX2-5), MADS box (como MEF2) e HAND (HAND1 e HAND2). A proteína TBX5 é capaz de interagir tanto com NKX2-5 quanto como GATA4, atuando como co-fator de transcrição para genes cardíacos (revisado por CUI et al., 2018). A expressão de *Mef2* é regulada positivamente por *Isl1* ou *Gata4*, sendo a proteína *Mef2* um regulador tanto de *Hand2* (PHAN et al., 2005) quanto de genes relacionados a manipulação de cálcio, importantes para o metabolismo de cardiomiócitos (BLAESER et al., 2000). *Hand1* e *Hand2* estão associadas a formação das cavidades do coração (revisado por GEORGE; FIRULLI, 2019) e a proliferação de cardiomiócitos imaturos, tendo diferentes padrões de

expressão de acordo com o subtipo celular (atrial e ventricular) (OKUBO et al., 2021).

A atuação coordenada dessa rede de regulação transcricional resulta na geração de cardiomiócitos fetais capazes de expressar os marcadores moleculares Nkx2-5 e troponina. Os tipos de troponina mais expressos nessas células imaturas são a isoforma fetal da troponina T cardíaca (TNNT2), bem como a isoforma de contração lenta de troponina I (TNNI). No período pós-natal, os cardiomiócitos sofrem maturação, expressando outras isoformas de troponina T cardíaca (TNNT3) e de troponina I (TNNI3) (AMES et al., 2013).

## 2.2 ATUAÇÃO DE RNAs NÃO CODIFICANTES E PROTEÍNAS LIGANTES DE RNA NA CARDIOMIOGÊNESE

Os RNAs não codificantes (ncRNAs) são RNAs que, embora não traduzidos, desempenham papéis regulatórios na diferenciação e homeostase celular. Considerados meros erros de transcrição no início do século XXI, os ncRNAs agora são objeto de um campo emergente de estudo que busca explorar a vasta extensão do genoma eucariótico que não codifica para proteínas. A grande diversidade de tamanho, estrutura e função celular dos ncRNAs dificulta sua classificação sistemática, sendo os micro RNAs e os RNAs longos não-codificantes (conceituados por DHAMIJA; MENON, 2018) os mais estudados no contexto do desenvolvimento embrionário de mamíferos.

Como revisado por Iwakawa e Tomari (2022), microRNAs ou miRNAs são RNAs pequenos (<200 pb) envolvidos na regulação negativa de mRNAs ou lncRNAs, integrando o complexo riboproteico RISC (*RNA-induced Silencing Complex*). Essas moléculas geralmente se anelam a porção 3'-UTR de seus RNAs-alvo, inibindo sua tradução ao interferir com o complexo eIF4F (*Eukaryotic Initiation Factor 4F*) e ao promover sua degradação. O processo de degradação mediada por miRNAs se inicia com o recrutamento das depoliadenilases PAN2/3 (Poly(A)-nuclease 2/3) e do complexo CCR4-NOT (*Carbon Catabolite Repression—Negative*

on *TATA-less*; O'BRIEN et al., 2018), que depoliadenila o transcrito-alvo e recruta proteínas decapeadoras. Por fim, o transcrito decapeado e depoliadenilado sofre hidrólise via nuclease XRN1 (5'-3' *Exoribonuclease 1*). Em animais, o anelamento miRNA-mRNA não é exato para a maioria dos RNAs-alvo, exceto nos nucleotídeos nas posições 2 a 8 do miRNA, a partir de sua extremidade 5', que formam a chamada região *seed*. MiRNAs com a mesma sequência *seed* são agrupados em uma família, pressupondo executar funções semelhantes no contexto celular (GARATE et al., 2018).

A relevância da atuação de microRNAs durante a cardiomiogênese foi elucidada por diversos estudos de biologia molecular. Um panorama geral do papel de microRNAs na regulação da diferenciação cardíaca pode ser vislumbrado nos estudos de Garate et al. (2018), que avaliaram o transcriptoma de células em 3 estágios (pluripotentes, mesodermais e cardiomiócitos), indicando que as famílias de miRNAs mais expressas em cardiomiócitos são miR-21, miR-10 (miR-10a, miR-10b, miR-99a, miR-99b, miR-100, miR-125a e miR-125b), miR-143, miR-30 (miR-30a, miR-30b, miR-30c e miR-30d), miR-26 (miR-26a e miR-26b) e miR-1 (miR-1 e miR-206). Por análises de ontologia gênica, os autores concluíram que essas famílias parecem atuar inibindo outras rotas de diferenciação cujo ponto de partida é a mesoderme, como células do sistema muscular estriado esquelético, o sistema urogenital e o esquelético. Um exemplo de interação desses miRNAs na regulação da cardiomiogênese é a expressão coordenada de miR-1/miR-133, que atuam de maneira antagônica na diferenciação de células de mesoderme em cardiomiócitos. A expressão de miR-1 favorece a diferenciação, enquanto a de miR-133 a inibe (WOJCIECHOWSKA et al., 2017).

A maioria do transcriptoma não-codificante humano é composto por RNAs longos não codificantes (lncRNAs), transcritos com mais de 200 pb que, de maneira semelhante a mRNAs, são poliadenilados e sintetizados pela RNA polimerase II (AMARAL et al., 2011). Os lncRNAs podem atuar como moléculas reguladoras de miRNAs, sequestrando esses pequenos nucleotídeos e inibindo seu efeito em mRNAs alvo. Como exemplo, Kallen et al. (2013) reportaram que o lncRNA H19

interage com miRNAs da família let-7, impedindo sua interação com um mRNA-alvo conhecido, o *Igf2* (*Insulin Like Growth Factor 2*). De maneira inversa, a expressão de alguns lncRNAs tem efeito positivo na função de determinados miRNAs, caso de outro mecanismo de ação do lncRNA *H19*, descrito por Liu et al. (2016). Em experimentos de atenuação (*knockdown*) e superexpressão, os autores provaram que o lncRNA *H19* serve como precursor de um miRNA envolvido na inibição da hipertrofia do coração, o miR-675.

Os RNAs longos também podem servir como reguladores de mRNAs, interferindo no seu *splicing* (GONZALEZ et al., 2015), tempo de meia-vida (LEE et al., 2016 e WU et al., 2019) localização celular (CAI et al., 2018) e tradução (BOOY et al., 2021). Um mecanismo possível de controle de mRNAs por lncRNAs envolve a formação de complexos ribonucleoproteicos (RNPs) entre determinado lncRNA e proteínas ligantes de RNA (RBPs).

As RBPs estão envolvidas em processos de síntese, armazenamento, tradução e estabilidade de transcritos (BALCERAK et al., 2019). Essas proteínas possuem motivos pequenos, contendo menos de 100 resíduos de aminoácidos, que são capazes de interagir com RNAs (CORLEY et al., 2020). Na maioria dos casos, as interações RNA-RBP dependem da sinergia de múltiplos sítios de interação proteicos (CLÉRY; ALLAIN, 2018), bem como da mudança da estrutura terciária do RNA em questão para conformações favoráveis (como exemplificado por FLORES; ATAIDE, 2018).

Um estudo de proteoma conduzido por Liao et al. (2016), que focou em RBPs expressas por cardiomiócitos murinos, identificou 1148 proteínas, sendo 391 consideradas RBPs comuns e 393 únicas para cardiomiócitos. Algumas RBPs reveladas possuem outras funções além da ligação a RNA, como enzimas envolvidas no ciclo do ácido tricarboxílico, proteínas mitocondriais e como a ATPase *Serca2*, responsável pelo influxo de cálcio no retículo endoplasmático de células musculares estriadas. Embora os domínios de ligação a RNA tenham sido identificados nas RBPs, o estudo não preservou dados relativos aos RNAs que interagem com essas regiões. Um exemplo que ilustra bem um mecanismo de ação

entre RBPs e RNAs em cardiomiócitos é relatado por Bertero et al. (2019), que estudaram as mudanças conformacionais da cromatina durante a cardiomiogênese de células de camundongo, culminando na proposição de um mecanismo molecular relacionando o pré-mRNA de uma proteína importante para a homeostase do coração, a titina (TNN), e um fator de *splicing* denominado RBM20 (*RNA-Binding Motif 20*). O pré-mRNA de TNN aparentemente atua como arcabouço para a interação de RBM20 com trechos do genoma contendo genes importantes para a contratilidade dos cardiomiócitos, como o CACNA1C (*Calcium Voltage-Gated Channel Subunit Alpha1 C*) e o CAMK2D (*Calcium/calmodulin-dependent protein kinase type II subunit delta*), recrutando cópias dessa RBP pelos sítios de ligação presentes em sua sequência. Outra RBP relevante ao desenvolvimento de cardiomiócitos é a RBM24, considerada um fator de *splicing* mestre para a diferenciação cardíaca. Estima-se que a RBM24 atua sobre 68 eventos de *splicing*, e seu nocaute em camundongos gerou más-formações cardíacas e tecido muscular sem sarcômeros aparentes (YANG et al., 2014).

A influência dos lncRNAs sobre mRNAs pode ocorrer via interações independentes de RBPs, ao interagir com outras proteínas reguladoras como fatores de transcrição e complexos de modificação epigenética de DNA. Em células de camundongo, o *Linc1405*, ou *Meteor*, se associa ao fator de transcrição Eomes para promover modificações epigenéticas no potenciador de um dos genes-chave para a diferenciação celular do estágio de mesoderme inicial para mesoderme cardíaca, o *Mesp1* (GUO et al., 2018). De maneira semelhante, o lncRNA humano *CARMEN* parece ser responsável por ativar a via de regulação transcricional mediada por MESP1, interagindo com um componente ainda desconhecido (OUNZAIN et al., 2015). Outro lncRNA atuante nesse intervalo da cardiomiogênese é o *Braveheart* murino, primeiramente descrito por Klattenhoff et al. (2013) como regulador de genes-alvo de *Mesp1*. Xue et al. (2016) demonstraram que esse transcrito também interage com uma proteína chaperona de nucleotídeos chamada Cnbp (*Cellular Nucleic Acid Binding Protein*) via uma sequência rica em guanossina (*G-rich RNA motif*, *AGIL*), dificultando sua ligação a outras estruturas *AGIL* presentes no genoma.

O encerramento do estágio de mesoderme cardíaca/lateral em camundongos é promovido pelo lncRNA *Fendrr*, que serve como um suporte para a ação do complexo de repressão epigenética Polycomb 2 (Prc2) em genes-chave da especificação em mesoderme lateral, como o *PITX2*, o *FOXF1* e *IRX3* (GROTE et al., 2013).

Em humanos, a expressão de um dos transcritos antisense do gene *GATA6*, denominado *MOSHE*, parece ativar o gene efetor *NKX 2-5* ao interagir com um potenciador localizado 3 kb à jusante da CDS. Simultaneamente, *MOSHE* aparenta inibir genes-alvo do fator de transcrição ISL-1, que estão relacionados à diferenciação de células de mesoderme cardíaca a progenitores do segundo campo cardíaco (KIM et al., 2021).

Apesar das funções descritas anteriormente para os lncRNAs associados à cardiomiogênese indicarem uma localização estritamente nuclear, há indícios que sua distribuição é dinâmica e regulada por fatores espaciais. Como exemplo, em experimentos *in vivo* que acompanharam o desenvolvimento embrionário do coração de camundongos, García-Padilla et al. (2019) perceberam que os lncRNAs *Fendrr* e *Carmen* localizam-se mais no núcleo de células do átrio esquerdo e dos ventrículos, enquanto encontram-se majoritariamente no citoplasma de células do átrio direito.

Em células humanas, estudos de Lyu et al. (2022) indicaram que o lncRNA *CPMER* (*Cytoplasmic Mesoderm Regulator*) interage diretamente com o transcrito do fator de transcrição *EOMES*, aumentando sua afinidade com uma proteína reguladora de tradução específica do coração, a *Eukaryotic Elongation Factor 1A2* (eEF1A2). Como consequência, há expressão da sequência controlada por *EOMES/Eomes*, *MESP1*, e comprometimento das células a progenitores cardíacos.

Adicionalmente, alguns lncRNAs citoplasmáticos podem conter pequenos quadros abertos de leitura (sORFs), codificantes para os denominados micropeptídeos. Os micropeptídeos possuem menos de 100 aminoácidos, não são processados a partir de polipeptídeos e não possuem sequência de endereçamento. Em decorrência disso, são majoritariamente citoplasmáticos (revisado por SOUSA;

FARKAS, 2018), embora já existam relatos de micropeptídeos localizados na membrana celular (ROSSI et al., 2019) e no núcleo (ZHANG et al., 2018).

Já há relatos de micropeptídeos codificados a partir de lncRNAs que atuam em processos envolvendo células musculares. Wang et al. (2020), por exemplo, estabeleceram uma correlação entre a expressão do micropeptídeo Lemp e a eficiência da diferenciação cardíaca de mioblastos murinos, bem como no alinhamento correto de fibras musculares esqueléticas em peixes-zebra. Especula-se que esse efeito na diferenciação de células musculares está relacionado ao metabolismo energético, uma vez que esse peptídeo se encontra associado a proteínas mitocondriais. Makarewich; Baskin et al. (2018) também estudaram um micropeptídeo mitocondrial, associado a membrana celular interna da organela. De acordo com o estudo, o chamado *Micropeptide regulator of  $\beta$ -oxidation* (Moxi) interage com a proteína mitocondrial trifuncional (Tfp) de maneira a viabilizar a  $\beta$ -oxidação de ácidos graxos em mitocôndrias de cardiomiócitos murinos, sendo a expressão de Moxi correlacionada a tolerância ao exercício em camundongos. Esses resultados foram complementados por Chugunova et al. (2019), que correlacionaram a expressão reduzida desse peptídeo com a redução da atividade do complexo I da cadeia respiratória em mitocôndrias, com o qual a proteína ligante de Moxi, a Tfp, interage através de um sítio de ligação a NADH (WANG et al., 2019).

Outro micropeptídeo digno de nota é a Miorregulina, conservada em camundongos e humanos e envolvida na inibição da bomba de cálcio Serca2 e, por consequência, no cessar da recaptação desse íon para o retículo endoplasmático especializado de células musculares esqueléticas (ANDERSON et al., 2015). De forma antagônica, o micropeptídeo Dwarf (*Dwarf Open Reading Frame*) estimula a captação de  $\text{Ca}^{2+}$  no retículo ao agir sobre Serca2, competindo com o peptídeo inibidor Fosfolambano (PLN) e, dessa maneira, exercendo influência positiva sobre a contratilidade de cardiomiócitos murinos primários (MAKAREWICH e MUNIR et al., 2018).

## 2.3 LncRNA CARDEL NA DIFERENCIAÇÃO CARDÍACA

A fim de elucidar a regulação pós-transcricional envolvida na cardiomiogênese, Pereira (2019) utilizou protocolos de diferenciação cardíaca (KATTMAN et al., 2011; LIAN et al., 2013 e WANG et al., 2011) para investigar padrões de expressão gênica em células originalmente pluripotentes (como a hESCs da linhagem H1) em diferentes estágios da cardiomiogênese, empregando a técnica de sequenciamento de RNA (RNA-seq). Buscou-se também determinar se os RNAs expressos estavam associados a ribossomos, através da técnica de perfil polissomal (*polysome profiling*) (MAŠEK et al., 2011 e SPANGENBERG et al., 2013).

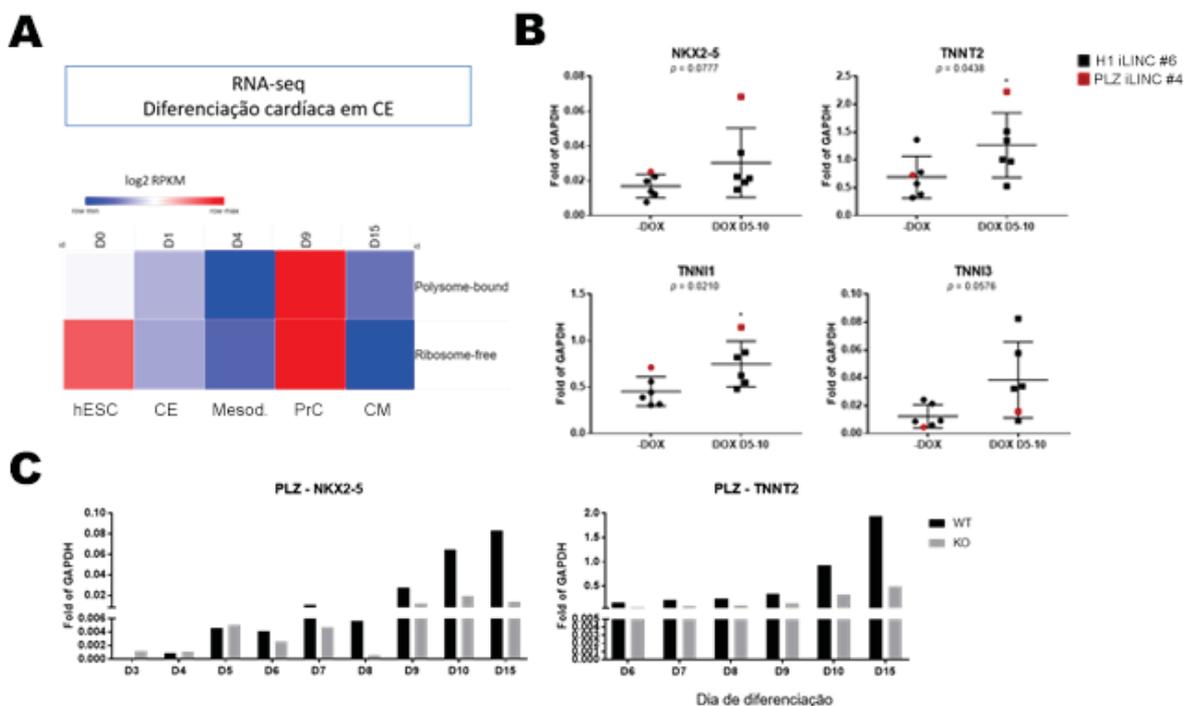
Pereira (2019) observou dados de RNA-seq do dia 0 (D0, células pluripotentes), dia 1 (D1, células de corpo embriode), dia 4 (D4, células de mesoderme), dia 9 (D9, progenitores cardíacos) e dia 15 (D15, cardiomiócitos imaturos) da diferenciação cardíaca por corpo embriode (KATTMAN et al., 2011), em frações livres de ribossomos e frações associadas a poliribossomos. Dentre os diversos transcritos diferencialmente expressos, um RNA longo não codificante intergênico (lincRNA), denominado *LINC00890*, se destacou por apresentar níveis de expressão variáveis durante a diferenciação cardíaca, como visível pelo mapa de calor da Figura 2A. Em ambas frações analisadas, o *LINC00890* apresentou um pico de expressão no dia correspondente ao estágio de progenitor cardíaco (D9) em comparação a expressão de células indiferenciadas (D0, fração livre de ribossomos), retornando a níveis semelhantes aos estágios anteriores (D1 e D4) na fase de cardiomiócitos imaturos (D15).

Em estudos funcionais posteriores, a superexpressão induzida do *LINC00890* nos estágios finais de um protocolo diferenciação cardíaca em monocamada (LIAN et al., 2013 e WANG et al., 2011; indução nos dias 5 e 10) promoveu um aumento significativo na expressão transcricional das isoformas fetais de troponina (*TNNT2* e *TNNT1*), além de uma tendência de aumento para o fator de transcrição *NKX2-5* e a isoforma adulta de troponina I (*TNNT3*), como demonstrado na Figura 2B. Além disso, o nocaute da expressão do *LINC00890* provocou uma

drástica diminuição de transcritos de genes marcadores de comprometimento cardíaco (*NKX2-5* e *TNNT2*) em comparação com a condição controle (Figura 2C).

De maneira geral, os dados levantados sugeriram uma relação entre a expressão do *LINC00890* e o comprometimento celular a linhagem cardíaca, estimulando a renomeação do transcrito para *CARDiac DEvelopment LncRNA (CARDEL)*. Essas observações instigaram a investigação do efeito de *CARDEL* no processo de cardiomiogênese em nível pós-transcricional, por via direta como transcrito regulador ou como parte de um complexo ribonucleoproteico (RNP).

FIGURA 2 - EXPRESSÃO E EFEITO DE CARDEL EM MARCADORES CARDÍACOS



NOTA: (A) Expressão do lncRNA *LINC00890* durante a diferenciação cardiomiogênica em corpo embriãoide (CE). Os valores de expressão baseados no RNA-seq estão expressos em  $\log_2$  de RPKM (*Reads per kilobase of transcript per Million reads mapped*) e representados no mapa de calor. Os dias de diferenciação estão indicados na figura, assim como as frações de RNA livre (*Ribosome-free*) ou associado a poliribossomos (*Polysome-bound*); (B) Expressão gênica de marcadores pan-cardíacos em 2 linhagens de superexpressão de *CARDEL* (iLINC) induzidas por doxiciclina (DOX), a H1 iLINC (pontos pretos) e a PLZ iLINC (pontos vermelhos); a superexpressão ocorreu nos dias 5 e 10 da diferenciação (DOX D5-10); (C) Expressão gênica de marcadores cardíacos durante a diferenciação cardíaca da linhagem H1 LINC KO, usando células parentais H1 WT como controle. A expressão foi quantificada por qPCR e os valores estão expressos relativos ao gene constitutivo GAPDH. WT, células selvagens parentais. KO, células nocauteadas.

Fonte: Pereira (2019).

Segundo o banco de dados Ensembl (CUNNINGHAM et al., 2022), sequência codificante para *CARDEL* está localizada no braço longo do cromossomo X, na banda citogenética 23. A contagem do número de leituras da região do *CARDEL* nos dados de RNA-seq por Pereira et al. (2019) indicou 2 variantes de *splicing* predominantes, ENST00000563467 (*SERTM2-201*), que compreende o éxon 2 e parte dos éxons 1 e 3 (Fig. 3, parte superior) e ENST00000569275 (*SERTM2-202*), que inclui os éxons 1, 2 e 3. A variante mais longa do *CARDEL* (*SERTM2-202*) parece ser a mais representativa e atual foco de estudo.

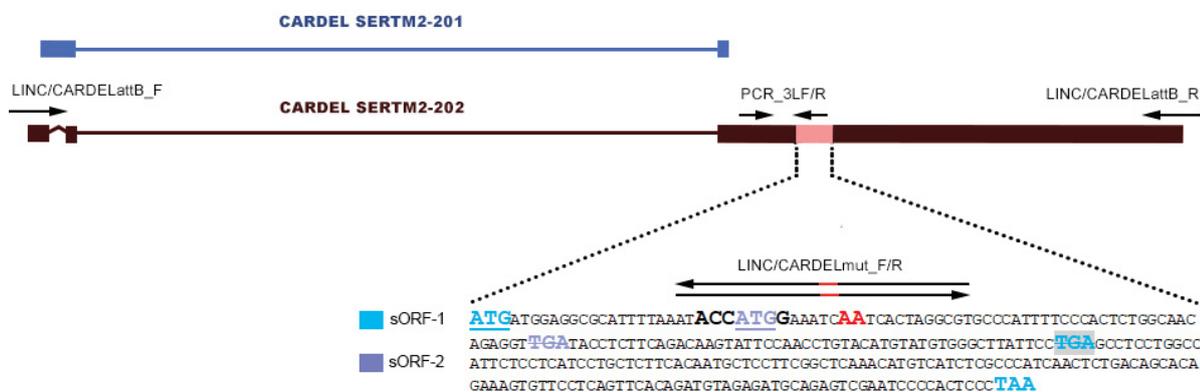
Apesar de primeiramente anotado como um lncRNA, uma sequência codificante (CDS) putativa foi anotada no éxon 3 da variante *CARDEL SERTM2-202* pelos sistemas *PAINT* (GAUDET et al., 2011) e *PhyloCSF* (MUDGE et al., 2019). O sistema *PAINT* utiliza dados de Ontologia Gênica e classificações filogenéticas para anotar sequências codificantes para proteína no genoma, bem como para estimar funções biológicas e localização celular das proteínas putativas. Como é um sistema de curadoria dependente de *inputs* do usuário, o *PAINT* tende a ser mais preciso do que sistemas automatizados, segundo testes realizados pelos próprios desenvolvedores. O *PhyloCSF*, por outro lado, é um sistema automatizado de anotação baseado em aprendizagem de máquina. Devido à independência dos algoritmos para a anotação, os autores comentam que algumas sequências não puderam ser bem discriminadas entre pseudogene e sequência codificante para proteína, sendo o *CARDEL SERTM2-202* parte integrante desse grupo (MUDGE et al., 2019, arquivo suplementar 2)

Análises no programa *ORFfinder* do *National Center for Biotechnology Information* (WHEELER et al., 2008) corroboraram o potencial codificante de *CARDEL*, indicando a presença de 2 sORFs principais. O *CARDEL* sORF-1 (Fig. 3, em *SERTM2-202*) possui potencial para gerar um micropeptídeo de 90 aminoácidos denominado SERTM2, já o sORF-2 (Fig. 3, em *SERTM2-202*) engloba parte da sequência da sORF-1 em um quadro de leitura distinto, e tem potencial para codificar um micropetídeo de 34 aminoácidos. Como indicativo adicional de seu potencial codificante, o sORF-2 apresenta ainda a mais conhecida sequência

potenciadora de tradução em vertebrados, o consenso de Kozak **ACCATGG**, sendo ATG o códon de início de tradução (KOZAK, 1989; NAKAGAWA et al., 2008).

A fim de gerar um modelo de superexpressão de *CARDEL SERTM2-202* sem potencial codificante, Pereira (2019) adicionou 2 adenosinas na CDS predita do DNA complementar (cDNA) de *CARDEL*, empregando PCRs (Reações em Cadeia de Polimerase) sequenciais e coordenadas para obtenção de amplicons de *CARDEL* mutado. A mutação causa uma mudança no quadro de leitura para sORF-1/2 e gera a terminação precoce das 2 sequências. Esses sORFs mutados, se traduzidos, dariam origem a peptídeos mais curtos – de 43 e 17 aminoácidos, respectivamente – de diferentes composições aminoácidas e, portanto, sem qualquer função biológica preservada. A sequência *CARDEL mut* foi inserida no vetor integrativo ao genoma humano AAVS1-GW-TRE-rtTA, dando origem à construção AAVS1-CARDEL mut. Detalhes da obtenção desse plasmídeo estão descritos na seção 4.3.

FIGURA 3 – VARIANTES DE *SPLICING* DE *CARDEL* E ESTRATÉGIA PARA MUTAÇÃO DA CDS PREDITA



NOTA: Os éxons de cada transcrito estão representados por caixas, enquanto os íntrons pelas linhas; a extensão de cada porção está aproximadamente proporcional a extensão das sequências. A variante de *splicing* *CARDEL SERTM2-201* compreende o éxon 2 e parte dos 1 e 3, não possuindo sequências codificantes anotadas, enquanto a variante *CARDEL SERTM2-202* compreende os éxons 1, 2 e 3, possuindo 2 pequenos ORFs em uma porção do éxon 3, o sORF-1 e o sORF-2, compondo uma CDS predita. A mutação (em vermelho) na CDS predita para *SERTM2-202* foi realizada por PCRs, utilizando de maneira cruzada os *primers* LINC/CARDELaatB\_F e LINC/CARDELmut\_R e vice-versa; os *primers* LINC/CARDELmut serviram para amplificar a CDS com 2 mutações do tipo inserção; e LINC/CARDELaatB, para incluir as sequências aatB necessárias à clonagem via sistema Gateway; a mutação causa uma mudança no quadro de leitura para sORF-1/2 e gera a terminação precoce das 2 sequências; os códons de início estão sublinhados e os códons de terminação precoce estão tachados; destacado em cinza está o códon de término original do sORF-2, que se sobrepõe ao códon de término precoce de sORF-1; a sequência de Kozak está destacada em negro.

FONTE: Adaptado de Pereira (2019).

## 2.4 CÉLULAS-TRONCO PLURIPOTENTES

O conhecimento profundo sobre as bases genéticas e moleculares da diferenciação cardíaca, demonstrado anteriormente, foi adquirido graças ao domínio e popularização da cultura de células-tronco *in vitro*. Células-tronco pluripotentes são unidades biológicas capazes de se diferenciar em qualquer célula somática, ou seja, qualquer célula do organismo com exceção aos anexos embrionários, e sofrer autorrenovação, que pode ser definida como a capacidade de proliferar sem sofrer alterações fenotípicas (YOUNG, 2011). As células-tronco pluripotentes podem ser classificadas ainda em células-tronco embrionárias (ESCs), oriundas da massa celular interior de um blastocisto (THOMSON, 1998), e células-tronco pluripotentes induzidas (iPSCs), derivadas de células somáticas submetidas a reprogramação *in vitro* para se tornar pluripotentes. Um grupo pioneiro na reprogramação celular foi Takahashi et al. 2007, que gerou iPSCs a partir de fibroblastos adultos pela superexpressão dos fatores de transcrição Oct3/4 (*Octamer-binding transcription factor 3/4*), Sox2 (*SRY-Box Transcription Factor 2*), Klf4 (*Krueppel-like factor 4*), e c-Myc (*Cellular myelocytomatosis oncogene*).

Essas células atualmente são modelo para estudo da cardiomiogênese e para a produção de cardiomiócitos, tipo celular de obtenção antes difícil e não adequado para múltiplas passagens em cultura primária (revisado por BEQQALI et al., 2009; HOES et al., 2019). Em laboratório, as células-tronco pluripotentes podem ser caracterizadas via diferenciação *in vitro* nos três folhetos embrionários (endoderme, mesoderme e ectoderme), pela quantificação de alguns transcritos necessários à manutenção da pluripotência, como o *OCT4A*, o *NANOG* e o *SOX2* ou pela detecção de alguns marcadores de superfície, como o *SSEA-3*, o *SSEA-4*, o *TRA-1-60* e o *TRA-1-81* (revisado por SMITH et al., 2009).

### 3 OBJETIVOS

#### 3.1 Objetivo geral

Investigar o potencial codificante do RNA *CARDEL* e possíveis interações desse transcrito com proteínas ligantes de RNA.

#### 3.2 Objetivos específicos

1. Acessar o potencial de tradução do transcrito de *CARDEL* por expressão transiente do sORF-1 de *CARDEL* fusionado a CDS de eGFP;
2. Investigar o efeito da superexpressão de *CARDEL* mutado, sem potencial codificante, nos estágios finais da diferenciação cardiomiogênica *in vitro*;
3. Avaliar potencial de interação de *CARDEL* com proteínas ligantes de RNA.

### 4 MATERIAL E MÉTODOS

Os meios de cultivo e kits utilizados para a realização desse trabalho encontram-se listados no Apêndice 1, Tabela 5.

#### 4.1 CONSTRUÇÃO DE PLASMÍDEO pcDNA 3.1-His B/myc-CARDEL1A.eGFP

A fim de construir plasmídeos pcDNA 3.1/myc His B (ThermoFisher) contendo fragmentos de *CARDEL* com potencial de tradução fusionados a CDS para a proteína fluorescente verde aperfeiçoada (eGFP, de *enhanced GFP*), foram utilizadas ferramentas de clonagem clássica via endonucleases de restrição.

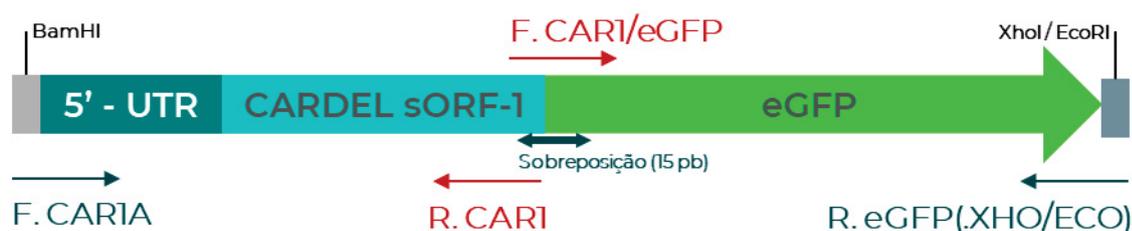
##### 4.1.1 Obtenção de amplicons CARDEL1A, sCARDEL1.eGFP e CARDEL1A.eGFP

A sequência fusionada *CARDEL1A.eGFP* a ser inserida no vetor de expressão pcDNA 3.1 consiste: no sORF-1 de *CARDEL* (*CARDEL1*), em uma

sequência de 100 pb sem potencial codificante localizada a montante do sORF-1 (A), denominada genericamente 5'-UTR; na CDS para eGFP; e nos sítios de restrição para as endonucleases BamHI e XhoI/EcoRI.

Essa sequência *CARDEL1A.eGFP* foi obtida pela amplificação por PCR de *CARDEL1A* e de *eGFP* contendo uma sequência de sobreposição à extremidade 3' do sORF-1 de *CARDEL* (*sCARDEL1-eGFP*). Como ilustrado na Figura 4, os *primers* F.CAR1A - R.CAR1 e F.CAR1/eGFP - R.eGFP foram desenhados para gerar os amplicons *CARDEL1A* e *sCARDEL1.eGFP*, respectivamente. Para otimizar os recursos disponíveis e permitir a amplificação de outras sequências do *CARDEL* em trabalhos futuros, o *primer* interno R.CAR1 foi desenhado de forma a não incluir sequência de sobreposição *eGFP* no amplicon *CARDEL1A*.

FIGURA 4– ESTRATÉGIA DE OBTENÇÃO DE *CARDEL1A.eGFP*



NOTA: Cada par de *primer* interno (em vermelho) e externo (em azul) permite a amplificação das sequências do sORF-1 de *CARDEL* (*CARDEL1A*) fusionado a uma porção 5'-UTR (*CARDEL1A*) e da etiqueta *eGFP* com sequência de sobreposição (*sCARDEL1.eGFP*) em separado, bem como da inclusão dos sítios de restrição BamHI e XhoI/EcoRI. Os *primers* externos são utilizados para a PCR de extensão por sobreposição (*Extension Overlap PCR*), que visa fusionar os fragmentos obtidos;

Os fragmentos gerados serviram então como molde para a chamada PCR de extensão por sobreposição, que se baseou no anelamento dos fragmentos individuais a partir da sequência de sobreposição de 15 pb presente no amplicon *sCARDEL1.eGFP*. Para isso, os *primers* externos a *CARDEL1A* (F.CAR1A) e *sCARDEL1.eGFP* (R.eGFP) foram empregados para a amplificação.

Todos os amplicons foram gerados com uma DNA polimerase de alta processividade e fidelidade, parte do mix *Phusion High-Fidelity PCR Master with HF Buffer* (ThermoFisher), com *primers* em concentração de 250 nM cada.

Para a amplificação de *CARDEL1A* e *sCARDEL1.eGFP*, 100 ng do vetor contendo cDNA de *CARDEL* (clone DKFZp686D0853) e 100 ng do vetor AAVS1-

caggs-eGFP foram utilizados respectivamente, em reações separadas. Para a PCR de extensão por sobreposição, os moldes foram *CARDEL1A* e *sCARDEL1.eGFP* na proporção de 100 ng/kb, como sugerido por Hilgarth; Lanigan, 2020.

Tanto para as PCR dos fragmentos quanto para a PCR de extensão por sobreposição, o ciclo de temperaturas foi: uma etapa de 98 °C por 30 seg., seguida por 20 ciclos de amplificação em 3 etapas (98 °C por 10 seg., 55-60 °C por 30 seg. e 72 °C por 30s/kb) e uma etapa de extensão final de 72°C por 5 min. A temperatura de anelamento foi estudada como ponto de otimização. As sequências dos *primers* podem ser visualizadas na Tabela 1.

TABELA 1 - PRIMERS PARA AMPLIFICAÇÃO DE SEQUÊNCIAS DE sORF *CARDEL*-eGFP

	PRIMER (5' > 3')
Amplificação do sORF-1 de <i>CARDEL</i>	
F.CAR1A*	AACGGATCCGAACACGTCCAATAGATTTAGA
R.CAR1	GGGAGTGGGGATTCTGACTCT
Amplificação de <i>eGFP</i> sequência de sobreposição a sORF (sORF-1.eGFP)	
F.CAR1.eGFP	CGAATCCCCACTCCCGTGAGCAAGGGCGAG
R.eGFP.XHO*	AACTCGAGTTACTTGTACAGCTCGTCCAT
R.eGFP.ECO*	GACATCGAATTCTTACTTGTACAGCTCGTCCAT

NOTA: Os *primers* indicados por um \* também foram utilizados para a PCR de extensão por sobreposição, sendo denominados *primers* externos (Figura 4, setas azuis).

#### 4.1.2 Clonagem de construção em *E. coli* DH5 $\alpha$ , seleção e *screening* de clones

Após a obtenção da sequência *CARDEL1A.eGFP*, cada reação de digestão utilizou: 10 U de endonuclease; 1x NEBuffer™ 3.1 (para digestões com BamHI e XhoI, BamHI e EcoRI ou digestões simples com essas enzimas) ou 1x NEBuffer™ 1.1 (para AgeI); 1  $\mu$ g de pcDNA 3.1 ou 500 ng de *CARDEL1A.eGFP*. As reações foram incubadas a 37°C por 1h. A amplificação das sequências fusionadas e as reações de digestão foram verificadas por eletroforese em gel de agarose 1%, utilizando como marcador o TrackIt 1 kb Plus (Invitrogen) e o SYBR™ Safe DNA Gel Stain (ThermoFisher) como corante intercalante. A revelação do gel foi realizada em

um transiluminador L-PIX EX (Loccus Ltda.). Os produtos de tamanho adequado foram purificados pelo protocolo de gel extração do QiAquick Gel Extraction Kit (Qiagen). Por fim, o pcDNA 3.1 myc/His B e o *CARDEL1A.eGFP* foram submetidos a reação de ligação com 1U da enzima T4 DNA ligase, 1x T4 DNA ligase *buffer*, 50 ng (14,72 fmol) de plasmídeo digerido, ATP em concentração final de 1 mM e quantidades variáveis de inserto (3/5/10:1 em proporção molar de inserto:vetor).

Para a transformação, 50 µL de bactérias *Escherichia coli* da cepa DH5α, quimiocompetentes, foram misturadas com a reação de ligação e incubadas em gelo por 10 min. Então, foi realizado o choque térmico a 42 °C por 30 seg., seguido de nova incubação no gelo por 2 min. As bactérias foram recuperadas em meio LB a 37 °C por 1 h, sob agitação a 200 rpm. A cultura foi plaqueada em meio LB sólido com ampicilina a 50 µg/mL e incubada a 37 °C por 16 h.

A fim de realizar uma varredura dos clones positivos para a construção pcDNA-CARDEL1A.eGFP, foram realizadas PCR exploratórias (como descrito na seção 4.1.1) com o propósito de amplificar a sequência CARDEL1A.eGFP, utilizando o esfregaço de cada colônia de *E. coli* DH5α como molde de DNA. Mini-preparações plasmidiais de clones positivos para a PCR de colônia foram obtidas seguindo as instruções do kit QiaPrep Spin MiniPrep (Qiagen). O mapa do vetor gerado encontra-se no Apêndice 2, Figura 18

#### 4.2 AVALIAÇÃO DO POTENCIAL CODIFICANTE DO sORF-1 DE CARDEL EM CÉLULAS HEK293

O potencial codificante da construção pcDNA 3.1 His B/myc – CARDEL1A-eGFP foi acessado por meio da transfecção de células HEK293, empregando os lipídeos catiônicos Lipofectamine 2000 e Lipofectamine 3000 (Invitrogen).

Células da linhagem HEK293 foram cultivadas a 37°C, 5% CO<sub>2</sub>, em meio DMEM suplementado com 10% de soro fetal bovino (SFB) e 1% de L-Glutamina (L-Gln), trocado a cada 2 dias até o fim do experimento ou até atingir confluência de 70-90%. Nas passagens, o meio de cultivo foi retirado e reservado, as células foram lavadas com tampão fosfato-salino (PBS), incubadas com tripsina-EDTA 0,05%

(Gibco) por 3 min. a 37 °C, transferidas para o meio reservado (1 vol. de DMEM para 1 vol. de tripsina) e centrifugadas por 3 minutos a 600 x g. As células foram ressuspensas em DMEM 10% SFB, 1% L-Gln e transferidas para garrafa de 75 cm<sup>2</sup> ou placas de 6 ou 12 poços.

Dois dias antes da transfecção, células foram transferidas para placas de 6 poços (70.000 células/poço). Para cada poço, 1 µg de vetor e 12 µL de Lipofectamine 2000 Reagent (Invitrogen) foram diluídos em 350 µL e 150 µL de Opti-MEM™ I Reduced Serum Medium (Gibco), respectivamente. As alíquotas de vetor e de lipofectamina foram misturadas e incubadas em temperatura ambiente por 15 min, a fim de formar complexos DNA-lipídio. O meio de cultura foi trocado e a mistura DNA-lipofectamina foi gotejada nos poços correspondentes.

Para as transfecções com Lipofectamine 3000, o mesmo protocolo de transfecção para Lipofectamine 2000 foi aplicado, com as seguintes alterações: foram utilizados 7,5 µL de Lipofectamine 3000 por poço e adicionados 10 µL do reagente Lipofectamine 3000 Enhancer para cada diluição de plasmídeo (1-3 µg/poço) em Opti-MEM. Nesses experimentos, foram testados os tempos de transfecção de 6h, 16h e 48h, e o controle do processo consistiu na transfecção com o vetor pX461-GFP (addgene code #48140), um vetor de superexpressão transiente não induzida da proteína Cas9 fusionada a GFP.

A fluorescência da cultura foi avaliada 24 e 48 h pós-transfecção, em microscópio óptico Leica DMIL LED, com cubo de fluorescência mantido no canal C (Ex. 482/35 nm), e imagens foram adquiridas pelo *software* LAS X. A presença do peptídeo fusionado a eGFP (SERTM2-eGFP), por outro lado, foi investigada por *Western Blot*.

#### 4.2.1 Eletroforese em gel de poliacrilamida-SDS (SDS-PAGE) e *Western Blot*

Para a obtenção dos extratos proteicos, as células em cultura foram lavadas com PBS gelado e desaderidas com espalhador de células (*scraper*) em 800 µL de tampão de lise [Tris-HCl 50 mM, SDS 4%, Triton X-100 1%, e o coquetel de

inibidores de protease cOmplete™]. As amostras foram mantidas em agitador com rotação de 360°, a 8-16°C por 30 min e depois homogeneizadas em seringa equipada com agulha de 3/4 pol., a 4°C. Os extratos foram centrifugados por 30 min a 16.000 x g e 4°C, e o sobrenadante coletado. A concentração de proteína foi dosada pelo fluorômetro Qubit 4 (ThermoFisher), após preparação de alíquotas de acordo com as instruções do kit Qubit Protein Assay (ThermoFisher).

Alíquotas correspondentes a 20 µg de proteína foram diluídas em tampão de amostra [trisaminometano 2,4 µM, SDS 0,8 g/L, glicerol 55 mM e β-mercaptoetanol 5% (V/V), azul de bromofenol 0,05% (m/V)], aquecidas a 98°C por 15 min e submetidas a SDS-PAGE, sob corrente constante de 25 mA por 2h30, utilizando o *Kaleidoscope Precision Plus* (Bio-Rad) como marcador de massa molecular.

Posteriormente, as proteínas do gel foram transferidas para membrana de nitrocelulose Amersham Protran® (Sigma-Aldrich) em sistema *Trans-Blot® SD Semi-Dry Transfer Cell* (Bio-rad), em diferença de potencial constante de 20V, por 1 h. A imunomarcação da membrana consistiu em 3 etapas: bloqueio com PBS-T 5% leite (0,05% Tween 20, 5% leite em pó desnatado), por 16 h a 8-16°C; incubação com o anticorpo primário diluído em PBS-T 5% leite, por 16 h a 8-16 °C; e incubação com o anticorpo secundário diluído em PBS-T 5% leite, por 2 h em temperatura ambiente. Ao final de cada etapa, a membrana foi submetida a 3 lavagens de 5 min. cada, com PBS-T. Por fim, a fluorescência da membrana foi lida no sistema de captura de imagem em infravermelho LiCor Odyssey 9120.

Os anticorpos primários utilizados foram o anti-vimentina (diluição 1:1000, BD Biosciences) derivado de camundongo, e o anti-GFP (diluição 1:2000, ThermoFisher), de coelho. Como anticorpo secundário, foi utilizado o Goat IRDye 800 anti-mouse (diluição 1:50000, Odyssey) e o Donkey IRDye 800 anti-rabbit (diluição 1:50000, Odyssey), respectivamente.

### 4.3 CONSTRUÇÃO DE FERRAMENTAS PARA AVALIAÇÃO DO EFEITO DO CARDEL COM MUTAÇÃO NOS sORFS PUTATIVOS

A construção AAVS1 CARDEL mut foi criada a partir do vetor AAVS1-GW-TRE-rtTA, esquematizado no Anexo 1, Figura 16, com o intuito de investigar se o *CARDEL* como transcrito é relevante para os efeitos observados durante a superexpressão de *CARDEL* com potencial codificante na diferenciação cardíaca. Pereira (2019) criou um vetor de superexpressão induzida desse transcrito (*SERTM2-202*) contendo uma mutação pontual na região da CDS predita, sequência chamada de *CARDEL mut* (Fig. 3). Foram gerados amplicons de *CARDEL* contendo uma inserção de 2 adenosinas no início dos 2 sORFs citados (Fig. 3, destacado em vermelho), alterando o quadro de leitura dessas sequências.

A obtenção da sequência mutada de *CARDEL SERTM2-202* consistiu na amplificação por PCR de fragmentos de *CARDEL*, utilizando de maneira cruzada os *primers* LINC/CARDELaatB\_F e LINC/CARDELmut\_R e vice-versa. Uma última PCR foi realizada para fusionar os amplicons de *CARDEL*, utilizando os *primers* externos CARDELaatB. Os *primers* LINC/CARDELmut serviram para amplificar a CDS com a mutação, e LINC/CARDELaatB para incluir as sequências *aatB* necessárias à clonagem via sistema *Gateway*.

O vetor AAVS1-GW-TRE-rtTA possui braços de homologia com o locus AAVS1 do genoma humano, localizado no cromossomo 19. A inserção do transgene de interesse ocorre após clivagem sítio-específica via nucleases *zinc fingers* recombinantes (ZFN) (HOCKEMEYER et al., 2009, LOMBARDO et al., 2011 e OCEGUERA-YANEZ et al., 2016). Como elemento para controle de expressão, há um sítio de elemento responsivo a tetraciclina (TRE, *Tetracyclin Responsive Element*), localizado no promotor do transgene, e uma sequência para expressão constitutiva da proteína auxiliar rtTA (*reverse tetracycline-controlled TransActivator*). O cassete de superexpressão induzida é do tipo Tet-ON, logo a rtTA constitutiva é ativada ao interagir com doxiciclina, tornando-se capaz de interagir com o sítio TRE e promover a transcrição do transgene, nesse caso o *CARDEL mut* (mecanismo

revisado por OCEGUERA-YANEZ et al., 2016). O vetor ainda contém sítios de resistência a puromicina, facilitando a seleção em etapas de geração de linhagens.

#### 4.3.1 Transfecção de células embrionárias de rim (HEK293)

A integridade e funcionalidade do plasmídeo AAVS1-CARDEL mut foram avaliadas através da transfecção de células HEK293, empregando o lipídeo catiônico Lipofectamine 3000, como descrito na seção 4.2. Após 48 h de transfecção, as células foram visualizadas em microscópio óptico Leica DMI6000B, com cubo de fluorescência para FITC 3450B (Ex. 482/35 nm) e imagens foram adquiridas pelo *software* LAS AF. Para teste da superexpressão induzida de *CARDEL* em condições com doxiciclina (+DOX) e sem doxiciclina (-DOX), o meio de 2 poços de HEK293 transfectados com AAVS1-CARDEL mut foram trocados por DMEM suplementado e incubados por mais 24 h, com adição do indutor doxiciclina em um dos poços (+DOX, 500 ng/mL). Em seguida, as células foram lisadas em TRIReagent® (Merck) e submetidas a protocolo de extração de RNA, síntese de cDNA e PCR quantitativa, em protocolo explicado em detalhes na seção 4.3.3.

#### 4.3.2 Transfecção de iPSC PLZ com AAVS1-CARDEL mut e seleção da linhagem

Células-tronco pluripotentes induzidas derivadas de fibroblastos de derme (PLZ- IPRN13.13, de DARABI et al., 2012) foram mantidas a 37°C, 5% CO<sub>2</sub>, em superfícies revestidas com Corning® Matrigel® Matrix (Gibco) diluído 1:100 em DMEM suplementado com o nutriente Ham F-12 (Gibco), durante todas as etapas de cultivo. Nas passagens, as células foram lavadas com PBS, incubadas com Accutase™ (STEMCELL Technologies) por 4 min. a 37 °C, transferidas para um tubo contendo PBS (5 volumes de PBS para 1 volume de accutase) e centrifugadas por 4 minutos a 220 x g. As células foram então ressuspensas em StemFlex™ (Gibco) suplementado com 10 µM de inibidor da via ROCK (iROCK, Y27632, Sigma-Aldrich) e transferidas para garrafa de 25 cm<sup>2</sup> ou placas de 6, 12 ou 24 poços previamente revestidas com Matrigel. No dia posterior à passagem, o meio foi

substituído por Stem Flex sem iROCK e trocado em dias alternados até o fim dos experimentos ou até a cultura atingir confluência de 70-80%.

No dia anterior à transfecção, células foram transferidas para placas de 24 poços (50.000 células/poço). Para cada poço, 0,5-1 µg de vetor (300-600 ng AAVS1 e, quando aplicável, 200-400 ng de vetores para *zinc fingers*) e 2 µL de Lipofectamine Stem Reagent (Invitrogen) foram diluídos em 25 µL de Opti-MEM cada. O vetor AAVS1-caggs-eGFP (addgene code #91565) foi utilizado como controle de transfecção. As alíquotas de vetor e de lipofectamina foram misturadas e incubadas em temperatura ambiente por 10 min. O meio de cultura foi trocado por 500 µL/poço de Opti-MEM e a mistura DNA-lipofectamina foi gotejada nos locais correspondentes. Após 4h, o meio foi trocado para 500 µL de Opti-MEM e as células foram incubadas por 16h. O meio de cultura foi trocado para Stem Flex, as células foram visualizadas em microscópio Leica DMI6000B, com cubo de fluorescência para FITC 3450B (Ex. 482/35) e imagens foram adquiridas pelo *software* LAS AF.

A seleção de clones a partir da transfecção ocorreu com a expansão da cultura com puromicina (200 ng/mL), em placas de 6 poços. Essas células pré-tratadas foram plaqueadas em baixa confluência (1:20/1:40). Sob visualização em microscópio invertido em cabine de segurança biológica, pequenas colônias isoladas foram transferidas, com o auxílio de um micropipetador, para poços de placa de 24 poços.

Alguns clones de PLZ CARDEL mut foram submetidos a testes de superexpressão induzida de *CARDEL* ao contrastar condições com adição do indutor doxíciclina (+DOX) e sem doxíciclina (-DOX). Para isso, células PLZ CARDEL mut foram semeadas 2 a 2 em placas de 24 poços (50.000 células/poço) e cultivadas até atingirem confluência de 50-60%. Na condição +DOX, as células foram então incubadas por 24 h em 500 µL de meio Stem Flex contendo 500 ng/mL de doxíciclina, enquanto na condição -DOX foram simplesmente mantidas por mais um dia em cultivo. Após esse período, as células foram lisadas e ressuspendidas em TRIReagent® (Merck) e submetidas a protocolo de extração de RNA, síntese de cDNA e PCR quantitativa, processos mais bem descritos na seção 4.3.3. Tanto as

células pré-selecionadas com puromicina quanto os clones foram submetidos a teste de integração do sistema AAVS1 no genoma por meio de PCR, descrita em mais detalhes na seção 4.3.4.

#### 4.3.3 Quantificação de expressão de RNA por PCR quantitativa (qPCR)

O RNA total da cultura de HEK293 ou PLZ foi extraído seguindo as instruções do kit *Direct-zol RNA Miniprep* (Zymo Research). Para a conversão do RNA em cDNA, *primers* oligodT foram adicionados a 1 µg de RNA extraído, incubados por 5 min a 70 °C e, posteriormente, por 5 min a 4 °C. Em seguida, a síntese do cDNA foi realizada segundo instruções do kit *ImProm-II™ Reverse Transcription System* (Promega). O cDNA foi então submetido a qPCR utilizando o kit *Go Taq® qPCR Master Mix* (Promega), que se baseia no SYBR Green como reagente para quantificação da fluorescência e amplificação do cDNA. GAPDH e RNA POL foram utilizados como genes normalizadores. A lista de todos os *primers* utilizados consta na Tabela 2.

TABELA 2 – PRIMERS UTILIZADOS PARA qPCR

GENE ALVO		PRIMER (5' > 3')	TAMANHO	REFERÊNCIA
GAPDH	F	GGCGATGCTGGCGCTGAGTAC	149 pb	(SPANGENBERG et al., 2013)
	R	TGGTTCACACCCATGACGA		
RNA POL	F	TACCACGTCATCTCCTTTGATGGCT	187 pb	Autor, 2022
	R	GTGCGGCTGCTTCATAA		
CARDEL (LINC00890)	F	GGGCCAGGAGAGGAAAATCAAT	145 pb	(PEREIRA, 2019)
	R	CACTGTAGGCTGGAGGGAAA		
MS2	F	TTCTGCAGATATCCAGCACAGT	N/A	(PEREIRA, 2019)
	R	CAACAGATGGCTGGCAACTAGAAG		
XIST	F	GCAACAACCCTAGGTCAGGA	149 pb	(PEREIRA, 2019)
	R	AGCTCTCTGCACTGCTTGTAG		
GFP	F	TGAAGTTCGAGGGCGACACC	74 pb	(PEREIRA, 2019)
	R	GATGTTGCCGTCCTCCTGAAGT		

O ciclo de temperaturas incluiu uma etapa *hot start* de 95 °C por 2 min, seguida por 40 ciclos de amplificação (95 °C por 15 seg. e 60 °C por 1 min) e, por fim, uma curva de fusão dos amplicons obtidos (95 °C por 10 seg., 65 °C por 1 min e 97 °C por 1 seg.). As análises das curvas de fluorescência e dos  $C_q$  (*quantification cycles*) foi realizada com o *software* LightCycler©SW 1.1 (Roche). O programa GraphPad Prism versão 7 foi empregado para construir os gráficos de expressão e realizar análises estatísticas.

Quanto aos tratamentos estatísticos, os dados de expressão relativa a RNA POL ou *Fold to RNA POL* de *CARDEL* para PLZ *CARDEL* mut nas condições com e sem doxiciclina foram primeiramente submetidos ao teste de normalidade pelo método de Shapiro-Wilk, ideal para amostras pequenas. Em seguida, foi aplicado o teste-t não pareado com correção de Welch, considerando que os 2 conjuntos de dados (-/+ DOX) possuem variâncias distintas.

#### 4.3.4 Detecção da sequência integrativa de AAVS1-GW-rtTA e outros testes

A detecção do cassete de superexpressão induzida (ilustrado como parte do vetor AAVS1-CARDEL mut, Anexo 1, Figura 16) integrado ao genoma de PLZ *CARDEL* mut foi realizada pela amplificação por PCR da porção final da sequência rtTA e do braço direito de homologia (RA-rtTA) do cassete, com 750 pb. Essa estratégia foi utilizada para confirmar a integração do sistema AAVS1 no genoma.

O DNA genômico (gDNA) de cultura de hiPSC PLZ foi extraído com o QIAmp DNA Extraction Kit (Qiagen), de acordo com o protocolo para células cultivadas *in vitro*, e usado como molde para gerar amplicons RA-rtTA. Para isso, foi empregado o par de *primers* rtTA-Fw (5' ACCTACCACCGATTCTATGCC 3') e AAVS1-Rv (5' TCCTCTCTGGCTCCATCGTAA 3'). Essa PCR foi realizada com o *mix* de PCR HPV/NAT (Biomanguinhos), e o ciclo de temperaturas constituiu em uma etapa *hot start* de 95 °C por 2 min, seguida por 40 ciclos de amplificação em 3 etapas (95 °C por 30 seg., 60 °C por 1 min e 72 °C a 60 s/kb).

#### 4.3.5 Preparo de células PLZ para cariotipagem

Primeiramente, as células foram expandidas em 3 poços de placa de 6 poços e mantidas até a cultura atingir 70-80% de confluência ( $2-3 \times 10^6$  células). A fim de interromper a metáfase, as células foram incubadas com colchicina (320  $\mu\text{g/mL}$ ) por 1 h 30 min. Em seguida, o meio de cultura foi retirado e reservado, as células foram lavadas 2 vezes com PBS 1x, e incubadas em PBS 1x por 10 minutos, a 37 °C.

As células foram retiradas mecanicamente com o auxílio de um *scraper* e transferidas para o meio reservado. A partir desse ponto, as células foram submetidas a ciclos de lavagem e centrifugação (800 x g, 5 min). Primeiro, o meio foi centrifugado, o *pellet* ressuspenso gentilmente em 7 mL de solução de KCl 57 mM e incubado por 10 min a 37 °C. Posteriormente, 1 mL de solução fixadora 1 (metanol 75%, ácido acético 25%) foi adicionada, o tubo foi incubado por mais 10 min a -20 °C e centrifugado. O *pellet* foi submetido a mais um ciclo lavagem-centrifugação com 5 mL de solução fixadora 1, depois mais 2 ciclos, cada qual com 5 mL da solução fixadora 2 (metanol 66,6%, ácido acético 33,3%). Por fim, as células foram ressuspendidas em 1 mL da solução fixadora 2 e mantidas a -20°C até a cariotipagem, realizada pelo Núcleo de Tecnologia Celular da Pontifícia Universidade Católica (PUC-PR).

#### 4.4 DETECÇÃO DE INTERAÇÕES CARDEL-PROTEÍNA POR incPRINT

##### 4.4.1 Teste de funcionalidade de pcDNA 3.1 – CARDEL10x MS2

O plasmídeo pcDNA 3.1 – CARDEL 10x MS2 foi construído por Pereira (dados não publicados), utilizando os sítios de restrição AgeI e PaeI, a partir do vetor pcDNA 3.1-10x MS2. Graindorge et al. (2019) incluíram 10 sequências do *stem loop MS2* entre os sítios de BamHI e BglIII do vetor, modificado com base no pcDNA 3.1 (ThermoFisher).

Comparou-se a expressão de *CARDEL* nas frações de RNA nuclear, citosólico e total, nas condições sem superexpressão (vetor pcDNA 3.1-10x MS2, pV) e com superexpressão (pcDNA 3.1 - **CARDEL**-10x MS2, pC). O teste de

superexpressão transiente foi realizado em células HEK293 como descrito na seção 4.2, sem indução com doxíciclina e utilizando Lipofectamine 2000. No segundo dia de transfecção, as células foram desaderidas por tripsina-EDTA 0,05%. Para as frações de RNA total, as células centrifugadas foram ressuspensas em TRIReagent® (Merck). Para as frações núcleo/citoplasma, por outro lado, o *pellet* foi ressuspensado em um tampão de lise [Tris-HCl (15mM) pH 7,4, MgCl<sub>2</sub> 15mM, NaCl 0,3M, Triton X-100 1%, RNase OUT (Invitrogen) 250 nM, DNase I (Invitrogen) 50 U]. Os extratos foram incubados em gelo por 10 min, centrifugados por mais 10 min a 12000 x g, 4 °C. O *pellet* formado foi considerado como a fração nuclear e o sobrenadante a fração citosólica. Cada fração foi misturada a 400 µL de TRI Reagent® (Merck) e submetida a extração de RNA pelo protocolo fenol-clorofórmio. Cada volume de amostra de lisado celular em TRIReagent® recebeu um volume equivalente de clorofórmio e foi centrifugada a 4°C, 12000 x g, por 10 min. Depois, o sobrenadante de cada mistura trifásica resultante foi coletado e submetido a precipitação com 3 volumes de isopropanol, sendo centrifugado a 4 °C, a 7500 x g por 5 min. Cada *pellet* resultante foi lavado com etanol absoluto e ressuspensado em água livre de nuclease.

A expressão de *CARDEL*, *MS2*, *XIST* (controle de RNA nuclear), *GAPDH* (controle de RNA citosólico) e *GFP* (normalizador) foi quantificada por qPCR como descrito na seção 4.3.3, com 2 diferenças: o RNA das frações totais, nucleares e citosólicas foi extraído pelo protocolo fenol-clorofórmio; e 5 pg de RNA *GFP* foram adicionados a cada alíquota de RNA extraído destinada à geração de cDNA, funcionando como normalizador do nível de expressão dos transcritos. A sequência dos *primers* utilizados para a qPCR consta na Tabela 2.

#### 4.4.2 Técnica de incPRINT

O experimento de *in cell protein-RNA interaction* (incPRINT) (GRAINDORGE et al., 2019) foi realizado pela Unidade de Genética e Biologia do Desenvolvimento (Instituto Curie, Paris), como parte da colaboração com a PhD Alena Shkumatava.

Resumidamente, células HEK293T expressando constitutivamente a luciferase NanoLuc fusionada a proteína de revestimento do bacteriófago MS2 (MS2CP) foram semeadas em placas de 96 poços e, no dia seguinte, transfectadas com pcDNA3.1-CARDEL10xMS2 e plasmídeo correspondente à proteína teste fusionada à etiqueta 3xFLAG. Após 48h de transfecção, a cultura foi lavada com PBS e ressuspendida em tampão de lise (20 mM HEPES-KOH [pH 7,9], 150 mM NaCl, 10mM MgCl<sub>2</sub>, 1mM CaCl<sub>2</sub>, 5% glicerol, 1% Triton X-100, suplementado com inibidores de protease). Os extratos celulares foram transferidos para placa de 324 poços revestida com o anticorpo anti-FLAG M2 e incubados por 3h a 4°C. Fumarizina foi adicionada a cada poço e a fluorescência foi medida em leitor de placa. Os dados foram discriminados em unidades relativas de luminescência (RLU), para um conjunto de proteínas consideradas ligantes comuns.

#### 4.5 ANÁLISES *IN SILICO* DO POTENCIAL DE INTERAÇÃO CARDEL-RBP

As sequências e anotações do transcrito curto (*SERTM2-201*) e longo (*SERTM2-202*) foram consultadas no banco de dados Ensembl (CUNNINGHAM et al., 2022), para a versão GRCh38.p13 do genoma humano. O programa RBPmap foi utilizado para anotar RBPs com potencial de interagir com as variantes de *CARDEL*, tendo como único dado de entrada as sequências de RNA dos transcritos. O algoritmo do programa identifica os sítios de ligação das RBPs presentes no RNA e calcula estatisticamente qual interação é mais provável, seguindo a abordagem de *Weighted-Rank* (WR) (PAZ et al., 2014). Nesse trabalho, a varredura de sítios de ligação foi realizada com as RBPs humanas/murinas listadas na base de dados do programa. Nas configurações avançadas, o nível de restrição (*stringency level*) foi mantido nas configurações padrão, e o filtro de conservação de motivos proteicos não foi aplicado.

O potencial de interação *SERTM2-201/202*-RBPs previamente selecionadas pelo RBPmap foi mensurada pelo programa RPI-Seq, baseado nos classificadores *Random Forest* (RF) e o *Support Vector Machine* (SVM), gerados por métodos de aprendizagem de máquina. Nesse contexto, RF e SVM classificam interações RNA-

proteína por probabilidade, com base em um RNA e uma lista de proteínas candidatas (MUPPIRALA et al., 2011). As proteínas candidatas processadas por RPISeq foram filtradas pelos seguintes critérios de exclusão (E):

- I. Ao menos uma interação putativa da RBP com *CARDEL* com Z-score menor que 2, correspondente a um p-value aproximado de 0,02;
- II. Valor de classificador RF menor a 0,9;
- III. Não ser uma proteína humana ou ortóloga a uma proteína humana.

A localização celular/subcelular, processos biológicos relacionados e sequência das proteínas selecionadas pelo RBPmap foi obtida do banco de dados UniProt (BATEMAN et al., 2021). Como critério de inclusão, selecionaram-se proteínas cujos processos biológicos descritos contivessem as seguintes palavras-chave: 1) “Muscle”, “Cardiac”, “Heart”; 2) “Translation” ou correlatos; 3) “Differentiation” ou correlatos; 4) “Embryonic development” ou correlatos.

## 5 RESULTADOS E DISCUSSÃO

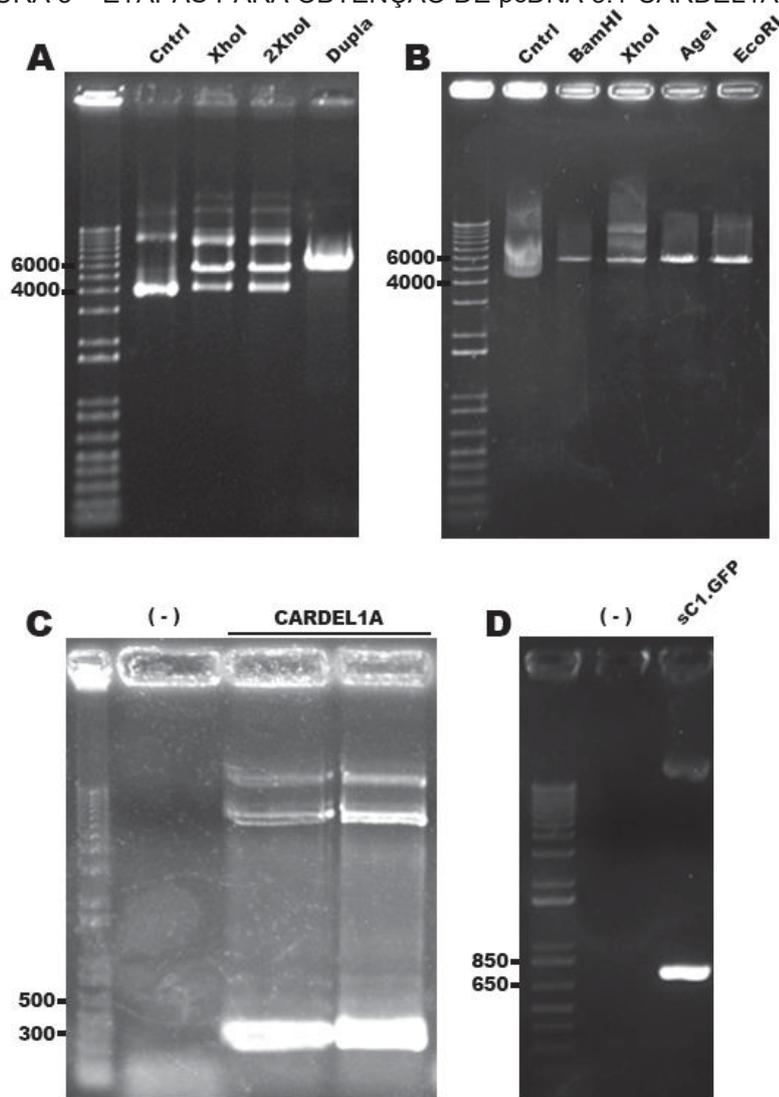
### 5.1 OBTENÇÃO DE CONSTRUÇÃO pcDNA 3.1-His B/myc-CARDEL1A.eGFP

O plasmídeo pcDNA 3.1-His B/myc-CARDEL1A.eGFP foi construído visando experimentos de avaliação do potencial de tradução de *CARDEL1A.eGFP* no peptídeo fusionado SERTM2-eGFP. Um dos obstáculos para a obtenção do plasmídeo foi a etapa de digestão do vetor pcDNA 3.1 com as endonucleases escolhidas num primeiro momento, a XhoI e a BamHI. A digestão do vetor pcDNA 3.1 com XhoI revelou-se ineficiente, mesmo em condições em que foi adicionado o dobro de unidades de enzima (Fig. 5A).

Logo, foi necessária a readequação da estratégia de clonagem através da escolha de outras endonucleases cujos sítios se localizassem a jusante do sítio de BamHI. Apesar de ambas as enzimas escolhidas para teste, AgeI e EcoRI, hidrolisarem eficientemente o vetor (Fig. 5B), a EcoRI foi selecionada por ser compatível para digestão dupla com BamHI. Os fragmentos *CARDEL1A* (370 pb) e *sCARDEL1-eGFP.ECO* (730 pb) foram amplificados a temperatura de anelamento

(Ta) de 60 °C (Fig.5C/D). A PCR de extensão por sobreposição foi realizada com uma Ta de 55 °C, visando facilitar o anelamento dos fragmentos *CARDEL1A* e *sCARDEL1-eGFP.ECO*, cuja sequência de sobreposição possui apenas 15 pb.

FIGURA 5 – ETAPAS PARA OBTENÇÃO DE pcDNA 3.1-CARDEL1A.eGFP

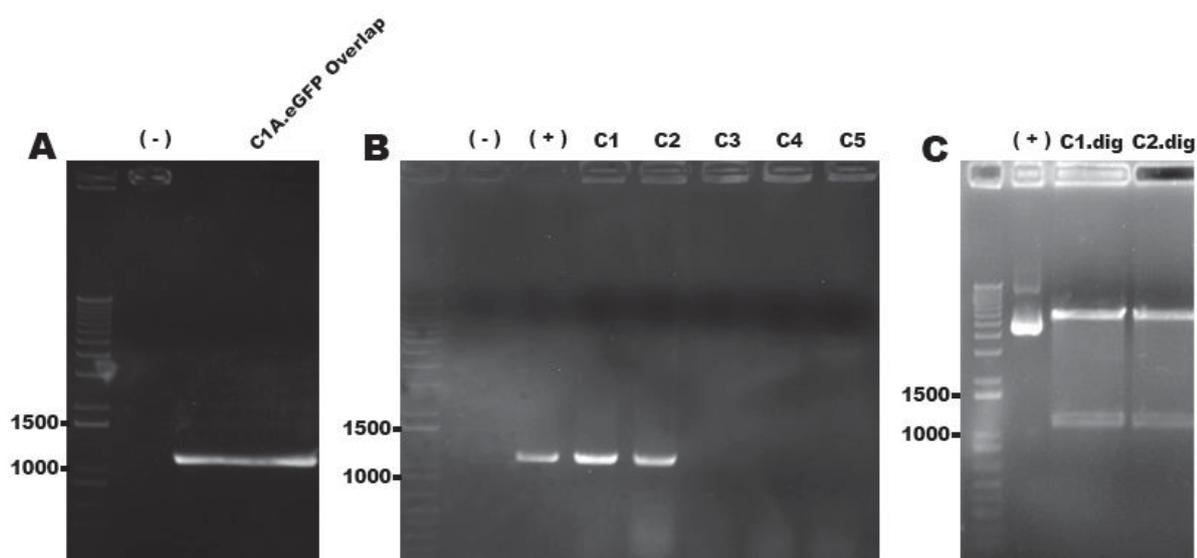


NOTA: Em (A), vetor pcDNA 3.1 myc/His B (Cntrl) e digerido com XhoI, o dobro de unidades de XhoI (2XhoI) e com BamHI/XhoI (Dupla); (B), teste de digestão de pcDNA 3.1 com AgeI e EcoRI, em comparação a BamHI e XhoI; (C), amplicons para *CARDEL1A*; (D) amplicon para eGFP com *overlap*, *sCARDEL1-eGFP* (sC1.GFP); (-) - controle de PCR sem DNA molde.

Não obstante amplificação concorrente dos fragmentos individuais, foi possível observar o amplicon estendido de cada reação, de aproximadamente 1090 pb: *CARDEL1A-eGFP* (Fig. 6A). Após reação de ligação do inserto *CARDEL1A-*

*eGFP* ao vetor pcDNA 3.1, em proporção molar de 3:1, e transformação por choque térmico em *E. coli* DH5 $\alpha$ , cinco clones foram testados para a presença do fragmento fusionado *CARDEL1A-eGFP*, por PCR convencional (Fig. 6B). O teste indicou 2 clones positivos para a sequência fusionada *CARDEL1A-eGFP* (C1 e C2).

FIGURA 6 – AMPLIFICAÇÃO DE SEQUÊNCIA FUSIONADA *CARDEL1A.eGFP*



NOTA: Em (A), amplicon da sequência fusionada *CARDEL1A.eGFP* (*C1A.eGFP Overlap*); (B), PCR de colônia de clones de *E. coli* DH5 $\alpha$  + pcDNA-*CARDEL1A.eGFP*; (+) - PCR a partir da sequência fusionada *CARDEL1A.eGFP*; (-) - controle de PCR sem DNA molde; (C), teste confirmatório de liberação do fragmento de *CARDEL1A.eGFP* dos clones positivos para a PCR de colônia (C1 e C2), via digestão dos plasmídeos com BamHI e EcoRI; (+) - plasmídeo pcDNA 3.1 não digerido (450 ng).

Com o intuito de confirmar os clones positivos para a PCR de colônia (Fig. 6B), foi realizada a digestão da mini-preparação plasmidial dos clones com as enzimas empregadas na clonagem, a BamHI e EcoRI (Fig. 6C), resultando na liberação de um fragmento de extensão condizente a *CARDEL1A-eGFP* (1090 pb).

## 5.2 ANÁLISE DE EXPRESSÃO DE CÉLULAS HEK293 TRANSFECTADAS COM pcDNA 3.1-His B/myc-*CARDEL1A.eGFP*

A expressão transiente e putativa do micropeptídeo SERTM2 (codificado pelo *sORF-1* de *CARDEL*) fusionado a *eGFP* em células HEK293 foi monitorada por

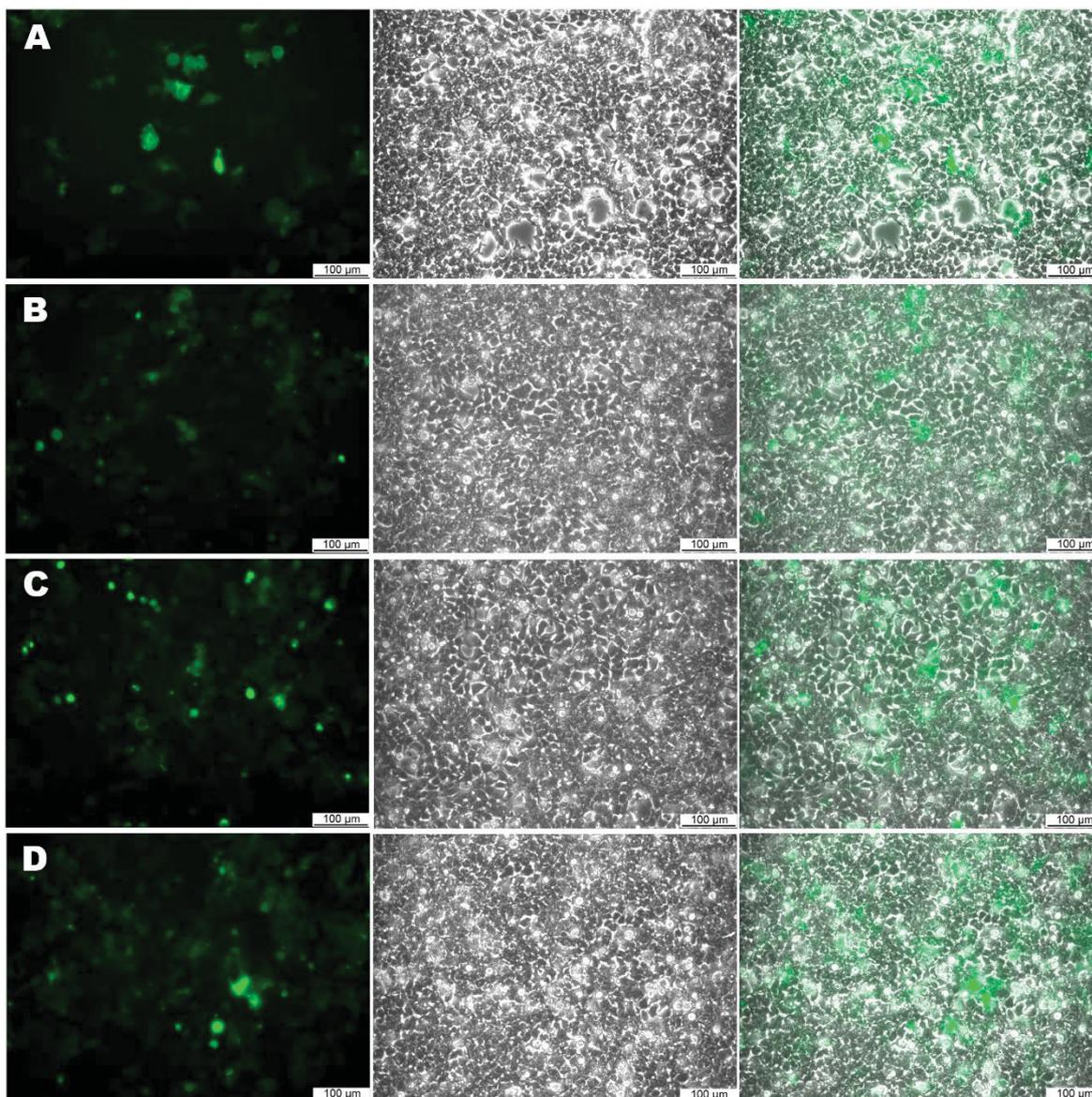
microscopia de fluorescência, comparando diferentes condições de transfecção com o plasmídeo pcDNA 3.1-His B/myc-CARDEL1A.eGFP.

A eficiência de transfecção nas condições do primeiro teste de expressão (Fig. 7A), utilizando lipofectamina 2000 e 1 µg de plasmídeo, estava aquém do desejado. Em decorrência disso, buscou-se ajustar alguns parâmetros da reação, como massa de DNA, meio de cultivo e reagente de transfecção. Shi et al. (2018) demonstraram que a eficiência de transfecção do vetor de expressão pLenti CMV GFP DEST contendo *eGFP* em HEK293 é maior utilizando lipofectamina 3000, em comparação aos reagentes FuGene 6 e lipofectamina 2000. Já Mori et al. (2020) inicialmente testam o efeito da massa de plasmídeo adicionada transfectando as células com poli-etilenamina (PEI) em DMEM com 10% de soro fetal bovino (SFB). Os autores desse estudo verificaram que, em placa de 6 poços, o aumento da massa de DNA transfectado se correlacionou de maneira linear com a eficiência de transfecção até um limite de 4 µg, promovendo um ganho de eficiência de 30%.

Considerando essas informações, foram testadas novas condições de transfecção de pcDNA 3.1-His B/myc-CARDEL1A.eGFP em placa de 6 poços, utilizando lipofectamina 3000 como reagente: 1 µg de vetor, em meio Opti-MEM, recomendado pelo fabricante (Fig. 7B); 3 µg de vetor, em meio Opti-MEM (Fig. 7C); e 3 µg de vetor, em meio DMEM 10% SFB (Fig. 7D).

O uso de lipofectamina 3000 parece aumentar a eficiência de transfecção sutilmente (Fig. 7A/B). Quando combinada com a adição de mais plasmídeo, a otimização torna-se mais distinta (Fig. 7A/C). O uso de DMEM suplementado durante a transfecção não parece ter impacto significativo na eficiência da reação, enquanto parece alterar a proliferação celular e, por consequência, aumentar o grau de compactação da monocamada (Fig. 7D). Logo, o protocolo de transfecção que utiliza 3 µg de plasmídeo com lipofectamina 3000 foi escolhido como padrão para experimentos futuros utilizando pcDNA 3.1-His B/myc-CARDEL1A.eGFP e células HEK293.

FIGURA 7 – EXPRESSÃO DE pcDNA 3.1-CARDEL1A.eGFP EM HEK293



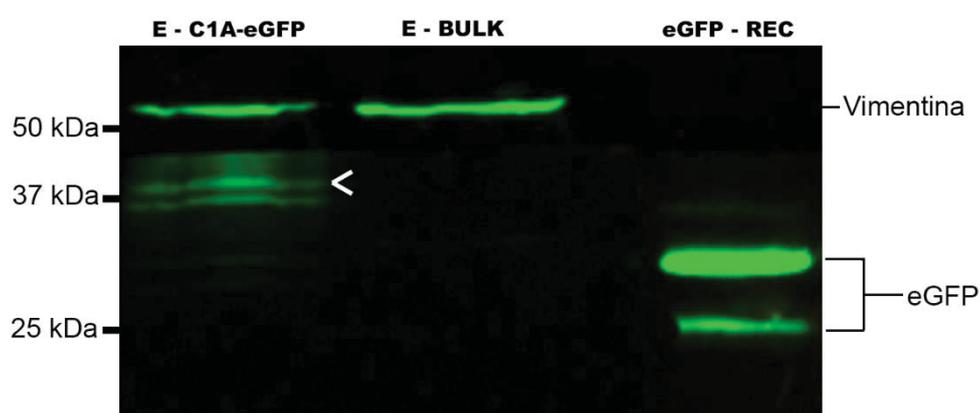
NOTA: Em (A), células transfectadas por 48h com lipofectamina 2000; em (B), (C) e (D) células transfectadas por 6h com lipofectamina 3000, 48h após início da transfecção; em (B), 1 µg de pcDNA 3.1-His B/myc-CARDEL1A.eGFP foi utilizado na transfecção, em meio Opti-MEM; em (C), 3 µg de plasmídeo, em Opti-MEM; e em (D), 3 µg de plasmídeo, em meio DMEM 10% SFB, 1% L-Gln. Imagens em aumento de 200x, obtidas com seguintes parâmetros de captura: (A) exposure 745 ms, gain 1; (B,C,D): exposure 300 ms, gain 1,5.

Em uma perspectiva mais geral, as células HEK293 transfectadas com construção pcDNA 3.1-His B/myc-CARDEL1A.eGFP foram capazes de emitir fluorescência verde em cultura, atestando a expressão de ao menos proteína eGFP. Uma vez que o micropeptídeo SERTM2 possui massa molecular predita de 10 KDa

(ProtParam Tool, ExPASy - DUVAUD et al., 2021), a proteína fusionada SERTM2-eGFP (36,9 KDa) poderia ser discriminada de eGFP (26,9 KDa, UniProt sequence P42212) por *Western Blot*.

A aplicação de *Western Blot* para o extrato de células HEK293 transfectadas com pcDNA 3.1 CARDEL1A.eGFP (Figura 8, E-CARDEL1A-eGFP) indicou a presença de proteínas de massas variáveis (50-37 KDa), reagentes aos anticorpos anti-eGFP e anti-vimentina.

FIGURA 8 – WESTERN BLOT DE EXTRATOS DE HEK293 TRANSFECTADAS COM pcDNA 3.1-CARDEL1A.eGFP



NOTA: Vimentina (54 kDa) foi utilizada como controle de *loading* para os extratos de HEK293; eGFP-REC indica solução purificada de eGFP recombinante; E-BULK indica extrato de células não transfectadas; E- CARDEL1A-eGFP indica extrato de células transfectadas com o vetor de expressão pcDNA 3.1 CARDEL1A.eGFP; a banda correspondente a SERTM2.eGFP está indicada pela seta branca.

O padrão de banda abaixo da marca de 50 KDa para o extrato de células transfectadas (E-C1A-eGFP) é visivelmente diferente do apresentado pelo extrato de células não transfectadas (E-BULK), sugerindo que a imunomarcação contra eGFP não foi inespecífica.

Percebe-se ainda uma marcação distinta alinhada ao topo do marcador de 37 KDa, não presente na solução de eGFP recombinante (eGFP-REC). Aparentemente, a proteína de fusão SERTM2-eGFP (Fig.8, seta branca) foi traduzida pelas células transfectadas. Esses resultados corroboram a hipótese de que o sORF-1 do lncRNA *CARDEL* apresenta potencial de tradução no micropeptídeo SERTM2, como anotado pelas bases de dados.

A tradução de sORFs de lncRNAs associados a polissomos é um mecanismo já descrito na literatura. Lara et al. (2021) detectaram, em nível proteico, um micropeptídeo de 68 aminoácidos, codificado pelo lincRNA *Linc013026*, que se revelou expresso de maneira endógena em células de carcinoma hepatocelular que apresentavam alta taxa de proliferação. Um exemplo ainda mais interessante de lncRNA associado a polissomos com potencial codificante é o *LINC00675*, que codifica para FORCP (*FOXA1-Regulated Conserved Small Protein*), um peptídeo com 79 aminoácidos (LI et al., 2020). Ao estudar a expressão dessa proteína em células diferenciadas de câncer colorretal, esse mesmo grupo elucidou algumas de suas funções, como a indução da apoptose e a inibição da proliferação das células cancerosas, tornando-as mais semelhantes a células com fenótipo normal. Logo, micropeptídeos apresentam potencial de ser traduzidos a partir de polissomos e exercer funções biológicas de maneira célula-específica.

Um sistema semi-automático de curadoria e anotação de sequências codificantes, o *Phylogenetic Annotation and INference Tool (PAINT)* (GAUDET et al., 2011), inferiu que o micropeptídeo SERTM2 estaria localizado na membrana de alguma organela ligada à membrana celular interna. Concordante com sua estrutura putativa, predições computacionais descrevem a conformação tridimensional do SERTM2 como composta por uma única hélice transmembrana e extremidades desordenadas (Fig. 9).

FIGURA 9 – PREDIÇÃO *IN SILICO* DE MICROPEPTÍDEO SERTM2



NOTA: Micropeptídeo posicionado com extremidade N-terminal à direita; a porção desordenada à esquerda (C-terminal) é anotada como um domínio citosólico; o padrão de cores indica a pontuação de confiança por resíduo de aminoácido (pLDDT) média em cada região, sendo azul-marinho: muito alta (pLDDT > 90), azul-ciano: confiável (90 > pLDDT > 70), amarelo: baixa (70 > pLDDT > 50) e laranja: muito baixa (pLDDT < 50).

FONTE: Jumper et al. (2021) e Varadi et al. (2022).

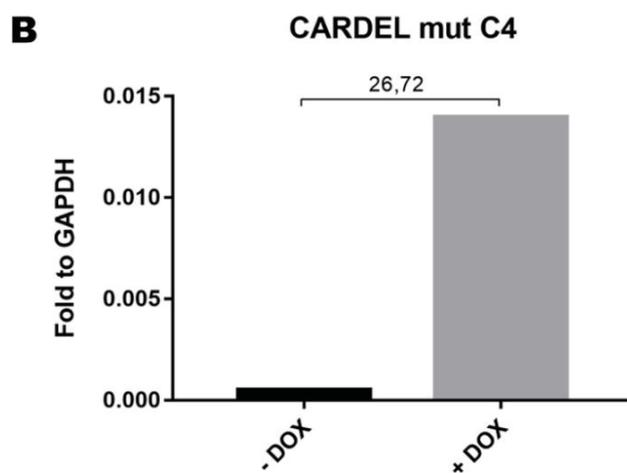
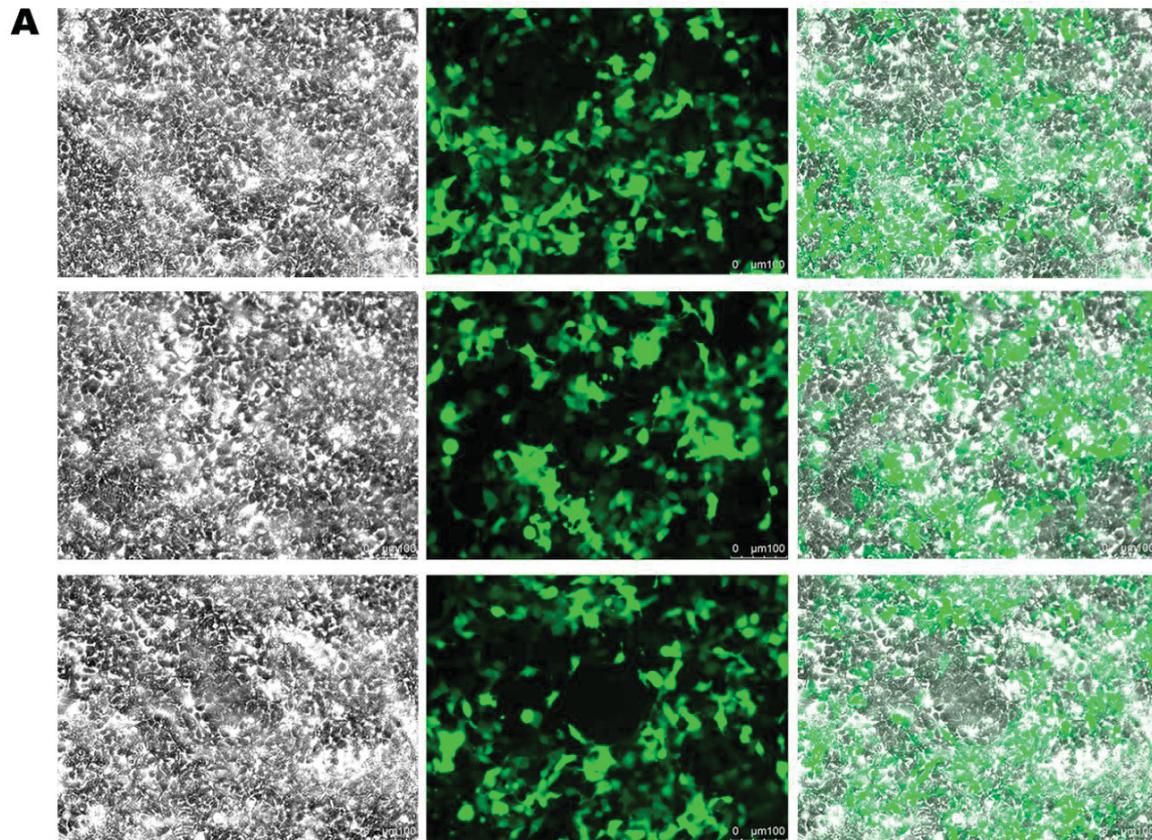
Embora a expressão de SERTM2-eGFP em HEK293 configure um marco importante na investigação do potencial codificante de *CARDEL*, os resultados obtidos possuem algumas limitações derivadas do delineamento experimental. A expressão do peptídeo ocorreu fora do contexto biológico da diferenciação cardíaca, em HEK293, linhagem celular modificada para otimizar a expressão de proteínas (como revisado por ABAANDOU et al., 2021). Algumas sequências presentes no vetor de expressão podem favorecer a geração de transcritos não expressos endogenamente, como o sinal de poliadenilação do fator de crescimento bovino (bGH) e o promotor de expressão eucariótica derivado do citomegalovírus humano (CMV), considerados sequências potenciadoras “fortes”, ou seja, geralmente associadas ao aumento da expressão do transgene quando comparadas a outros sinais de poliadenilação (REAL et al., 2011 e XU et al., 2002) ou promotores (BOSHART et al., 1985 *apud* ANDERSEN et al., 2011). Portanto, análises de expressão endógena do SERTM2 durante a diferenciação cardíaca são necessárias para consolidar as evidências levantadas no presente trabalho.

### 5.3 CARACTERIZAÇÃO DE LINHAGEM DE iPSC PLZ iCARDEL mut

A fim de testar a funcionalidade do sistema de superexpressão induzida do plasmídeo AAVS1-CARDEL mut, células HEK293 foram transfectadas separadamente com essa construção e com o vetor controle AAVS1-caggs-eGFP. A fluorescência das células transfectadas com o vetor controle foi registrada após 48h de transfecção, como mostrado na Fig. 10A. A transfecção com AAVS1-caggs-eGFP foi bem-sucedida, apresentando fluorescência em 3 campos distintos. Por isso, estimou-se que a transfecção das células com AAVS1-CARDEL mut também deveria ter sido um sucesso.

Para fins de confirmação, a expressão de *CARDEL* em células transfectadas com AAVS1-CARDEL mut foi monitorada por PCR quantitativa, utilizando a expressão de *GAPDH* como normalizador (Figura 10B).

FIGURA 10 – TRANSFECCÃO COM AAVS1--eGFP E EXPRESSÃO DE CARDEL EM CÉLULAS HEK293



NOTA: Em (A), microscopia de 3 campos representativos da cultura de células HEK293 após 48h de transfecção, em aumento de 200x; em (B), expressão relativa de *CARDEL* nas alíquotas purificadas de cDNA AAVS1 *CARDEL* mut C4 (alíquota do clone 4 de *E. coli* DH5 $\alpha$ ), comparando a expressão de *CARDEL* com o indutor doxíciclina (+DOX) e sem indutor (-DOX); resultados correspondentes a uma replicata.

Percebeu-se que a expressão de *CARDEL* em células incubadas com o indutor doxíciclina (+ DOX) foi quase 27 vezes mais alta quando comparada às células somente cultivadas por mais 24h pós transfecção (-DOX).

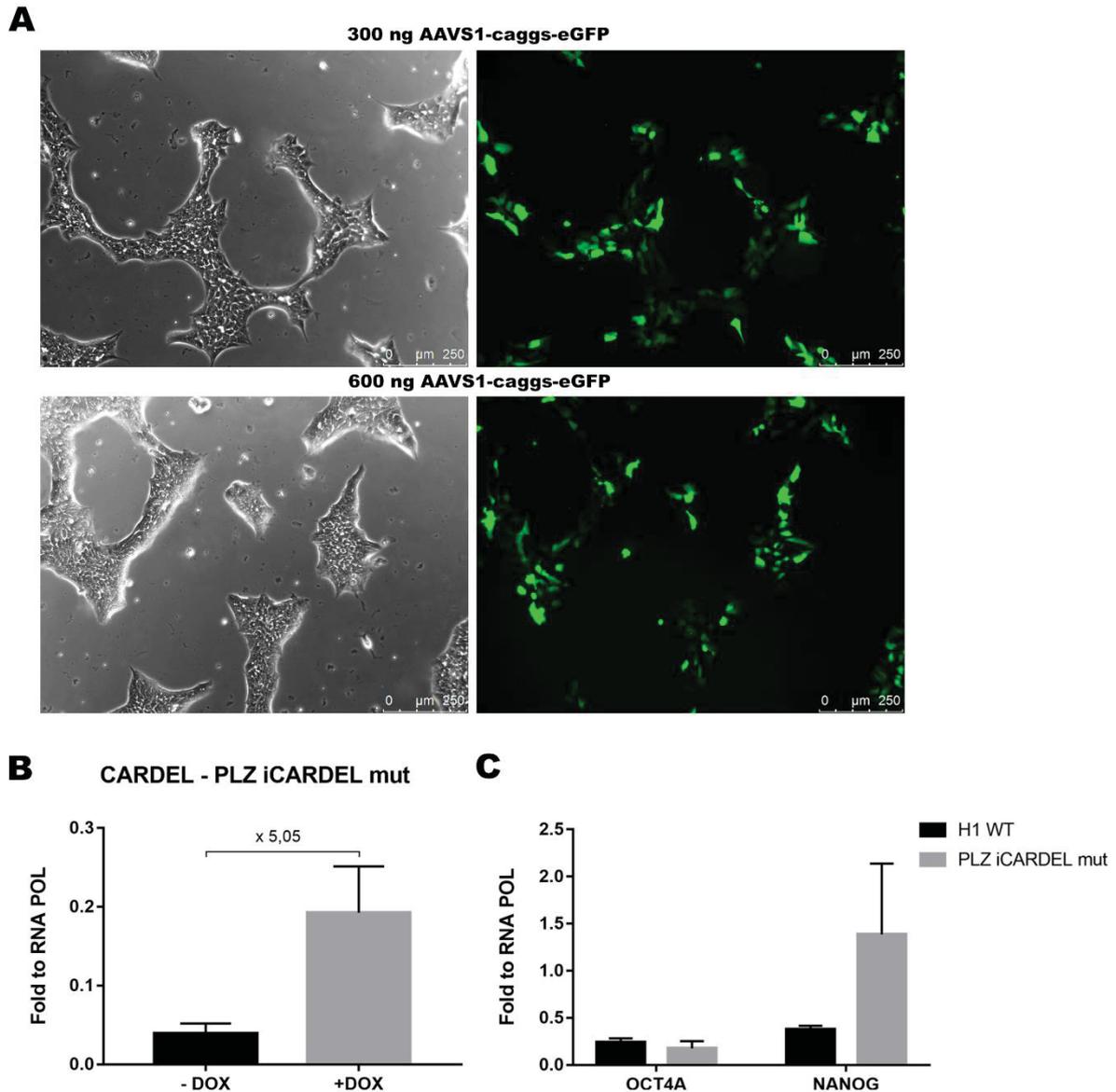
Como o sistema de superexpressão do plasmídeo AAVS1 *CARDEL* mut aparentava estar funcional, células-tronco pluripotentes PLZ foram transfectadas com o intuito de gerar uma nova linhagem, a PLZ *CARDEL* mut. Novamente, a transfecção das células com os plasmídeos AAVS1 foi estimada pela fluorescência verde de uma cultura controle (Fig.11A), transfectada com o plasmídeo AAVS1—eGFP. Visando avaliar indiretamente a eficiência da transfecção com AAVS1-*CARDEL* mut, 2 quantidades de plasmídeos AAVS1 foram testadas: 300 e 600 ng. A microscopia das células transfectadas com o plasmídeo controle sugere que a eficiência de transfecção com 300 e 600 ng de plasmídeo foi similar.

Apesar da semelhança de fluorescência e, portanto, de eficiência de transfecção nas células AAVS1--eGFP, era previsto que as células transfectadas com AAVS1-*CARDEL* mut apresentassem uma eficiência um pouco menor, devido a maior extensão desse vetor (14,7 Kb), em contraste com o AAVS1--eGFP (5,9 Kb). Por isso, optou-se por prosseguir com a cultura transfectada por maior massa de plasmídeo e, conseqüentemente, maior número de cópias de AAVS1-*CARDEL* mut.

Após passagens e rodadas de seleção com puromicina, 9 clones de PLZ *CARDEL* mut foram gerados (C1-7, C8.1 e C8.2), e submetidos a uma verificação preliminar de fenótipo pluripotente. Essa triagem inicial revelou-se importante para o avanço do trabalho, já que 2 dos clones demonstraram células de fenótipo divergente às pluripotentes. Essas colônias, derivadas dos clones C4 e C7, tornaram-se menos compactas, com bordas difusas e núcleo não visível (dados não mostrados), características discutidas por Wakui et al. (2017) e Tigges et al. (2021) como critérios associados a perda de pluripotência e redução da qualidade da cultura de células-tronco pluripotentes induzidas (iPSC). Sete clones (C1-3, C5, C6, C8.1 e C8.2) apresentaram fenótipo condizente a iPSCs que, de maneira semelhante a agrupamentos de células-tronco embrionárias (TAKAHASHI et al.,

2007) são caracterizadas por sua alta razão núcleo/citoplasma e são organizadas em colônias compactas e com bordas refringentes (THOMSON et al., 1998).

FIGURA 11 – TESTES DE TRANSFEÇÃO E CARACTERIZAÇÃO DE PLZ CARDEL mut



NOTA: Em (A), imagens representativas da cultura de células PLZ após 16h de transfeção com 2 quantidades de AAVS1-eGFP, 300 e 600 ng. Em (B), quantificação da expressão de *CARDEL* em PLZ *CARDEL* mut C8.1, em condições com doxiciclina (+DOX) e sem (-DOX), usando *RNA POL* como normalizador; o valor indicado nos colchetes corresponde ao  $\Delta\Delta Ct$  da condição +DOX; as 2 condições diferem com significância estatística ( $p$ -value = 0,0405), seguindo teste-t com correção de Welch; em (C), expressão de marcadores de pluripotência *OCT4*, isoforma A e *NANOG* em células embrionárias controle (H1 WT) e células PLZ *CARDEL* mut em condições -DOX; resultados de expressão gênica correspondentes a comparação de triplicatas experimentais.

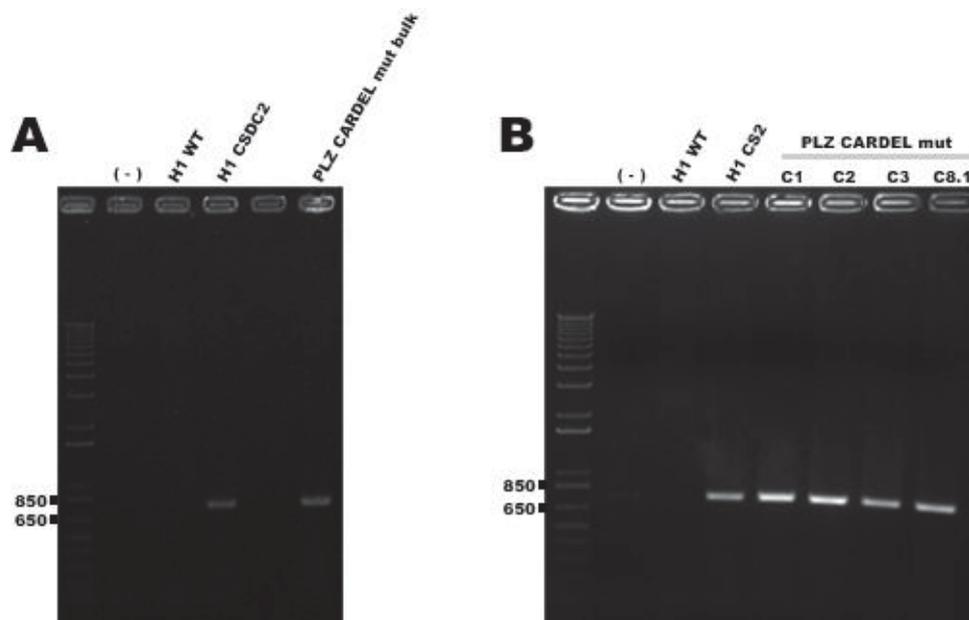
Para a seleção de um clone a ser utilizado em experimentos de diferenciação cardíaca, a expressão de *CARDEL* em células PLZ *CARDEL* mut C2 e C8.1 foi mensurada por qPCR, sem e com adição do indutor doxíciclina no cultivo. O clone PLZ *CARDEL* mut C2 não aparentou sofrer alterações de expressão de *CARDEL* com ou sem doxíciclina, enquanto o clone C8.1 apresentou aumento de expressão de *CARDEL* quando tratado com o indutor (APÊNDICE 2, Figura 19). É importante pontuar que o nível de expressão de *CARDEL* em células PLZ *iCARDEL* mut não induzidas por doxíciclina (-DOX) é ao menos 30 vezes menor do que o pico de expressão endógena observado durante a diferenciação cardíaca em monocamada realizada por Pereira (2019), indicando a baixa probabilidade do vazamento do sistema de superexpressão gerar um artefato em futuros experimentos de diferenciação cardíaca.

A funcionalidade da superexpressão induzida no clone PLZ *CARDEL* mut C8.1 foi confirmada por qPCR empregando outro gene normalizador (RNA POL), evidenciado na Fig. 11B. Células PLZ *CARDEL* mut C8.1 expressaram em média 5 vezes mais *CARDEL* quando expostas a doxíciclina (Fig. 11B).

Com o intuito de verificar a manutenção da pluripotência da linhagem obtida, a expressão de marcadores de pluripotência também foi averiguada (Fig. 11C). A PLZ *CARDEL* mut C8.1 foi capaz de expressar *OCT4A* e *NANOG*, em nível similar ou maior quando comparadas a células previamente caracterizadas como pluripotentes, como a linhagem de célula-tronco embrionária H1. Adicionalmente, o cariótipo desse clone não apresentou alterações (Anexo 2).

Com o objetivo de testar a integração do cassete de superexpressão na linhagem de PLZ *CARDEL* mut e validar os testes de expressão de *CARDEL* (Fig. 11B), buscou-se amplificar uma porção da sequência supostamente inserida no genoma, compreendendo o sítio rtTA e o braço direito de homologia ao sítio AAVS1 (amplicon de 750 pb), a partir do DNA genômico (gDNA) da linhagem. Os testes indicaram que tanto o gDNA de células recém transfectadas e não selecionadas com puromicina (*bulk*) (Fig. 12A) quanto o gDNA de 4 clones selecionados (C1, C2, C3 e C 8.1, Fig. 12B) são positivos para o cassete de superexpressão.

FIGURA 12 – TESTES DE INTEGRAÇÃO DO SISTEMA DE SUPEREXPRESSÃO DE CARDEL mut EM CÉLULAS PLZ



NOTA: Em (A), teste de integração para linhagem PLZ CARDEL mut recém transfectada (*bulk*); em (B), teste para clones C1, C2, C3 e C8.1. (-) corresponde a PCR sem cDNA. Como controle negativo, gDNA de hESC H1 *wild type* (H1 WT) e como positivo, gDNA de H1 iCSDC2 (H1 CS2).

Uma ressalva quanto a esse resultado é a porção amplificada para detecção não abranger parte do genoma do cromossomo 19 para além da sequência de homologia. Por isso, a técnica atual de detecção é demasiado sensível a contaminação da amostra de gDNA com cópias restantes do plasmídeo AAVS1—CARDEL mut, que poderiam igualmente servir como moldes para gerar os amplicons de 750 pb. Contudo, é relevante ressaltar que há degradação progressiva de plasmídeos não integrados ao genoma de células de mamífero associada a proliferação celular, consistindo muitas vezes na divisão do plasmídeo entre as células-filhas sem replicação adequada (WANG et al., 2016). Carapuça et al. (2007), por exemplo, determinaram que o tempo de meia-vida do plasmídeo de 3,7 kb pVAX1GFP em células de Ovário de Hamster Chinês (CHO) era de apenas 20h. Logo, estima-se que a quantidade de cópias de AAVS1-CARDEL mut em células após a seleção com puromicina era muito baixa para se tornar contaminante no extrato de gDNA.

A mutação para mudança de quadro de leitura dos sORFs de *CARDEL* não impede a transcrição e tradução de sequências no 3 quadro de leitura do transcrito, gerando peptídeos circunscritos à sequência do sORF-1. A sequência de um sORF localizado na porção final da sequência da sORF-1 (768 pb – 1027 pb do éxon 3 de *CARDEL*) é preservada, codificando para um peptídeo putativo pequeno de 19 aminoácidos. A mudança do quadro de leitura também origina uma sORF não antes lida no éxon 3 de *CARDEL*, que engloba os códons finais da sequência do sORF-1 e é codificante para um peptídeo que corresponde aos 54 aminoácidos da porção C-terminal de SERTM2. Apesar da similaridade com SERTM2, estima-se pelos modelos gerados via *software* SWISS-Model (ExpASy, DUVAUD et al., 2021) que o peptídeo truncado não apresentaria o domínio desordenado não citosólico do micropeptídeo completo (domínio visível na Fig. 9), muito provavelmente impactando sua funcionalidade e mantendo o *status* de PLZ *CARDEL* mut como um modelo de nocaute de proteína funcional.

#### 5.4 ESTUDO DA INTERAÇÃO DO CARDEL COM RBPs

O potencial de interação de *CARDEL* com RBPs foi acessado pela equipe da Dra. Alena Shkumatava, da Unidade de Genética e Biologia do Desenvolvimento (Instituto Curie, Paris), utilizando a técnica de incPRINT.

A incPRINT detecta proteínas que interagem com determinado RNA de interesse, sendo por isso classificada como uma técnica RNA-cêntrica. Essa técnica depende de 3 componentes expressos intracelularmente: a proteína a ser testada fusionada com cauda FLAG (proteína-FLAG); o RNA de estudo fusionado ao *stem loop* do bacteriófago MS2 (RNA-*stem loop* MS2; RNA-MS2); e a proteína de revestimento do bacteriófago MS2 (MS2CP) fusionada com a luciferase NanoLuc (MS2CP-NanoLuc). A porção MS2CP é capaz de interagir de maneira específica com o *stem loop* MS2, formando complexos de RNA-MS2 e MS2CP-NanoLuc (GRAINDORGE et al., 2019).

O processo se inicia com a expressão dual e transiente da proteína-FLAG e do RNA-MS2 em cultura de células modificadas para expressar constitutivamente a MS2CP-NanoLuc. Caso as proteínas testadas pudessem interagir com o lncRNA *CARDEL*, formar-se-ia um complexo de proteína-FLAG, RNA-MS2 e MS2CP-NanoLuc que poderia ser capturado por anticorpo anti-FLAG e detectado por ensaio de fluorescência.

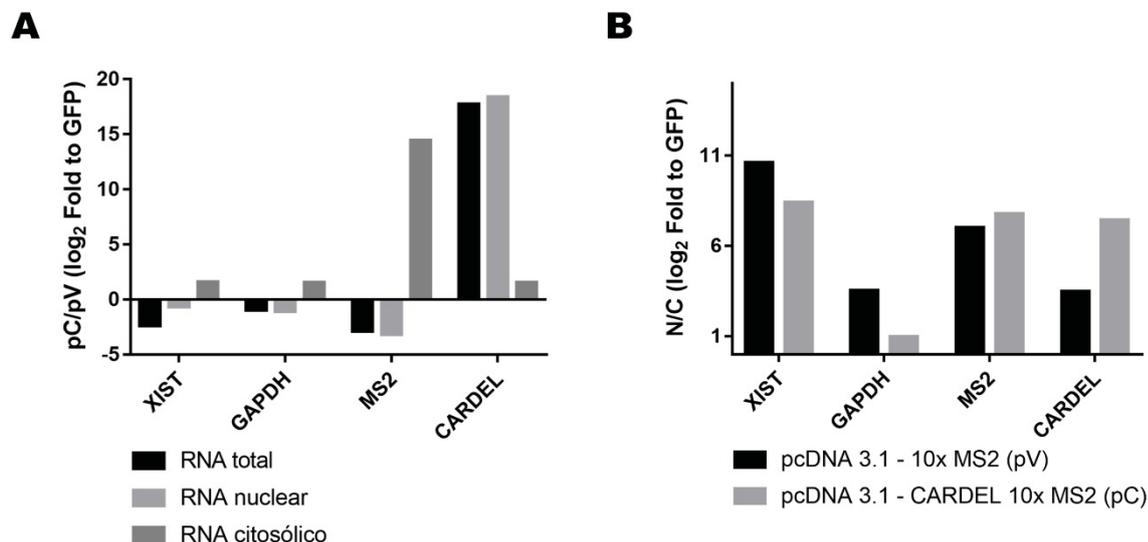
#### 5.4.1 Teste de funcionalidade de construção pcDNA 3.1 – *CARDEL* 10x MS2

Para a análise de incPRINT a partir de *CARDEL*, o único recurso fornecido para o grupo da Dra. Shkumatava foi um plasmídeo de expressão transiente de *CARDEL*, o pcDNA-*CARDEL*-10xMS2 (Anexo 1, Fig. 17), construído por Pereira a partir do vetor base cedido por Dra. Shkumatava et al. (dados não publicados). Restava confirmar se esse vetor era capaz de expressar *CARDEL* no citosol, como observado na expressão endógena durante a cardiomiogênese (PEREIRA, 2020).

Essa questão foi abordada pelo fracionamento núcleo/citoplasma de células HEK293 e análise da expressão de *CARDEL*, *MS2*, *XIST* e *GAPDH* por qPCR (Fig.13). *XIST* e *GAPDH* atuaram como controles de localização celular, sendo *XIST* predominantemente nuclear, e *GAPDH* citosólico. A análise das razões de expressão entre células transfectadas com o vetor pcDNA 3.1-*CARDEL* 10x MS2 (pC) e células contendo o vetor pcDNA 3.1 (pV) (razão pC/pV), mostradas na Figura 13A, indicam que a expressão de *CARDEL* na condição de superexpressão (pC) foi maior que a expressão em células abarcando o vetor vazio (pV), tendência observada em todas as frações de RNA (total, nuclear e citosólico). A superexpressão de *MS2* foi indicada apenas na fração citosólica, enquanto os controles *XIST* e *GAPDH* não sofreram variações tão drásticas de expressão quando comparados com *CARDEL*.

Os transcritos de *XIST*, *MS2* e *CARDEL* apresentam uma razão núcleo/citoplasma (razões N/C, Fig. 13B) maior do que 1, logo parecem se localizar prevalentemente na fração nuclear, independentemente da condição.

FIGURA 13 – PERFIS DE EXPRESSÃO DE HEK293 TRANSFECTADAS COM pcDNA3.1-CARDEL-10XMS2



NOTA: Gráficos de expressão do lncRNA nuclear *XIST*, do transcrito de *GAPDH*, uma proteína citosólica, do *stem loop MS2* e do *CARDEL*, por qPCR. Em (A), relação entre condições de superexpressão de *CARDEL*-10XMS2 (pC) e sem superexpressão (pV), nas diferentes frações de RNA; (B), expressão das razões núcleo/citoplasma (N/C), nas condições com e sem superexpressão. RNA de *GFP* foi utilizado como normalizador de expressão (*spike GFP*); resultados correspondentes a uma replicata.

O transcrito de *GAPDH*, por outro lado, é mais presente na fração nuclear na condição sem superexpressão de *CARDEL* (pV), tornando-se mais distribuído na condição de superexpressão (pC). O efeito reverso pode ser observado na retenção nuclear de *CARDEL*, sendo maior na condição de superexpressão quando comparada à ausência de superexpressão. *XIST* e *MS2*, por fim, parecem ter razões N/C semelhantes nas 2 condições.

É possível especular que a maior retenção nuclear de *CARDEL* em contraste a de *GAPDH* deve-se à expressão de *CARDEL* endógeno estimulada pelo sistema de superexpressão transiente, uma vez que os níveis de retenção da cauda *MS2* não se alteram entre as condições. Diferentemente da transcrição do *CARDEL* ectópico, a expressão do *CARDEL* endógeno ainda está atrelada ao correto processamento do pré-RNA por *splicing* alternativo. A comunidade científica já notou que processos de transcrição e *splicing* de lncRNAs tendem a ser mais ineficientes quando comparados a mRNAs, pela atuação de RNA polimerase II ativa, mas em conformação subótima (SCHLACKOW et al., 2017) e pela presença de sinais de

*splicing* mais fracos, como por exemplo a maior distância entre a extremidade 3' do íntron e ponto de forquilha, e a presença de menos pirimidinas (C/U) nessa região (MELÉ et al., 2017).

Devido a essas questões o *CARDEL* endógeno, quando expresso em excesso por algum estímulo associado à superexpressão, poderia ser incorretamente processado, impedindo sua exportação ao citosol e causando seu acúmulo no núcleo.

Apesar disso, o aumento conjunto de expressão de *CARDEL* e de *MS2* nas frações citosólicas (Fig. 13A) sugere que o transcrito *CARDEL-MS2* pôde ser superexpresso e exportado para o citosol, como desejado.

#### 5.4.2 Teste de interação de *CARDEL* com proteínas ligantes comuns

As intensidades de fluorescência dos complexos NanoLuc-*CARDEL-MS2*-proteína foram comparadas com complexos NanoLuc-*Menhir-MS2*-proteína (controle positivo) e NanoLuc-*MS2*-proteína (controle negativo), por meio da técnica de incPRINT.

Para ser considerado um candidato aos experimentos de incPRINT, os complexos com *CARDEL* deveriam apresentar intensidade de fluorescência em RLU igual ou maior aos complexos de *Menhir*, ao interagir com proteínas consideradas ligantes comuns (HNRPc, RBM15B, Mov10, SNRNP70). A fluorescência também deveria ser superior aos complexos formados com a proteína GFP, um controle negativo adicional do incPRINT.

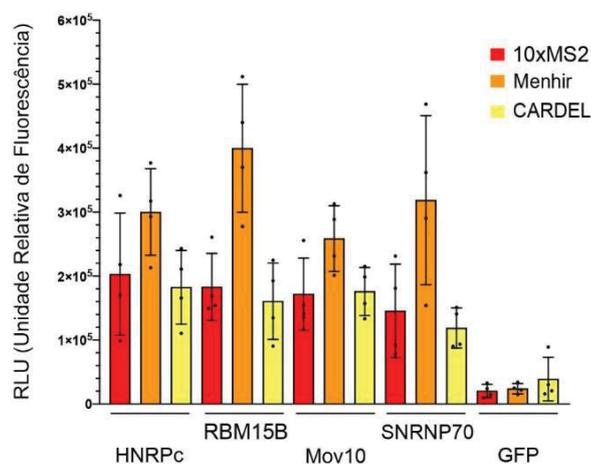
A unidade relativa de fluorescência (RLU) dos complexos de *CARDEL*, entretanto, não atingiu esses patamares de fluorescência com nenhuma das proteínas ligantes comuns, em padrão mais similar ao controle negativo 10xMS2. (Fig.14). Por isso, os experimentos com a biblioteca incPRINT não avançaram para além dessa etapa.

A falta de resposta do incPRINT para interações do *CARDEL* com as proteínas testadas pode estar mais relacionada às escolhas de construção do vetor

de expressão de *CARDEL* do que a sua capacidade de interagir com as RBPs. As proteínas utilizadas no teste de incPRINT com o *CARDEL* estão relacionadas a processos de *splicing*, como a HNRPC (LIU et al., 2015) e a proteína codificada por SNRNP70 (KONDO et al., 2015; KRUMMEL et al., 2009). De forma semelhante, as proteínas RBM15B e MOV-1 estão envolvidas, respectivamente, na metilação N6-metiladenosina (m6A) e exportação de RNA (PATIL et al., 2016; URANISHI et al., 2009) e na degradação de RNA via microRNAs e complexo RISC (CHENDRIMADA et al., 2007). A maioria desses eventos requerem a presença de sequências 5'/3'-UTR e/ou íntrons para reconhecimento e ligação ao RNA.

Como a sequência incorporada ao vetor pcDNA 3.1-10x MS2 corresponde ao transcrito já processado *CARDEL*, formulou-se a hipótese de que a incorporação de um íntron artificial ou um íntron próprio iria favorecer a resposta do *CARDEL* às proteínas testadas.

FIGURA 14 – INTERAÇÃO DE *CARDEL* COM LIGANTES COMUNS



NOTA: HNRPC, RBM15B, Mov10, SNRNP70 são proteínas ligantes comuns; as emissões de fluorescência do lncRNA *Menhir* são considerados limites mínimos para a sensibilidade do método de incPRINT; complexos formados pela proteína GFP e/ou pela cauda 10XMS2 são os controles negativos de interação RNA-proteína; resultados obtidos por uma única replicata.

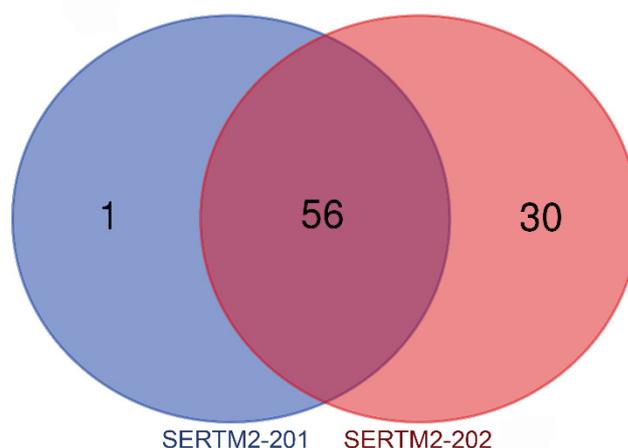
#### 5.4.3 Análise *in silico* de interações de *CARDEL* com RBPs

O programa RBPmap revelou sítios de interação potenciais para RBPs tanto para o transcrito curto (*SERTM2-201*) quanto para o transcrito longo (*SERTM2-202*)

de *CARDEL*, resultando num total de 57 proteínas candidatas para *SERTM2-201* e 86 para o *SERTM2-202*. Como as 2 variantes possuem sequências em comum, grande parte dos pontos de interação de *SERTM2-201* mostraram-se sobrepostos aos de *SERTM2-202*, e somente uma proteína (*TRA2A*) apresentou interação exclusiva com o transcrito curto (Fig. 15).

As proteínas levantadas pelo RBPmap se tornaram dados de entrada para o programa RPISeq, que retornou os indicadores de probabilidade RF (*Random Forest*) e SVM (*Support Vector Machine*). A lista de proteínas investigadas está contida na Tabela 3.

FIGURA 15 – DIAGRAMA DE VENN DE RBPs COM POTENCIAL DE LIGAÇÃO PARA *CARDEL* *SERTM2-201* E *SERTM2-202*



NOTA: *SERTM2-201* corresponde à variante curta de *CARDEL* e *SERTM2-202*, à variante longa. Detalhes sobre as RBPs selecionadas por RBPmap estão listadas na Tabela 3. Uma proteína é exclusiva a *SERTM2-201*: *TRA2A*; 30 são exclusivas a *SERTM2-202*: *CELF6*, *HNRH2*, *HNRPL*, *MSI1H*, *RBMS1*, *RFOX1*, *SFPQ*, *HNRPK*, *HNRPM*, *NOVA1*, *RBM28*, *RBM38*, *SRSF6*, *ANKH1*, *FMR1*, *FXR1*, *FXR2*, *HNRPU*, *IF2B2*, *IF2B3*, *KHDR1*, *MATR3*, *PCBP1*, *PCBP2*, *RBM3*, *RBMS3*, *ENOX1*, *HNRPF*, *RA1L2* e *HNRL*.

Durante o treinamento do algoritmo baseado em aprendizagem de máquina (*machine learning*), os desenvolvedores Muppirala et al., 2011 calcularam parâmetros de precisão, *recall* e acurácia de conjuntos de 10 predições do RPISeq, para cada proteína de 2 bancos de dados validados experimentalmente. A precisão foi medida como a parcela de interações genuinamente corretas segundo os bancos de dados (verdadeiros-positivos), o *recall* como a porcentagem de interações

verdadeiras que puderam ser detectadas pelo método e a acurácia como a proporção de predições corretas frente ao total de predições. Em ambos os bancos de dados, os 3 parâmetros calculados para o classificador *Random Forest* (RF) se mostraram mais altos do que os mensurados para o *Support Vector Machine* (SVM). Como exemplo, no banco de dados RPI2241 a acurácia do classificador RF foi de 89,6%, enquanto do SVM foi de 87,1%. Em decorrência disso, o valor do classificador RF foi utilizado como critério de exclusão para o presente trabalho.

As proteínas obtidas foram submetidas aos critérios de exclusão: Z-score menor do que 2 (EI), RF menor a 0,9 (EII) e ausência de homologia de sequência com humanos (EIII).

TABELA 3 – PROTEÍNAS LIGANTES DE RNA COM POTENCIAL DE INTERAÇÃO COM CARDEL

ID UniProtKB	Proteína	RBPmap		RPISeq	
		Nº de Sítios	Maior Z- score	Classificador RF	Classificador SVM
Q9H9Z2	LN28A	5	3.203	0.95	0.971
Q96J87*	<b>CELF6*</b>	2	1.773	0.90	0.992
Q96EP5	DAZP1	24	3.896	0.90	0.988
P55795	<b>HNRH2</b>	1	2.733	0.90	0.990
P14866	<b>HNRPL</b>	76	4.893	0.90	0.984
Q5VWX1	KHDR2	63	3.529	0.90	0.987
O43347	<b>MSI1H</b>	20	3.242	0.90	0.986
Q86U42	PABP2	57	3.250	0.90	0.967
Q13310	PABP4	118	4.027	0.90	0.993
Q9UKM9	RALY	64	3.491	0.90	0.973
P29558	<b>RBMS1</b>	65	4.452	0.90	0.992
Q9NWB1	<b>RFOX1</b>	4	2.974	0.90	0.987
P23246	<b>SFPQ*</b>	1	1.759	0.90	0.994
Q9UPU9	SMAG1	2	2.778	0.90	0.996
Q07955	SRSF1	54	3.333	0.90	0.984
Q01130	SRSF2	55	3.080	0.90	0.978
P67809	YBOX1	64	3.740	0.90	0.974
Q9Y2T7	YBOX2	28	3.595	0.90	0.979
Q9NQ94	A1CF	52	3.711	0.85	0.996
O95628	CNOT4	31	3.707	0.85	0.995
Q7Z5Q1	CPEB2	53	3.786	0.85	0.997
P07910	HNRPC	92	3.509	0.85	0.953
P61978	<b>HNRPK</b>	17	3.522	0.85	0.993

P52272	<b>HNRPM</b>	10	3.762	0.85	0.977
O75525	KHDR3	110	3.330	0.85	0.990
P51513	<b>NOVA1</b>	21	3.380	0.85	0.991
P57721	PCBP3	25	3.270	0.85	0.980
P26599	PTBP1	207	3.777	0.85	0.992
Q96PU8	QKI	20	3.054	0.85	0.995
Q9BX46	RBM24	15	2.039	0.85	0.975
Q9NW13	<b>RBM28</b>	4	3.292	0.85	0.995
Q9H0Z9	<b>RBM38</b>	8	3.128	0.85	0.991
Q9BTD8	RBM42	39	3.405	0.85	0.968
P09651	ROA1	19	3.541	0.85	0.989
P08621	RU17	15	4.233	0.85	0.979
O75494	SRS10	31	4.169	0.85	0.981
Q13243	SRSF5	83	3.258	0.85	0.984
Q13247	<b>SRSF6</b>	2	2.596	0.85	0.980
Q13148	TADBP	70	3.495	0.85	0.993
P26368	U2AF2	68	3.390	0.85	0.996
Q8IWZ3	<b>ANKH1</b>	5	2.768	0.80	0.996
Q9BZC1	CELF4	25	3.892	0.80	0.994
Q8N6W0	CELF5	32	3.987	0.80	0.996
Q17RY0	CPEB4	82	3.574	0.80	0.997
Q15717	ELAV1	44	3.549	0.80	0.979
Q06787	<b>FMR1</b>	5	2.250	0.80	0.983
P51114	<b>FXR1</b>	3	2.443	0.80	0.992
P51116*	<b>FXR2*</b>	2	1.986	0.80	0.991
O60812	HNRC1	92	3.509	0.80	0.918
P31943	HNRH1	9	2.240	0.80	0.989
Q00839	<b>HNRPU</b>	11	3.312	0.80	0.990
Q9Y6M1	<b>IF2B2</b>	41	4.041	0.80	0.992
O00425	<b>IF2B3</b>	53	4.000	0.80	0.989
Q07666	<b>KHDR1</b>	47	3.680	0.80	0.988
P43243	<b>MATR3</b>	29	3.500	0.80	0.995
Q9NR56	MBNL1	135	4.013	0.80	0.992
P11940	PABP1	86	3.725	0.80	0.996
Q96DU9	PABP5	61	3.330	0.80	0.985
Q15365	<b>PCBP1</b>	16	3.623	0.80	0.988
Q15366	<b>PCBP2</b>	18	3.713	0.80	0.990
P98179	<b>RBM3</b>	8	2.772	0.80	0.990
Q8TBY0	RBM46	26	2.934	0.80	0.994
Q6XE24	<b>RBMS3</b>	99	4.583	0.80	0.984

P22626	ROA2	14	2.908	0.80	0.987
P09012	SNRPA	16	2.958	0.80	0.985
P84103	SRSF3	521	3.765	0.80	0.988
Q16629	SRSF7	28	3.412	0.80	0.984
Q13242	SRSF9	9	2.462	0.80	0.981
P62995	TRA2B	175	3.516	0.80	0.984
Q61464**	ZN638**	18	3.969	0.80	1.000
Q92879	CELF1	47	3.354	0.75	0.989
Q8TC92	<b>ENOX1</b>	14	3.219	0.75	0.982
Q9UN86	G3BP2	13	2.667	0.75	0.992
P52597	<b>HNRPF</b>	6	3.143	0.75	0.984
Q32P51	<b>RA1L2</b>	9	3.394	0.75	0.991
P52756	RBM5	12	3.064	0.75	0.995
P78332	RBM6	27	4.549	0.75	0.990
Q9H6E5	STPAP	18	3.574	0.75	0.986
P31483	TIA1	70	3.461	0.75	0.994
Q6PJT7	ZC3HE	57	3.711	0.75	0.996
Q8TBF4	ZCRB1	61	3.634	0.75	0.980
Q8WVV9	<b>HNRL</b>	40	5.228	0.70	0.992
Q9H361	PABP3	58	3.807	0.70	0.996
Q8TB72	PUM2	62	4.034	0.70	0.991
Q96IZ5	RBM41	42	3.790	0.70	0.982
Q15020	SART3	86	3.735	0.70	0.978
Q13595***	TRA2A***	1	1.739	0.75	0.870

NOTA: Número de interações se refere à variante longa de *CARDEL* (*SERTM2-202*), com exceção a TRA2A. As RBPs em negrito tiveram interação predita apenas com *CARDEL SERTM2-202*. (\*) indicam proteínas excluídas por apresentarem sítio de ligação a *CARDEL* com Z-score menor do que 2; (\*\*) indica proteína murina; (\*\*\*) indica proteína com ligação exclusiva à variante curta de *CARDEL*, excluída pelo critério do Z-score.

As 16 proteínas remanescentes foram melhor investigadas em nível funcional/molecular pela presença de palavras-chave relacionadas à cardiomiogênese ou, em específico, à regulação da tradução, como resumido na Tabela 4. Apenas 7 proteínas atenderam a esse critério de inclusão: LN28A, DAZP1, HNRPL, MSI1, RBFOX1, SMAG1/SMAUG1 e YBOX1.

De acordo com a literatura, a proteína LIN28 é relacionada à manutenção da pluripotência, sendo inclusive utilizada para a geração de células-tronco pluripotentes induzidas (YU et al., 2007). Todavia, uma isoforma de LIN28A também regula processos como a osteogênese, via interação com o lncRNA TUG1 (HE et al.,

2018) e a diferenciação *in vitro* de progenitores cardíacos, ao interagir com o mRNA da proteína TNFR2 (XIANG et al., 2019). Das proteínas candidatas, a LN28A é a única que poderia ser analisada por incPRINT.

DAZAP1 (*DAZ-Associated Protein 1*) é mais bem estudada no contexto da espermatogênese e geração de folhetos embrionários. Alguns mecanismos de regulação da tradução consistem na interação dessa RBP com a porção 3'-UTR de mRNAs seguida pelo recrutamento do fator de iniciação de tradução eIF4G (SMITH et al., 2011) e, mais recentemente, na inibição do *splicing* do pré-mRNA COX6C, tendo como resultado a manutenção do crescimento celular (SASAKI et al., 2018).

A Hnrp-I (*Heterogeneous Nuclear Ribonucleoprotein L*) é um fator de *splicing* e uma proteína transportadora de mRNAs que parece estar associada ao desenvolvimento do músculo esquelético em moscas *Drosophila* e peixes-zebra, sendo seu nocaute suficiente para provocar más-formações e morte precoce em peixes-zebra (ALEXANDER et al. , 2021). Zhao et al. (2019) indicaram que essa RBP pode formar complexos com RNAs não codificantes cuja expressão é estimulada por pelo fator de transcrição MyoD, preponderante durante a diferenciação de células de músculo esquelético. Esses complexos interagem com genes relevantes à diferenciação e recrutam a RNA polimerase II para iniciar sua transcrição.

MSI1 (Musashi1) é parte fundamental do processo de neurogênese, promovendo a proliferação e autorrenovação de progenitores neurais ao inibir uma série de mRNAs-alvo. O controle desse processo aparenta ser refinado pela presença de proporções variáveis das formas ativas – ou fosforiladas - e não ativas de Musashi1 (MACNICOL et al., 2015). A proliferação celular exacerbada em células-tronco tumorais e células de diversos cânceres, como o glioblastoma e o carcinoma de células escamosas, muitas vezes estão correlacionados a maior expressão de Musashi1 ativa (revisado por FOROUZANFAR et al., 2020).

A RBFOX1 é uma proteína nuclear reguladora de *splicing* expressa em abundância na medula óssea, coração e próstata humanos (FAGERBERG et al., 2014). Atua na regulação da contração e homeostase de músculos esqueléticos

(PEDROTTI et al., 2015) e na integridade do coração (GAO et al., 2016). Essa RBP também possui uma isoforma citoplasmática, que em neurônios murinos interage com a porção 3'-UTR de mRNAs relacionados ao desenvolvimento do córtex cerebral, elevando sua taxa de tradução (LEE et al., 2016). O banco de dados Ensembl caracteriza a RBFOX1 como potencial regulador na diferenciação de células musculares esqueléticas.

A proteína Smaug1 é conhecida por reprimir a tradução e degradar RNAs e está presente em grânulos de estresse em várias espécies de mamíferos, incluindo humanos. Sua atuação como inibidor de tradução em mamíferos foi primeiramente descrita por Baez; Boccaccio (2005), através de experimentos em linhagens de células humanas. Mais recentemente, um mecanismo de inibição de tradução via Smaug1 humana foi proposto, no qual a Smaug1 sequestraria o fator de iniciação de tradução eIF4E (FERNÁNDEZ-ALVAREZ et al., 2016).

Por fim, a YBOX1 (*Y-Box Binding Protein 1*) é uma proteína multifuncional capaz de interferir na tradução ao se associar a polissomos, interagindo com as 5'-UTR de mRNAs, caso do mRNA para a G3BP1 (proteína de grânulos de estresse), cuja tradução aumenta quando em interação com a YBOX1 (SOMASEKHARAN et al., 2015). Além de mRNA, sabe-se que a YBOX1 é capaz de se ligar a lncRNAs, como o *CAR10*, evitando sua própria degradação via proteossoma em células de câncer de pulmão (WEI; ZHOU, 2016), e como o *TP53TG1*, que inibe sua atividade ao impedir sua translocação para o núcleo em células de câncer gastrointestinal (DIAZ-LAGARES et al., 2016).

Considerando esse corpo de literatura, as proteínas selecionadas *in silico* seriam boas candidatas para experimentos exploratórios que pudessem confirmar as interações RBP-CARDEL preditas, com níveis variáveis de potencial. LIN28A, HNRPL e YBOX1 parecem ser as RBPs mais promissoras, devido a sua associação já descrita com lncRNAs visando o controle do processo de tradução. RBFOX1 também se destaca, devido aos papéis que desempenha em células musculares cardíacas e esqueléticas.

TABELA 4 – INFORMAÇÕES FUNCIONAIS DE PROTEÍNAS SELECIONADAS APÓS APLICAÇÃO DOS CRITÉRIOS DE EXCLUSÃO

ID UniProtKB	Proteína	Localização Subcelular	Funções Biológicas
Q9H9Z2*	LN28A*	núcleo; polissomos; corpos-P; grânulos de estresse; RE	glial cell differentiation (-); neuron cell differentiation (+); proliferation in kidney development; stem cell differentiation; TOR signalling (+)
Q96EP5*	DAZP1*	núcleo; <b>citoso</b>	cell differentiation; cell proliferation; maternal placenta development; spermatogenesis; spliceosome
P14866*	HNRPL*	núcleo; grânulos no citosol	RNA degradation by "poison exon"; spliceosome; translation
O43347*	MSI1H*	núcleo; citosol	mRNA translation interference; nervous system development
Q9NWB1*	RFOX1*	núcleo; grânulos de estresse; rede trans-Golgi	alternative splicing; erythropoiesis; nervous system/muscle development; spliceosome
Q9JUP9*	SMAG1*	núcleo; junção celular; corpos-P	regulation of RNA translation (+); translational repressor
P67809*	YBOX1*	citoplasma, em células proliferativas	DNA repair and transcription regulation; embryonic morphogenesis; epidermis development; Notch pathway; mRNA splicing; RNA stabilization; translational repression
Q9Y2T7	YBOX2	núcleo; citosol	germ layer expression pattern; messenger ribonucleoprotein particles (mRNPs); thermogenesis
P55795	HNRH2	núcleo	metabolic processes; pre mRNA substrate
Q5VWX1	KHDR2	núcleo	regulation of inclusion/deletion of exons; spliceosome
Q86U42	PABP2	<b>núcleo</b> ; citosol	degradation of RNA (-); MAPK cascade; muscle contraction; nonsense degradation of RNA; poly(A), 3'-UTR binding
Q13310	PABP4	núcleo; grânulos de estresse	blood coagulation; myeloid cell development; poly(A) binding; RNA processing; translation
Q9UKM9	RALY	núcleo	LeXis lncRNA-mediated cholesterologenesis; splicing
P29558	RBMS1	núcleo; citosol	DNA replication; RNA processing
Q07955	SRSF1	núcleo (nuclear speckles)	mRNA export; RNA (alternative) splicing
Q01130	SRSF2	núcleo (nuclear speckles); nucleoplasma	response to vitamin E; mitotic cycle; splicing; RNA export

NOTA: (\*) indicam proteínas selecionadas pela busca de palavras-chave que sugerissem funções relacionadas ao coração ou tecido muscular, à tradução, diferenciação celular ou ao desenvolvimento embrionário. Na coluna "Localização Subcelular", os termos em negrito indicam a localização subcelular predominante da proteína analisada, segundo o banco de dados UniProt. Na coluna "Funções Biológicas", (+) significa uma regulação positiva do processo biológico descrito, enquanto (-) corresponde a uma regulação negativa. Corpos-P é uma abreviatura de corpos de processamento; RE é uma abreviatura de retículo endoplasmático

## 6 CONSIDERAÇÕES FINAIS E PERSPECTIVAS

A associação do transcrito *CARDEL* a polissomos durante a diferenciação cardíaca foi investigada considerando 2 alternativas: que *CARDEL* seria traduzido em um micropeptídeo e/ou que formaria complexos com RBPs para regular a expressão de mRNAs. O potencial codificante de um sORF de *CARDEL* foi fortalecido pela tradução do peptídeo correspondente SERTM2 a partir de um vetor de expressão, motivando estudos futuros que possam indicar sua localização celular e atestar a expressão endógena desse peptídeo durante a diferenciação cardíaca.

Como a localização celular está altamente correlacionada à função de uma proteína, um próximo passo natural para o estudo de SERTM2 seria a análise da proteína de fusão em microscopia de fluorescência confocal, capaz de fornecer uma visão tridimensional de células e de biomoléculas marcadas pela aquisição de imagens em diferentes planos focais (CONCHELLO; LICHTMAN, 2005).

Já a detecção endógena de SERTM2 durante a diferenciação poderia ser feita via anticorpos específicos, embora sua produção não seja trivial, devido ao tamanho reduzido do antígeno. Apesar disso, existem experimentos bem-sucedidos utilizando anticorpos contra micropeptídeos, como o conduzido por Rossi et al. (2019), que puderam avaliar a expressão do micropeptídeo codificado pelo lncRNA *EPR* por *Western Blot* em células murinas imortalizadas NMuMG.

Outra abordagem provável, devido ao conhecimento técnico disponível internamente, seria a inserção de uma etiqueta a jusante do sORF-1 ou sORF-2 pelo sistema CRISPR/Cas9, de maneira similar aos trabalhos de Anderson et al. (2015), com a fusão de uma etiqueta *FLAG* ao sORF do micropeptídeo miorregulina, e Wang et al. (2020), com a fusão de etiquetas *FLAG* e *GFP* ao micropeptídeo codificado pelo transcrito *ASRPS*. Dessa maneira, a expressão endógena do peptídeo poderia ser monitorada com o uso de anticorpos comuns e padronizados.

A geração e caracterização da linhagem PLZ *CARDEL* mut abre oportunidades para a avaliação dos efeitos da superexpressão do transcrito de *CARDEL* na diferenciação cardíaca em monocamada, utilizando pequenas

moléculas moduladoras da via de sinalização Wnt (adaptado de LIAN et al., 2013 e WANG et al., 2011). A PLZ iCARDEL mut C8.1 já está sendo empregada em experimentos de superexpressão que objetivam investigar o efeito do transcrito de *CARDEL*, sem potencial codificante, na diferenciação cardíaca *in vitro*, avaliando parâmetros como frequência de batimentos dos cardiomiócitos diferenciados a partir da linhagem e o nível de expressão do marcador cardíaco TNNT2, codificante para a proteína troponina T cardíaca. Esses testes preliminares de diferenciação cardíaca indicam que a linhagem é capaz de diferenciar para células contráteis e/ou troponina T cardíaca positivas, em condições com e sem superexpressão induzida por doxiciclina (dados não mostrados).

Uma vez observado o efeito da superexpressão do transcrito de *CARDEL* sem potencial codificante (*CARDEL mut*), a localização e o montante de expressão de *CARDEL* podem ser esmiuçados por técnicas como o RNA-FISH (*RNA-fluorescence in situ hybridization*) que envolve a hibridização de sondas fluorescentes ao transcrito (revisado por YOUNG et al., 2020). Como alternativa ao RNA-FISH, há uma técnica emergente que facilita a visualização de RNA *in vivo* e é baseada na fusão do transcrito de interesse a sequências repetitivas de RNA, denominadas aptâmeros. Esses pequenos RNAs formam estruturas secundárias capazes de interagir com moléculas pouco tóxicas, chamadas fluorocromos, ampliando sua fluorescência em várias ordens de grandeza. Dessa forma, os aptâmeros atuam como etiquetas para RNAs de interesse (revisado por NEUBACHER; HENNIG, 2019).

Por fim, o estudo *in silico* do potencial de interação do *CARDEL* com RBPs conseguiu filtrar centenas de proteínas em uma lista consistente de candidatas, de maneira RNA-cêntrica e utilizando um *pipeline* sistemático e passível de ser aplicado, com mínimas alterações, ao estudo de outros lncRNAs em outros modelos de diferenciação *in vitro*.

Considerando que a técnica de incPRINT apresenta alto rendimento de detecção de interações *CARDEL*-proteína em sistemas *in vivo*, medidas para o ajuste do protocolo para a detecção de RNPs de *CARDEL* poderiam ser

implementadas. A escolha mais simples parece ser a geração de outra construção com *CARDEL-10xMS2*, preservando o íntron localizado entre os éxons 1 e 2 do pré-RNA de *CARDEL*, com comprimento aproximado de 160 pb.

A associação do lncRNA a RBPs poderia ser investigada pela técnica de EMSA (*Electrophoretic Mobility Shift Assay*), na qual a diferença de mobilidade de determinado RNA em gel de poliacrilamida é indicativa de interação com proteínas. A natureza das proteínas integrantes da RNP poderia ser confirmada, a partir de uma lista de candidatas, pela imunomarcação das bandas alteradas (HELLMAN; FRIED, 2007).

Uma alternativa laboriosa, mas robusta de busca por RNPs de *CARDEL* seria a co-imunoprecipitação de proteínas candidatas a interagir com *CARDEL* (seção 5.4.3) com RNA (GAGLIARDI; MATARAZZO, 2016). De maneira inversa, RNAs poderiam sofrer marcação por proximidade com cada proteína candidata fusionada a uma enzima biotinizase, como a TurboID (BRANON et al., 2018) ou a AirID (KIDO et al., 2020), podendo ser capturados diretamente por uma coluna de adsorção de estreptavidina.

Apesar do mecanismo de regulação pós-transcricional de *CARDEL* não ter sido definitivamente elucidado, o presente trabalho conseguiu fortalecer hipóteses e desenvolver ferramentas para estudos mais aprofundados sobre o papel desse transcrito na cardiomiogênese.

## REFERÊNCIAS

ABAANDOU, L.; QUAN, D.; SHILOACH, J. Affecting HEK293 Cell Growth and Production Performance by Modifying the Expression of Specific Genes. **Cells** **2021**, Vol. 10, Page 1667, v. 10, n. 7, p. 1667, 2021. Multidisciplinary Digital Publishing Institute. Disponível em: <<https://www.mdpi.com/2073-4409/10/7/1667/htm>>. Acesso em: 12/7/2022.

ALBERTS, B.; JOHNSON, A.; LEWIS, J.; et al. **Biologia molecular da célula**. 2017.

ALEXANDER, M. S.; HIGHTOWER, R. M.; REID, A. L.; et al. hnRNP L is essential for myogenic differentiation and modulates myotonic dystrophy pathologies. **Muscle & nerve**, v. 63, n. 6, p. 928–940, 2021. NIH Public Access. Disponível em: <[pmc/articles/PMC8131270/](https://pubmed.ncbi.nlm.nih.gov/3511270/)>. Acesso em: 29/11/2022.

AMARAL, P. P.; CLARK, M. B.; GASCOIGNE, D. K.; DINGER, M. E.; MATTICK, J. S. lncRNADB: a reference database for long noncoding RNAs. **Nucleic Acids Research**, v. 39, p. D146–D151, 2011. Disponível em: <<http://www.lncrnadb.org/>>. Acesso em: 1/4/2020.

AMES, E. G.; LAWSON, M. J.; MACKAY, A. J.; HOLMES, J. W. Sequencing of mRNA identifies re-expression of fetal splice variants in cardiac hypertrophy. **Journal of molecular and cellular cardiology**, v. 62, p. 99, 2013. NIH Public Access. Disponível em: <[pmc/articles/PMC3735658/](https://pubmed.ncbi.nlm.nih.gov/23735658/)>. Acesso em: 21/8/2022.

ANDERSEN, C. R.; NIELSEN, L. S.; BAER, A.; TOLSTRUP, A. B.; WEILGUNY, D. Efficient expression from one CMV enhancer controlling two core promoters. **Molecular Biotechnology**, v. 48, n. 2, p. 128–137, 2011. Springer. Disponível em: <<https://link.springer.com/article/10.1007/s12033-010-9353-7>>. Acesso em: 16/8/2022.

ANDERSON, D.; ANDERSON, K.; CHANG, C.; et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. **Cell**, v. 160, n. 4, p. 595–606, 2015. Cell Press. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/25640239/>>. Acesso em: 6/4/2020.

BAEZ, M. V; BOCCACCIO, G. L. Mammalian Smaug Is a Translational Repressor That Forms Cytoplasmic Foci Similar to Stress Granules \* □ S. , 2005. JBC Papers in Press. Disponível em: <<http://www.jbc.org/>>. Acesso em: 21/10/2022.

BALCERAK, A.; TREBINSKA-STRYJEWSKA, A.; KONOPINSKI, R.; WAKULA, M.; GRZYBOWSKA, E. A. RNA-protein interactions: Disorder, moonlighting and junk contribute to eukaryotic complexity. **Open Biology**, v. 9, n. 6, 2019. Royal Society Publishing. Disponível em: <[pmc/articles/PMC6597761/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/3511270/)>. Acesso em: 7/10/2020.

BATEMAN, A.; MARTIN, M. J.; ORCHARD, S.; et al. UniProt: the universal protein knowledgebase in 2021. **Nucleic Acids Research**, v. 49, n. D1, p. D480–D489, 2021. Oxford Academic. Disponível em:

<<https://academic.oup.com/nar/article/49/D1/D480/6006196>>. Acesso em: 12/1/2022.

BEQQALI, A.; VAN ELDIK, W.; MUMMERY, C.; PASSIER, R. Human stem cells as a model for cardiac differentiation and disease. **Cellular and Molecular Life Sciences**, 2009.

BERTERO, A.; FIELDS, P. A.; RAMANI, V.; et al. Dynamics of genome reorganization during human cardiogenesis reveal an RBM20-dependent splicing factory. **Nature Communications**, v. 10, n. 1, 2019. Nature Publishing Group.

BLAESER, F.; HO, N.; PRYWES, R.; CHATILA, T. A. Ca<sup>2+</sup>-dependent gene expression mediated by MEF2 transcription factors. **Journal of Biological Chemistry**, v. 275, n. 1, p. 197–209, 2000. Elsevier. Disponível em: <<http://www.jbc.org/article/S0021925819528526/fulltext>>. Acesso em: 21/8/2022.

BLECH-HERMONI, Y.; LADD, A. N. RNA binding proteins in the regulation of heart development. , 2013.

BOOY, E. P.; GUSSAKOVSKY, D.; CHOI, T.; MCKENNA, S. A. The noncoding RNA BC200 associates with polysomes to positively regulate mRNA translation in tumor cells. **Journal of Biological Chemistry**, v. 296, p. 100036, 2021. Elsevier B.V. Disponível em: <<https://doi.org/10.1074/jbc.RA120.015775>>. .

BOSHART, M.; WEBER, F.; JAHN, G.; et al. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. **Cell**, v. 41, n. 2, p. 521–530, 1985. Elsevier. Disponível em: <<http://www.cell.com/article/S0092867485800258/fulltext>>. Acesso em: 16/8/2022.

BRADÉ, T.; PANE, L. S.; MORETTI, A.; CHIEN, K. R.; LAUGWITZ, K. L. Embryonic heart progenitors and cardiogenesis. **Cold Spring Harbor Perspectives in Medicine**, 2013. Cold Spring Harbor Laboratory Press.

BRANON, T. C.; BOSCH, J. A.; SANCHEZ, A. D.; et al. Efficient proximity labeling in living cells and organisms with TurboID. **Nature Biotechnology** 2018 **36:9**, v. 36, n. 9, p. 880–887, 2018. Nature Publishing Group. Disponível em: <<https://www.nature.com/articles/nbt.4201>>. Acesso em: 16/7/2022.

CAI, R.; SUN, Y.; QIMUGE, N.; et al. Adiponectin AS lncRNA inhibits adipogenesis by transferring from nucleus to cytoplasm and attenuating Adiponectin mRNA translation. **Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids**, v. 1863, n. 4, p. 420–432, 2018. Elsevier B.V.

CARAPUÇA, E.; AZZONI, A. R.; PRAZERES, D. M. F.; MONTEIRO, G. A.; MERGULHÃO, F. J. M. Time-course determination of plasmid content in eukaryotic and prokaryotic cells using Real-Time PCR. **Molecular Biotechnology**, v. 37, n. 2, p. 120–126, 2007. Springer. Disponível em: <<https://link.springer.com/article/10.1007/s12033-007-0007-3>>. Acesso em: 24/11/2022.

CHAHAL, G.; TYAGI, S.; RAMIALISON, M. Navigating the non-coding genome in heart development and Congenital Heart Disease. **Differentiation**, 1. maio 2019. Elsevier Ltd.

CHENDRIMADA, T. P.; FINN, K. J.; JI, X.; et al. MicroRNA silencing through RISC recruitment of eIF6. **Nature**, v. 447, n. 7146, p. 823–828, 2007. Nature Publishing Group.

CHUGUNOVA, A.; LOSEVA, E.; MAZIN, P.; et al. LINC00116 codes for a mitochondrial peptide linking respiration and lipid metabolism. **Proceedings of the National Academy of Sciences of the United States of America**, v. 116, n. 11, p. 4940–4945, 2019. National Academy of Sciences. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/30796188/>>. Acesso em: 2/2/2021.

CLÉRY, A.; ALLAIN, F. H. T. From Structure to Function of RNA Binding Domains. **RNA Binding Proteins**. 2º ed, p.137–152, 2018. CRC Press.

CONCHELLO, J. A.; LICHTMAN, J. W. Optical sectioning microscopy. **Nature methods**, v. 2, n. 12, p. 920–931, 2005. Nat Methods. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/16299477/>>. Acesso em: 19/8/2022.

CORLEY, M.; BURNS, M. C.; YEO, G. W. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. **Molecular Cell**, v. 78, n. 1, p. 9–29, 2020. Elsevier Inc. Disponível em: <<https://doi.org/10.1016/j.molcel.2020.03.011>>. .

COSTELLO, I.; PIMEISL, I. M.; DRÄGER, S.; et al. The T-box transcription factor Eomesodermin acts upstream of Mesp1 to specify cardiac mesoderm during mouse gastrulation. **Nature Cell Biology**, v. 13, n. 9, p. 1084–1092, 2011. Nat Cell Biol.

CUI, M.; WANG, Z.; BASSEL-DUBY, R.; OLSON, E. N. Genetic and epigenetic regulation of cardiomyocytes in development, regeneration and disease. **Development (Cambridge)**, 15. dez. 2018. Company of Biologists Ltd.

CUNNINGHAM, F.; ALLEN, J. E.; ALLEN, J.; et al. Ensembl 2022. **Nucleic Acids Research**, v. 50, n. D1, p. D988–D995, 2022. Oxford Academic. Disponível em: <<https://academic.oup.com/nar/article/50/D1/D988/6430486>>. Acesso em: 23/11/2022.

DAS, A.; TENENBAUM, L.; BERKHOUT, B. Tet-On Systems For Doxycycline-inducible Gene Expression. **Current Gene Therapy**, v. 16, n. 3, p. 156–167, 2016. Bentham Science Publishers Ltd.

DHAMIJA, S.; MENON, M. Non-coding transcript variants of protein-coding genes - what are they good for? **RNA biology**, v. 15, n. 8, p. 1025–1031, 2018. RNA Biol. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/30146915/>>. Acesso em: 19/7/2021.

DIAZ-LAGARES, A.; CRUJEIRAS, A. B.; LOPEZ-SERRA, P.; et al. Epigenetic inactivation of the p53-induced long noncoding RNA TP53 target 1 in human cancer. **Proceedings of the National Academy of Sciences of the United States of America**, v. 113, n. 47, p. E7535–E7544, 2016. National Academy of Sciences. Disponível em: <<https://www.pnas.org/content/113/47/E7535>>. Acesso em: 25/1/2022.

DUVAUD, S.; GABELLA, C.; LISACEK, F.; et al. Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. **Nucleic acids research**, v.

49, n. W1, p. W216–W227, 2021. *Nucleic Acids Res.* Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/33849055/>>. Acesso em: 15/8/2022.

ELGUINDY, M. M.; MENDELL, J. T. NORAD-induced Pumilio phase separation is required for genome stability. **Nature** **2021** **595:7866**, v. 595, n. 7866, p. 303–308, 2021. Nature Publishing Group. Disponível em: <<https://www.nature.com/articles/s41586-021-03633-w>>. Acesso em: 27/1/2022.

FAGERBERG, L.; HALLSTROM, B. M.; OKSVOLD, P.; et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. **Molecular and Cellular Proteomics**, v. 13, n. 2, p. 397–406, 2014. *Mol Cell Proteomics*. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/24309898/>>. Acesso em: 23/3/2020.

FERNÁNDEZ-ALVAREZ, A. J.; PASCUAL, M. L.; BOCCACCIO, G. L.; THOMAS, M. G. Smaug variants in neural and non-neuronal cells. <http://www.tandfonline.com/action/authorSubmission?journalCode=kcib20&page=instructions>, v. 9, n. 2, p. 1–8, 2016. Taylor & Francis. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/19420889.2016.1139252>>. Acesso em: 21/10/2022.

FLORES, J. K.; ATAIDE, S. F. Structural changes of RNA in complex with proteins in the SRP. **Frontiers in Molecular Biosciences**, v. 5, n. FEB, p. 7, 2018. Frontiers Media S.A.

FOROUZANFAR, M.; LACHINANI, L.; DORMIANI, K.; et al. Intracellular functions of RNA-binding protein, Musashi1, in stem and cancer cells. **Stem Cell Research & Therapy**, v. 11, n. 1, 2020. BioMed Central. Disponível em: <[pmc/articles/PMC7245930/](https://pmc/articles/PMC7245930/)>. Acesso em: 29/11/2022.

GAGLIARDI, M.; MATARAZZO, M. R. Chapter 7. , v. 1480, p. 73–86, 2016.

GAO, C.; REN, S.; LEE, J. H.; et al. RBFox1-mediated RNA splicing regulates cardiac hypertrophy and heart failure. **The Journal of clinical investigation**, v. 126, n. 1, p. 195–206, 2016. *J Clin Invest*. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/26619120/>>. Acesso em: 25/1/2022.

GARATE, X.; LA GRECA, A.; NEIMAN, G.; et al. Identification of the miRNAome of early mesoderm progenitor cells and cardiomyocytes derived from human pluripotent stem cells. **Scientific Reports**, v. 8, n. 1, p. 1–14, 2018.

GARCÍA-PADILLA, C.; DOMÍNGUEZ, J. N.; ARÁNEGA, A. E.; FRANCO, D. Differential chamber-specific expression and regulation of long non-coding RNAs during cardiac development. **Biochimica et Biophysica Acta - Gene Regulatory Mechanisms**, v. 1862, n. 10, p. 194435, 2019. Elsevier. Disponível em: <<https://doi.org/10.1016/j.bbagr.2019.194435>>. .

GAUDET, P.; LIVSTONE, M. S.; LEWIS, S. E.; THOMAS, P. D. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. **Briefings in bioinformatics**, v. 12, n. 5, p. 449–462, 2011. *Brief Bioinform*. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/21873635/>>. Acesso em: 31/7/2022.

GEORGE, R. M.; FIRULLI, A. B. Hand Factors in Cardiac Development. **Anatomical record (Hoboken, N.J. : 2007)**, v. 302, n. 1, p. 101–107, 2019. Anat Rec (Hoboken). Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/30288953/>>. Acesso em: 21/8/2022.

GONZALEZ, I.; MUNITA, R.; AGIRRE, E.; et al. A lncRNA regulates alternative splicing via establishment of a splicing-specific chromatin signature. **Nature Structural and Molecular Biology**, v. 22, n. 5, p. 370–376, 2015. Nature Publishing Group.

GRAINDORGE, A.; PINHEIRO, I.; NAWROCKA, A.; et al. In-cell identification and measurement of RNA-protein interactions. **Nature Communications**, v. 10, n. 1, p. 1–11, 2019.

GROTE, P.; WITTLER, L.; HENDRIX, D.; et al. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. **Developmental Cell**, v. 24, n. 2, p. 206–214, 2013. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC23369715/?tool=EBI>>. Acesso em: 19/8/2022.

GUO, X.; XU, Y.; WANG, Z.; et al. A Linc1405/Eomes Complex Promotes Cardiac Mesoderm Specification and Cardiogenesis. **Cell Stem Cell**, v. 22, n. 6, p. 893–908.e6, 2018. Cell Stem Cell. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/29754779/>>. Acesso em: 26/7/2021.

HE, Q.; YANG, S.; GU, X.; et al. Long noncoding RNA TUG1 facilitates osteogenic differentiation of periodontal ligament stem cells via interacting with Lin28A. **Cell Death and Disease**, v. 9, n. 5, 2018. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/29674645/>>. Acesso em: 22/3/2021.

HELLMAN, L. M.; FRIED, M. G. Electrophoretic Mobility Shift Assay (EMSA) for Detecting Protein-Nucleic Acid Interactions. **Nature protocols**, v. 2, n. 8, p. 1849, 2007. NIH Public Access. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/1677439/>>. Acesso em: 16/7/2022.

HILGARTH, R. S.; LANIGAN, T. M. Optimization of overlap extension PCR for efficient transgene construction. **MethodsX**, v. 7, p. 100759, 2020. Elsevier.

HOCKEMEYER, D.; SOLDNER, F.; BEARD, C.; et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. **Nature Biotechnology**, v. 27, n. 9, p. 851–857, 2009. Disponível em: <<https://www.nature.com/articles/nbt.1562>>. Acesso em: 2/4/2020.

HOES, M. F.; BOMER, N.; VAN DER MEER, P. Concise Review: The Current State of Human In Vitro Cardiac Disease Modeling: A Focus on Gene Editing and Tissue Engineering. **STEM CELLS Translational Medicine**, v. 8, n. 1, p. 66–74, 2019. John Wiley and Sons Ltd. Disponível em: <<http://doi.wiley.com/10.1002/sctm.18-0052>>. Acesso em: 25/3/2020.

HUANG, Y. The novel regulatory role of lncRNA-miRNA-mRNA axis in cardiovascular diseases. **Journal of Cellular and Molecular Medicine**, 1. dez. 2018. Blackwell Publishing Inc.

IWAKAWA, H. OKI; TOMARI, Y. Life of RISC: Formation, action, and degradation of RNA-induced silencing complex. **Molecular Cell**, v. 82, n. 1, p. 30–43, 2022. Cell Press.

JUMPER, J.; EVANS, R.; PRITZEL, A.; et al. Highly accurate protein structure prediction with AlphaFold. **Nature** **2021** **596:7873**, v. 596, n. 7873, p. 583–589, 2021. Nature Publishing Group. Disponível em: <<https://www.nature.com/articles/s41586-021-03819-2>>. Acesso em: 16/7/2022.

KALLEN, A. N.; ZHOU, X. B.; XU, J.; et al. The Imprinted H19 LncRNA Antagonizes Let-7 MicroRNAs. **Molecular Cell**, v. 52, n. 1, p. 101–112, 2013. Mol Cell.

KATTMAN, S. J.; WITTY, A. D.; GAGLIARDI, M.; et al. Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. **Cell Stem Cell**, v. 8, n. 2, p. 228–240, 2011. Cell Stem Cell.

KIDO, K.; YAMANAKA, S.; NAKANO, S.; et al. Airid, a novel proximity biotinylation enzyme, for analysis of protein–protein interactions. **eLife**, v. 9, 2020. eLife Sciences Publications Ltd. Disponível em: </pmc/articles/PMC7302878/?report=abstract>. Acesso em: 18/8/2020.

KIM, N. J.; LEE, K. H.; SON, Y. S.; et al. Spatiotemporal expression of long noncoding RNA Moshe modulates heart cell lineage commitment. **RNA biology**, v. 18, n. sup2, p. 640–654, 2021. RNA Biol. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/34755591/>>. Acesso em: 19/8/2022.

KLATTENHOFF, C. A.; SCHEUERMANN, J. C.; SURFACE, L. E.; et al. Braveheart, a long non-coding RNA required for cardiovascular lineage commitment. **Cell**, v. 152, n. 3, p. 570, 2013. NIH Public Access. Disponível em: </pmc/articles/PMC3563769/>. Acesso em: 19/8/2022.

KONDO, Y.; OUBRIDGE, C.; VAN ROON, A. M. M.; NAGAI, K. Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition. **eLife**, v. 4, 2015.

KOZAK, M. The scanning model for translation: an update. **Journal of Cell Biology**, v. 108, n. 2, p. 229–241, 1989. The Rockefeller University Press. Disponível em: <<http://rupress.org/jcb/article-pdf/108/2/229/1057345/229.pdf>>. Acesso em: 31/7/2022.

LARA, D.; ALLISTER, S. ; NGUYEN, A. B. ; et al. Identification of Novel Micropeptides Derived from Hepatocellular Carcinoma-Specific Long Noncoding RNA. **International Journal of Molecular Sciences** **2022**, **Vol. 23**, **Page 58**, v. 23, n. 1, p. 58, 2021. Multidisciplinary Digital Publishing Institute. Disponível em: <<https://www.mdpi.com/1422-0067/23/1/58/htm>>. Acesso em: 16/8/2022.

LEE, J. A.; DAMIANOV, A.; LIN, C. H.; et al. Cytoplasmic Rbfox1 Regulates the Expression of Synaptic and Autism-Related Genes. **Neuron**, v. 89, n. 1, p. 113–128, 2016. Neuron. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/26687839/>>. Acesso em: 25/1/2022.

LEE, S.; KOPP, F.; CHANG, T. C.; et al. Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. **Cell**, v. 164, n. 1–2, p. 69–80, 2016. Cell Press. Disponível em: <<http://dx.doi.org/10.1016/j.cell.2015.12.017>><<http://dx.doi.org/10.1016/j.cell.2015.12.017>>. Acesso em: 22/3/2021.

LI, X. L.; PONGOR, L.; TANG, W.; et al. A small protein encoded by a putative lncRNA regulates apoptosis and tumorigenicity in human colorectal cancer cells. **eLife**, v. 9, p. 1–44, 2020. Elife. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/33112233/>>. Acesso em: 16/8/2022.

LIAN, Xiaojun; ZHANG, Jianhua; AZARIN, Samira M; et al. Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/ $\beta$ -catenin signaling under fully defined conditions. **Nature Protocols**, v. 8, n. 1, p. 162–175, 2013. Disponível em: <<https://www.nature.com/articles/nprot.2012.150>>. Acesso em: 5/6/2022.

LIAO, Y.; CASTELLO, A.; FISCHER, B.; et al. The Cardiomyocyte RNA-Binding Proteome: Links to Intermediary Metabolism and Heart Disease. **Cell Reports**, v. 16, n. 5, p. 1456–1469, 2016. Elsevier B.V. Disponível em: <<http://www.cell.com/article/S2211124716308531/fulltext>>. Acesso em: 9/12/2021.

LIU, L.; AN, X.; LI, Z.; et al. The H19 long noncoding RNA is a novel negative regulator of cardiomyocyte hypertrophy. **Cardiovascular research**, v. 111, n. 1, p. 56–65, 2016. Cardiovasc Res. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/27084844/>>. Acesso em: 22/8/2022.

LIU, N.; DAI, Q.; ZHENG, G.; et al. N6 -methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. **Nature**, v. 518, n. 7540, p. 560–564, 2015. Nature Publishing Group.

LOMBARDO, A.; CESANA, D.; GENOVESE, P.; et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. **Nature Methods**, v. 8, n. 10, p. 861–869, 2011. Nat Methods.

LYU, Y.; JIA, W.; WU, Y.; et al. Cpmer: A new conserved eEF1A2-binding partner that regulates Eomes translation and cardiomyocyte differentiation. **Stem Cell Reports**, v. 17, n. 5, p. 1154–1169, 2022. Cell Press.

MACNICOL, A. M.; HARDY, L. L.; SPENCER, H. J.; MACNICOL, M. C. Neural stem and progenitor cell fate transition requires regulation of Musashi1 function. **BMC Developmental Biology**, v. 15, n. 1, 2015. BioMed Central. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/26444444/>><<https://pubmed.ncbi.nlm.nih.gov/26444444/>>. Acesso em: 29/11/2022.

MAKAREWICH, C. A.; BASKIN, K. K.; MUNIR, A. Z.; et al. MOXI Is a Mitochondrial Micropeptide That Enhances Fatty Acid  $\beta$ -Oxidation. **Cell reports**, v. 23, n. 13, p. 3701–3709, 2018. Cell Rep. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/29949755/>>. Acesso em: 16/11/2022.

MAKAREWICH, C. A.; MUNIR, A. Z.; SCHIATTARELLA, G. G.; et al. The DWORF micropeptide enhances contractility and prevents heart failure in a mouse model of dilated cardiomyopathy. **eLife**, v. 7, 2018. eLife Sciences Publications Ltd.

MAŠEK, T.; VALÁŠEK, L.; POSPÍŠEK, M. Polysome analysis and RNA purification from sucrose gradients. **Methods in molecular biology (Clifton, N.J.)**, v. 703, p. 293–309, 2011. Methods Mol Biol.

MEILHAC, S. M.; BUCKINGHAM, M. E. The deployment of cell lineages that form the mammalian heart. **Nature Reviews Cardiology**, 1. nov. 2018. Nature Publishing Group.

MELÉ, M.; MATTIOLI, K.; MALLARD, W.; et al. Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. **Genome Research**, v. 27, n. 1, p. 27–37, 2017. Cold Spring Harbor Laboratory Press. Disponível em: <<https://genome.cshlp.org/content/27/1/27.full>>. Acesso em: 16/8/2022.

MORI, Y.; YOSHIDA, Y.; SATOH, A.; MORIYA, H. Development of an experimental method of systematically estimating protein expression limits in HEK293 cells. **Scientific Reports 2020 10:1**, v. 10, n. 1, p. 1–11, 2020. Nature Publishing Group. Disponível em: <<https://www.nature.com/articles/s41598-020-61646-3>>. Acesso em: 5/4/2022.

MUDGE, J. M.; JUNGREIS, I.; HUNT, T.; et al. Discovery of high-confidence human protein-coding genes and exons by whole-genome PhyloCSF helps elucidate 118 GWAS loci. **Genome Research**, v. 29, n. 12, p. 2073–2087, 2019. Cold Spring Harbor Laboratory Press. Disponível em: <[/pmc/articles/PMC6886504/?report=abstract](https://pmc/articles/PMC6886504/?report=abstract)>. Acesso em: 21/10/2020.

MUPPIRALA, U. K.; HONAVAR, V. G.; DOBBS, D. Predicting RNA-Protein Interactions Using Only Sequence Information. **BMC Bioinformatics**, v. 12, n. 1, p. 1–11, 2011. BioMed Central. Disponível em: <<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-489>>. Acesso em: 12/1/2022.

NAKAGAWA, S.; NIIMURA, Y.; GOJOBORI, T.; TANAKA, H.; MIURA, K. ICHIRO. Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. **Nucleic acids research**, v. 36, n. 3, p. 861–871, 2008. Nucleic Acids Res. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/18086709/>>. Acesso em: 31/7/2022.

NEUBACHER, S.; HENNIG, S. RNA Structure and Cellular Applications of Fluorescent Light-Up Aptamers. **Angewandte Chemie (International Ed. in English)**, v. 58, n. 5, p. 1266, 2019. Wiley-Blackwell. Disponível em: <[/pmc/articles/PMC6391945/](https://pmc/articles/PMC6391945/)>. Acesso em: 16/7/2022.

O'BRIEN, J.; HAYDER, H.; ZAYED, Y.; PENG, C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. **Frontiers in Endocrinology**, v. 9, n. AUG, p. 402, 2018. Frontiers Media S.A.

OCEGUERA-YANEZ, F.; KIM, S. IL; MATSUMOTO, T.; et al. Engineering the AAVS1 locus for consistent and scalable transgene expression in human iPSCs and their differentiated derivatives. **Methods**, v. 101, p. 43–55, 2016. Academic Press Inc.

OKUBO, C.; NARITA, M.; INAGAKI, A.; et al. Expression dynamics of HAND1/2 in in vitro human cardiomyocyte differentiation. **Stem Cell Reports**, v. 16, n. 8, p. 1906–1922, 2021. Cell Press. Disponível em: <<http://www.cell.com/article/S2213671121003222/fulltext>>. Acesso em: 21/8/2022.

OUNZAIN, S.; MICHELETTI, R.; ARNAN, C.; et al. CARMEN, a human super enhancer-associated long noncoding RNA controlling cardiac specification, differentiation and homeostasis. **Journal of Molecular and Cellular Cardiology**, v. 89, n. Pt A, p. 98–112, 2015. Academic Press.

PARIKH, A.; WU, J.; BLANTON, R. M.; TZANAKAKIS, E. S. Signaling Pathways and Gene Regulatory Networks in Cardiomyocyte Differentiation. **Tissue Engineering - Part B: Reviews**, 1. ago. 2015. Mary Ann Liebert Inc. Disponível em: <[pmc/articles/PMC4533091/?report=abstract](http://pmc/articles/PMC4533091/?report=abstract)>. Acesso em: 4/9/2020.

PATIL, D. P.; CHEN, C. K.; PICKERING, B. F.; et al. M6 A RNA methylation promotes XIST-mediated transcriptional repression. **Nature**, v. 537, n. 7620, p. 369–373, 2016. Nature Publishing Group.

PAZ, I.; KOSTI, I.; ARES, M.; CLINE, M.; MANDEL-GUTFREUND, Y. RBPmap: a web server for mapping binding sites of RNA-binding proteins. **Nucleic Acids Research**, v. 42, n. Web Server issue, p. W361, 2014. Oxford University Press. Disponível em: <[pmc/articles/PMC4086114/](http://pmc/articles/PMC4086114/)>. Acesso em: 31/7/2022.

PEDROTTI, S.; GIUDICE, J.; DAGNINO-ACOSTA, A.; et al. The RNA-binding protein Rbfox1 regulates splicing required for skeletal muscle structure and function. **Human molecular genetics**, v. 24, n. 8, p. 2360–2374, 2015. Hum Mol Genet. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/25575511/>>. Acesso em: 25/1/2022.

PEREIRA, I. T. **Caracterização Molecular da Regulação Pós-transcricional durante a Diferenciação Cardiomiogênica**, 2019. Fundação Oswaldo Cruz.

PHAN, D.; RASMUSSEN, T. L.; NAKAGAWA, O.; et al. BOP, a regulator of right ventricular heart development, is a direct transcriptional target of MEF2C in the developing heart. **Development**, v. 132, n. 11, p. 2669–2678, 2005. The Company of Biologists. Disponível em: <<https://journals.biologists.com/dev/article/132/11/2669/52360/BOP-a-regulator-of-right-ventricular-heart>>. Acesso em: 21/8/2022.

PHILIPPEN, L. E.; DIRKX, E.; DA COSTA-MARTINS, P. A.; DE WINDT, L. J. Non-coding RNA in control of gene regulatory programs in cardiac development and disease. **Journal of Molecular and Cellular Cardiology**, 1. dez. 2015. Academic Press.

POMERANZ KRUMMEL, D. A.; OUBRIDGE, C.; LEUNG, A. K. W.; LI, J.; NAGAI, K. Crystal structure of human spliceosomal U1 snRNP at 5.5 resolution. **Nature**, v. 458, n. 7237, p. 475–480, 2009.

REAL, G.; MONTEIRO, F.; BURGER, C.; ALVES, P. M. Improvement of lentiviral transfer vectors using cis-acting regulatory elements for increased gene expression. **Applied Microbiology and Biotechnology**, v. 91, n. 6, p. 1581–1591,

2011. Springer. Disponível em: <<https://link.springer.com/article/10.1007/s00253-011-3392-2>>. Acesso em: 16/8/2022.

ROSSI, M.; BUCCI, G.; RIZZOTTO, D.; et al. LncRNA EPR controls epithelial proliferation by coordinating Cdkn1a transcription and mRNA decay response to TGF- $\beta$ . **Nature Communications**, v. 10, n. 1, 2019. Nature Publishing Group.

SASAKI, K.; ONO, M.; TAKABE, K.; SUZUKI, A.; KURIHARA, Y. Specific intron-dependent loading of DAZAP1 onto the cox6c transcript suppresses pre-mRNA splicing efficacy and induces cell growth retardation. **Gene**, v. 657, p. 1–8, 2018. Elsevier.

SCHLACKOW, M.; NOJIMA, T.; GOMES, T.; et al. Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. **Molecular Cell**, v. 65, n. 1, p. 25–38, 2017. Cell Press.

SHI, B.; XUE, M.; WANG, YI; et al. An improved method for increasing the efficiency of gene transfection and transduction. **International Journal of Physiology, Pathophysiology and Pharmacology**, v. 10, n. 2, p. 95, 2018. e-Century Publishing Corporation. Disponível em: </pmc/articles/PMC5943608/>. Acesso em: 1/8/2022.

SKELTON, R. J. P.; KAMP, T. J.; ELLIOTT, D. A.; ARDEHALI, R. Biomarkers of Human Pluripotent Stem Cell-Derived Cardiac Lineages. **Trends in Molecular Medicine**, 1. jul. 2017. Elsevier Ltd.

SMITH, K. P.; LUONG, M. X.; STEIN, G. S. Pluripotency: Toward a gold standard for human ES and iPS cells. **Journal of Cellular Physiology**, jul. 2009. J Cell Physiol.

SMITH, R. W. P.; ANDERSON, R. C.; SMITH, J. W. S.; et al. DAZAP1, an RNA-binding protein required for development and spermatogenesis, can regulate mRNA translation. **RNA**, v. 17, n. 7, p. 1282, 2011. Cold Spring Harbor Laboratory Press. Disponível em: </pmc/articles/PMC3138565/>. Acesso em: 29/11/2022.

SOMASEKHARAN, S. P.; EL-NAGGAR, A.; LEPRIVIER, G.; et al. YB-1 regulates stress granule formation and tumor progression by translationally activating G3BP1. **Journal of Cell Biology**, v. 208, n. 7, p. 913–929, 2015. Rockefeller University Press. Disponível em: <[www.jcb.org/cgi/doi/10.1083/jcb.201411047JCB913](http://www.jcb.org/cgi/doi/10.1083/jcb.201411047JCB913)>. Acesso em: 25/1/2022.

SOUSA, M. E.; FARKAS, M. H. Micropeptide. **PLOS Genetics**, v. 14, n. 12, p. e1007764, 2018a. Public Library of Science.

SOUSA, M. E.; FARKAS, M. H. Micropeptide. **PLOS Genetics**, v. 14, n. 12, p. e1007764, 2018b. Public Library of Science. Disponível em: <<https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1007764>>. Acesso em: 17/7/2022.

SPANGENBERG, L.; SHIGUNOV, P.; ABUD, A. P. R.; et al. Polysome profiling shows extensive posttranscriptional regulation during human adipocyte stem cell differentiation into adipocytes. **Stem Cell Research**, v. 11, n. 2, p. 902–912, 2013. Stem Cell Res.

SPÄTER, D.; HANSSON, E. M.; ZANGI, L.; CHIEN, K. R. How to make a cardiomyocyte. **Development (Cambridge)**, v. 141, n. 23, p. 4418–4431, 2014. Company of Biologists Ltd.

TAKAHASHI, K.; TANABE, K.; OHNUKI, M.; et al. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. **Cell**, v. 131, n. 5, p. 861–872, 2007. Elsevier. Disponível em: <<http://www.cell.com/article/S0092867407014717/fulltext>>. Acesso em: 25/7/2022.

THOMSON, J. A. Embryonic stem cell lines derived from human blastocysts. **Science**, 1998.

THOMSON, J. A.; ITSKOVITZ-ELDOR, J.; SHAPIRO, S. S.; et al. Embryonic Stem Cell Lines Derived from Human Blastocysts. **Science**, v. 282, n. 5391, p. 1145–1147, 1998. American Association for the Advancement of Science. Disponível em: <<https://www.science.org/doi/10.1126/science.282.5391.1145>>. Acesso em: 25/7/2022.

TIGGES, J.; BIELEC, K.; BROCKERHOFF, G.; et al. Academic application of Good Cell Culture Practice for induced pluripotent stem cells. **ALTEX - Alternatives to animal experimentation**, v. 38, n. 4, p. 595–614, 2021. ALTEX Edition. Disponível em: <<https://www.altex.org/index.php/altex/article/view/2246>>. Acesso em: 25/7/2022.

URANISHI, H.; ZOLOTUKHIN, A. S.; LINDTNER, S.; et al. The RNA-binding motif protein 15B (RBM15B/OTT3) acts as cofactor of the nuclear export receptor NXF1. **Journal of Biological Chemistry**, v. 284, n. 38, p. 26106–26116, 2009.

VARADI, M.; ANYANGO, S.; DESHPANDE, M.; et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. **Nucleic Acids Research**, v. 50, n. D1, p. D439–D444, 2022. Oxford Academic. Disponível em: <<https://academic.oup.com/nar/article/50/D1/D439/6430488>>. Acesso em: 31/7/2022.

WAKUI, T.; MATSUMOTO, T.; MATSUBARA, K.; et al. Method for evaluation of human induced pluripotent stem cell quality using image analysis based on the biological morphology of cells. <https://doi.org/10.1117/1.JMI.4.4.044003>, v. 4, n. 4, p. 044003, 2017. SPIE. Disponível em: <<https://www.spiedigitallibrary.org/journals/journal-of-medical-imaging/volume-4/issue-4/044003/Method-for-evaluation-of-human-induced-pluripotent-stem-cell-quality/10.1117/1.JMI.4.4.044003.full>>. Acesso em: 25/7/2022.

WANG, X.; LE, N.; DENOTH-LIPPUNER, A.; BARRAL, Y.; KROSCHEWSKI, R. Asymmetric partitioning of transfected DNA during mammalian cell division. **Proceedings of the National Academy of Sciences of the United States of America**, v. 113, n. 26, p. 7177–7182, 2016. National Academy of Sciences. Disponível em: <<https://www.pnas.org/doi/abs/10.1073/pnas.1606091113>>. Acesso em: 24/11/2022.

WANG, Hanmin; HAO, Jijun; HONG, Charles C. Cardiac Induction of

- Embryonic Stem Cells by a Small Molecule Inhibitor of Wnt/ $\beta$ -Catenin Signaling. **ACS Chemical Biology**, v. 6, n. 2, p. 192–197, 2011. Disponível em: <<https://pubs.acs.org/doi/10.1021/cb100323z>>. Acesso em: 5/6/2022.
- WANG, Y.; PALMFELDT, J.; GREGERSEN, N.; et al. Mitochondrial fatty acid oxidation and the electron transport chain comprise a multifunctional mitochondrial protein complex. **The Journal of biological chemistry**, v. 294, n. 33, p. 12380–12391, 2019. J Biol Chem. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/31235473/>>. Acesso em: 22/11/2022.
- WANG, Y.; WU, S.; ZHU, X.; et al. LncRNA-encoded polypeptide ASRPS inhibits triple-negative breast cancer angiogenesis. **The Journal of Experimental Medicine**, v. 217, n. 3, 2020. Rockefeller University Press. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/31816634/>>. Acesso em: 2/2/2021.
- WANG, Lantian; FAN, Jing; HAN, Lili; et al. The micropeptide LEMP plays an evolutionarily conserved role in myogenesis. **Cell Death & Disease**, v. 11, n. 5, p. 357, 2020.
- WEI, M. M.; ZHOU, G. B. Long Non-coding RNAs and Their Roles in Non-small-cell Lung Cancer. **Genomics, Proteomics & Bioinformatics**, v. 14, n. 5, p. 280–288, 2016. Elsevier.
- WHEELER, D. L.; BARRETT, T.; BENSON, D. A.; et al. Database resources of the National Center for Biotechnology Information. **Nucleic acids research**, v. 36, n. Database issue, 2008. Nucleic Acids Res. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/18045790/>>. Acesso em: 24/11/2022.
- WOJCIECHOWSKA, A.; BRANIEWSKA, A.; KOZAR-KAMIŃSKA, K. MicroRNA in cardiovascular biology and disease. **Advances in clinical and experimental medicine : official organ Wroclaw Medical University**, v. 26, n. 5, p. 865–874, 2017. Adv Clin Exp Med. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/29068585/>>. Acesso em: 28/3/2020.
- WU, Y.; YANG, X.; CHEN, Z.; et al. M 6 A-induced lncRNA RP11 triggers the dissemination of colorectal cancer cells via upregulation of Zeb1. **Molecular Cancer**, v. 18, n. 1, 2019. BioMed Central Ltd.
- XIANG, Q.; YANG, B.; LI, L.; et al. Critical role of Lin28-TNFR2 signalling in cardiac stem cell activation and differentiation. **Journal of cellular and molecular medicine**, v. 23, n. 4, p. 2943–2953, 2019. J Cell Mol Med. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/30734494/>>. Acesso em: 24/1/2022.
- XU, Z. L.; MIZUGUCHI, H.; ISHII-WATABE, A.; et al. Strength evaluation of transcriptional regulatory elements for transgene expression by adenovirus vector. **Journal of Controlled Release**, v. 81, n. 1–2, p. 155–163, 2002. Elsevier.
- XUE, Z.; HENNELLY, S.; DOYLE, B.; et al. A G-Rich Motif in the lncRNA Braveheart Interacts with a Zinc-Finger Transcription Factor to Specify the Cardiovascular Lineage. **Molecular Cell**, v. 64, n. 1, p. 37–50, 2016. Cell Press.
- YANG, J.; HUNG, L. H.; LICHT, T.; et al. RBM24 Is a Major Regulator of Muscle-Specific Alternative Splicing. **Developmental Cell**, v. 31, n. 1, p. 87–99,

2014. Cell Press.

YOUNG, A. P.; JACKSON, D. J.; WYETH, R. C. A technical review and guide to RNA fluorescence in situ hybridization. **PeerJ**, v. 8, 2020. PeerJ Inc. Disponível em: [/pmc/articles/PMC7085896/](https://doi.org/10.21963/peerj.10007)>. Acesso em: 16/7/2022.

YOUNG, R. A. Control of the Embryonic Stem Cell State. **Cell**, v. 144, n. 6, p. 940–954, 2011. Elsevier. Disponível em: <http://www.cell.com/article/S0092867411000717/fulltext>>. Acesso em: 15/8/2022.

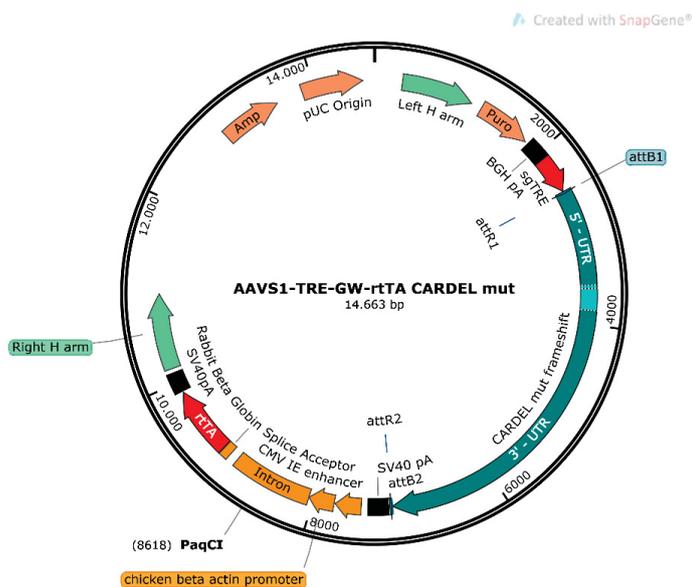
YU, J.; VODYANIK, M. A.; SMUGA-OTTO, K.; et al. Induced pluripotent stem cell lines derived from human somatic cells. **Science (New York, N.Y.)**, v. 318, n. 5858, p. 1917–1920, 2007. Science. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/18029452/>>. Acesso em: 24/1/2022.

ZHANG, M.; ZHAO, K.; XU, X.; et al. A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. **Nature Communications** 2018 9:1, v. 9, n. 1, p. 1–17, 2018. Nature Publishing Group. Disponível em: <https://www.nature.com/articles/s41467-018-06862-2>>. Acesso em: 27/1/2022.

ZHAO, Y.; ZHOU, J.; HE, L.; et al. MyoD induced enhancer RNA interacts with hnRNPL to activate target gene transcription during myogenic differentiation. **Nature Communications** 2019 10:1, v. 10, n. 1, p. 1–17, 2019. Nature Publishing Group. Disponível em: <https://www.nature.com/articles/s41467-019-13598-0>>. Acesso em: 29/11/2022.

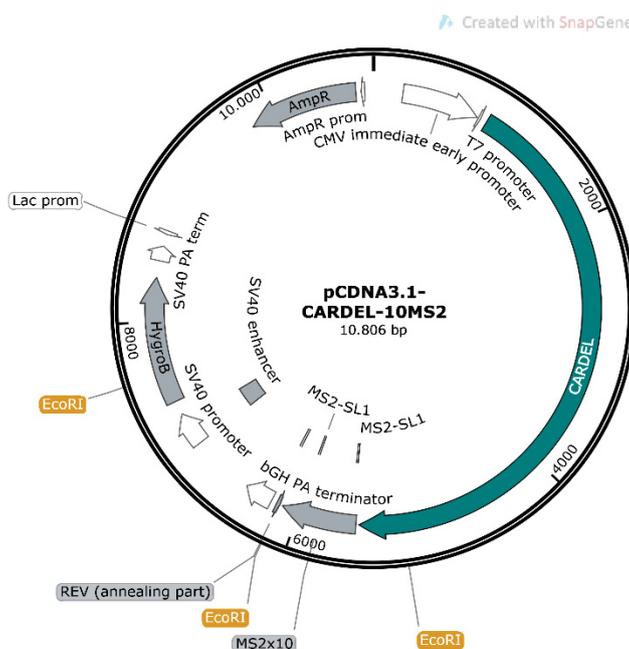
## ANEXO 1 – MAPA DE VETORES

FIGURA 16 – MAPA DO VETOR AAVS1-TRE-GW-rtTA COM CARDEL COM MUTAÇÃO FRAMESHIFT



NOTA: Criado com SnapGene.

FIGURA 17 – MAPA DO VETOR pcDNA3.1-CARDEL10xMS2



NOTA: Criado com SnapGene.

## ANEXO 2 – LAUDO DO CARIÓTIPO DE PLZ CARDEL MUT C8.1



**PUCPR**

ASSOCIAÇÃO PARANAENSE DE CULTURA  
PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ  
NÚCLEO DE TECNOLOGIA CELULAR

**CITOGENÉTICA**  
**TÉCNICA: CARIOTIPAGEM COM BANDEAMENTO GTG**

IDENTIFICAÇÃO: PLZ i CARDEL mut C8.1

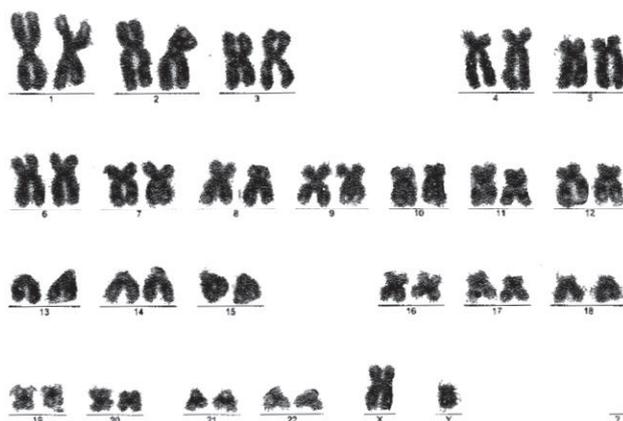
MATERIAL: Células Pluripotentes induzidas em

DATA DA ENTRADA: 03/05/2022

SOLICITANTE: Rhaiza S. V. França

RESULTADO: Vinte e duas metáfases foram completamente analisadas e não foram observadas aberrações cromossômicas clonais;

CARIOGRAMA REPRESENTATIVO DA AMOSTRA



**CARIÓTIPO: 46,XY[22]**

CÓDIGO: IP074

DATA DO LAUDO: 11/05/2022

ANALISTA: Valderez Ravaglio Jamur

REVISOR: Isadora May Vaz

SUPERVISOR:

*Valderez Ravaglio Jamur*  
**Valderez Ravaglio Jamur**

CITOGENETICISTA  
CRF9 1752

## APÊNDICE 1 – REAGENTES

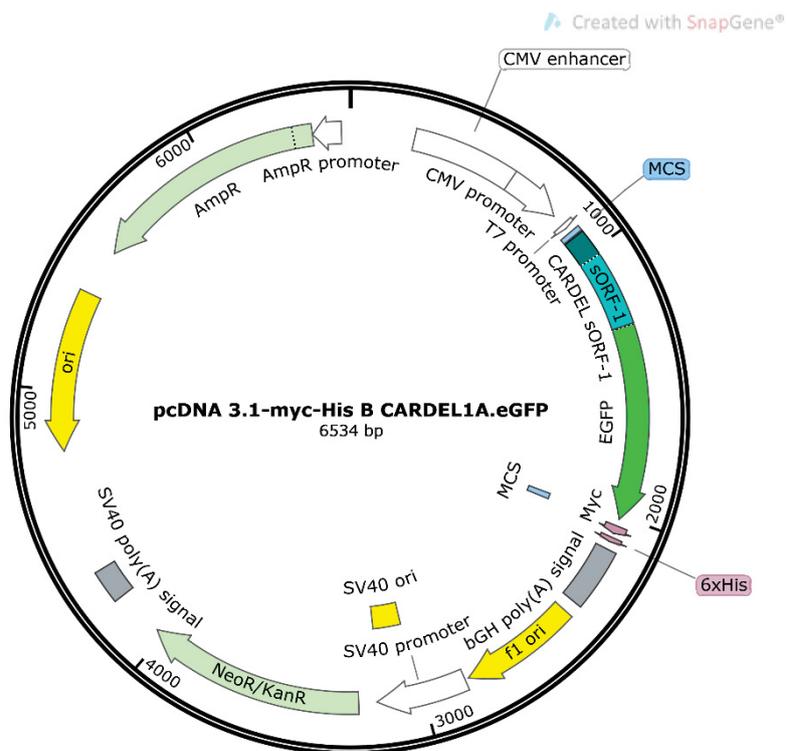
TABELA 5 – MEIOS DE CULTIVO E KITS COMERCIAIS UTILIZADOS

Reagente	Fornecedor	Código
Phusion High-Fidelity PCR Master/QG Buffer	Thermo Scientific	F532S
NEBuffer™ Set (r1.1/3.1)	New England BioLabs	B6003S
TrackIt 1 kb Plus	Invitrogen	10787018
SYBR™ Safe DNA Gel Stain	ThermoFisher	S33102
QIAquick Gel Extraction Kit	Qiagen	28706X4
T4 DNA ligase	Thermo Scientific	15224025
T4 DNA ligase buffer 5x	Invitrogen	15224017
ATP 10 mM	NEB Labs	B0756S
QiaPrep Spin MiniPrep	Qiagen	27106
Lipofectamine™ 2000 Transfection Reagent	Invitrogen	11668019
Lipofectamine™ 3000 Transfection Reagent	Invitrogen	L3000015
Opti-MEM™ I Reduced Serum Medium	Gibco	31985088
cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail	Roche	04639159001
Qubit Protein Assay	ThermoFisher	Q33211
Kaleidoscope Precision Plus	Bio-rad	1610375
Protran® Western Blotting Membrane	Amersham,	GE10600002
Purified Mouse Anti-Vimentin	BD Biosciences	550513
GFP Rabbit Policlonal Antibody	ThermoFisher	A-11122
IRDye® 800CW Goat anti-Mouse IgG Secondary Antibody	Odyssey	926-32210
IRDye® 800CW Donkey anti-Rabbit IgG Secondary Antibody	Odyssey	926-32213
Corning® Matrigel® Matrix	Gibco	356234
DMEM Mix	Gibco	21013-024
Ham's F-12 Nutrient Mix	Gibco	21700075
Accutase™	STEMCELL Technol.	07920_C
StemFlex™	Gibco	A3349401
iROCK	Sigma-Aldrich	SCM075
Lipofectamine Stem Reagent	Invitrogen	STEM00001
TRIReagent® RNA Isolation Reagent	Merck	T9424
Direct-zol RNA Miniprep Kit	Zymo Research	R2052

ImProm-II™ Reverse Transcription System	Promega	A3800
Go Taq® qPCR Master Mix	Promega	A6020
QIAmp DNA Extraction Kit	Qiagen	51304
Mix de PCR HPV/NAT	Biomanguinhos/Fiocruz	N/A
pcDNA™3.1 (+) Mammalian Expression Vector	Invitrogen	V79020
TRI Reagent	Sigma-Aldrich	T9424-25ML
RNaseOUT™ Recombinant Ribonuclease Inhibitor	Invitrogen	10777019
DNase I	Invitrogen	18047-019

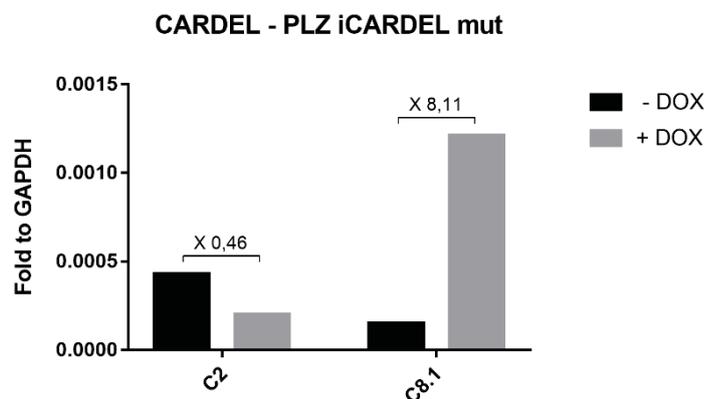
## APÊNDICE 2 – INFORMAÇÕES SUPLEMENTARES DE PLZ CARDEL MUT

FIGURA 18 – MAPA DO VETOR pcDNA 3.1/myc-His B CARDEL1A.eGFP



NOTA: Criado com SnapGene.

FIGURA 19 – TESTE DE SUPEREXPRESSIONE DE TRANSCRITO DE CARDEL EM PLZ CARDEL mut



NOTA: -DOX e +DOX indicam condições sem e com indução de superexpressão de CARDEL por doxiciclina, respectivamente; os números posicionados em cima das colunas são os  $\Delta\Delta Ct$  das condições -DOX e +DOX, para os 2 clones de PLZ CARDEL mut; resultados obtidos a partir de única replicata.