Polycomb protein RYBP regulates transcription factor *Plagl1* during *in vitro* cardiac differentiation of mouse embryonic stem cells

SURYA HENRY

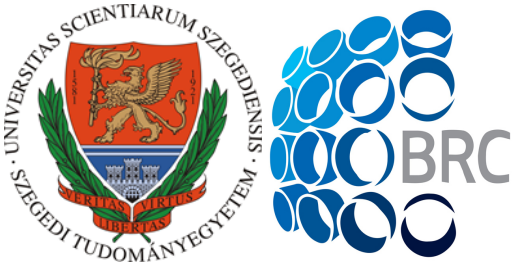
Ph.D. THESIS

SUPERVISOR:

DR. MELINDA K. PIRITY

Laboratory of Embryonic and Induced Pluripotent Stem Cells
Institute of Genetics
Biological Research Centre

Doctoral School of Biology
Faculty of Science and Informatics
University of Szeged



Szeged
2021

TABLE OF CONTENTS

LIST OF ABBAREVTIONS..... 6
1. INTRODUCTION..... 9
1.1 Stem cells as model systems of developmental process and congenital heart disorders 9
1.2 Process of mouse embryonic cardiac differentiation..... 10
1.3 Markers of cardiac differentiation..... 10
1.4 *In vitro* cardiac differentiation methods..... 13
1.5 Utilisation of luciferase reporter system to study gene regulation..... 15
1.6 Polycomb repressive complexes and their regulatory roles in mammalian development 16
1.7 Role of RYBP in mouse embryogenesis..... 19
1.8 The role of *Rybp* in cardiac development..... 20
1.9 Overview of the regulatory activities of RYBP..... 20
1.10 Relativeness between the functions of RYBP and PLAGL1..... 22
1.11 *Plagl1*, a key cardiac transcription factor..... 24
1.12 Relevance of *Plagl1* towards diseases..... 25
2. AIMS..... 26
3. MATERIALS AND METHODS..... 27
3.1 Cell culture techniques..... 27
3.1.1 Cell lines and culture conditions..... 27
3.1.2 *In vitro* cardiac differentiation of mouse embryonic stem cells..... 28
3.1.3 Calcium Phosphate transient transfection method..... 29
3.1.4 The luciferase reporter assay system..... 30
3.1.5 Inhibition of PRC1 activity..... 31
3.2 Molecular biology techniques..... 31
3.2.1 Quantitative real-time polymerase chain reaction (qRT-PCR)..... 31
3.2.2 Chromatin immunoprecipitation and qRT-PCR (ChIP-qRT-PCR)..... 32
3.2.3 Molecular cloning, transformation and confirmation..... 33
3.2.4 Site directed mutagenesis..... 37
3.3 Biochemical assays..... 39
3.3.1 Western blot analysis..... 39
3.3.2 Protein stability assays..... 39

3.3.3 Co-Immunoprecipitation (Co-IP) 39
3.3.4 Immunocytochemistry (ICC) analysis 40
3.4 Bioinformatic analysis 41
3.4.1 Transcriptome analysis 41
3.4.2 Analysis of the reported ESTs of the *Plagl1* splice variants 41
3.4.3 Analysis of functional domains and degron sites in PLAGL1 41
3.4.4 Analysis of *Plagl1* promoters for CpG islands and TATA box 42
3.4.5 Transcription factor binding analysis in *Plagl1* promoters 42
3.4.6 Motif search in gene promoter regions 42
3.4.7 Metadata analysis in ES cells and CMCs 43
3.5 Statistical analysis 43
4. RESULTS 44
4.1 Hierarchical gene cluster analysis of the whole transcriptome 44
4.1.1 Calcium homeostasis, the JAK-STAT pathway and cell adhesion are amongst the most affected mechanisms in the *Rybp*$^{-/-}$ ES cells and derived CMCs 44
4.1.2 Key cardiac transcription factors and sarcomeric components are downregulated in the *Rybp* null mutant CMCs 47
4.2 *Plagl1* is the most strikingly downregulated gene in the *Rybp*$^{-/-}$ ES cells and CMCs 48
4.3 PLAGL1 is not detectable at protein level either in the *Rybp* null mutant cells at all time points of *in vitro* cardiac differentiation 49
4.4 Overview of the *Plagl1* genomic locus 50
4.5 The expression of the two ncRNAs *Hymai* and *Plagl1it* are also affected in the *Rybp* null mutant CMCs during cardiac differentiation 52
4.6 Multiple splice variants of *Plagl1* can be transcribed from monoallelic and biallelic promoters 53
4.7 Splice variants of *Plagl1* are transcribed from *P1* and *P3* promoters during *in vitro* cardiac differentiation 54
4.8 Two isoforms of the PLAGL1 protein are detectable during *in vitro* cardiac differentiation 56
4.9 PLAGL1 undergoes post translation degradation during CMC formation 57
4.10 PLAGL1 has a degron site immediately after the Serine-Throenine phosphorylation sites at the N-terminal 58
4.11 PLAGL1 and RYBP are co-expressed in the nuclei of the differentiating cardiac cultures 59

4.12 *Plagl1* expression is first detected at early progenitor stage during *in vitro* cardiac differentiation 62
4.13 *Plagl1* promoters contain distinctive regulatory elements 64
4.14 RYBP activates the *Plagl1 P1* and *P3* promoters 65
4.15 The activation of *Plagl1 P1* and *P3* promoters by RYBP is polycomb independent 66
4.15.1 A small molecule- PRC1 inhibitor did not attenuate the activation by RYBP at the *P1* and *P3* promoter 67
4.15.2 RING1 does not synergistically function with RYBP at the *Plagl1 P1* and *P3* promoters 68
4.15.3 Existing ChIP-seq analysis reveal polycomb independent binding of RYBP at the *Plagl1 P3* promoter 70
4.16 RYBP did not activate the *P1* and *P3* promoters synergetic with E2F and YY1 transcription factors 73
4.17 *Hymai* and *Plagl1it* overexpression did not affect the activity of RYBP at the *P1* and *P3 Plagl1* promoters 75
4.18 RYBP activates the *P3* promoter via NKX2-5 76
4.18.1 RYBP activates the 3' region of the *P3* promoter containing binding sites for NKX2-5 and MEF2C 77
4.18.2 RYBP does not activate the *P3* promoter via MEF2C consensus sites 79
4.18.3 RYBP and NKX2-5 can synergistically activate the *P3* promoter 80
4.18.4 RYBP can interact with NKX2-5 protein 81
4.18.5 RYBP bound at the NKX2-5 consensus at the *P1* and *P3* promoters in CMCs but not in ES cells 83
4.19 *Hymai* and *Plagl1it* synergistically functions with NKX2-5 to activate the *P3* promoter 84
4.20 PLAGL1 is important for the formation of terminally differentiated CMCs and sarcomere organisation 85
4.20.1 PLAGL1 co-expressed with CTNT in the differentiating CMCs 86
4.20.2 Sarcomere gene promoter regions contain consensus binding sites for PLAGL1 87
4.20.3 PLAGL1 can activate the *Tnnt2 promoter* 88
5. DISCUSSION 90
6. ACKNOWLEDGEMENTS 96

7. REFERENCES 98
8. SUMMARY OF THE PHD THESIS 116
9. ÖSSZEFOGLALÓ 118
10. APPENDICES 121

LIST OF ABBREVIATIONS

aa	amino acids
Actc1	Actin alpha cardiac muscle 1
Adora2	Adenosine A2a receptor
AGE	Agarose gel electrophoresis
Agtr1	Angiotensin II receptor type 1
Avpr1a	Arginine vasopressin receptor 1a
BAC	Bacterial artificial chromosome
bFGF	basic fibroblast growth factor
BMP	Bone morphogenetic protein
bp	base pair
BSA	bovine serum albumin
$CaCl_2$	Calcium chloride
Cacna1g	Calcium voltage-gated channel T type alpha 1G subunit
Cacng5	Calcium channel voltage-dependant gamma subunit 5
Calcr	Calcitonin receptor
Casr	Calcium-sensing receptor
CBX	Chromobox domain containing transcription factors
CDC6	Cell division cycle 6
Cdh17	Cadherin 17
Cdh6	Cadherin 6
Cdh7	Cadherin 7
Cdkn1a	Cyclin-dependent kinase inhibitor 1A
CHDs	Congenital heart disorders
ChIP	Chromatin immunoprecipitation
CHX	Cyclohexamide
CMC	Cardiomyocytes
Co-IP	Co-immunoprecipitation
CTNT	Cardiac Muscle Troponin T
DAPI	4’,6-diamidino-2-phenylindole
DEDAF	Death effector domain-associated factor
DMEM	Dulbecco’s modified eagle’s medium
DMR	differentially methylated region
E	embryonic
E2F2/3	E2F transcription factor 2/3
EB	embryoid body
Ednra	Endothelial receptor type A
EDTA	Ethylenediaminetetraacetic acid
EED	Embryonic ectoderm development
EMC	Extraskeletal myxoid chondrosarcoma
END2	Endoderm like cell
Eomes	Eomesodermin

ES cells	embryonic stem cells
EST	expressed sequenced tag
EYFP	Enhanced Yellow Fluorescent Protein
EZH1/2	Enhancer of zeste homolog 1/2
FBS	fetal bovine serum
FC	fold change
Gata4	Gata binding protein 4
gDNA	genomic DNA
Gjb2	Gap junction protein beta 2
HBS	HEPES buffered saline
HCN2	Hyperpolarization-activated cyclic nucleotide-gated potassium and sodium channel 2
HCN3	Hyperpolarization-activated cyclic nucleotide-gated potassium and sodium channel 3
HD	hanging drop
HEK293T	Human Embryonic Kidney (HEK) 293T
HPRT	Hypoxanthine-guanine phosphoribosyltransferase I
hr	hour
Hymai	Hydatidiform mole associated and imprinted transcript
ICC	Immunocytochemistry
ICM	inner cell mass
Isl1	Islet-1
JAK-STAT	Janus Kinase-Signal transducer and activator of transcription proteins
kb	kilobase
Kcne1	Potassium voltage-gated channel Isk-related subfamily member 1
Kcne3	Potassium voltage-gated channel Isk-related subfamily member 3
Kcnj5	Potassium inwardly rectifying channel subfamily J member 5
Kdm2b	Lysine (K)-specific demethylase 2B
Kdr	Kinase insert domain protein receptor
LAR II	Luciferase Assay Reagent II
LIF	Leukemia inhibitory factor
MEF	Mouse embryonic fibroblast
Mef2c	Myocyte enhancer factor 2c
Mesp1	Mesoderm posterior 1
MPCs	multipotent progenitor cells
Myh7	Myosin heavy peptide 7 cardiac muscle beta
Myl2	Myosin light chain 2
Mylk3	Myosin light chain kinase 3
Myom1	Myomesin 1
ncPRC1	non-canonical PRC1s
ncRNA	non-coding RNA
ND	No data
Nkx2-5	Nk2 Homeobox 5
NLS	Nuclear localization signal

PBS	phosphate buffered saline
PcG	Polycomb group
PCGF2	Polycomb group ring finger 2
Plagl1	Pleiomorphic adenoma gene like 1
Plagl1it	*Plagl1* intronic transcript
POU5F1	POU domain, class 5 transcription factor 1
PRC1	Polycomb repressive complex 1
PRC2	Polycomb repressive complex 2
PVDF	Polyvinylidene fluoride
qRT-PCR	quantitative real-time polymerase chain reaction
RBBP4/7	Retinoblastoma binding protein 4/7
RING1	Ring finger protein 1
RNF2	Ring finger protein 2
RT	room temperature
RYBP	RING1 and YY1 binding protein
Scn1a	Sodium channel voltage-gated type I alpha
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
Shh	Sonic hedgehog
STLV	Slow turning lateral vessel bioreactor
SUZ12	Suppressor of Zeste
T	Brachyury
Tbx-18	T-box 18
Tbx-3	T-box 3
Tbx5	T-box 5
TF	transcription factor
TFBS	transcription factor binding sites
TGF-β	Transforming growth factor beta 1
TNDM	Transient neonatal diabetic mellitus
Tnni3	Troponin I cardiac 3
Tnnt2	Cardiac troponin T2
Tpm1	Tropomyosin 1
Tpm4	Tropomyosin 4
TRP53	Transformation related protein 53
Trpv4	Transient receptor potential cation channel subfamily V member 4
TrxG	Trithorax
Ttn	Titin
Vcam1	Vascular cell adhesion molecule 1
WNT	Wingless-type MMTV integration site family
YAF2	YY1 associated factor 2
YY1	Ying-Yang-1
ZAC1	Zinc finger protein inducer of apoptosis and cell cycle arrest

1. INTRODUCTION

During mammalian embryogenesis the development of cell lineages and multiple organ systems are tightly associated with the finely tuned and orchestrated functions of transcription factors (Adamson & Gardner, 1979). The resulting differential gene expression is the basis of tissue specific protein production, which enables conduction of diverse functions of the organism constructed by myriad of cells harbouring the same genetic information (Adamson & Gardner, 1979). Epigenetic factors are crucial in modulating these mechanisms via multiple interacting factors and regulatory networks (Artyomov et al., 2010). In my thesis, I determined the connections between a polycomb group epigenetic factor RING1 and YY1 binding protein (RYBP) also called as Death effector domain-associated factor (DEDAF) and cardiac transcription factor Pleiomorphic adenoma gene like 1 (PLAGL1) in the regulation of cardiac differentiation using *in vitro* model system applying mouse embryonic stem (ES) cells.

1.1 Stem cells as model systems of developmental process and congenital heart disorders

During embryogenesis, the inner cell mass (ICM) of the blastocyst contains the pluripotent ES cells that have the capability to differentiate towards all lineages of the body (Marikawa & Alarcón, 2009). These cells undergo several cellular events leading towards organogenesis (Sasai et al., 2012). ES cells have distinctive features such as- (i) they can divide perpetually and self-renew by maintaining the pluripotency of the cells even in cultures; (ii) by changing the culture conditions, they can be differentiated into special cell types such as muscle cells, nerve cells etc. (Zakrzewski et al., 2019). ES cells also have the ability to the formation of germ layers when re-introduced into the early-stage embryo (Gardner & Brook, 1997). Due to these abilities, stem cells based *in vitro* differentiation systems have been widely used as excellent model systems to recapitulate the early events of organ development and related disease conditions (Keller, 1995; Levinson & Benvenisty, 1995).

Congenital heart disorders (CHDs) occur due to the structural and functional anomalies during heart development (Mckusick, 1964; McCulley & Black, 2012). CHDs are commonly caused due to the atrial septal defects, ventricular septal defects, atrioventricular canal defect and valve stenosis which often led to serious conditions of contractility disorders. The loss of function of several transcription factors have been demonstrated to contribute to CHD conditions both using *in vivo* mouse and *in vitro* model systems. Improper expression of key cardiac transcription factors such as the NK2 homeobox 5 (NKX2-5), Myocyte enhancer factor 2c

(MEF2C) and T-box 5 (TBX5) caused serious malformations of the developing heart and contractility functions (McCulley & Black, 2012; Weerd et al., 2011). The effect of *Plagl1* mutation is also implicated to heart malformations and CHD conditions as the *Plagl1* heterozygous mutant mice formed atrial and ventricular septal defects (Yuasa et al., 2010).

Stem cells based *in vitro* cardiac differentiation methods offer a unique platform to model CHD conditions, to dissect and study the molecular mechanisms involved in regulating contractility as well as the underlying role of specific transcription factors that might also potentiate to future therapeutics (Moretti et al., 2013).

1.2 Process of mouse embryonic cardiac differentiation

Since ES cell based *in vitro* cardiac differentiation system can mimic the *in vivo* heart formation morphologically, functionally and electrophysiologically (Hescheler et al., 1997; Fijnvandraat et al., 2003) the different stages of *in vitro* cardiomyogenesis can be related to the knowledge gained from the plethora of studies carried *in vivo*. A series of cardiac markers identified *in vivo* are generally used to characterize the differentiation state of *in vitro* cultures.
Mammalian cardiac system is one of the first functional organ that develops in an early embryo (Savolainen et al., 2009). The contraction of the embryonic heart and initiation of the rudimentary circulatory system is essential for mammalian embryonic development. The heart originates from the embryonic mesoderm that further differentiates into mesothelium, endothelium and myocardium (DeRuiter et al., 1992; Yutzey et al., 1995). Mesothelial pericardium forms the outer lining of the heart (Madani & Golts, 2014). The inner lining of the heart as well as the lymphatic and blood vessels, develop from the endothelium. The early multipotent progenitor cells (MPCs) give rise to the atrial and ventricular cell types, fibroblast cells, endocardial and epicardial cells, cells of the conductive system (sinoatrial, atrioventricular, Purkinje fiber cells), the smooth muscle cells of the aorta, artery and the autonomic nerve cells (Weerd et al., 2011). The cardiac sarcomere is the critical unit of cardiac muscle fibres that functions in contraction (Sweeney & Hammers, 2018). The formation of all these cell types is crucial in maintaining the structure and functions of the developing heart.

1.3 Markers of cardiac differentiation

Series of mouse knockouts and *in vitro* cell culture models have shown the precise timeline and exact spatiotemporal expression of key cardiac transcription factors (Table 1). Mesodermal

lineage specification is guided by the exit of pluripotency and induction of the T-box transcription factors Brachyury (T) and Eomesodermin (EOMES) by graded Transforming growth factor beta 1 (TGF-β)/NODAL and the canonical Wingless-type MMTV integration site family (WNT) signalling (Figure 1) (Arnold et al., 2009; Watabe & Miyazono, 2009; Tosic et al., 2019). *T* and *Eomes* are expressed from the early gastrulation stage embryos from E6 in the primitive streak of the early mouse embryo (Wilkinson et al., 1990; Chesley, 1935; Russ et al., 2000; Nowotschin et al., 2013). The *T* homozygous mice showed serious implications in the morphogenesis of mesoderm derived structures such as the heart (Table 1) (Yanagisawa et al., 1981). Both *Eomes* and *T* can induce the expression of the Mesoderm posterior 1 (*Mesp1*) which is expressed in the developing heart tube and specifies cardiovascular lineage (Saga et al., 1999; David et al., 2011; Guo et al., 2018; Ameele et al., 2012). The expression of *Mesp1* defines the earliest step of cardiac lineage commitment (Figure 1) (Saga et al., 2000; Lescroart et al., 2018). MESP1 can target and induce the expression of several cardiac transcription factors that specifies progenitor formation (Bondue et al., 2008; Soibam et al., 2015) such as the Kinase insert domain protein receptor (*Kdr* also called as *Flk1*), multipotent cardiac progenitor marker Islet-1 (*Isl1*) (Cai et al., 2003 ; Moretti et al., 2006) the early cardiac progenitor markers which implicates formation of first and second heart fields *Nkx2.5, Mef2c* and cardiac endothelial progenitor marker GATA binding protein 4 (*Gata4*) (Christoforou et al., 2008; Lyons et al., 1995; Tanaka et al., 1999; Terada et al., 2011) (Figure 1). The non-canonical WNT pathway and NODAL are determined to function upstream to the cardiac progenitor formation during the first and second heart field derivation (Figure 1) (Brade et al., 2006; Gessert & Kühl, 2010; Kamps, 2016). Ascorbic acid is shown to promote cardiac differentiation through the induction of BMP, SMAD1 signalling and inhibition of the TGF-β signalling (Ivanyuk et al., 2015; Perino et al., 2017). Ascorbic acid is also determined as a potent inducer of several cardiac progenitor expression such as *Nkx2-5, Mef2c* and *Gata4* (Figure 1) (Takahashi et al., 2003; Lin et al., 1997; Kamps, 2016; Molkentin et al., 1997. During the cardiac progenitor formation, the expression of T-Box proteins T-Box 3 (*Tbx3*), T-Box 5 (*Tbx5*) and T-Box 18 (*Tbx18*) are required for the generation of pacemaker cells that function in the conduction system of the heart (Mori et al., 2006). As a result of these finely tuned events governed by series of key transcription factors, the developing heart starts beating as early as E7.5–8 in mouse (Tyser et al., 2016) (Figure 1).

Gene Name	Mutation type	Mutation at	Lethality time	Cardiac Phenotype	Source
T	complete null	ND	E10	Improper notochord formation	Chesley., 1935
Eomes	complete null	exon 2 & intron 2	E6.0	Improper primitive streak formation	Russ et al., 2000
MesP1	complete null	exon 1 & 2	E10.5	Altered heart morphology, two abnormally symmetrical heart tubes and the beat periodicity was not coordinated between one tube and the other.	Saga., 1999
Isl1	hypomorph	exon 3- second LIM domain	E8.5- E11.5	No outflow tract, no right ventricule and formation of hypoplastic atria	Cai et al., 2003
Nkx2-5	hypomorph	exon 2	E9	No looping, lack of endocardial cushion &trabeculae	Lyons et al., 1995
Gata4	hypomorph (only 2 ZNF binding domains deleted)	exon 3 & 4	E10.5	No ventral folding, position of developing heart is lateral & dorsal to the neural tube.	Molkentin et al., 1997
Mef2c	complete null	exon 2	E10.5	Pericardial effusion, very slowly beating, formation of hypoplastic ventricular chamber with no looping of cardiac tube	Lin et al., 1997
Tbx-5	hypomorph	exon 3	E10.5	Dilated heart atrium, formation of hypoplastic left ventricle and bifurcated atrial tube	Mori et al., 2006
Plagl1	complete null	intron 3	E10.5	Atrium Septum defect, ventricular septum defect and formation of thin ventricular wall	Yuasa et al., 2010

Table 1: Mutant phenotypes for key cardiac genes

Key cardiac genes used in this study are enlisted in the rows of the table. The corresponding nature of the analysed mutations, lethality time, the cardiac phenotypes and the reference of the studies are presented in the indicating columns of the table. Lethality time of mutant mice corresponds to the stage of embryonic heart development when the effect of the loss of the respective markers are implicated, and the represented alterations specify the significance of these genes in cardiac development. ***Abbreviations****: E-embryonic, ND-no data.*

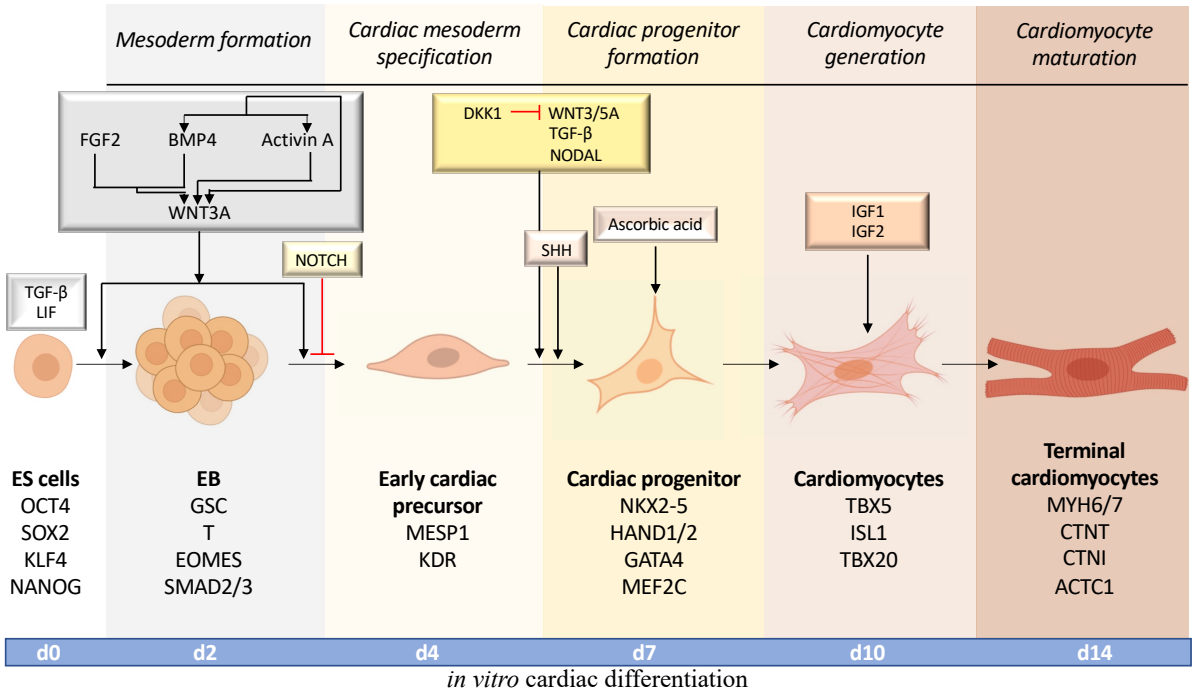


Figure 1: Schematic illustration of the markers involved in the stages of in vitro cardiac differentiation (modified from Kamps, 2016)

Factors that influence the progression of the five stages of cardiomyocyte differentiation: mesoderm formation (grey background), cardiac mesoderm specification (beige background), cardiac progenitor formation (pale yellow background), cardiomyocyte generation (rose background) and cardiomyocyte maturation (peach colour). Transcription factors associated with each of the six cell types during cardiomyocyte differentiation are presented below. The relating time points of in vitro cardiac differentiation is presented at the bottom. Signalling pathways and chemical inducers that guide the expression of the described transcription factors are highlighted in boxes. Pointed arrows (in black) indicates positive regulation and bar headed arrows (in red) indicates inhibition of differentiation. ***Abbreviations****: ES- embryonic stem, EB- embryoid body, d- day, LIF- Leukaemia inhibitory factor.*

1.4 *In vitro* cardiac differentiation methods

Over the years, several *in vitro* differentiation methods have been described for the differentiation of ES cells to form contractile cardiac cultures. ES cells cultured with stromal cells can differentiate to functional cardiomyocytes (CMCs) by (i) the formation of embryoid bodies (EB's), (ii) by monolayer formation in Matrigel or by, (iii) generation of cardiac organoids (Figure 2).

(i) The EB based methods (Evans & Kaufman, 1981) of cardiac differentiation are more robust and simpler, therefore they are widely used for cardiac differentiation using mouse ES cells. These methods generally start with the EB formation step by either the hanging drop (HD) method, suspension cultures or employing specific culture vessels such as the slow turning lateral vessel bioreactor (STLV) (Rungarunlert et al., 2013). As the EBs cultured in HDs and STLVs are more uniform and evenly sized, these methods are preferred over the suspension culture method for cardiac differentiation in which the formed EBs are irregularly shaped and diverse in their size (Wang & Yang, 2008). The EBs are cultured in humidified conditions for 48 hours, plated out and cultured further to form beating CMCs with no chemical inducers (Wang & Yang, 2008). Some EB based methods for cardiac differentiation also use inducing factors of signalling pathways that influence cardiac lineage commitment such as basic fibroblast growth factor-2 (bFGF-2) (Kawai et al., 2004).

(ii) In the monolayer system of ES cells are cultured along with supporting cell layers (e.g., mouse embryonic fibroblasts (MEF) in monolayers seeded on matrigel or surface treated tissue culture plates) (Batalov & Feinberg, 2015). The cells are then induced with growth factors such as basic fibroblast growth factor (bFGF) and ascorbic acid (Kokkinopoulos et al., 2016) to generate beating CMCs. Another method depicts the induction of BMP signalling pathway for the differentiation of monolayer ES cells. Induction of these growth factors promote the signalling for cardiac lineage commitment and CMCs formation (Zhang et al., 2012).

(iii) Another previously used method involves co-culture of ES cells with isolated mesenchymal cells (Pucéat, 2008) or endoderm like cell (END2 cells) which resulted in contractile CMCs (Mummery et al., 2012).

(iv) In the last few years, organoid culture has become the new trend for *in vitro* differentiation of stem cells as they replicate key spaciotemporal features of the *in vivo* organ. In these methods, ES cells are let to form EBs and further cultured in media containing FGF4, BMP and ascorbic acid for several days to form beating cardiac organoids (cardioids) with chamber like specifications (Lee et al., 2020).

Due to the reproducibility and the simplicity of the technique, the hanging drop method of EB generation was employed in this thesis study and no specific inducers were utilized (described in detail in Methods 3.1.2).

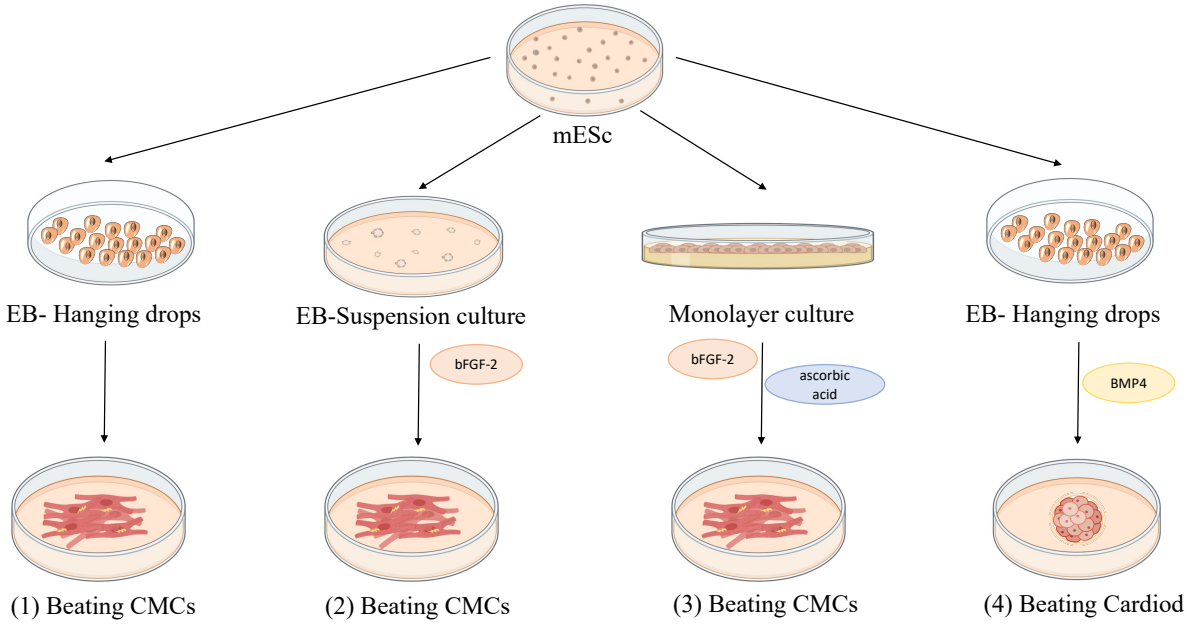


Figure 2: In vitro cardiac differentiation methods

CMCs can be differentiated in vitro from ES cells through (EB formation by (1) Hanging drops or (2) suspension culture, (3) monolayer culture or (4) formation of cardioids. ***Abbreviations:*** *mES cells: Mouse embryonic stem cells, EB: Embryoid body, CMCs: Cardiomyocytes.*

1.5 Utilisation of luciferase reporter system to study gene regulation

Luciferase reporter systems are widely used for studying promoter activities influenced by the regulatory functions of transcription factors which also relate to distinct cellular responses. This system is based on the activity of a bioluminescent protein Firefly, which produces light when reacted with a substrate D-luciferin (Figure 3) (Marques & Da Silva, 2009). This system is mainly exhausted to establish a functional relationship between the presence and concentration of specific regulatory proteins and the level of transcriptional activation (Firefly reporter) of the promoter analysed. The promoter of interest is cloned upstream to the *firefly* coding region; therefore, it effectively controls its expression. In a typical experimental assay, the cloned promoter construct is subjected to the influence of presumptive regulatory proteins upon overexpression (Figure 3). The relative expression level of the resulting firefly reporter is directly proportional to the activation/repression levels of the monitored promoter (Fan & Wood, 2007). Since Firefly protein is not natively present in mammalian cells, it does not have any detrimental effect on the endogenous cellular process (Keller et al., 1987), thereby this system provides reliable and reproducible opportunity to study regulatory activities of specific transcription factors.

Although luciferase reporter systems do not determine whether a regulatory protein directly interacts with a putative promoter region, they can be used to establish a functional connection between the protein and the amount of promoter activity induced by the protein (Carter & Shieh, 2015). Many proteins could indirectly affect transcription of the examined promoter by activating or repressing other regulatory proteins, assembly of different protein complexes, or signalling mechanisms that in return can affect the regulation of the promoter.

To distinguish between direct regulators and secondary effects, during my PhD studies I used luciferase reporter systems and firefly reporter constructs to molecularly analyse the detailed regulation of *Plagl1* promoters. I made several mutations of the predicted transcriptional binding sites and coupled these experiments with Chromatin immunoprecipitation (ChIP) assays (details in methods 3.1.5, 3.2.2, 3.2.4 and 3.4.5) to identify the molecular mechanism by which RYBP could mediate its effects during cardiac differentiation.

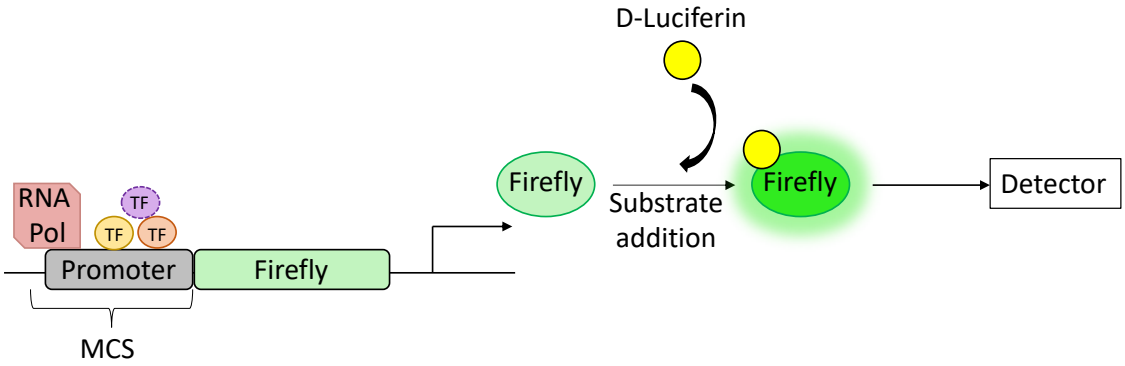


Figure 3: **Promoter regulation analysis using luciferase reporter system**
Schematic representation of the luciferase reporter system. Promoter of interest is cloned upstream of the firefly coding region. Adding a specific substrate, D-luciferin, the illumination of the firefly protein is triggered and detected using a luminometer. The amount of bioluminescent signal generated by firefly is directly proportional to the activity of the promoter. **Abbreviations**: RNA Pol-RNA Polymerase, TF-Transcription factor.

1.6 Polycomb repressive complexes and their regulatory roles in mammalian development

Epigenetic maintenance of differential gene expression is essential for proper differentiation and lineage commitment. The interplay between the Polycomb (PcG) and Trithorax (TrxG)

group protein complexes orchestrate these mechanisms by regulating genes as early as the trophoblast differentiation of a developing embryo (Schuettengruber et al., 2007; Kuroda et al., 2020). Polycomb group proteins (PcG) are a family of epigenetic silencers implicated in growth and development (Gould, 1997); cancer progression and suppression (Laugesen et al., 2016); stem cell maintenance and regulation (Aloia et al., 2013) and X-chromosome inactivation (Simon & Kingston, 2013). PcG proteins form two principal complexes, named Polycomb repressive complex 1 (PRC1) and Polycomb repressive complex 2 (PRC2) (Figure 4A) (Vidal, 2009). Each complex has constant and variable protein subunits that lead to distinctive cell and tissue specific regulation (Gao et al., 2012). The PRC2 complex is a highly conserved multi-subunit protein complex that consists of Enhancer of zeste subunit 1/2 (EZH1/2), Suppressor of zeste (SUZ12), Embryonic ectoderm development (EED) and Retinoblastoma binding protein 4/7 (RBBP4/7) which form the minimum core of the complex (Czermin et al., 2002; Kuzmichev et al., 2002; Müller et al., 2002). The SET domain containing EZH1/2 factor is the catalytic subunit of the complex which specifically deposits the trimethyl mark on lysine 27 of histone 3 (H3K27me3), a major chromatin repressive modification (Figure 4A) (Shen et al., 2008). SUZ12 consists of the ZnB-Zn domain that is able to bind to several interacting partners which provides target specificity to the PRC2 complex (Chen et al., 2018). RBBP4/7 factors are dispensable for the catalytic activity of the complex (Cao & Zhang, 2004; Ketel et al., 2005). EED functions in the interaction and recruitment of the chromobox domain containing transcription factors (CBX) containing PRC1 complex (also called as the canonical PRC1s) to the H3K27me3 modified loci for further enhancement of the PRC1 mediated mono-ubiquitination mark on lysine 119 of histone 2a (H2AK119ub1) an alternative repression mark (Figure 4B) (Cao et al., 2014).

The canonical PRC1 (cPRC1) complexes consists of the CBX transcription factors, which are capable of specific binding to H3K27me3. The catalytic subunit of the cPRC1 is an E3 ubiquitin ligase imparting factor Ring finger protein 1 (RING1) or its homolog Ring finger protein 2 (RNF2), that can deposit H2AK119ub1 (Aranda et al., 2015). The RING1/RNF2 subunits are always bound to either Polycomb group ring finger 2 (PCGF2) or Polycomb group ring finger 4 (PCGF4) which together with CBX factors constitute to the core of the cPRC1 complex's (Figure 4B). The cPRC1s are categorized based on the presence of PCGF2 and PCGF4 as cPRC1.2 and cPRC1.4 complexes respectively (Geng & Gao, 2020).

Another variant of the PRC1 complexes, the non-canonical PRC1s (ncPRC1s) (also called as the variant PRC1(vPRC1s)) were also identified in later studies, in which the canonical core subunit of the cPRC1s CBX is replaced by RYBP or its homolog YY1 associated factor 2 (YAF2) (Gao et al., 2012). As the CBX subunits are replaced by RYBP and YAF2 which lack the chromobox domains in the ncPRC1s, these complexes are not capable to recognise and bind to H3K27me3. RYBP is able to recognise and bind to the H2AK119ub1 mark (Zhao et al., 2020) at the repressed genes and deposits further H2AK119ub1 through the RING proteins, aiding the compaction of the chromatin and enabling more stable gene repression of the targeted loci by the ncPRC1s (Blackledge et al., 2014; Martinez et al., 2020) (Figure 4C). The ncPRC1s are categorized based on the presence of either of the 6 PCGF factors to form the core of the complex along with RING1/RNF2 and the RYBP/YAF2.

Early models of maintained epigenetic repression of genes are generally referred as the "hierarchical recruitment of the PRC complexes" (Dorafshan et al., 2017). The proposed idea involves the initial assembly and binding of the PRC2 complex at the promoters of repressed genes. EZH2 imparts the catalytic functions of PRC2 by depositing H3K27me3 which is recognised and bound by the chromobox domain of the CBX factors from the cPRC1. The RING proteins RING1/RNF2 deposits H2AK119ub1 which facilitates repression of the chromatin by histone compaction. The ncPRC1 comprising RYBP then recognises the histone ubiquitination and exerts H2AK119ub1 for further chromatin compaction and more stable repression. These controlled mechanisms maintain a gene repressed thoroughly during lineage commitment thus conserving the identity of the cells during differentiation.

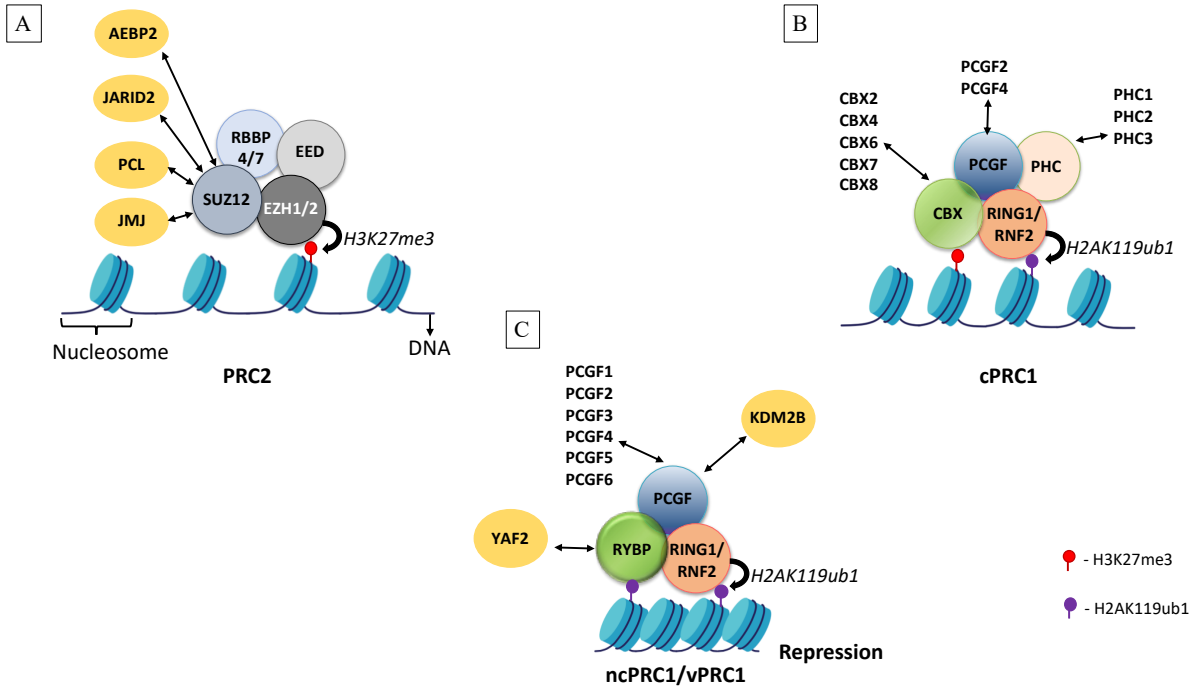


Figure 4: Composition and activities of different PRC complexes. core-members and their main interacting partners are indicated

Subunit compositions of the core of ***(A)*** *PRC2,* ***(B)*** *cPRC1 and* ***(C)*** *ncPRC1 are represented in the schematic illustration. SUZ12, EZH2, EED and RBBP4/7 form the core of the PRC2 complex. The PRC2 complex deposits H3K27me3 for gene repression. The cPRC1 complex contains CBXs, PCGFs, RING1/RNF2 and PHCs as the core complex members. cPRC1 can deposit H2AK119ub1 for chromatin compaction and gene repression. The core of the ncPRC1 complex includes ubiquitination binding RYBP, PCGFs and RING1/RNF2. The RING1/RNF2 in the cPRC1 and ncPRC1 can deposit H2AK119ub1 and are capable to cause chromatin compaction and gene repression.* ***Abbreviations:*** *PRC2: Polycomb repressive complex 2, cPRC1: canonical Polycomb repressive complex 1, ncPRC1/vPRC1: non canonical/variant Polycomb repressive complex 1.*

1.7 Role of RYBP in mouse embryogenesis

In the last few years, the loss of function mutations of several *PcG* genes and the use of high-throughput experiments like RNA-seq and ChIP-seq have revealed the major target genes of different PRCs in the regulation of developmental genes relating to different lineages.

Rybp is essential for mammalian development as the *Rybp* knock out homozygous mice were embryonic lethal during the peri-implantation stages and a portion of the *Rybp* heterozygous mutant mice presented neural tube defects and exencephaly (Pirity et al., 2005). Further *in vivo*

studies have demonstrated the important role of RYBP in the development of organ systems such as the central nervous system (Pirity et al., 2005), hematopoietic system (Calés et al., 2016), testis development (Tian et al., 2020) and the formation of the eye (Pirity et al., 2007). Due to the limitations poised by early embryonic lethality of the homozygous mice, *in vitro* based differentiation model systems were preferred to be utilized for analysing the role of RYBP during early lineage commitment (Ujhelly et al., 2015; Kovacs et al., 2016; Henry et al., 2020).

1.8 The role of *Rybp* in cardiac development

In my thesis work, I utilized wild type and *Rybp* null mutant ES cells. The mutant ES cells proliferate normally, maintain pluripotency and initiate differentiation towards multiple lineages (Ujhelly et al., 2015; Kovacs et al., 2016; Henry et al., 2020) making this cell line suitable for differentiation-based studies to elucidate the functions of RYBP.

We have previously identified that mouse ES cells lacking *Rybp* could not form beating CMCs upon *in vitro* cardiac differentiation (Ujhelly et al., 2015). The expression of several key cardiac transcription factors including cardiac progenitor formation markers *Isl1* and *Tbx5* were deficient in the $Rybp^{-/-}$ CMCs in comparison to the wild type during the time course of *in vitro* cardiac differentiation. The deficient expression of *Isl1* and *Tbx5* are connected to the formation of CHD conditions *in vivo*, in mice (detailed in chapter 1.1). Moreover, Cardiac troponin T2 (*Tnnt2*), a major sarcomere component of wild type CMCs was amongst the most downregulated genes in the *Rybp* null mutant, suggesting that these gene expression changes were likely to contribute to the contractility defect of the *Rybp* mutant cell line (Ujhelly et al., 2015). One of the most strikingly downregulated genes in the *Rybp* null mutant cells was *Plagl1*, a key cardiac transcription factor identified to affect chamber specification in the developing mouse heart (Yuasa et al., 2010).

1.9 Overview of the regulatory activities of RYBP

RYBP is a moonlighting protein, which exerts different functions based on its versatile interacting partners (Neira et al., 2009). As specified earlier, RYBP is a member of the ncPRC1s, which functions as a repressor of genes distinctive to multiple lineages during developmental process (Figure 5A) (Garcia et al., 1999). Although as part of the ncPRC1.3 and the ncPRC1.5 complex, RYBP can also exert activation functions (Figure 5B) (Gao et al.,

2014). The interaction of Autism related Autism susceptibility candidate 2 (AUTS2) and Casein kinase 2 (CK2) with the ncPRC1.3 and ncPRC1.5 complex was key in exerting the activation functions of the complexes in central nervous system (Gao et al., 2014). Further, in ES cells, the ncPRC1.3 and ncPRC1.5 complexes were identified to interact with Testis expressed 10 (TEX10) and E1a binding protein p300 (P300) to activate gene expression (Zhao et al., 2017).

ChIP-seq experiments displayed the binding of RYBP at various genomic loci independent to the binding of its PRC1 co-factor RNF2 indicating that polycomb independent regulatory activities of RYBP does also exist (Morey et al., 2015; Bajusz et al., 2018). Recent studies have revealed that the repressive activities of RYBP depends on the ability of RYBP to recognise and bind to H2AK119ub1- a repression mark and the initiation of further compaction upon binding (detailed in chapter 1.5) (Rose et al., 2016; Zhao et al., 2020; Barbour et al., 2020). No consensus DNA binding has been established for RYBP yet, but the protein is able to associate with DNA binding transcription factors such as Pluripotency factor POU domain, class 5 transcription factor 1 (POU5F1, also called as OCT4), E2F transcription factors 2 and 3 (E2F2 and E2F3) and YY1 transcription factor (YY1). The association of RYBP with these transcription factors generally lead to the activation of the targeted gene loci. For example, RYBP associated with OCT4 to activate lysine (K)-specific demethylase 2B (*Kdm2b*), a histone demethylase which can recruit PRCs to developmental genes in ES cells (Figure 5C) (Li et al., 2017; He et al., 2013). RYBP is also demonstrated to bridge the interaction between E2F and YY1 transcription factors to activate Cell division cycle 6 (*Cdc6*) (Figure 5D) (Schlisio et al., 2002).

Since the emerging studies showed the connections between the expression of RYBP and lineage commitment, the diverse associations of RYBP could impact its distinct roles during differentiation processes.

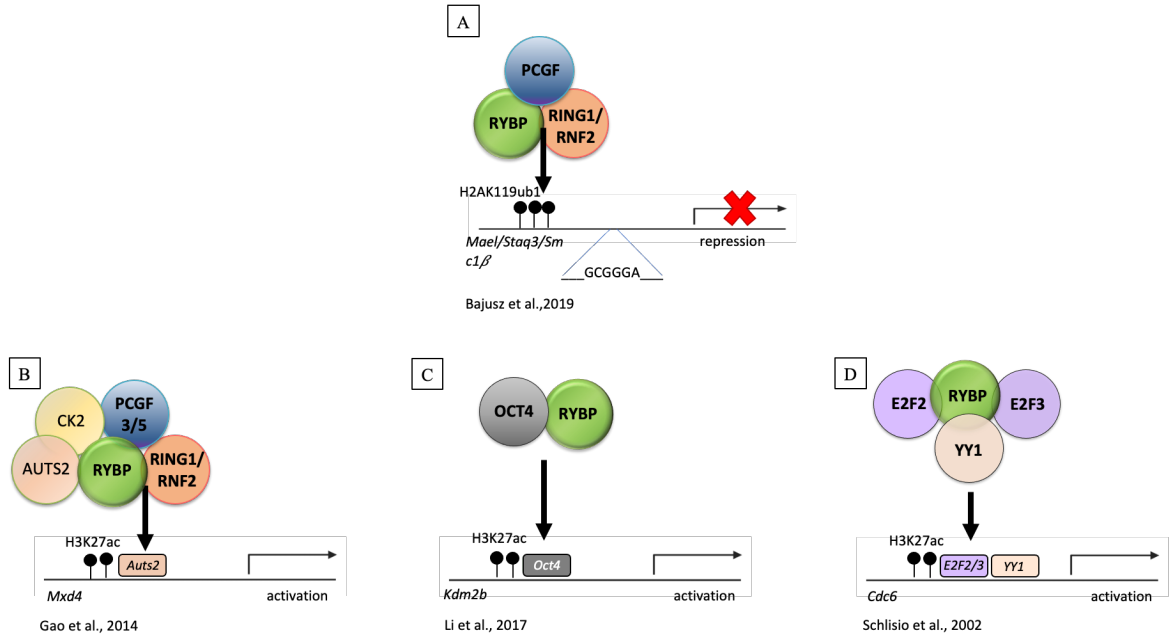


Figure 5**: **Different regulatory activities of RYBP

*Schematic illustration representing: **(A)** the polycomb dependent repression function of RYBP (Bajusz et al., 2019), **(B)** the ncPRC1.3 and ncPRC1.5 complex in which RYBP is a member of can exert activation mechanism (Gao et al., 2014), **(C)** RYBP association with OCT4 activated pluripotency genes (Li et al., 2017) and **(D)** RYBP associates with YY1 and either E2F2 or E2F3 transcription factors to activate Cdc6 expression (Schlisio et al., 2002).*

1.10 Relativeness between the functions of RYBP and PLAGL1

Genome wide transcriptomics of the wild type and *Rybp* null mutant ES cells and derived CMCs revealed altered expression of several cardiac genes crucial for the functional morphogenesis of a developing heart (Ujhelly et al., 2015). *Plagl1* was one of the most down regulated genes in both the *Rybp* null mutant ES cells and derived CMCs. Intriguingly, the *Plagl1* homozygous mice was also embryonic lethal as *Rybp* and the heterozygous mice resembled the neural tube defects of the *Rybp* heterozygous mice (Yuasa et al., 2010). *Plagl1* is also shown to be expressed in the cerebellum of the brain and showed exencephaly defects which were also seen in the *Rybp* heterozygous mice (Pirity et al., 2005; Yuasa et al., 2010). During organogenesis, RYBP and PLAGL1 are co-expressed in the developing organ systems such as the central nervous system, the heart and the eye (Table 2) (Valente & Auladell, 2001; Miró et al., 2009; Pirity et al., 2005; Garcia et al., 1999; Pirity et al., 2007).

PLAGL1 is also called as Zinc finger protein inducer of apoptosis and cell cycle arrest (ZAC1) due to its roles in apoptosis (Spengler et al., 1997) similar to RYBP (Stanton et al., 2007). Both RYBP (Tan et al., 2017; Voruganti et al., 2015; Zhu et al., 2017) and PLAGL1 (Abdollahi et al., 1999; Bilanges et al., 1999) have been identified to function as tumor suppressors as well. Both RYBP and PLAGL1 are previously identified to physically interact with tumor suppressor Transformation related protein 53 (TRP53, also called as P53). RYBP can modulate the stability of P53 by inhibiting ubiquitination of the protein (Chen et al., 2009). On the other hand, PLAGL1 can interact with p53 to activate the expression of Cyclin-dependent kinase inhibitor 1A (*Cdkn1a*, also called as p21) to regulate cell cycle exit (Liu et al., 2008; Benedetti et al., 2017). Taken together, these suggested a possible genetic or biochemical connection between the two proteins.

	Organ	*Rybp*		*Plagl1*		*Reference for Rybp*	*Reference for Plagl1*
		Mouse Embryo	Mouse Adult	Mouse Embryo	Mouse Adult		
CENTRAL NERVOUS SYSTEM	Cerebral cortex	+	ND	+	±	*Pirity et al, 2005*	*Valente et al, 2001, Miró X et al,2009*
	Cerebellum	+	ND	-	+	*Garcia et al, 1999*	*Valente et al, 2001*
	Forebrain	++	ND	+	ND	*Garcia et al, 1999*	*Valente et al, 2001*
	Midbrain	+	ND	++	ND	*Pirity et al, 2005*	*Valente et al, 2001; Alam et al, 2005*
	Hindbrain	++	ND	++	ND	*Garcia et al, 1999*	*Valente et al, 2001; Alam et al, 2005*
	Cortical plate	++	ND	+	ND	*Pirity et al, 2005*	*Valente et al, 2001; Alam et al, 2005*
	Marginal zone	++	ND	+	-	*Pirity et al, 2005*	*Valente et al, 2001*
	Hippocampus	++	ND	+	++	*Pirity et al, 2005*	*Valente et al, 2001; Alam et al, 2005*
	Ventricular zone	-	ND	++	±	*Pirity et al, 2005;*	*Valente et al, 2001; Alam et al, 2005*
	Sub-ventricular zone	+	ND	+	±	*Pirity et al, 2005*	*Valente et al, 2001; Alam et al, 2005*
	Choroid plexus	ND	ND	++	+		*Valente et al, 2001*
HEART	Atrium and Ventricle	++	+	++	++	*Pirity et al, 2005; Garcia et al, 1999; Ujhelly et al, 2015*	*Yuasa et al, 2010; Tsuda et al, 2004; Alam et al, 2007*
EYE	Retina	++	++	+	+	*Pirity et al, 2007*	*Alam et al, 2005*
	Lens	++	-	+	±	*Pirity et al, 2007*	*Alam et al, 2005*
	Cornea	++	±	ND	ND	*Pirity et al, 2007*	*Alam et al, 2005*
	Optic nerve	±	+	ND	+	*Pirity et al, 2007*	*Alam et al, 2005*

Table 2: Gene expression of Rybp and Plagl1 in mouse embryonic tissues

The expression of Rybp and Plagl1 in in vivo mouse tissues based on previously analysed data from RNA-in situ hybridisation experiments are presented in the table. Rybp and Plagl1 are co-expressed in the same tissue types in the CNS, heart and the eye. + denotes presence of the respective factors, ++ denotes stronger expression, - denotes no expression and ± denotes weak expression of Rybp or Plagl1. ***Abbreviation****: ND-no data.*

1.11 *Plagl1*, as a key cardiac transcription factor

Mouse *Plagl1* has been identified as a transcription factor with diverse functions, expressed at various developing tissues during embryonic developmental and adult stages (Table 2) (Valente & Auladell, 2001; Alam et al., 2005). *Plagl1* expressed strongly in the forelimb, hindlimb, liver primordium, neural tube, neural retina, primordial heart, epithalamus, pituitary lobe, choroid plexus, cortical plate, marginal zone, hippocampus, atrium, ventricle in the developing embryo (Valente & Auladell, 2001; Valente et al., 2005; Alam et al., 2005) revealing the role of *Plagl1* in the organogenesis of various lineages.

PLAGL1 is determined to be a cardiac transcription factor with chamber specific expression in the developing heart (Tsuda et al., 2004; Yuasa et al., 2010). *Plagl1* heterozygous exhibited atrial and ventricular septal defects and improper chamber specification in the E15.5 hearts (Yuasa et al., 2010). *Plagl1* is shown to be regulated by the cardiac transcription factors NKX2-5 in mouse and by MEF2C in rat mesenchymal cells (Yuasa et al., 2010; Czubryt et al., 2010). These studies established PLAGL1 as a cardiac transcription factor, regulated by cardiac progenitor transcription factors during mammalian heart development.

1.12 Relevance of *Plagl1* towards diseases

Plagl1 encodes for zinc finger type transcription factor with anti-proliferative activity and is a presumptive tumour suppressor gene on 10q24 which expression is frequently lost in various neoplasms. Alterations of *Plagl1* expression were profoundly classified in various cancers such as breast, ovarian primary tumors and also in tumor derived cell lines, basal cell carcinoma and Extraskeletal myxoid chondrosarcoma (EMC) (Cvetkovic et al., 2004; Jacobs et al., 2013; Kowalczyk et al., 2015; Li et al., 2014; Ribarska et al., 2014). Allelic deletions of *Plagl1* have been implicated in different cancers as well (Kowalczyk et al., 2015). Like most imprinted genes, a differentially methylated region (DMR), rich in CpG sequences, are influencing *Plagl1* transcription. An *in vitro* model for *Plagl1* gene regulation demonstrated that methylation of the CpG islands induces heterochromatin modification that represses gene transcription (Varrault et al., 2001).

Furthermore, the biallelic expression of *Plagl1* from an alternate promoter is associated to transient neonatal diabetes mellitus (TNDM) (Hoffmann, 2015). The expression of both *Plagl1* and the non-coding RNA (ncRNA) in its locus, Hydatiform mole associated and imprinted (*Hymai*) were identified to be higher in TNDM conditions (Arima et al., 2001).

These studies indicated the important role of *Plagl1* not only in normal mammalian development but also in disease conditions as well.

2. AIMS

The aims of this thesis study were to understand the connections between epigenetic factor RYBP and cardiac transcription factor *Plagl1* during *in vitro* cardiomyogenesis and to broaden our knowledge about the functions of RYBP during cardiac development. Our focus was directed towards unravelling the specific molecular mechanisms by which RYBP affected the regulation of key cardiac transcription factors such as *Plagl1* and to understand the critical role of *Plagl1* in the formation of contractile CMCs.

The detailed aims of the thesis were:

i. To examine the expression of *Plagl1* and compare it with the expression of *Rybp* during *in vitro* cardiac differentiation.

ii. To characterize and compare the protein localization of RYBP and PLAGL1.

iii. To identify putative regulatory elements in the *Plagl1* genomic locus.

iv. To identify the nature of regulatory mechanism which RYBP exerts on the *Plagl1* locus.

3. MATERIALS AND METHODS

3.1 Cell culture techniques

3.1.1 Cell lines and culture conditions

Mouse (129SV/Ola) R1 ES cells (Nagy et al., 1993) (mentioned as wild type or *Rybp*$^{+/+}$) and D11 ES cells (mentioned as *Rybp* null mutant or *Rybp*$^{-/-}$) (M K Pirity et al., 2005)(Figure 6) were thawed on mitomycin C (Mit C; Sigma, Cat.No M0503) inactivated mouse embryonic fibroblast (MEF) layer and cultured on 0.1% gelatin (Gelatin from bovine skin, Sigma, Cat.No G-9391) coated tissue culture plates as described in Magin et al (Magin et al., 1992). The cells were maintained in Dulbecco's modified eagle's medium (DMEM (1x) + Gluta MAX^{TM-1} Dulbecco's Modified Eagle Medium, Gibco, Cat.No 31966-021) containing 15% Fetal Bovine Serum (FBS) (APS, Cat.No S-001A-USDA grade), 0.1 mM non-essential amino acids (MEM Non-Essential Amino Acids (100x), Corning, Cat. No 34319012), 0.1 mM ß-mercaptoethanol (2-Mercaptoethanol, Gibco, Cat.No 31350-010), 1% sodium pyruvate (Sodium Pyruvate (100mM) (100x), Gibco, Cat. No 11360-039), 1% glutamine (L-Glutamine (200 mM) Gibco, Cat.No 25030-024), 50 U/ml penicillin/streptomycin (Penicillin/ Streptomycin (100x), Gibco, Cat.No 15140-122) and 100 U/ml Leukemia inhibitory factor (LIF, ESGRO, Chemicon/ Millipore, Billerica, MA, USA). The cells were passaged prior to reaching 70% confluence (approximately every second day). ES cells were cultured on gelatin coated dishes for at least three passages prior to differentiation to deplete potentially present MEF cells from the ES cell culture. Cells were cultured in humidified conditions containing 5% CO_2 at 37°C. The cells were grown with fresh ES cell media supplemented every day.

Human Embryonic Kidney (HEK) 293T cells was used for PLAGL1 protein assays. HEK293T cells were maintained in Dulbecco's modified eagle's medium (DMEM (1x) + Gluta MAX^{TM-1} Dulbecco's Modified Eagle Medium, Gibco, Cat.No 31966-021) containing 10% FBS (Gibco, Cat.No 10500-064), 0.1mM non-essential amino acids (MEM Non-Essential Amino Acids (100x), GIBCO, Cat.No 11140-035), 1% sodium pyruvate (Sodium Pyruvate (100 mM), Gibco, Cat.No 11360-039) and 50 U/ml penicillin/streptomycin (Penicillin/Streptomycin (100x), GIBCO, Cat.No 15140-122). The cells were passaged before the confluency reached 90% (approximately every 2-3 days). Medium was changed every second day. Cells were cultured in humidified conditions containing 5% CO_2 at 37°C.

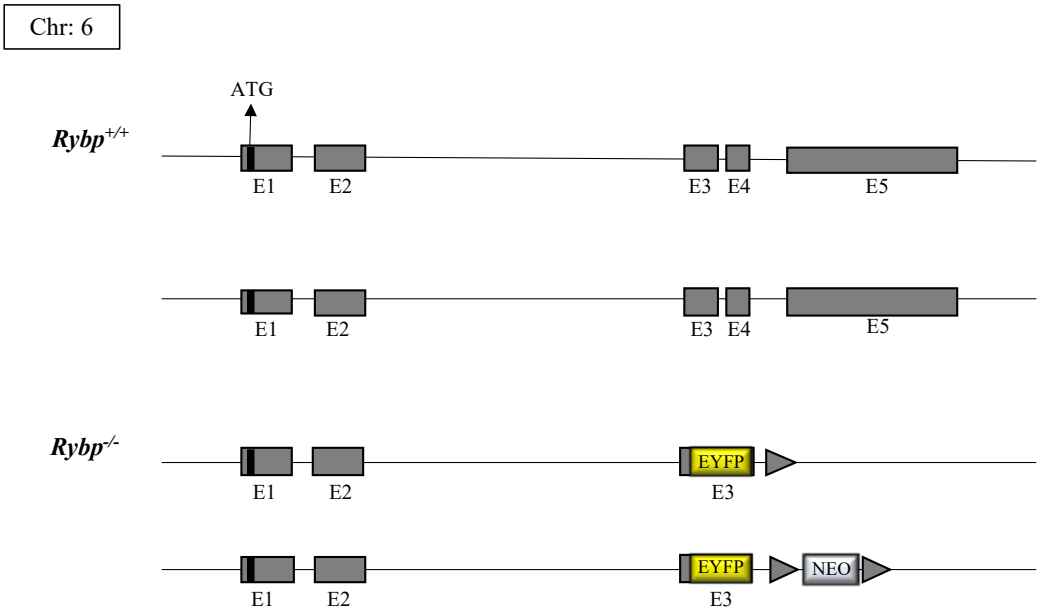


Figure 6: ES cell lines used in this study

Rybp genomic locus (Chr6: 100228565-100287358) contains 5 exons in the wild type (Rybp$^{+/+}$) ES cells. In the Rybp null mutant (Rybp$^{-/-}$) ES cells, the 3' of exon 3, exon 4 and exon 5 are replaced by a donor cassette containing Enhanced Yellow Fluorescent Protein (EYFP) followed by a floxed neomycin-phosphotransferase (NEO) cassette (Pirity et al., 2005).

3.1.2 *In vitro* cardiac differentiation of mouse embryonic stem cells

Mouse ES cells were harvested as single cell suspension using 0.05% (wt/vol) trypsin (Trypsin-EDTA (1x) 0,05% / 0,02% in D-PBS, GIBCO, Cat.No 15400-054) and then the cell number was calculated using a Burker chamber. The cell number was diluted to 50 cells/µl in suspension and 20 µl droplets of cell suspension were dispensed to lids of bacterial dishes where each droplet contained around 1000 cells, and then the cells were let to form EBs by the HD method as described in Keller et al. (Keller, 1995) (Figure 7). The EBs were harvested on the second day and plated into cell culture dishes (60 mm, Corning, Cat.No 430196) coated with gelatin containing ES medium (described in 3.1.1) without LIF. The medium was changed every second day and the cells were cultured to a maximum of 21 days. The cells were harvested for further analysis at different time points of cardiac differentiation: day 0, 2, 7, 10, 14 and 21 (labelled as d0, d2, d7, d10, d14 and d21). Day 0 represents pluripotent stem cell stage, day 2 represents the EB stage, day 7 and day 10 represents early and late cardiac progenitor stages respectively and day 14 and day 21 represents the terminal stage of *in vitro* cardiac differentiation.

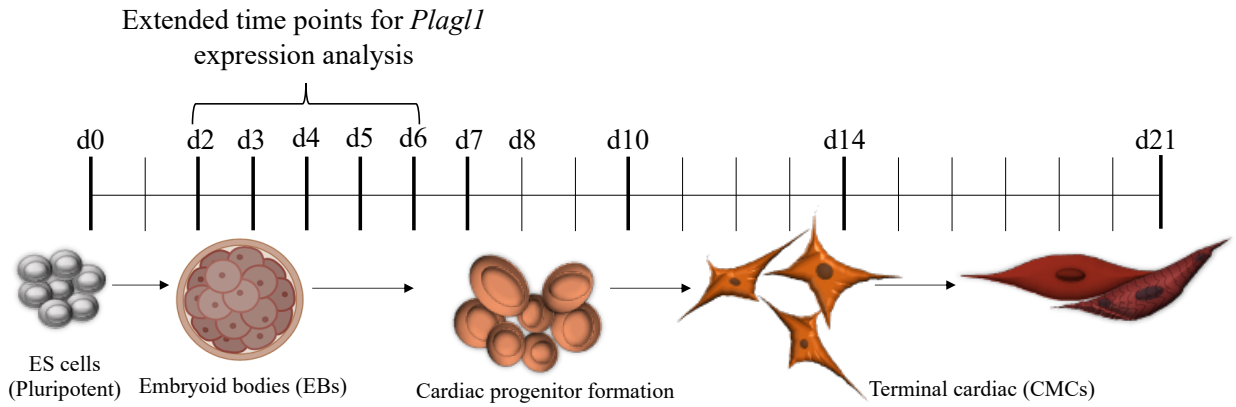


Figure 7: In vitro cardiac differentiation

CMCs were differentiated in vitro from ES cells through EB formation by using the HD method. Cardiac colonies were grown for maximum 21 days, sampled for mRNA expression analysis (qRT-PCR) and fixed for ICC analysis at day (d) 0, 2, 7, 10, 14 and 21 (indicated in bold). Samples were derived earlier at d0, d8 and d14 for whole genome transcriptomics as described previously (Henry et al., 2020; Ujhelly et al., 2015). For the analysis of the initial time points of Plagl1 expression, samples were derived at an extended interim time points between day 0 till day 7 (i.e., d2, 3, 4, 5, and 6) for qRT-PCR and ICC analysis. ***Abbreviations:*** *ES cells: Embryonic stem cells, EBs: Embryoid bodies, HD: Hanging drops, CMCs: Cardiomyocytes, ICC: Immunocytochemistry, qRT-PCR: quantitative real-time polymerase chain reaction.*

3.1.3 Calcium Phosphate transient transfection method

Calcium Phosphate method (Kingston et al., 2003) was used to transiently transfect HEK293T cells for reporter assays and protein overexpression for protein stability assays (methods 3.3.2) and co-immunoprecipitation (Co-IP, methods 3.3.3) analysis. HEK293T cells were seeded at a density of 1 x 10^6 cells per 6 cm tissue culture dishes and maintained as described above. 5 hours before transfection the cells were fed with fresh medium. The transfection mix were prepared by diluting the required plasmids in 0.1 mM Tris-EDTA (Trizma base, Sigma, Cat.No T1503) buffer and 2.5 M Calcium chloride ($CaCl_2$, Sigma, C-3881) and 2X HEPES buffered saline (HBS, Sigma, Cat.No H3375) dropwise by bubbling the solution using Pasteur pipette to provide oxygen for the mixture. The transfection mix was added to the cells dropwise and the cells were then maintained with the transfection mix in humidified conditions. 16 hours after the transfection, fresh media was provided to the cells and after 40 hours the cells were washed twice with 2ml of 1X PBS on ice and then harvested for whole cell protein lysate using cell lysis buffer (Cell culture lysis 5X reagent, Promega, Cat.No E153).

3.1.4 The luciferase reporter assay system

HEK293T cells was transfected with Calcium Phosphate transient transfection method as mentioned above (3.1.4) (Figure 8). The transfected cells were harvested for their protein lysates 40 hours after transfection with 1X Passive lysis buffer (1X PLB) provided by the luciferase assay kit (Dual Luciferase Reporter Assay System, Promega, Cat.No E1910) (Figure 8). Concentration of the whole cell lysate was determined by the Bradford's method (5X Bio-Rad Protein Assay Dye reagent concentrate, Cat.No 5000006) according to the manufacturer's instructions. Protein concentrations were measured from OD_{600} taken in UV spectrophotometer (WPA Photometer UV110 Cambridge, UK, Cat.No RS232). The concentration of the lysates was then determined by Bradford's method (Bradford, 1976) using Bovine Serum Albumin (BSA, VWR, Cat.No G22361V) as the standard. 20 µg of the protein lysates were measured from each transfection with 100 µl of Luciferase Assay Reagent II (LAR II, provided with the kit). Luciferase activity was recorded with Perkin Elmer TopCount NXT Luminometer in dark conditions. Each measurement was recorded in triplicates.

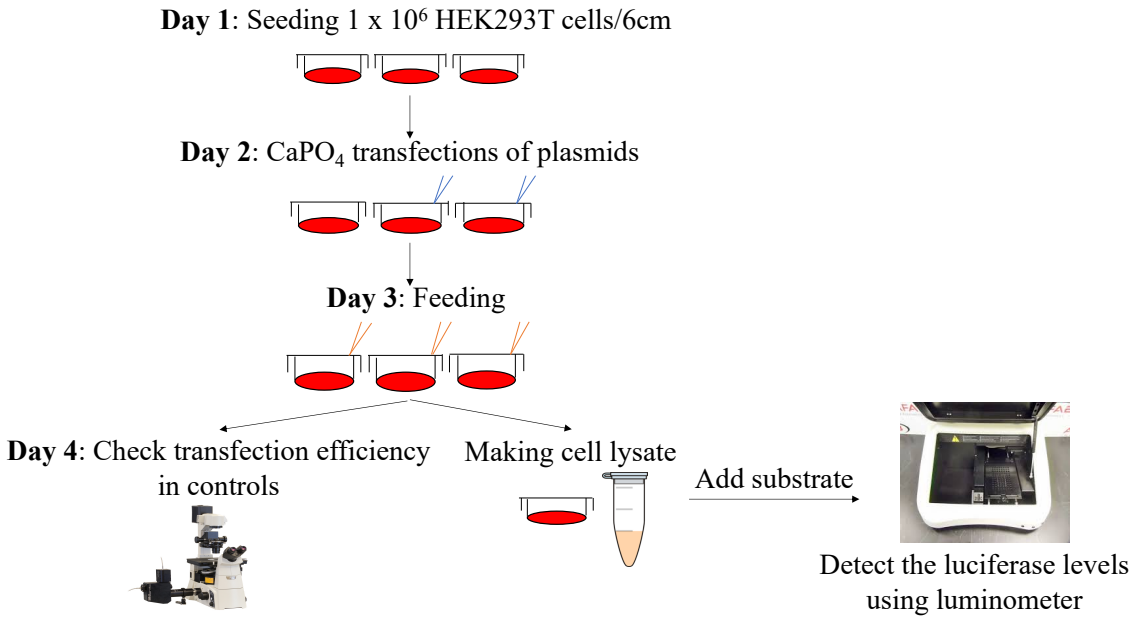


Figure 8: Flow chart of the working model for luciferase reporter assay

Schematic representation of the workflow for luciferase reporter assay. 1 x 10^6 HEK293T cells were seeded in 6 cm petri dishes and on d2 were transfected with the required plasmids by Calcium Phosphate method. 16 hours after transfection the cells were fed with fresh media and after 40 hours the cells were harvested for the protein cell lysates. The cells transfected with EGFP was checked to measure the transfection efficiency and the samples were prepared for luciferase measurement as described in 3.1.4.

3.1.5 Inhibition of PRC1 activity

Inhibition of PRC1 activity was performed to analyse the PRC1 dependent and independent activities of RYBP in promoter assays. 16 hours after transfection of the required plasmids by Calcium Phosphate method (detailed in 3.1.4), HEK293T cells were fed with growth media supplemented with 50 µM of PRC1 inhibitor, PRT4165 (PRT4165, Sigma, Cat.No NSC600157) as previously reported by Ismail et al. and Gracheva et al. (Ismail et al., 2013; Gracheva et al., 2016). The cells were maintained with PRT4165 supplemented media for further 1 hour after treatment and the whole cell lysates were procured. The cell lysates were then prepared for luciferase reporter assay as described in 3.1.4.

3.2 Molecular biology techniques

3.2.1 Quantitative real-time PCR (qRT-PCR)

Relative quantification of mRNA expression during *in vitro* cardiac differentiation was performed using quantitative real-time PCR (qRT-PCR). Total RNA was isolated from the harvested cells at the required time points of *in vitro* cardiac differentiation (described in 3.1.2) using GeneJET RNA Purification Kit (Thermo Scientific, Cat.No K0732) according to the manufacturer's instruction. Reverse transcription PCR for the cDNA synthesis from the isolated RNA was performed using Applied Biosystems High-capacity cDNA Reverse Transcription Kit (Invitrogen Life Technologies, Cat.No 4368814) according to the manufacturer's instructions. qRT-PCR analysis was performed with SYBR green master mix (SYBR® Select Master Mix for CFX, Applied Biosystems, Cat.No 4472942) using Bioer LineGene Real-time PCR system (Bioer, China).

Relative mRNA expression changes were determined using the ΔΔCt method. The threshold cycle (Ct) values for each gene were normalized to the expression level of *Hprt* (Hypoxanthine guanine phosphoribosyl transferase I) as internal control. The data is presented as fold expression changes normalized to wild type d0. The primers used in this study are listed in Table 3.

Gene Name	Forward Primer sequence	Reverse primer sequence
Hprt	5'- AGTCCCAGCGTCGTGATTAG-3'	5'-GCAAGTCTTTCAGTCCTGTCC-3'
Rybp	5'-TTAGGAACAGCGCCGAAG-3'	5'-GCCACCAGCTGAGAATTGAT-3'
Plagl1 ex 1/2	5'-AGCAAGGCTTCTCACAGGC-3'	5'-GTGAGGTACTTCCTTCAGCATCTTG-3'
Plagl1 ex 4/5	5'-GATTGCTTCAGCGTGCCATCG-3'	5'-ACTCCTCTGACTCCTATGCAAA-3'
Plagl1 ex 10/11	5'ATGGCTCCATTCCGCTGTC-3'	5'-CTCAGCCTTCGAGCACTTGAA-3'
Hymai	5'-AAGTAGTGACAACCGGGGCCAT-3'	5'-GAACACAAATCACCTCTTCCC-3'
Plagl1it	5'-GCAACCCCACACATCCTTAAGC-3'	5'-GAACATTCACAGAAACTCAAGG-3'

Table 3: Primers used in qRT-PCR reactions

3.2.2 Chromatin Immunoprecipitation and qRT-PCR (ChIP-qRT-PCR)

Chromatin immunoprecipitation (ChIP) was performed by using EpiXplore ChIP kit, (Clonetech, Cat.No 632011) according to manufacturer's instructions. In brief nuclear extraction from ES cells and d7 cardiac differentiated cells from 10 cm plates was carried out by carefully lysing the cytoplasm and nuclei using the lysis buffers (provided in the kit) and subsequent shearing of the DNA was performed using an ultrasonicator (Ultrasonic homogenizer 3000, BioLogics) at 4x30 s cycles, 60 pulse and 20 kHz. The sheared DNA was loaded into 1 % Agarose gel electrophoresis (AGE) and the size of the sheared chromatin was seen between 200 bp to 800 bp (ideal for IP and qRT-PCR). The sheared DNA was then incubated with prewashed magnetic beads (Mag Capture beads, Clonetech, Cat.No 632577) under gentle rocking for 4 hours at 4°C. The wash steps were carried out according to the manufacturer's instructions with the help of a magnetic stand. The eluted immunoprecipitated chromatin was then treated with RNase A and Proteinase K (provided in the kit).

The immunoprecipitated chromatin was then used for qRT-PCR using SYBR green as described in 3.2.1.

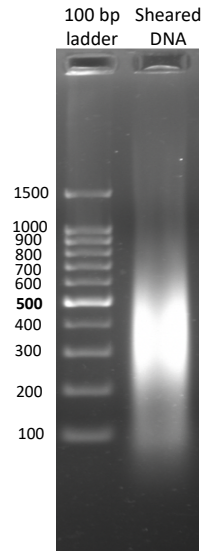


Figure 9: Sheared chromatin of d0 wild type ES cells used for ChIP-qPCR

Wild type ES cells were harvested from 10 cm petri dishes and the isolated nuclear fractions with the chromatin was sheared by sonication. The sonicated sheared chromatin was loaded in 2% AGE with 100 bp ladder (GelPilot 100 bp ladder, Qiagen, Cat.No 239045) in the left. The size of the ladder bands is labelled accordingly. The isolated chromatin was sheared at an average size of 200-800 bp in length which is ideal to use in qRT-PCR. ***Abbreviations:*** *kb: kilobase, bp: base pair.*

3.2.3 Molecular cloning, transformation and confirmation

All enzymes required for molecular cloning of promoter and cDNA constructs of interest were performed using NEB enzymes. Amplification of the promoter regions was done using BAC clone (RP23-259L24 BAC clone for *Plagl1* promoters) or using wild type genomic DNA (gDNA) isolated from ES cells (for *Tnnt2* promoter) as the template. One Taq Hot Start DNA Polymerase kit (NEB, Cat.No M0481S) was used for the amplification of DNA following the manufacturer's instructions.

The *Plagl1 P1* promoter (4612 bp) was PCR amplified using 5' and 3' HindIII site containing primers from RP23-259L24 BAC (RPCI23-259L24, BACPAC resource, RPCI) construct (Table 4). The amplified DNA was gel eluted using QIAquick Gel Extraction kit (Qiagen, Cat.No 28706) according to the manufacturer's instructions and cloned into the HindIII site of the MCS (multiple cloning site) in the pGL4.20 vector (Figure 10A) (pGL4.20 (luc2/Puro) vector, Promega, Cat.No E6751).

The *Plagl1 P2* promoter (1821 bp) was also amplified and cloned using 5' and 3' HindIII site containing primers (Table 4). The *P2* promoter was cloned into pGL4.20 vector (Figure 10B) as mentioned above.

The *Plagl1 P3* promoter containing pGL3-mZac1pr (henceforth called as the *Plagl1 P3* promoter) construct (Figure 10C) was a kind gift from Dr. Michael Czubryt, Institute of Cardiovascular Sciences, University of Manitoba, Canada.

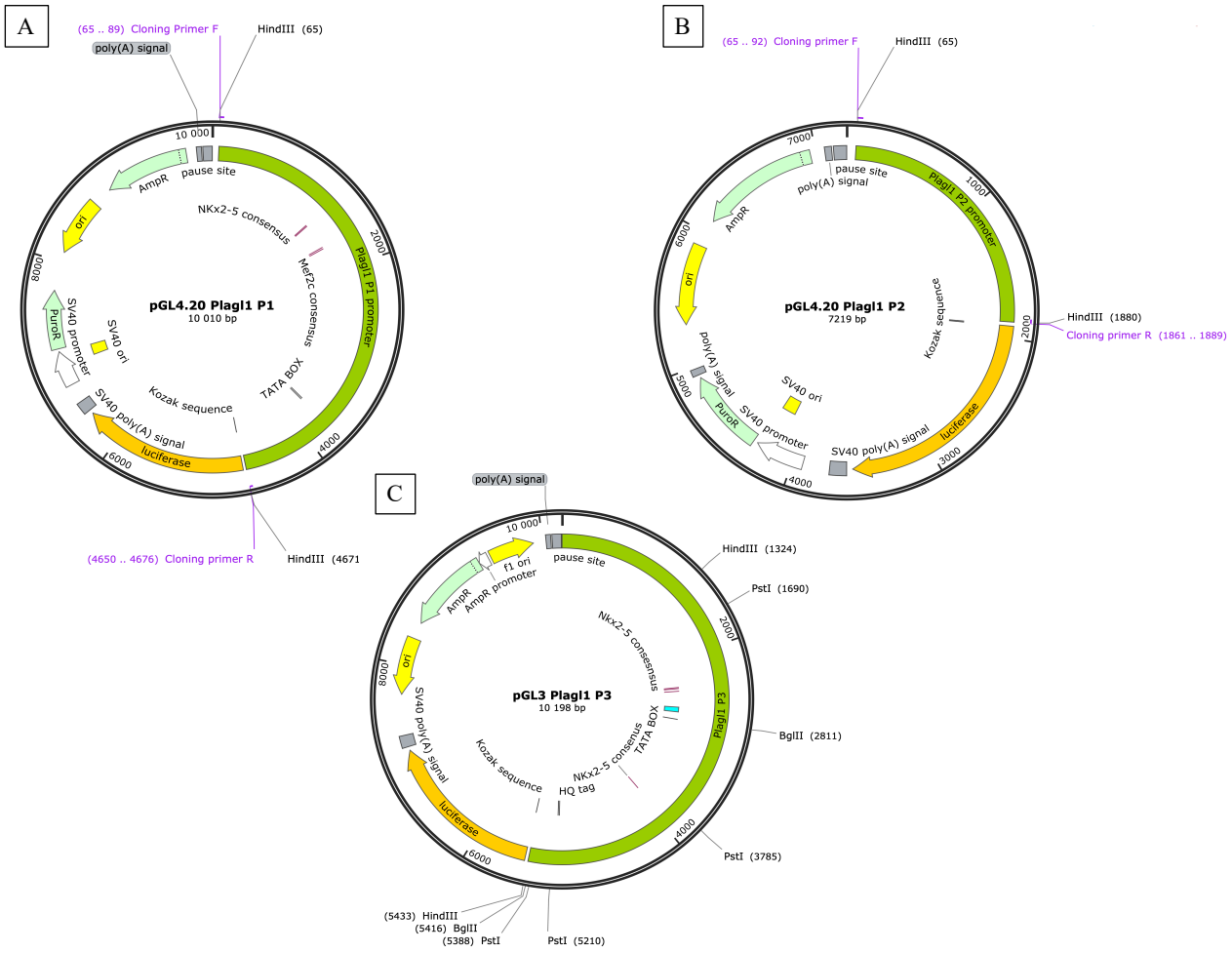


Figure 10**: **Schematic representation of the generated Plagl1 promoters containing luciferase reporter constructs

*The **(A)** Plagl1 P1 and **(B)** P2 promoter regions were cloned at the HindIII sites immediately upstream to the luciferase coding region in pGL4.20 Luc2 vector. Plagl1 P3 promoter **(C)** incorporating luciferase reporter construct was a kind gift from Dr. Michael Czubryt. The constructs were labelled as pGL4.20 Plagl1 P1, pGL4.20 Plagl1 P2 and pGL3 Plagl1 P3 according to the encompassing promoter region. The Plagl1 promoter regions are represented in green colour along with indicating cloning restriction sites, regulatory elements such as TATA box (blue box), consensus binding sites for NKX2-5 and MEF2C, and vector elements.*

The subcloning of the *P3* promoter was performed as follows. Clone a (1-2.8 kb) and f (2.8-5.4 kb) were produced by cleaving the *P3* with BglII (Figure 10C). Clone a (1-2.8 kb) was self-ligated after digestion with BglII and the 2.8-5.4 kb band was eluted and re-cloned into pGL3 empty vector at the BglII site. Clones b (1-1.3 kb) and d (1.3- 2.8 kb) were generated by HindIII digestion of clone a. Clone c (1-1.6 kb) was generated by digesting clone a with PstI and self-ligating the 6.5 kb band. Clone e (1.6-3.7 kb) construct was generated by digesting the *Plagl1 P3* promoter by PstI (Figure 10C). The 2.1 kb band after digestion with PstI was gel eluted and re-cloned into the same sites in pGL3 empty vector. Clone g (2.8-3.7 kb) and h (3.7-5.4 kb) were generated by digesting clone f with PstI and performing self-ligation and insert ligation of fragments as mentioned earlier.

The *Tnnt2* promoter (2688 kb) was PCR amplified using wild type gDNA from ES cells as template. The PCR amplicon was gel eluted and cloned into KpnI sites (Table 4) and cloned into pGL4.20 vector (Figure 11) as described above.

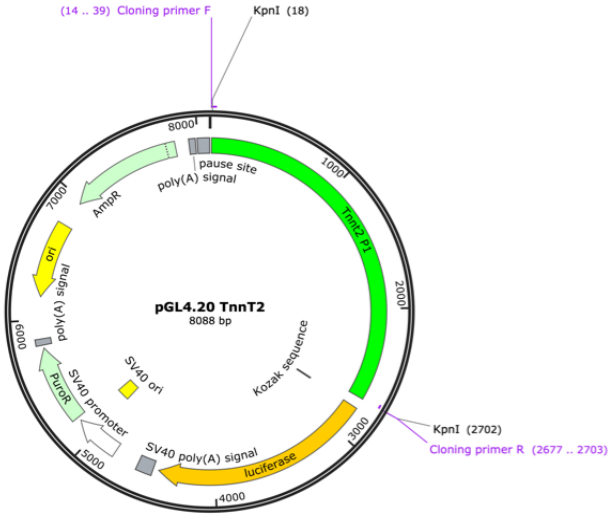


Figure 11**: **Schematic representation of the generated Tnnt2 promoter containing luciferase reporter construct

The Tnnt2 promoter was cloned at the KpnI sites immediately upstream to the luciferase coding region in pGL4.20 Luc2 vector.

Tnnt2 promoter is represented in green colour along with indicating cloning restriction sites.

cDNA overexpression constructs for *Hymai* (Figure 12A) and *Plagl1it* ncRNA (Figure 12B) were generated by PCR amplifying the ncRNAs from d14 cardiac differentiated wild type cells. The PCR amplicons were gel eluted and cloned into the XbaI site in pcDNA3.1- vector (Figure 12).

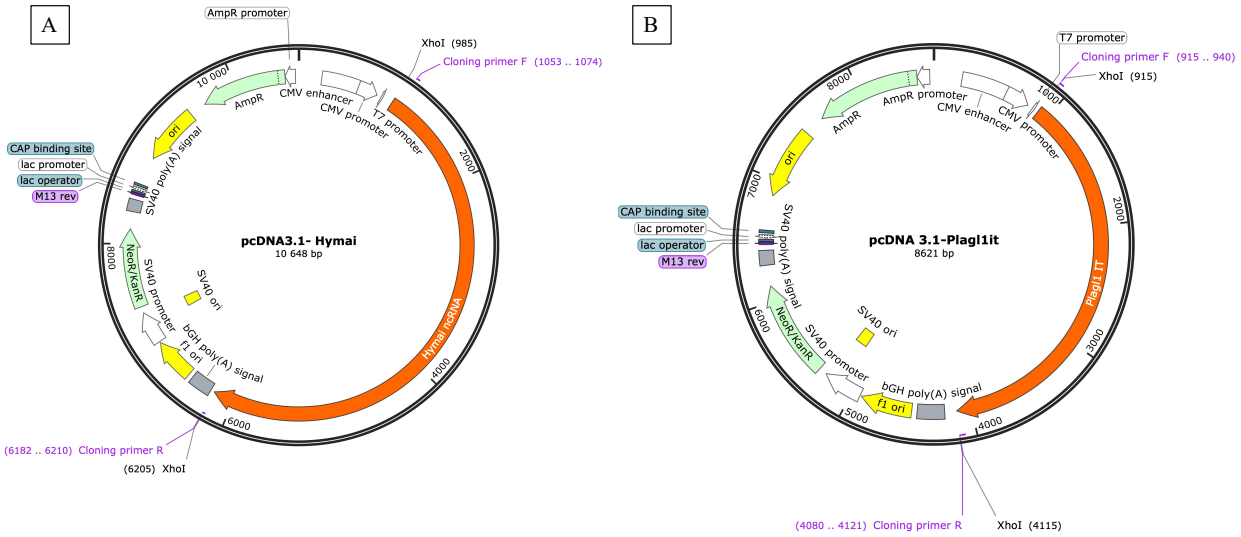


Figure 12**: **Schematic representation of the Hymai and Plagl1it cDNA overexpression constructs

*Both **(A)** Hymai and **(B)** Plagl1it ncRNA were cloned at the XhoI sites in the pcDNA3.1- vector. The cloned cDNA region is represented in orange colour along with indicating vector elements.*

cDNA overexpression constructs were generated by PCR amplifying *Nkx2-5* and *Mef2c* from cDNA pool generated from whole cell RNA isolated from d10 cardiac differentiated wild type cells. The PCR amplicons were gel eluted and cloned into the BamHI site in pRK7 FLAG vector in frame with the N-terminal FLAG tag. Both FLAG-NKX2-5 (Figure 13A) and FLAG-MEF2C (Figure 13B) constructs produced N-terminally FLAG tagged proteins.

Further confirmation of the cloned constructs was performed by orientation check of the ligated insert by restriction digestion of the plasmids and by sequencing the plasmids (Deltagene, Szeged, Hungary).

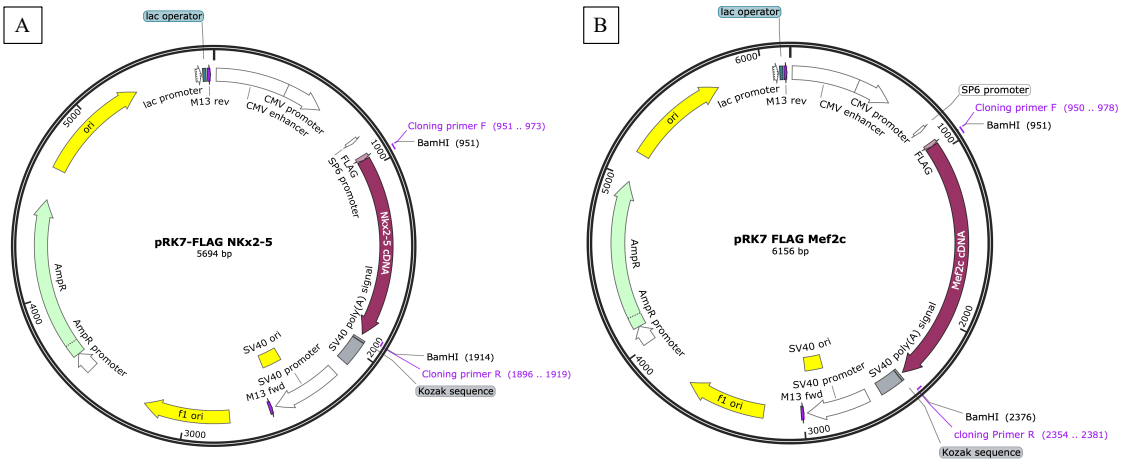


***Figure 13**: Schematic representation of the Nkx2-5 and Mef2c cDNA overexpression construct*

*Both **(A)** Nkx2-5 and **(B)** Mef2c cDNA were cloned at the BamHI sites in frame with a N-terminal FLAG tag. The cloned cDNA region is represented in maroon colour along with indicating vector elements.*

Gene Name	Forward Primer sequence	Reverse primer sequence
Plagl1 P1	5'-GCTGAAGCTTATTAACCGCCTCATTCTCA-3'	5'-TACTAAGCTTTGGGTCTGATGGTTCCATAGA-3'
Plagl1 P2	5'-TGTAAGCTTCACTTTTCCTTTTGCAAGGCAT-3'	5'-TGTCAAGCTTAAGTGTGCAGAGGGAACTT-3'
Tnnt2 promoter	5'-TGATGGTACCGGAATCTAACAGTGTCTGGA-3'	5'-TATTGGTACCCCTCCCACAAGCTTACAATCA-3'
Hymai cDNA	5'-TATTCTCGAGCCCACGGCATCTGCGATTTG-3'	5'-ACGCTCGAGAGCATGTGAGGCAAATGACAAAC-3'
Plagl1it cDNA	5'-TATTCTCGAGCCTTGCTGCACGGACAGACT-3'	5'-GAGCTCGAGAGCAGCAACTGGGTGACATGC-3'
Nkx2-5 cDNA	5'-TAATTAGGATCCATGTTCCCCAGCCCTGC-3'	5'-TATTAGGATCCCTACCAGGCTCGGATGCC -3'
Mef2c cDNA	5'-AGCAGGATCCATGGGGAGAAA AAAGATTCAGA-3'	5'-TAATGGATCCTCATGTTGCCCATCCTTCAGAG-3'

Table 4: Table of the primers used to clone the Plagl1 P1, P2, Tnnt2 promoters and cDNA overexpression constructs in this study (3.2.3).

3.2.4 Site directed mutagenesis

Site directed mutagenesis was performed using Q5 site directed mutagenesis kit (NEB, Cat.No E0554S) following the manufacturer's instructions. Primers were designed to mutate consensus sites for *Nkx2-5* and *Mef2c* at the *P3* promoter by using NEBase Changer tool (https://nebasechanger.neb.com) provided by NEB (Table 5). The primers were designed to mutate the consensus of 3 *Nkx2-5* and one *Mef2c* sites by introducing BamHI and HindIII sites respectively at the consensus to assist with screening positive mutants harbouring the right mutation. The PCR reaction was set according to the corresponding primer annealing temperature suggested by NEBase Changer tool. The KLD (kinase, ligase and DpnI digestion) enzyme (provided in the kit) was used to digest template DNA and ligation for rapid generation of mutant constructs carrying mutation for *Nkx2-5* and *Mef2c* consensus. The transformed

colonies were then screened and confirmed by BamHI and HindIII digestions for *Nkx2-5* and *Mef2c* consensus sites respectively. Seven different mutants were generated harbouring single and multiple mutants of *Nkx2-5* and *Mef2c* consensus (Figure 14). Further confirmation was performed by sequencing the plasmids (Deltagene, Szeged, Hungary) and checked for carrying the mutation with no off-target mutations in the constructs.

Gene Name	Forward Primer sequence	Reverse primer sequence
Nkx2-5 (1)	5'-CTTGAATATCCATCTTGGAAGACCAAAATG-3'	5'-CTTTTGGGTCTTTGGGGGTGG
Nkx2-5 (2)	5'-TCCCATTTCCAAGCTTGTGGGCCTCAC-3'	5'-TTTCCATTTTGGTCTTCCAAG
Nkx2-5 (3)	5'-CTACACCATGAAGCTTGGCCTTTATTC-3'	5'-CTAATGGTTCCTAGATATTG
Mef2c	5'-TCCTTAGAATGGGGGGACACTGAAAATGAAATGAAATCCTGAGACTTTGG-3'	5'-TCCTGAGTGAGTGATAGAGATCTGCCAATTGAGCCATCTGCTTCATTC-3'

Table 5: Table of the primers used to mutate Nkx2-5 and Mef2c sites at the P3 promoter in this study (3.2.4).

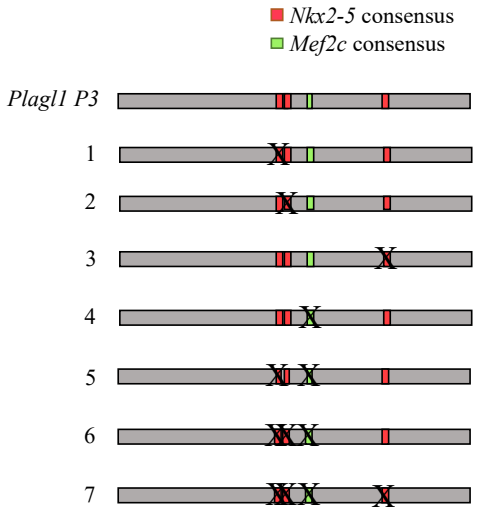


Figure 14: Schematic illustration of the position of Nkx2-5 and Mef2c sites and the generated mutants of the P3 promoter.

The P3 promoter containing 3 Nkx2-5 (red colour) and 1 Mef2c (green colour) binding sites were mutated by the site directed mutagenesis method. The single mutants (1-4) and multiple mutants of the Nkx2-5 and Mef2c sites (5-7) were generated and labelled accordingly to the left of the schematic representation.

3.3 Biochemical assays

3.3.1 Western blot analysis

Analysis of proteins during *in vitro* cardiac differentiation was carried out by the Western blot technique. Whole cell lysates were isolated from differentiated samples by using 1x Passive lysis buffer (5x Passive lysis buffer, Promega, Cat.No E1941). Concentration of the whole cell lysate was determined by the Bradford's method (detailed in 3.1.4). The protein samples were stored in 6X Laemmli dye (Laemmli, 1970) and 20 µg of the quantified total protein was then loaded in 10% sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) using Bio-Rad Mini-Protean® 3 cell, Cat.No 67S/11919. The protein was then transferred to Polyvinylidene fluoride (PVDF transfer membrane, Immobilon®-P, Millipore, Cat.No IPVH00010) membrane and was hybridised with RYBP antibody, (anti-DEDAF, Merck Millipore, Cat.No AB3637, 1:1000) and PLAGL1 Antibody (anti-Zac1 C-7, Santa Cruz, Cat.No sc-166944, 1:1000). Bio-Rad Goat-anti-mouse IgG-HRP conjugate, (Cat.No 172-101, 1:2000) and Merck Millipore Goat-anti-Rabbit IgG-HRP conjugate, (Cat.No AP132P, 1:2000) were used as the secondary antibodies. The membranes were washed with TBST buffer for 5 times with 5 minutes of gentle shaking and then hybridised with Immobilon™ Western, Chemiluminescent HRP Substrate, Millipore, Cat.No WBKLS0500. Alliance Q9 system (UVITECH) was used to capture the chemiluminescent signals.

3.3.2 Protein stability assays

The wild type d14 differentiated cardiomyocytes were treated with 75 µg/ml concentration of Cyclohexamide, (CHX, Sigma, Cat.No C7698) and 10 µM MG132 (MG132, Cayman Chemicals, Cat.No 133407-82-6) for up to 6hours and the cells were then lysed in a time dependant manner between 1 hour and 6 hours of the treatment by using 1X Passive lysis buffer (5X Passive lysis buffer, Promega, Cat.No E1941) respectively.

3.3.3 Co-Immunoprecipitation (Co-IP)

HEK293T cells were transiently transfected with 5 µg of pcDNA3.1-RING1A FLAG, pRK7-FLAG NKX2-5, pRK7-FLAG MEF2C and pRK7-FLAG PLAGL1 (a kind gift from Dr. Dietmar Spengler, Max Plank Institute of Psychiatry, Germany) in combination with 5 µg of pcDNA3.1 RYBP cDNA containing expression vectors. Transient transfection and protein lysis were performed as mentioned above (detailed in 3.1.3). The whole cell lysates were

incubated in ice for 15 minutes and were spun at 15,000 x g for 10 minutes at 4°C. The supernatant was separated and pre-cleaned with 30 µl of Protein A-Agarose beads (Roche, Ref.No 11134515001) at 4°C under gentle rocking for 20 minutes. The precleared supernatant with agarose beads were spun at 500 x g for 2 minutes at 4°C. 80 µl of the supernatant was collected and mixed with 6X Laemmli dye, boiled for 10 minutes at 100°C to use as input lysates for Western blot analysis. The remainder of the supernatant was incubated overnight at 4°C under gentle rocking with 30 µl of RYBP antibody (anti-DEDAF, Millipore, Cat.No AB3637) bound agarose beads (5 µl of RYBP antibody (1 µg/ml) was bound to 100 µl of agarose beads for 4 hours at 4°C under gentle rocking). To wash the immunoprecipitated proteins, the protein bound FLAG-tagged beads were centrifuged for at 500 x g for 2 minutes at 4°C and washed with 1X PBS for 5 times. The immunoprecipitated proteins bound to the RYBP-tagged beads were then mixed with 30 µl of 6X Laemmli dye, boiled for 10 minutes at 100°C and stored in -20°C until further use. 20 µl of the input lysates and 20 µl of the immunoprecipitated proteins were loaded in 10 % SDS-PAGE and Western blot analysis (detailed in 3.3.2) was carried out. The Western transferred membrane was immunoblotted with anti-FLAG antibody (Monoclonal anti-FLAG M2 Peroxidase (HRP), Sigma, Cat.No A8592) at 4°C under gentle shaking overnight. The membranes were processed as mentioned in methods 3.3.2.

3.3.4 Immunocytochemistry (ICC) analysis

Immunofluorescence staining of *in vitro* cardiac cell cultures was achieved by culturing the cells over glass coverslips in 24 well plates (24 well Cell Culture Cluster Costar, Cat.No 3524) as described before (detailed in 3.1.2) and fixed with 4% Paraformaldehyde (PFA, Sigma, Cat.No 158127) for 20 minutes at room temperature (RT). Cells were permeabilized by 0.2% Triton X-100 (Triton® X-100, Sigma, Cat.No T8787) in Phosphate Bovine Saline (Dulbecco's PBS (1x), Gibco, Cat.No 14190-144) for 20 minutes in gentle shaking at RT. 5% BSA in PBS was used to block the cells for 1 hour at RT. The cells were incubated with RYBP antibody (Anti-DEDAF antibody, Merck Millipore, Cat.No AB3637, 1:1000 dilution) and PLAGL1 antibody (anti-Zac1 C-7, Santa Cruz, Cat.No sc-166944, 1:1000 dilution) in 5% BSA overnight at 4°C under gentle shaking. The cells were washed for 5 times with PBS and incubated with fluorescent labelled secondary antibody (Alexa Fluor 488 Goat-Anti-Rabbit, Invitrogen, Cat.No A-21206; Alexa Fluor 647 Donkey-Anti-Mouse, Invitrogen, Cat.No A-31571) at 1:2000 dilution in BSA for 1 hour at 4°C. The cells were then washed 3 times with PBS. The

cells were then incubated for 20 minutes with DAPI (Vector Laboratories, Cat.No H-1200) diluted at 1:2500 in PBS. The cells were then washed 3 times with PBS and mounted in Fluoromount-G™, (eBioscience, Cat.No 00-4958-02). The images were taken in Olympus LSM confocal microscopy (Olympus Corporation, Japan).

3.4 Bioinformatic analysis

3.4.1 Transcriptome analysis

Microarray analysis of the genome wide transcriptomics (Ujhelly et al., 2015) was carried out by mapping the sequenced reads from RNA-seq experiment by TopHat1 and the $\log^2$ fold change (FC) counts were calculated by read counts after normalization using DESeq1 package in R programming and deposited in GEO (Gene expression omnibus-https://www.ncbi.nlm.nih.gov/geo/) accession ID GSM4575880 (Henry et al., 2020).
The sorting of the genes based on their corresponding $\log^2$ FC values was performed in Microsoft Excel using the VLookup and sorting functions. Hierarchical clustering of the values ($\log^2$ FC ≥ 2) for upregulation and ($2 \leq \log^2$ FC) for downregulation of genes between wild type and *Rybp* null mutant ES cells and differentiated CMCs was performed by the k-means method using the XLSTAT extension tool in Microsoft Excel. Representative heatmaps were generated by transferring the clustered gene sets into Prism GraphPad 8 software.

3.4.2 Analysis of the reported ESTs of the *Plagl1* splice variants

Complete CDS (coding sequence) of *Plagl1* mRNA and deposited transcript variants were downloaded in FASTA format from NCBI-Nucleotide database. Each variant sequence was BLASTed with the *Plagl1* genomic locus from Ensembl (https://www.ensembl.org/index.html) ID: ENSMUSG00000019817 as the reference file with indicating exon positions. The exons transcribed in each splice variant was identified and the splice variant sequences were aligned using BioEdit software. The corresponding position of the promoter region from which the splice variants were transcribed were presumed based on the coding exons and the relative position of the promoter regions.

3.4.3 Analysis of functional domains and degron sites in PLAGL1

The analysis for the functional domains in the PLAGL1 protein was determined by uploading the PLAGL1 amino acid (aa) sequence (NCBI ID: NP_033564.2) in the PROSITE ExPasy

Motif search tool (https://prosite.expasy.org) maintained by the Swiss Institute of Bioinformatics (SIB). Serine-threonine phosphorylation analysis was done by using kinase specific phosphorylation site prediction using GPS 5.0 online tool (http://gps.biocuckoo.cn/online.php) (Wang et al., 2020). Serine-threonine and tyrosine kinases sites in the PLAGL1 protein was predicted by setting high threshold cut-off.

3.4.4 Analysis of the *Plagl1* promoter for CpG islands and TATA box

The CpG islands in the *Plagl1 P1*, *P2* and *P3* promoters were analysed by uploading the FASTA sequence in the DBCAT online tool (http://dbcat.cgm.ntu.edu.tw). DBCAT uses methylation microarray data to analytically identify the CpG islands in the query sequence.
TATA box prediction was done by uploading the FASTA sequence of *Plagl1 P1*, *P2* and *P3* promoters into YAPP Eukaryotic core promoter prediction webtool (http://www.bioinformatics.org/yapp/cgi-bin/yapp.cgi).

3.4.5 Transcription factor binding analysis in *Plagl1* promoters

Transcription factors binding (TFB) analysis was performed using TRANSFAC webtool (https://genexplain.com/transfac/). TRANSFAC is a widely used TFB analysis tool which identifies TFB sites based on the experimentally proven consensus of several transcription factors and ChIP binding (Wingender et al., 1996; Wingender, 2008; Kaplun et al., 2016) . The amplified and cloned *Plagl1* promoters *P1*, *P2* and *P3* promoter sequences were analysed for TFB sites by choosing either muscle specific, cell cycle specific or for all eukaryotic transcription factors.

3.4.6 Motif search in gene promoter regions

Promoter regions of sarcomeric genes- *Actc1*, *Tnnt2*, *Tnni3*, *Tpm1*, *Tpm4*, *Myh7*, *Myom1* and *Ttn* were download from ENSEMBL (https://www.ensembl.org/index.html). ENSEMBL database identifies promoter regions in the chromatin based on metagenomic index containing pre-selected set of ChIP-Seq assays for CTCF, H3K4me1, H3K4me2, H3K4me3, H3K9ac, H3K27ac, H3K27me3, H3K36me3, H4K20me1 (Zerbino et al., 2015). The downloaded promoter regions in FASTA format were uploaded into JASPAR (http://jaspar.genereg.net/) database for motif search (Sandelin et al., 2004). The analysis was done by choosing Mus musculus PLAGL1, NKX2-5, MEF2C and TBX5. Predicted consensus sites and binding scores were used to generate Manhattan plot. Binding scores of over 10 was considered significant.

3.4.7 Metadata analysis in ES cells and CMCs

Metadata analysis for existing ChIP-seq analysis was performed by downloading pre-existing ChIP-seq data from GEO database (https://www.ncbi.nlm.nih.gov/geo/) under the following IDs. In ES cells, RYBP ChIP- GSM4052120, RNF2 ChIP- GSM4052131 and input ChIP- GSM4052104 (Zepeda-Martinez et al., 2020), In differentiated CMCs, RYBP ChIP- GSM1657391, RNF2 ChIP- GSM1657390 and input ChIP- GSM1657392 (Morey et al., 2015). The downloaded BigWig files were uploaded into IGV (Integrative Genomics Viewer) choosing specific annotations i.e., mm9 or mm10 according to the original analysis and the binding peaks were visualized by setting the data range of the peaks using the input file as the reference.

3.8 Statistical analysis

All experiments were repeated three times. Experiments were evaluated by using two-way ANOVA for significance in qRT-PCR data and one-way ANOVA for significance in luciferase reporter assays using GraphPad Prism version 7. All data mentioned in this thesis are expressed as mean ± standard deviation (SD). Values of $p \leq 0.05$ were accepted as significant (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$).

4. RESULTS

4.1 Hierarchical gene cluster analysis of the whole genome transcriptome

In order to get a global view of the transcriptional changes in the *Rybp* null mutant cells during cardiac differentiation, a detailed comparison of the mRNA transcriptomes across wild type and *Rybp* null mutant ES cells (d0) and derived CMCs (d8, d14) was previously performed (Ujhelly et al., 2015; Henry et al., 2020); GEO acc. GSE151349). In brief, wild type and *Rybp* null mutant ES cells were let to form EBs by the hanging drop method. On the second day the EBs were collected, plated on cell culture plates and cultured up to 14 days. The samples were collected from the designated time points of cardiac differentiation (methods 3.1.2) where d0 represented the pluripotent stem cells stage, d8 the progenitor stage and d14 the terminal cardiac stage (Figure 7). In this analysis genes expression changes revealed that the levels of several key cardiac transcription factors part of signalling pathways and genes that code for proteins indispensable for contractility were downregulated (Ujhelly et al., 2015). In the frame of the current study, further analyses of the whole genome transcriptome data with functional annotations studies were carried out to identify the mechanisms that possibly led to the contractility defect of the *Rybp* null mutant CMCs.

4.1.1 Calcium homeostasis, the JAK-STAT pathway and cell adhesion are amongst the most affected mechanisms in the *Rybp*$^{-/-}$ ES cells and derived CMCs

Hierarchical clustering of the values ($2 \leq \log^2$ fold change ≥ 2) between wild type and *Rybp* null mutant ES cells and differentiated CMCs was performed by the k-means method using XLSTAT tool revealed 8 distinct gene clusters (detailed in methods 3.4.1) (Figure 15A).
From the analysis, clusters of genes with discrete fold change patterns during the analysed time points (methods 3.1.2) were procured (Figure 15A). Cluster 1 contained genes that were profoundly upregulated ($\log^2$ fold change ≥ 4) in ES (d0) at both examined stages of *in vitro* cardiac differentiation (i.e., d8 and d14) in the *Rybp* null mutant cells in comparison to the wild type (Figure 15B and C). These include Hyperpolarization-activated cyclic nucleotide-gated potassium and sodium channel 2 (*Hcn2*) and Hyperpolarization-activated cyclic nucleotide-gated potassium and sodium channel 3 (*Hcn3*) (Figure 15B). High expression of *Hcn2* and *Hcn3* are associated to cause sinoatrial node dysfunction ultimately leading to heart failure (Yanni et al., 2011).

Other cardiac ion channel genes such as Potassium voltage-gated channel Isk-related subfamily member 1 (*Kcne1*), Potassium voltage-gated channel Isk-related subfamily member 2 (*Kcne2*), Calcium channel voltage-dependant gamma subunit 5 (*Cacng5*), Calcium-sensing receptor (*Casr*), Transient receptor potential cation channel subfamily V member 4 (*Trpv4*) and gap junction genes such as Gap junction protein beta 2 (*Gjb2*) were identified to be part of cluster 1 (Figure 15B and C). These genes play essential roles in the maintenance of ion homeostasis in the developing CMCs.

In cluster 2 we identified several genes with significantly upregulated expression level in the *Rybp* null mutant ES cells ($\log^2$ fold change ≥ 4) and decreased expression level at d8 and d14 ($\log^2$ fold change ≤ 2) (Figure 15D and E). Genes that contribute to vascular smooth muscle contraction such as Angiotensin II receptor type 1 (*Agtr1*), Endothelin receptor type A (*Ednra*), Arginine vasopressin receptor 1a (*Avpr1a*), Myosin light chain kinase 3 (*Mylk3* also called as *Mlck*) and Myosin light chain 2 (*Myl2*) that function in vasoconstriction and Adenosine A2a receptor (*Adora2*) that plays role in vasodilation were all part of the same cluster. Genes essential in maintaining calcium homeostasis in the developing CMCs such as Potassium inwardly-rectifying channel subfamily J member 5 (*Kcnj5*), Calcium voltage-gated channel T type alpha 1G subunit (*Cacna1g*), Calcitonin receptor (*Calcr*) and Sodium channel voltage-gated type I alpha (*Scn1a*) were also identified in the same cluster showing that key cardiac genes were upregulated from the ES cell stage and these gene expression changes together could potentially lead to the loss of ion equilibrium which is required for the normal formation of CMCs (Figure 15D).

Cluster 3 contained genes that were extensively downregulated at d8 and upregulated by d14 in the *Rybp* null mutant cells (Figure 15F and G). By Gene Ontology (GO) analysis we identified 16 genes that acts on the JAK-STAT (Janus Kinase-Signal transducer and activator of transcription proteins) signalling pathway (GO:0046425) that contributes to the normal proliferation and apoptosis of the differentiating cells.

In cluster 4, 5, 6 and 7 we did not identify genes that significantly related to any function connected to cardiac development. In cluster 8, cell adhesion markers such as Cadherin protein 6, 7 and 17 (*Cdh6*, *Cdh7* and *Cdh17*, respectively) and Vascular cell adhesion molecule 1 (*Vcam1*) were identified to be downregulated in the *Rybp* null mutant ES cells (Figure 15H and I). Cell adhesion is a key feature which is required for the proper proliferation and differentiation of various cell types during mammalian heart development.

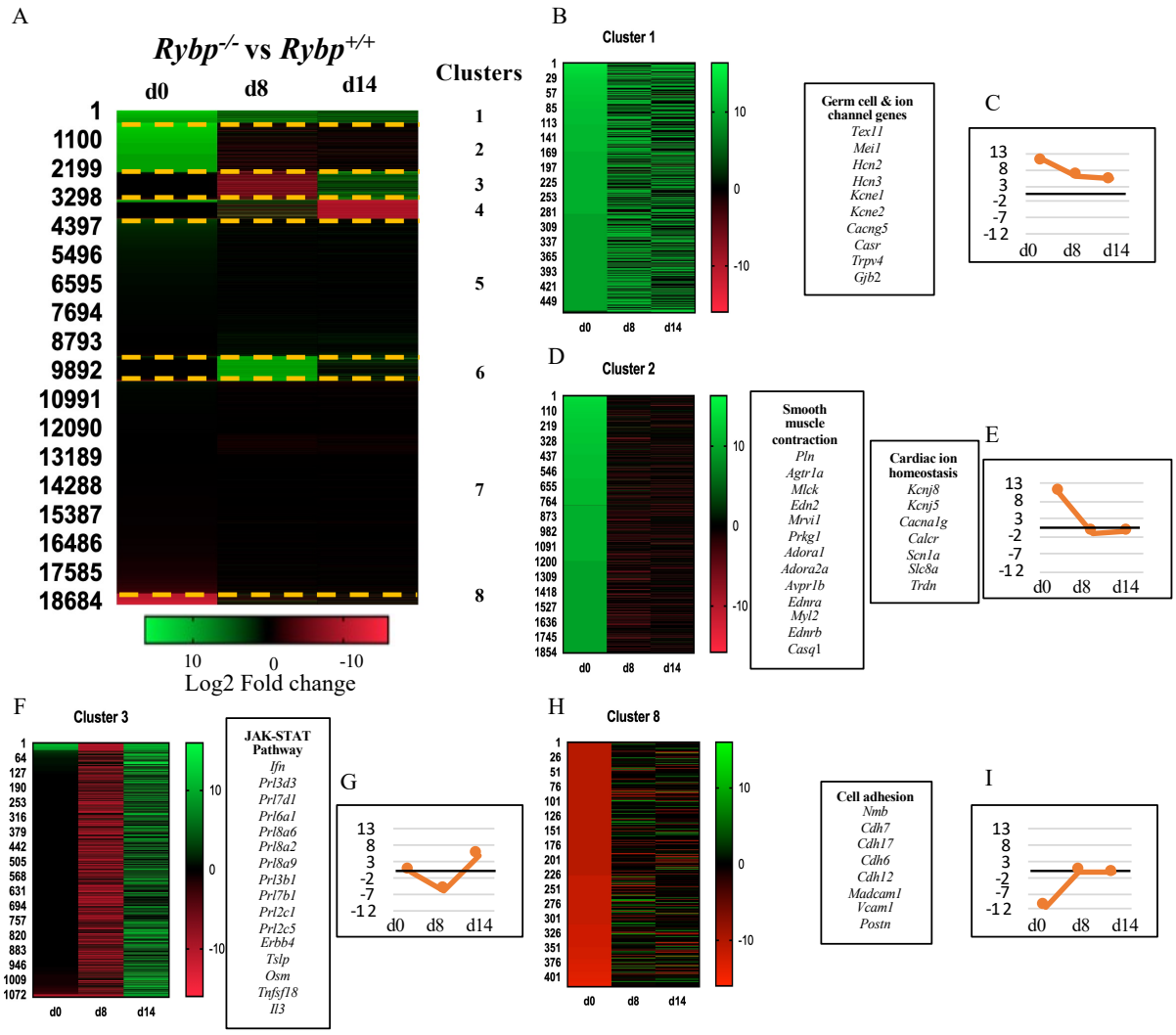


Figure 15: Hierarchical gene clustering of transcriptome data from wild type and Rybp null mutant cells during in vitro cardiac differentiation.

(A) *Heat map from hierarchical clustering of RYBP regulated gene expression changes with significant upregulated (log^2 fold change ≥2; green colour) and downregulated (log^2 fold change ≤2; red colour) genes in the Rybp null mutant cells. The cluster numbers are listed on the right side of the heat map.* ***(B and C)*** *Cluster 1 heat map and tendency graph highlight the upregulated gene set at all three time points i.e., d0, d8 and d14.* ***(D and E)*** *Cluster 2 heat map and tendency graph highlight the genes upregulated in d0 only and downregulated in d8 and d14.* ***(F and G)*** *Cluster 3 heat map and tendency graph highlight the genes downregulated in d8 and upregulated in d14.* ***(H and I)*** *Cluster 8 heat map and tendency graph highlight the genes highly downregulated in d0. The tendency graphs are represented as an average of the overall log^2 fold change for each time point pertaining to each cluster respectively.*

4.1.2 Key cardiac transcription factors and sarcomeric components are downregulated in the *Rybp* null mutant CMCs

To gain more insights about the expression of key cardiac genes which are vital for the formation of beating CMCs, we further dissected the transcriptome. Our analysis revealed that genes required for cardiac progenitor formation and sarcomere organization were remarkably downregulated at d8 and d14 in the mutant cells (Figure 16A). Cardiac progenitor markers *Shh* (d8 Log^2 FC: -5.40, d14 Log^2 FC: -2.42), *Isl1*(d8 Log^2 FC: -1.14, d14 Log^2 FC: -2.82), *Nkx2-5* (d8 Log^2 FC: -9.14, d14 Log^2 FC: -8.84) and *Mef2c* (d8 Log^2 FC: -1.27, d14 Log^2 FC: -1.02) displayed severe downregulation at d8 and d14 in *Rybp*$^{-/-}$ CMCs (Figure 16A). Cardiac transcription factors with known roles in first and second heart field specification such *Hand2* (d8 Log^2 FC: -0.61, d14 Log^2 FC: -0.08), *Gata4* (d8 Log^2 FC: -0.55, d14 Log^2 FC: -0.22), *Tbx5* (d8 Log^2 FC: -0.58, d14 Log^2 FC: -0.98) and *Tbx20* (d8 Log^2 FC: -0.34, d14 Log^2 FC: -0.13) also displayed faint downregulation at d8 and d14 in *Rybp*$^{-/-}$ CMCs.

Sarcomeric genes such as Myomesin 1 (*Myom1*), Titin (*Ttn*), Actin alpha cardiac muscle 1 (*Actc1*), Myosin heavy peptide 6 cardiac muscle alpha and Myosin heavy peptide 7 cardiac muscle beta (*Myh6* and *Myh7*, respectively) were highly downregulated in the mutant cells (Figure 16B). This analysis shed light on the impairment of sarcomere formation, which can immensely contribute towards the non-contractility phenotype of the *Rybp* null mutant cells as well.

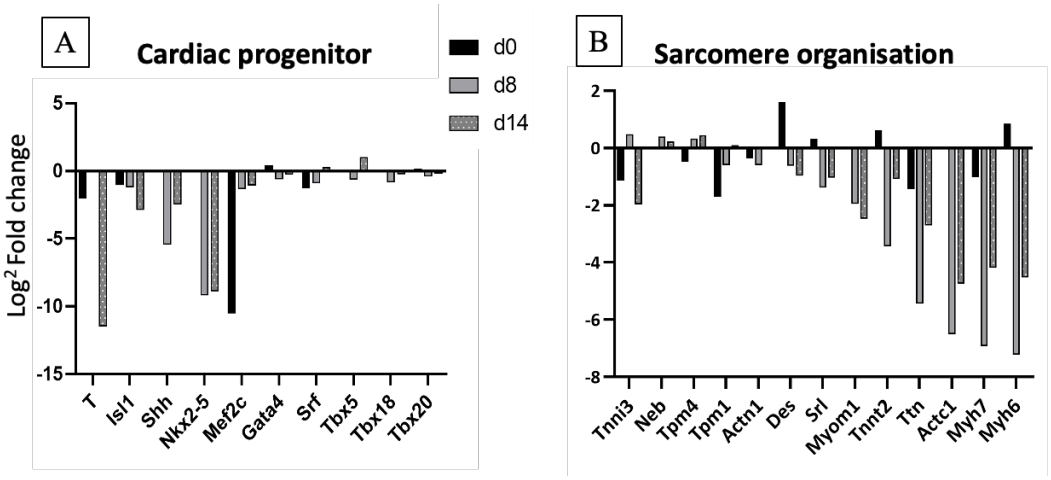


Figure 16: Genes required for cardiac progenitor formation and sarcomere organisation are downregulated in the Rybp null mutant CMCs

Log^2 Fold change values from the transcriptome were used to generate bar graphs. ***(A)*** *Bar graph representing the downregulation of genes functioning in cardiac progenitor formation in d8 and d14 differentiated Rybp null mutant CMCs.* ***(B)*** *Bar graph representing the*

downregulation of sarcomeric genes in d8 and d14 differentiated Rybp null mutant CMCs. ***Abbreviations:*** *d:day.*

4.2 *Plagl1* is the most strikingly downregulated gene in the *Rybp*$^{-/-}$ ES cells and CMCs

In the whole genome transcriptome analysis, *Plagl1* a cardiac transcription factor with transactivation functions during mammalian embryonic development (Yuasa et al., 2010) was one of the most strikingly down regulated gene in the *Rybp* null mutant ES cells as well as derived CMCs (d0 Log^2 FC: -3.79, d8 Log^2 FC: -6.26, d14 Log^2 FC: -6.53). In order to further characterize the expression of *Plagl1* during an extended time course of *in vitro* cardiac differentiation (methods 3.1.2), we performed gene expression analysis by qRT-PCR. In brief, whole cell RNA was extracted from d0, d2, d7, d10, d14 and d21 time points of *in vitro* cardiac differentiation, reverse transcribed and qRT-PCR analysis was performed (details in methods 3.2.1) (Primer list in Table 3). Gene expression changes were analysed using wild type and *Rybp* null mutant cells from the designated time points of *in vitro* cardiac differentiation. Our results showed that in the wild type cells (*Rybp*$^{+/+}$), *Plagl1* mRNA was first detectable from d7 (cardiac progenitor formation stage) and its expression peaked by d14 and d21 (CMC stage) (Figure 17). *Plagl1* was not expressed at any analysed time points of *in vitro* cardiac differentiation in the *Rybp* null mutant cultures (Figure 17).

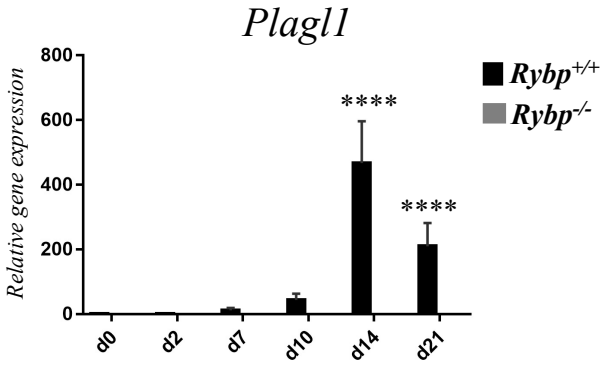


Figure 17: Plagl1 is not expressed in the Rybp$^{-/-}$ during in vitro cardiac differentiation

*Relative gene expression analysis of Plagl1 during in vitro cardiac differentiation by qRT-PCR analysis. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed in the Rybp$^{-/-}$ compared to Rybp$^{+/+}$ by the statistical method two-way ANOVA (****p < 0.0001).*

4.3 PLAGL1 is not detectable at protein level either in the *Rybp* null mutant cells at all time points of *in vitro* cardiac differentiation.

In order to investigate whether PLAGL1 is detectable during any time point of cardiac differentiation, we performed immunocytochemical analysis with samples derived from d0, d2, d7, d10, d14 and d21 of *in vitro* cardiac differentiation from the wild type and *Rybp*$^{-/-}$ cultures. The samples were stained for PLAGL1 with anti-PLAGL1 antibody (methods 3.3.4), and the pictures were taken using Olympus LSM confocal microscope. Our results defined that PLAGL1 protein levels corelated to the mRNA levels detected by qRT-PCR (Figure 17) in the wild type cells. PLAGL1 staining was detectable at d7 (Figure 18A) from the analysed time points and the PLAGL1 signal was the highest at d14 CMC stage (Figure 18A). As expected, PLAGL1 was not detected at all time points of *in vitro* cardiac differentiation in the *Rybp* null mutant cultures revealing that PLAGL1 expression was absent in the *Rybp* null mutant cultures in both the mRNA and protein levels during the time course of *in vitro* cardiac differentiation (Figure 18B).

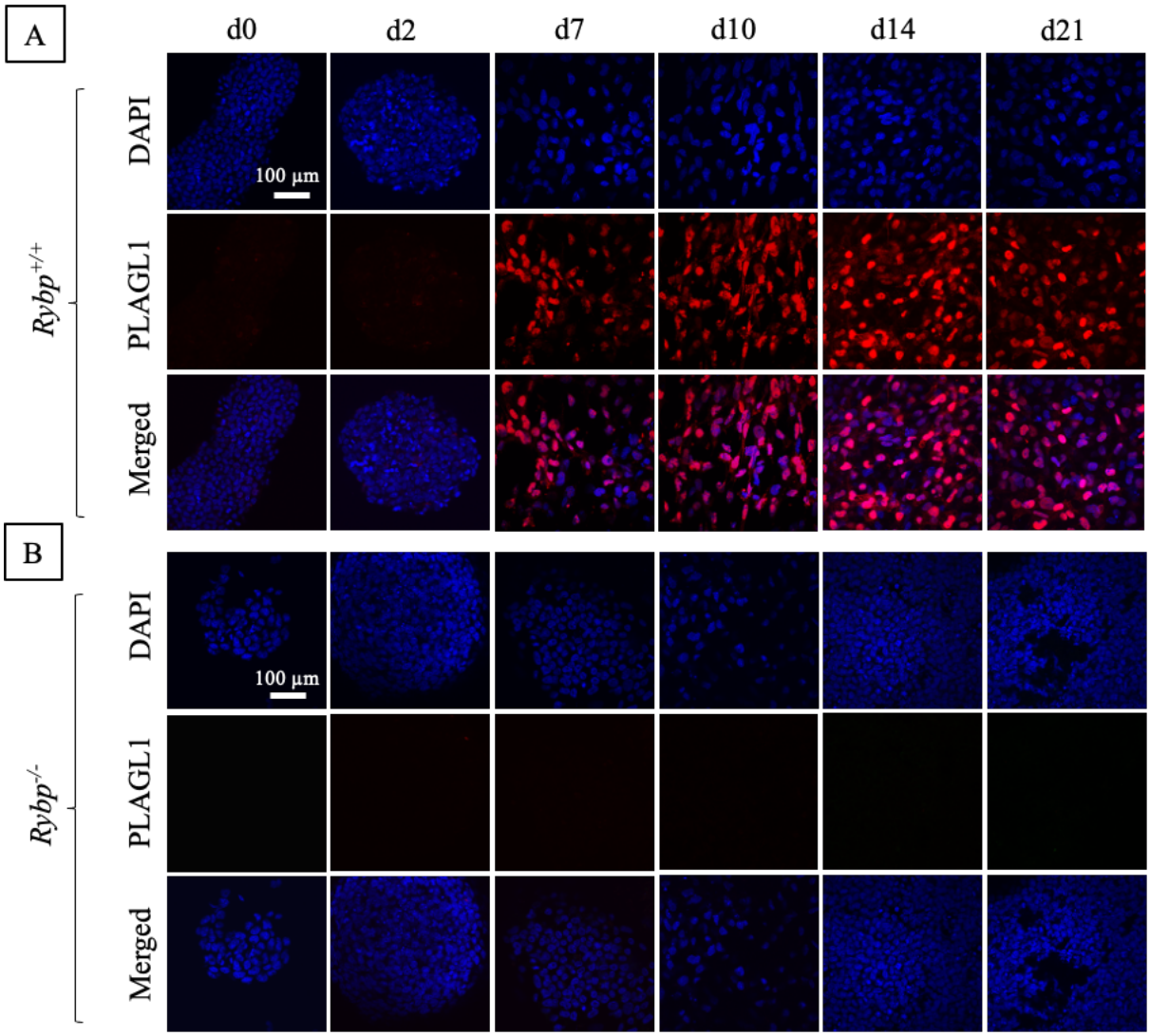


Figure 18: PLAGL1 is not detectable in the RYBP null mutant cells during in vitro cardiac differentiation

Immunocytochemical analysis of PLAGL1 (Red) in the ***(A)*** *wild type and* ***(B)*** *Rybp$^{-/-}$ null mutant cells derived from d0, d2, d7, d10, d14 and d21. DAPI (blue) was used to stain the nuclei. The indicating time points are represented in top. Olympus Confocal IX 81, Obj: 60x; Scale bar: 100 µm.*

4.4 Overview of the *Plagl1* genomic locus

To unravel the molecular mechanism behind the downregulation of *Plagl1* in the *Rybp* null mutant cells during cardiac differentiation, we analysed the *Plagl1* genomic locus (Chr10:13090832-13131694 bp) for the position of promoters, regulatory RNAs and potential splice variants. The *Plagl1* genomic locus was downloaded in FASTA format from ENSEMBL

(ENSMUSG00000019817) and the respective position of the coding exons were mapped according to ENSEMBL and the downloaded mRNA sequences from ESTs. Based on previous publications and by carefully mapping the regulatory regions for promoter regions and different ncRNAs we have reconstructed the *Plagl1* genomic locus. The *Plagl1* genomic locus consists of three promoter regions *P1*, *P2* and *P3*, eleven exons and two ncRNAs *Hymai* and Plagl1 intronic transcript (*Plagl1it*) (Figure 19). The *P1* promoter, which harbours demethylated CpG islands is the site of imprinting of the *Plagl1* locus. The *P2* promoter which lies 30 kb upstream to the transcription start site (TSS) is previously identified to express *Plagl1* biallelically and functions only in disease conditions. The *P3* promoter contains variable enhancer elements like a TATA box and several consensus binding-sites for key lineage specific transcription factors such as NKX2-5, MEF2C and TBX5. *Hymai* and *Plagl1it* ncRNAs are imprinted and expressed downstream to the *P1* promoter. Exons 10 and 11 code for the full length PLAGL1 protein. PLAGL1 protein contains seven C2H2-type zinc finger domains at the amino terminal of the protein, from the amino acids 1 to 210. This region also encompasses two nuclear localization signals (NLS) along the zinc finger domains. PLAGL1 also contains proline and glutamine rich regions at the carboxyl terminal (residues 220 to 444) (Figure 19).

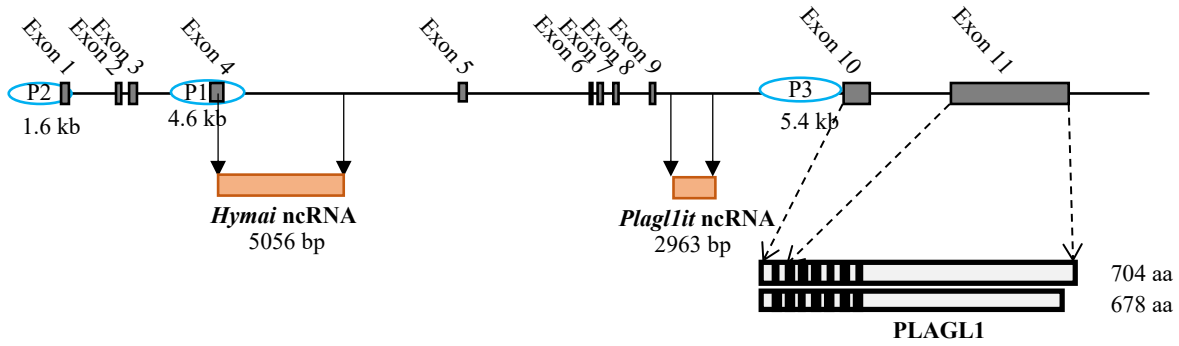


Figure 19: Schematic illustration of the Plagl1 genomic locus

Schematic representation of the Plagl1 genomic locus. Exons are represented with grey bars; The three promoters P1, P2 and P3 are marked in blue ovals; The two ncRNA, Hymai ncRNA and Plagl1it are represented with orange rectangle. ***Abbreviations:*** *kb: kilobase, bp: base pair, ncRNA: noncoding RNA, aa: amino acid.*

4.5 The expression of the two ncRNAs *Hymai* and *Plagl1it* are also affected in the *Rybp* null mutant CMCs during cardiac differentiation

Since our results determined that all splice variants of *Plagl1* are not detected in the *Rybp* null mutant cells during the time course of *in vitro* cardiac differentiation, we wondered if the two ncRNAs *Hymai* and *Plagl1it* are also affected in the *Rybp* null mutant CMCs. Gene expression analysis using qRT-PCR determined that the expression kinetics of both *Hymai* and *Plagl1it* in the wild type cultures resembled the expression kinetics of *Plagl1* (Figure 20A and B). *Hymai* (Figure 20A) and *Plagl1it* (Figure 20B) expression could be first detected by d2, EB formation stage and the expression peaked by d14 suggesting that the two ncRNAs might function in the regulation of *Plagl1*. The high expression levels of both ncRNAs at d14 is also indicative of their potential functions during CMC development.

Both *Hymai* and *Plagl1it* expression in the *Rybp* null mutant cells were also affected at all the analysed time points (Figure 20A and B). Unlike *Plagl1*, *Hymai* and *Plagl1it* ncRNA expression was detected at low levels in the *Rybp* null mutant CMCs further indicating the significant effect of the loss of *Rybp* in *Plagl1* regulation.

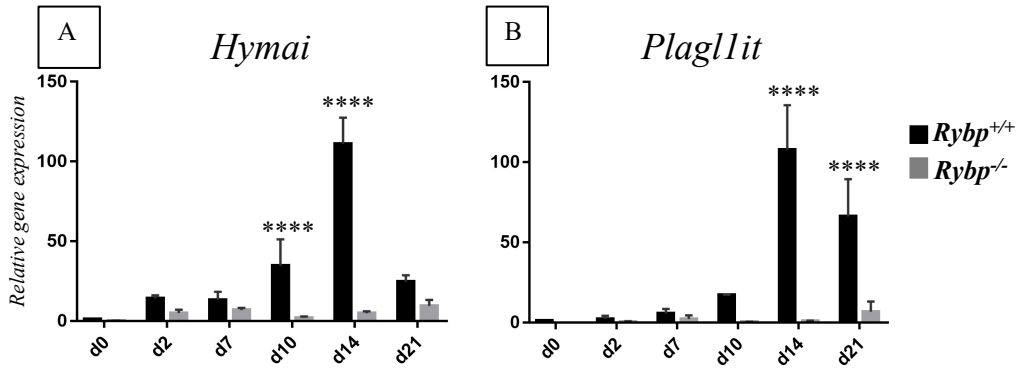


Figure 20: Hymai and Plagl1it expressed at low levels in Rybp$^{-/-}$ at all examined stages of in vitro cardiac differentiation

Relative gene expression analysis of Hymai ***(A)*** *and Plagl1it* ***(B)*** *during in vitro cardiac differentiation by qRT-PCR analysis. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed in the Rybp$^{-/-}$ compared to Rybp$^{+/+}$ by the statistical method two-way ANOVA (****p < 0.0001).*

4.6 Multiple splice variants of *Plagl1* can be transcribed from monoallelic and biallelic promoters

As the *Plagl1* mRNA was not expressed in the *Rybp* null mutant cells at any of the examined time points of *in vitro* cardiac differentiation (results 4.2), we further performed detailed *in silico* analysis of the various splice variants that are transcribed from the *Plagl1* genomic locus. We have also identified the corresponding promoters producing protein coding transcripts and the promoters that can be active during cardiac differentiation. cDNA sequences were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/nuccore/) and the transcripts were aligned with ClustalW tool in BioEdit programme.

The schematic representation presented on Figure 21 was generated based on the analyses with representative NCBI accession I.D of the cDNA transcripts. The position of the *Plagl1* promoter regions is mutually exclusive to the emerging cDNAs. Our analysis showed that FJ425893.1 emerge from the *P2* promoter, NM_009538.2, NM_009538.3, NM_001364643.1, NM_001364644.1, NM_001364645.1, BC141284.1 and AF147785.1 emerge from the *P1* promoter and X95504.1, AA919394.1 and AF324471.1 emerge from the *P3* promoter. NM_009538.2 and NM_009538.3 are mostly similar and differ only in their 5' region of exon 8 coding. NM_001364643.1, NM_001364644.1 and NM_001364645.1 are also similar splice variants with NM_001364644.1 harbouring an alternate splice site within exon 11. NM_001364645.1 variant differs from NM_001364644.1 by harbouring exon 8 additionally. The identified splice variants may have tissue and disease specific expression. We also conclude that all the three promoters can produce protein coding transcripts since the last two exons (exon 10 and 11) code for the full length PLAGL1 protein.

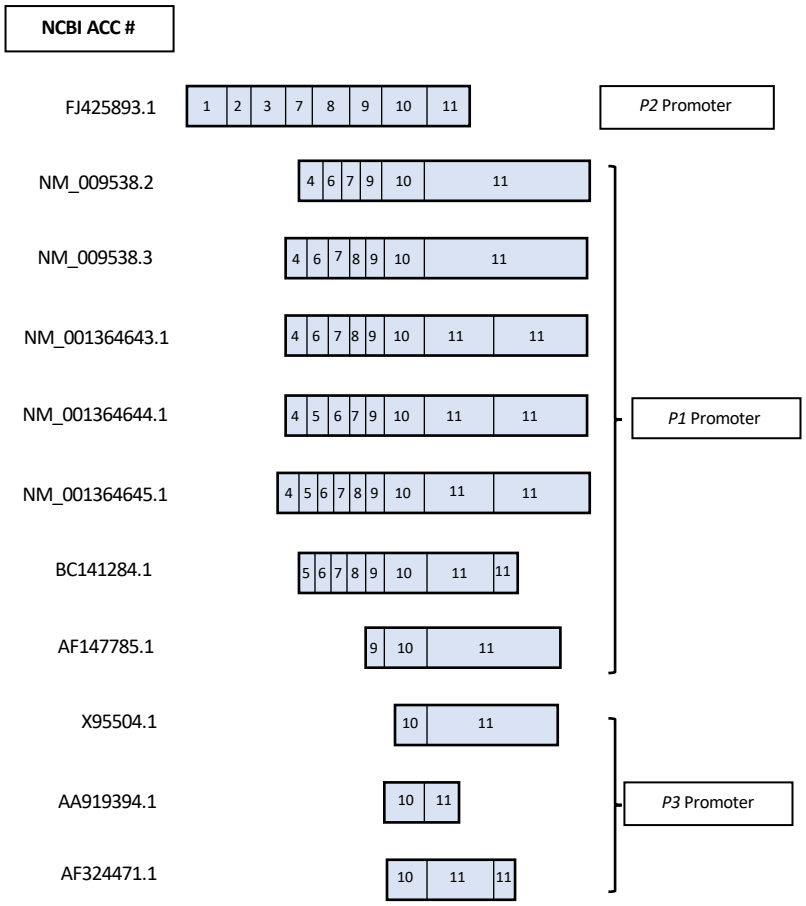


Figure 21: Plagl1 can be expressed by mono and biallelic promoters

Schematic representation of the various splice variants of Plagl1 based on ESTs deposited in EST database (methods 3.4.2). The NCBI accession numbers (NCBI ACC #) are presented on the left side, the corresponding promoters from where the transcripts are transcribed from are presented at the right side. Splice variants are shown in grey boxes in the middle. The numbers in the boxes represent corresponding exons.

4.7 Splice variants of *Plagl1* are transcribed from *P1* and *P3* promoters during *in vitro* cardiac differentiation

To gain further insights about which promoters are active during the time course of *in vitro* cardiac differentiation we performed gene expression analysis using primers specific to the exons that are distinctive to the transcripts produced from the alternative promoter regions. We used primers specific to exon 1 and 2 (hereafter mentioned as *Plagl1 1/2*) to check the expression from *Plagl1 P2* promoter, primers specific to exon 6 and 7 (hereafter mentioned as *Plagl1 6/7*) to check the expression from *Plagl1 P1* promoter. The *P2* promoter produces biallelic expression of *Plagl1* whereas the *P1* and *P3* promoters regulate monoallelic

expression of *Plagl1* as the *P1* promoter is the site of imprinting and all genomic products downstream to *P1* are imprinted. The *Plagl1 P3* promoter is situated immediately upstream to the last two exons which code for the full length PLAGL1 protein. As these two exons are present in all the transcripts, we couldn't make primers specific to check the activity of only *P3* promoter. We used primers specific to exon 10 and 11 as a universal primer which can detect the expression of all the splice variants of *Plagl1* together.

QRT-PCR analysis using *Plagl1 6/7* primers in the wild type cultures showed that *Plagl1* expressed weakly until day 7 and its expression levels induced to over 100 folds when compared to the wild type d0 (Figure 22A). The expression of *Plagl1* using *Plagl1* 10/11 (Figure 22B) showed up to 400 folds change increase in d14 as opposed to the 100 folds increase in the d14 *Plagl1 6/7* suggesting that both *Plagl1 P1 and P3* promoters could be presumably active during *in vitro* cardiac differentiation (Figure 22B). Using primers specific to *Plagl1 1/2* we did not get any signal in the wild type cells (data not shown) suggesting that the *P2* promoter might not be active during cardiac development and could only cause biallelic expression of *Plagl1* in disease states such as the Transient neonatal diabetes mellitus (TNDM) (Valleley et al., 2007). In the case of the *Rybp*$^{-/-}$ cells, we could not detect *Plagl1* expression using any primer sets.

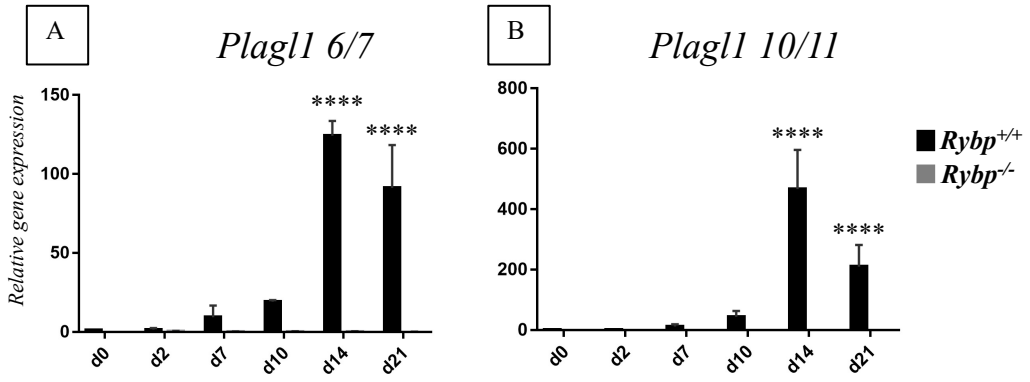


Figure 22: Plagl1 is expressed from both P1 and P3 promoters during in vitro cardiac differentiation

Relative gene expression analysis of Plagl1 using primers specific to ***(A)*** *exon 6/7 and* ***(B)*** *exon 10/11 during in vitro cardiac differentiation. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed in the Rybp*$^{-/-}$ *compared to Rybp*$^{+/+}$ *by the statistical method two-way ANOVA (****p < 0.0001).*

4.8 Two isoforms of the PLAGL1 protein are detectable during *in vitro* cardiac differentiation

In order to further characterize the isoforms of the PLAGL1 protein, that are expressed during *in vitro* cardiac differentiation, we performed Western blot analyses. Protein cell lysates were derived from the indicated points of *in vitro* cardiac differentiation and 20 µg of the total protein from each time point was loaded into each well for Western blot analysis (methods 3.3.1). Our results showed that PLAGL1 signals correlated to the mRNA expression levels (methods 3.1.1). PLAGL1 was hardly detectable until d2. The first time point when PLAGL1 was clearly detectable was at day 7, which corresponds to the stage of cardiac progenitor formation stage (Figure 23). The expression peaked by d14, corresponding to the time of cardiomyocyte formation. At d7 and d10 the two major isoforms of PLAGL1: PLAGL1 a (NCBI Accession: NP_033564.2, 79 kDa) and PLAGL1 b (NCBI Accession: NP_001351572.1, 76 kDa) can be seen indicating that just two isoforms of PLAGL1 are expressed during *in vitro* cardiac differentiation (Figure 23). In d14 and d21 several bands of PLAGL1 protein were obtained which we further analysed for potential post-translation modification of PLAGL1 (Figure 23).

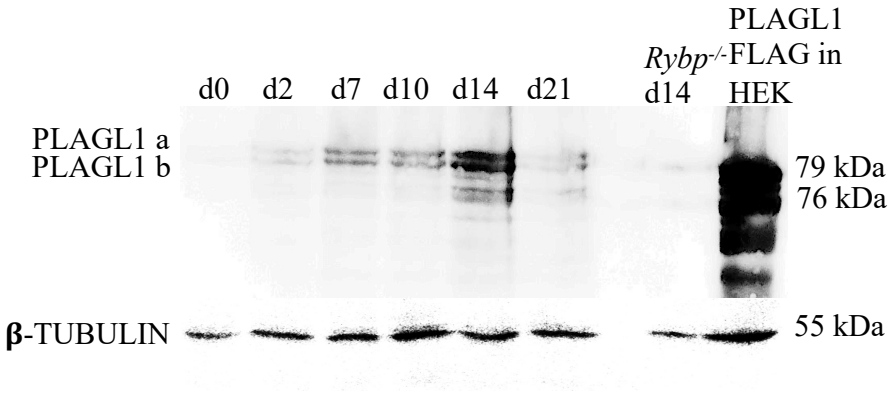


Figure 23: Two isoforms of PLAGL1 are expressed during in vitro cardiac differentiation
Western blot analysis detected the two isoforms of PLAGL1 protein: PLAGL1 a and PLAGL1 b can be seen clearly in d7 and d10. Protein lysates from d21 Rybp null mutant CMCs was used as the negative control. Lysate from flag tagged PLAGL1 over-expresser in HEK293T cells was used as a positive control. β Tubulin was used as an internal loading control to monitor the kinetics of PLAGL1 during the time course of cardiac differentiation. The respective molecular weights of both PLAGL1 and β Tubulin are indicated at the right. ***Abbreviations****: d: day, HEK: Human embryonic kidney cells, β TUB: β Tubulin, kDa- kilo Daltons.*

4.9 PLAGL1 undergoes post translational degradation during CMC formation

Analysis of the PLAGL1 protein in wild type differentiated CMCs by Western blot displayed multiple bands of PLAGL1 protein in d14 and d21 cardiac differentiated cell lysates (Figure 23). The highest two bands correspond to the isoforms of PLAGL1 protein: PLAGL1 a and PLAGL1 b which can be seen at 79 kDa and 76 kDa respectively (Figure 23). We wondered if the additional bands between 76 kDa and 51 kDa were post translational modifications of the PLAGL1 isoforms. In order to identify if PLAGL1 undergoes post translational modifications during cardiac differentiation we performed experiments with Cyclohexamide (CHX) a translational inhibitor and MG132 a protease inhibitor (methods 3.3.2). In brief, we differentiated ES cells to form CMCs *in vitro* for 14 days and treated the cells with media supplemented with 75 µg/ml concentration of CHX. Cells were harvested and lysed at every hour until 6 hr. Our results determined that PLAGL1 was undergoing post translational degradation and formed a stable 51 kDa size protein upon CHX treatment (indicated in red arrow, Figure 24A). Next, we performed the same experiment by treating the cells with 10 µM MG132. Upon MG132 treatment the 51 kDa degraded band was not formed due to the inhibition of protease activity and the two isoforms PLAGL1 a and PLAGL1 b were detected stronger (Figure 24B). These results demonstrated that PLAGL1 underwent degradation by d14, the time point with strongest PLAGL1 expression.

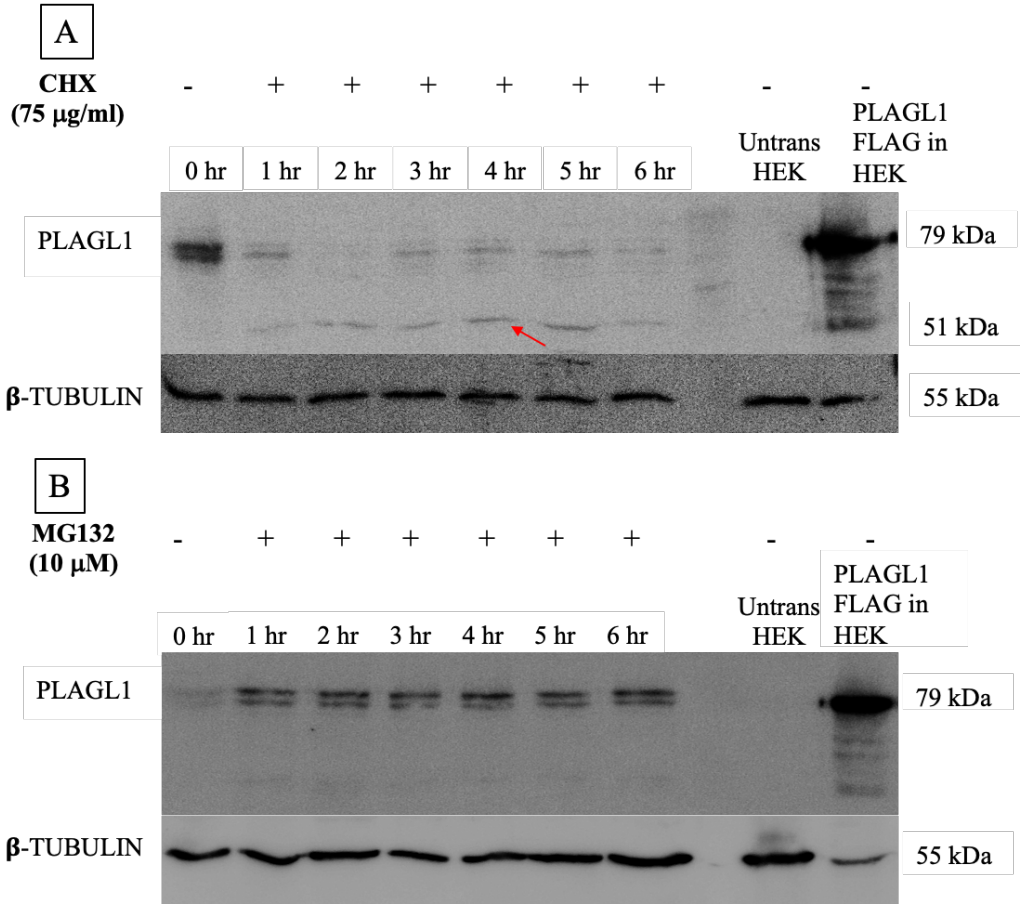


Figure 24: PLAGL1 undergoes degradation upon CHX treatment

(A)*Western blot analysis displaying PLAGL1 modifications after treatment with 75 µg/ml CHX suplemented to wild type d14 CMCs along with media. Both isoforms of the PLAGL1 protein degraded to a stable 51 kDa size protein.* ***(B)*** *Western blot analysis displaying PLAGL1 modifications after treatment with 10 µM MG132 suplemented to wild type d14 CMCs along with media. MG132 treatment caused inhibition of protease activity resulting in accumulation of the two PLAGL1 isoforms. β-Tubulin was used as an internal loading control. The respective molecular weights of both PLAGL1 and β -Tubulin are labelled in kDa at the right.* ***Abbreviations****: CHX: Cyclohexamide, hr: hour, Untrans HEK: Untransfected human embryonic kidney cells, kDa- kilo Daltons, β TUB: β Tubulin.*

4.10 PLAGL1 has a degron site immediately after the Serine-Threonine phosphorylation sites at the N-terminal

In order to identify the site of degradation and undertsand the machanism by which PLAGL1 protein undergoes post-translational degradation, we analysed the PLAGL1 amino acid sequence for possible degron sites and motifs. Our analysis with bioinformatic tools based on

previously indicated conserved degron motifs (methods 3.4.3) determined that PLAGL1 has 2 NLS in the amino terminal of the protein. This region also has numerous serine-threonine sites that can get phosphorylated and harbour polyubiqutination for protease activity (Figure 25A). Our analysis also determined that S223-T232-S233 posses a degron site immediately after the rapid serine-throeonine phosphorylation sites which can aid protease activity (Figure 25A). The supposed cleavage of PLAGL1 protein at the S223-T232-ST233 degron site will result in 33 kDa amino terminal region and 51 kDa carboxyl end of PLAGL1 (Figure 25B).

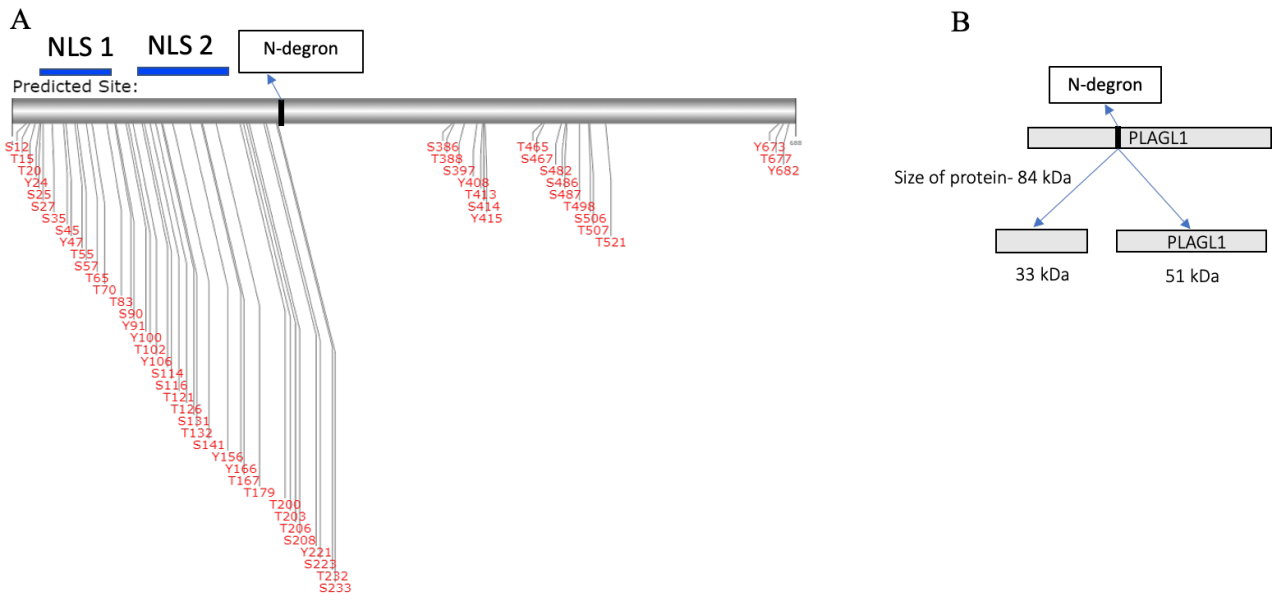


Figure 25: NLS and degron motifs in the N-terminal of PLAGL1

(A) *Schematic representation of NLS, serine-threonine phosphorylation and degron sites in the PLAGL1 amino acid sequence. The two NLS sequences are indicated in blue boxes and the degron site is indicated in black bar. The serine-threonine phosphorylation sites in PLAGL1 protein are indicated in red.* ***(B)*** *Schematic representation of PLAGL1 protein fragmentation after post-translational degradation at the degron site (S223-S233).* ***Abbreviations****: NLS: nuclear localization signal, s: serine, t: threonine, y: tyrosine, kDa: kilo Daltons.*

4.11 PLAGL1 and RYBP are co-expressed in the nuclei of the differentiating cardiac cultures

The fact, that there is no *Plagl1* in the *Rybp* null mutant ES cells and differentiated CMCs made us think whether *Plagl1* is regulated by RYBP. To deepen our understanding about the relationship between *Rybp* and *Plagl1,* we next analysed available *in vivo* evidence to see if

Rybp and *Plag11* are expressed in the same tissue types of the developing mouse embryo. We collected data containing RNA *in situ* hybridization experiments from existing publications about the expression pattern of *Rybp* and *Plagl1* in various tissue types during mouse embryonic development. This analysis revealed that *Rybp* and *Plagl1* co-expressed in several tissue types specific to the central nervous system, heart and eye (Table 2). We previously established that *Rybp* expressed moderately in the E8.5 and E9.5 heart (Ujhelly et al., 2015). RNA *in situ* hybridization and Northern blot analysis showed that *Plagl1* expressed immensely from as early as E7.5 heart until adulthood in a chamber-restricted pattern (Yuasa et al., 2010). These data suggested that the two proteins might be present in the same cells during CMC differentiation.

Therefore, to examine the subcellular localization of RYBP and PLAGL1 we co-stained the wild-type cardiac cultures, with anti-RYBP and anti-PLAGL1 antibodies. We performed immunostaining on cells derived during *in vitro* cardiac differentiation (i.e., d0, d2, d7, d10, d14 and d21). At d0, RYBP was observed both in the nuclei and cytoplasm of the wild type cells (Figure 26). At later time points RYBP was predominantly seen in the nuclei and expressed persistently at all the observed time points of *in vitro* cardiac differentiation (Figure 26). On the other hand, in the wild type cells PLAGL1 was not detected at early time points d0 and d2 (Figure 26). PLAGL1 signals were first observed from d7 and its expression gradually increased as differentiation proceeded with highest observed expression at d14 (Figure 26). At d7, which represents an early cardiac stage showed a mixed population of both PLAGL1 expressing and non-expressing cells suggesting that the cells are in heterogenous state of differentiation and PLAGL1 could start to be expressing in the differentiating cells only (Figure 26). These data suggested that the RYBP and PLAGL1 prominently co-expressed when cells start to differentiate presumably towards CMCs (Figure 26).

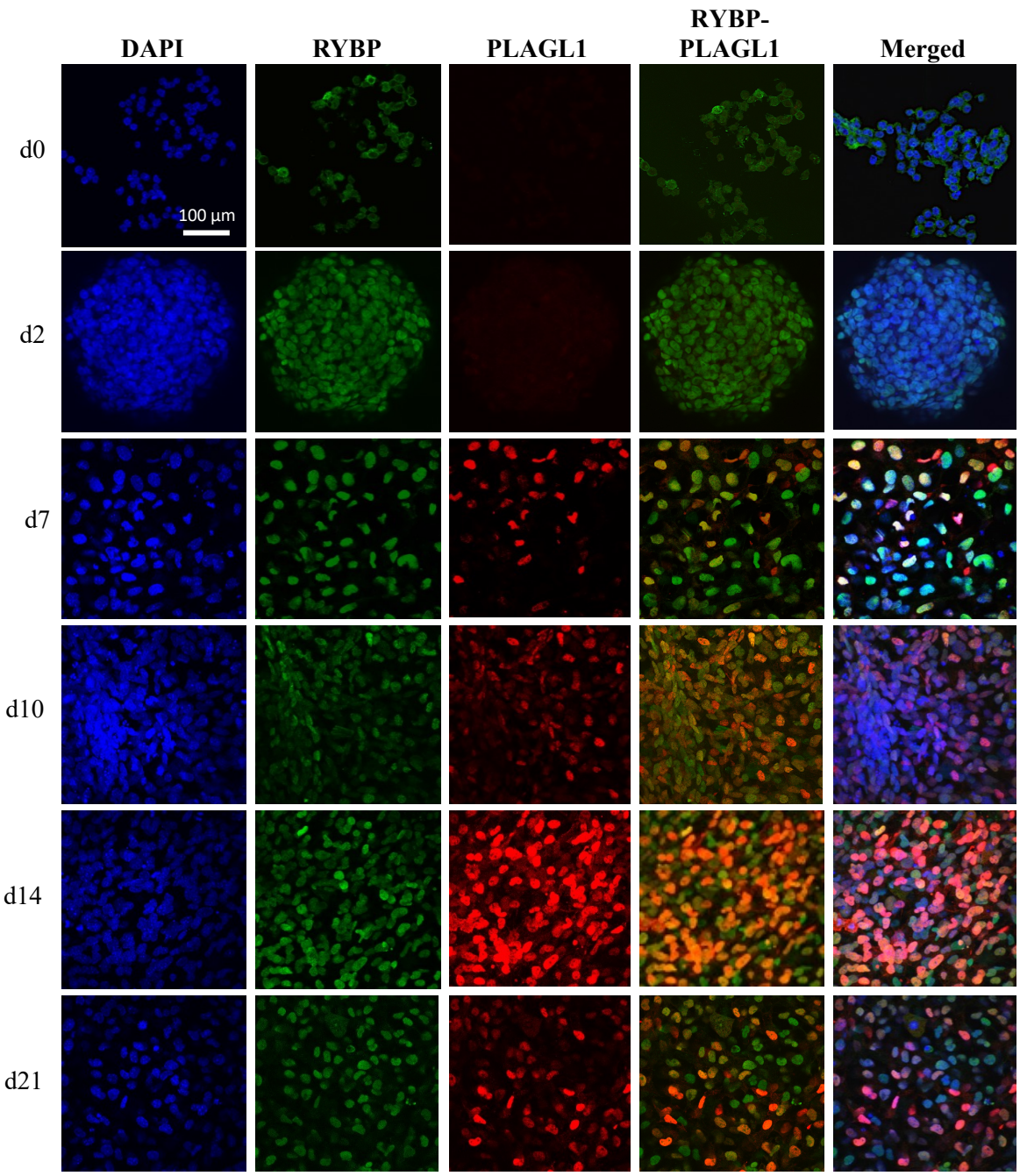


Figure 26: RYBP colocalized with PLAGL1 in the nuclei of differentiating cardiac cultures *Immunocytochemical analysis for the subcellular localisation of RYBP and PLAGL1 of wild type cultures during d0, d2, d7, d10, d14 and d21 time points of in vitro cardiac differentiation. Immunostainings: blue: DAPI (nuclei), green: RYBP, red: PLAGL1. Olympus Confocal IX 81, Obj.: 60 x; Scale bar: 100 µm.* ***Abbreviations:*** *d: day.*

4.12 *Plagl1* expression is first detected at early progenitor stage during *in vitro* cardiac differentiation

To determine the earliest time point, when *Plagl1* first appears in the differentiated cardiac cultures and to gain insights into the possible relationship between RYBP and PLAGL1, we checked the expression of *Rybp* and *Plagl1* between d2 and d7 time points. We performed a new experimental setup (methods 3.1.2) differentiating mouse ES cells until the early phase of *in vitro* cardiac differentiation deriving samples every day from d3, d4, d5 and d6 of *in vitro* cardiac differentiation. The EBs from hanging drops on d2 were plated into 6 cm plates for further gene expression analysis by qRT-PCR and in 24 well plates for protein localisation studies by ICC (methods 3.1.2). *Rybp* expressed persistently while *Plagl1* expression elevated after d3, and the expression increased for over 2 folds from d4 in the wild type cultures. As expected, *Plagl1* expression was not detected at any time point in the $Rybp^{-/-}$ cells (Figure 27A, B). From our previous results, we knew that RYBP was detected uniformly in the nuclei of d2 EBs (Figure 26). ICC experiments using d3, d4, d5 and d6 EBs revealed that RYBP was detected strongly in the outgrowth of the attaching EBs after d3 (Figure 27C e, f, g and h). PLAGL1 was more explicitly detected from d4 in the wild type cells, which is likely to correspond to the earliest days of progenitor formation stage during *in vitro* cardiac differentiation (Figure 27C j). The number of PLAGL1 positive cells gradually increased from d4 and more PLAGL1 positive cells were detected in d5 and d6 (Figure 27C k and l). From d4, RYBP and PLAGL1 co-expressed in the nuclei and the intensity of PLAGL1 signals varied in the heterogenous population of differentiating cells. The PLAGL1 expressing cells were mostly detected in the outgrowth of the attaching EBs, the place from where differentiation proceeds (Figure 27C n, o and p).

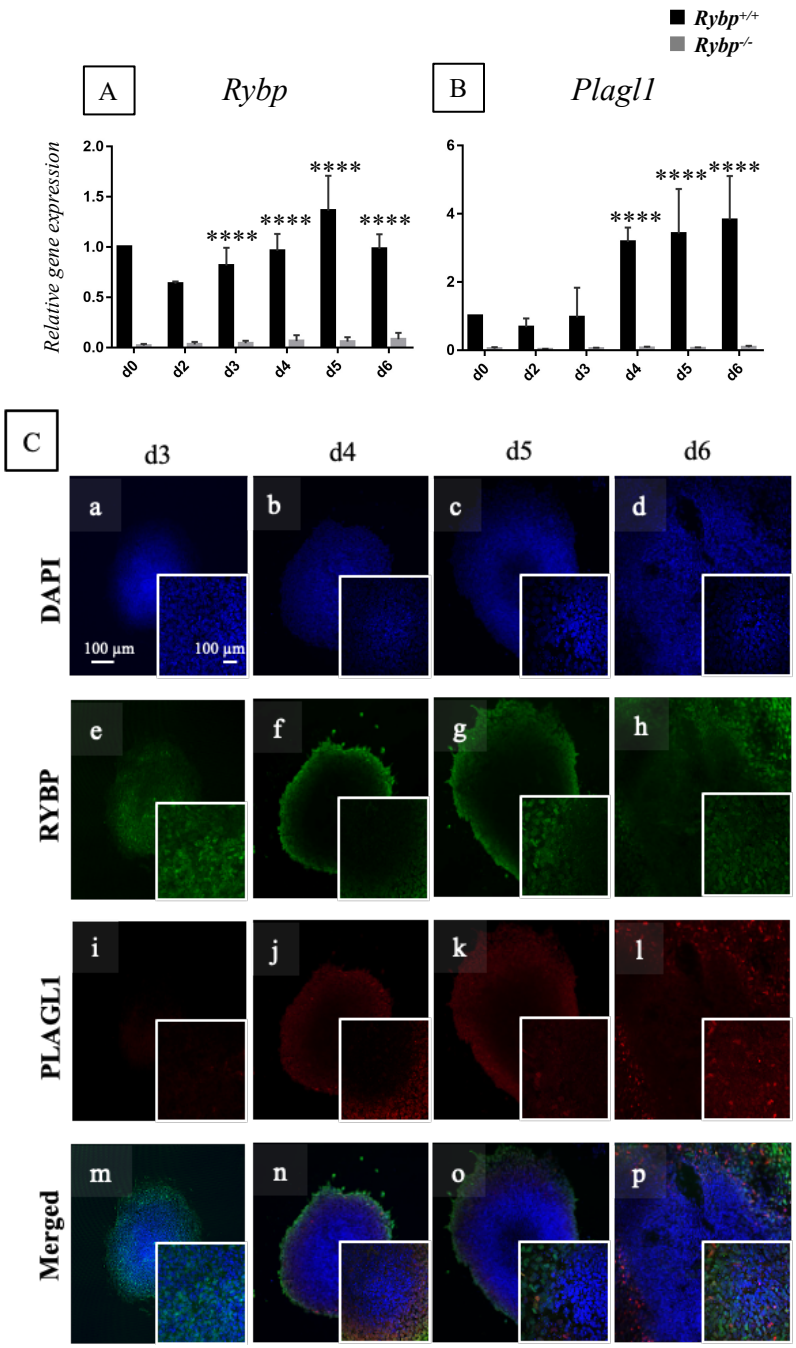


Figure 27: PLAGL1 expression is induced during the early progenitor stages of cardiac differentiation

(A, B) *Relative gene expression analysis of Rybp and Plagl1 by qRT-PCR in samples derived from in vitro cardiac differentiation at d0, 2, 3, 4, 5 and 6 days of cardiac differentiation. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed in the* $Rybp^{-/-}$ *compared to* $Rybp^{+/+}$ *by the statistical method two-way ANOVA (****p < 0.0001).* ***(C)*** *Immunocytochemical analysis of RYBP and PLAGL1 in day 0, 2, 3, 4, 5 and 6 samples of in vitro cardiac*

differentiated samples. Immunostainings: blue: DAPI (nuclei), green: RYBP, red: PLAGL1. Olympus Confocal IX 81, Obj.: 60 x; Scale bar: a-p: 100 µm. ***Abbreviations****: d: day.*

4.13 *Plagl1* promoters contain distinctive regulatory elements

Since our preliminary results suggested that *Plagl1* might be a downstream target of RYBP, we next investigated the regulatory elements in the *Plagl1* promoters. The *Plagl1* promoter sequences were downloaded from ENSEMBL (ENSMUSG00000019817) based on the previously identified sequences in Platas et al., 2012 and Platas et al., 2013. The promoter sequences were analysed for presence of CpG islands and TATA box (Figure 28). The *P1* promoter (4612 bp), which is the site of imprinting contains long stretch of CpG island (1673 bp) that covers 27.5% of the promoter region. The *P2* promoter which produces biallelic *Plagl1* transcripts and active only in disease conditions, has a 655 bp long CpG island whereas the *P3* promoter does not contain any CpG islands. The *P1* promoter contains a 14 bp long TATA box (TACAGTTTTTTATAC) away from the CpG island (Figure 28). The *P2* promoter does not contain any TATA box but contains three E-box consensus sequences (CACGTG and CAGCTG) which is not found in *P1* and *P3* promoters (Figure 28). The *P3* promoter contains a 67 bp long TATA box at the middle of the promoter region (Figure 28). The identification of these regulatory positions indicated potential regulatory mechanisms at these promoters also emphasizing the possible mechanism by which RYBP could regulate these promoters. RYBP containing ncPRC1s were previously identified to bind at the CpG islands (Farcas et al., 2012) and RYBP was shown to associate with consensus binding E-box binding homeobox factors (Zhu et al., 2017).

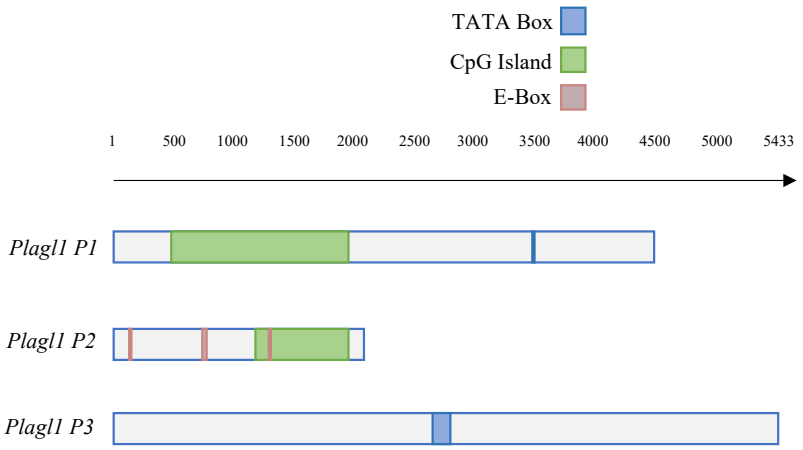


Figure 28: P1 and P2 promoters contain CpG islands and only P1 and P3 promoters have TATA box

Schematic representation of the regulatory elements in the Plagl1 P1, P2 and P3 promoters. The CpG islands (Green box) positions were identified from DBCAT (http://dbcat.cgm.ntu.edu.tw) and the positions of TATA box (Blue box) and E-box (light brown box) motifs were identified by TRANFAC (https://genexplain.com/transfac/). The labels for the identified regulatory elements in the Plagl1 promoters are presented at the top.

4.14 RYBP activates the *Plagl1 P1* and *P3* promoters

In order to elucidate if RYBP can directly influence the activation of *Plagl1* expression via its promoters, we performed luciferase reporter assays using reporter constructs containing *Plagl1 P1*, *P2* and *P3* promoters. To investigate this hypothesis, we cloned a 4612 bp long *Plagl1 P1* promoter region encompassing exon 4 and *P2* promoter encompassing exon 1 into pGL4.20 vectors. The promoter constructs were then transiently transfected into HEK293T cells in combination with RYBP cDNA constructs (Arrigoni et al., 2006) (methods 3.1.4). The protein cell lysates from the transfected cells were harvested 48 hrs after transfections and the luciferase levels were measured using a luminometer after inducing the luciferase signals with substrate. All three promoters *P1*, *P2* and *P3* containing luciferase constructs produced modest level of luciferase signals due to the endogenous transcription factors in HEK293T cells. The luciferase activity of the three promoters co-transfected with RYBP were normalized to the base level luciferase signal of the single transfected *P1*, *P2* and *P3* promoter constructs, respectively in all consecutive experiments. Our results showed that when co-transfected with RYBP, *Plagl1 P1* and *P3* promoter luciferase levels increased for up to 1.6-fold and 2.5-fold respectively when compared to the single transfected *P1* and *P3* promoter controls indicating

activation of the promoters by RYBP (Figure 29). In the contrary, the luciferase levels produced by *P2* promoter construct got marginally reduced to 0.8-fold under the influence of RYBP indicating that this *P2* promoter is not activated, but perhaps mildly repressed by RYBP.

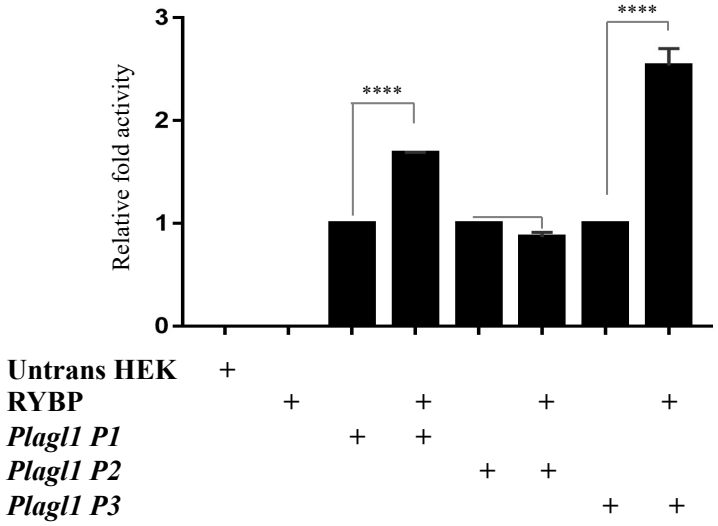


Figure 29: RYBP overexpression activated P1 and P3 promoters and repressed the P2 promoter

*HEK293T cells were transfected with 2.5 µg of Rybp and 5 µg of Plagl1 P1, P2 and P3 promoter containing luciferase reporter constructs. Luciferase activity of the transfected cell lysates were measured 48 hours after transfection. Values are expressed as fold changes of luciferase activity normalized to P1, P2 or P3 single transfected signals. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed from the value taken as 1 according to the statistical method one-way ANOVA (****p < 0.0001).*

4.15 The activation of *Plagl1 P1* and *P3* promoters by RYBP is polycomb independent

Since RYBP was originally identified as a polycomb protein and purified as a core member of the ncPRC1s, we wondered whether RYBP activates *Plagl1 P1* and *P3* promoters in a polycomb dependent mechanism.

4.15.1 A small molecule- PRC1 inhibitor did not attenuate the activation by RYBP at the *P1* and *P3* promoters

In order to unravel whether RYBP activates *Plagl1 P1* and *P3* promoters in a polycomb dependant or independent manner, HEK293T cells were transiently transfected with all three *Plagl1* promoter containing luciferase reporter constructs in combination with RYBP overexpression construct. 16 hrs after transfection the cells were treated with 50 μM of PRT4165 (PRC1 inhibitor) supplemented with media (methods 3.1.5). To let the inhibitory effect of PRT4165 develop, the transfected cells were cultured further for another 24 hours under humidified conditions and finally harvested for their protein cell lysates. Luciferase reporter assay was carried out using single transfected promoter constructs as normalization controls as described earlier (results 4.1.4). Our results showed that RYBP overexpression could still activate *Plagl1 P1* (1.5-fold) and *P3* (4.43-fold) promoters in the presence of the PRC1 inhibitor (Figure 30). Moreover, in these conditions the originally repressed *Plagl1 P2* promoter did not present any significant activity when compared to the single transfected control suggesting that the repressive activity of RYBP at the *P2* promoter was reversed by PRT4165. These results suggested that RYBP activated *Plagl1 P1* and *P3* promoters in a polycomb independent mechanism.

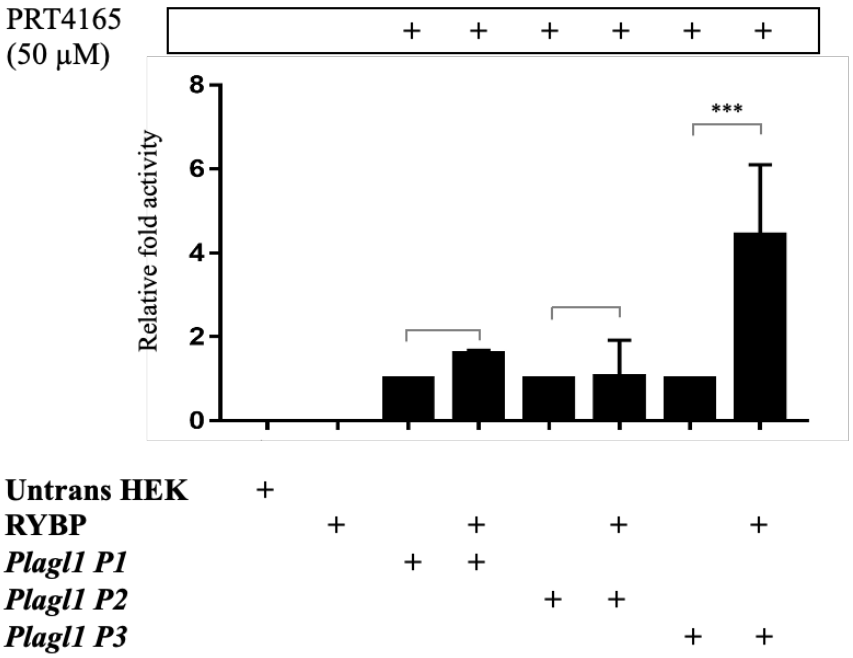


Figure 30: RYBP overexpression activates P1 and P3 promoters in a PRC1 independent mechanism.

*HEK293T cells were transfected with 2.5 µg of Rybp and 5 µg of Plagl1 P1, P2 and P3 promoter containing luciferase reporter construct. Transfected cells were treated with 50 µg of PRT4165 supplemented with media after 16 hrs after transfection. Luciferase activity of the transfected cell lysates were measured 48 hours after transfection. Values are expressed as fold changes of luciferase activity normalized to P1, P2 or P3 single transfected signals respectively for samples containing the respective promoters. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed from the value taken as 1 according to the statistical method one-way ANOVA (***p < 0.001).*

4.15.2 RING1 does not synergistically function with RYBP at the *Plagl1 P1* and *P3* promoters

RYBP and RING1 are core members of the ncPRC1s, physically interacting with each other (Gao et al., 2012). After elucidating that a potent PRC1 inhibitor PRT4165 did not affect the activation of the *P1* and *P3* promoters by RYBP, indicating that the activation of the two promoters by RYBP acts in a polycomb independent manner, we further confirmed this by checking if RING1 can act synergistically with RYBP. In these experiments, HEK293T cells

were transiently transfected with all three *Plagl1* promoter containing luciferase reporter constructs in combination with a concomitant RYBP and RING1 overexpression. Luciferase reporter assay was carried out using single transfected promoter constructs as normalization controls as described earlier (results 4.1.4). Our results determined that RING1 could not elevate the activation signals of both *P1* and *P3* promoters by RYBP (Figure 31A and B). The activity of RING1 itself caused mild increase in the luciferase levels in both *P1* (1.7-fold) and *P3* (1.6-fold) promoters although the measured levels never reached the level achieved by RYBP itself. In case of the *P2* promoter, which is in a maintained repressed state during normal developmental conditions, showed an unexpected increase of 2.8-fold by RING1 itself when compared to the *P2* promoter control disclosing the possible role of RING1 in PLAGL1 related disease conditions connected to *P2* promoter (Figure 31 C). These results further validate the previous results suggesting a polycomb independent activation of the *P1* and *P3* promoters by RYBP and at the same time pinpointing that the mild repressive effect of RYBP on the *P2* promoter is not a result of synergism between RYBP and RING1.

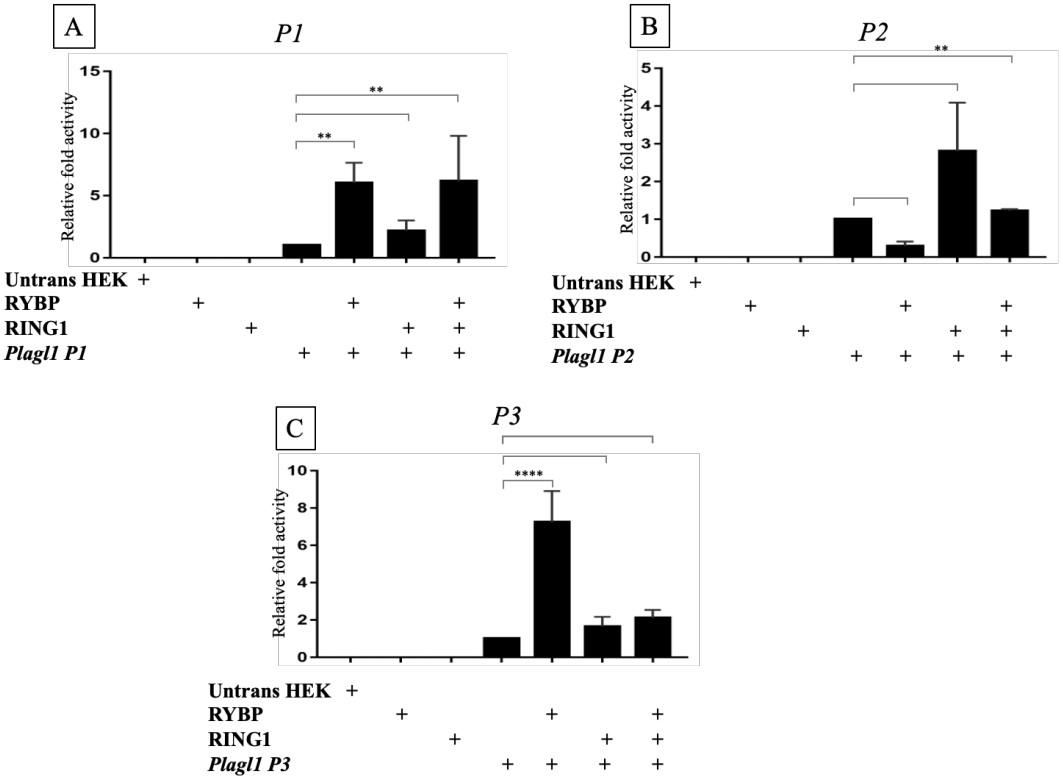


Figure 31: RING1 did not enhance the activation potential of RYBP at P1 and P3 promoters

HEK293T cells were transfected with 2.5 µg of Rybp, 2.5 µg of pcDNA3.1 Ring1 FLAG and 5 µg of Plagl1 P1, P2 and P3 promoter containing luciferase reporter construct. Luciferase activity of the transfected cell lysates were measured 48 hours after transfection. Values are expressed as fold changes of luciferase activity normalized to P1, P2 or P3 single transfected signals respectively for **(*A*, *B*** and ***C*)** *respectively. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed from the value taken as 1 according to the statistical method one-way ANOVA (**p < 0.01; ****p < 0.0001).*

4.15.3 Existing ChIP-seq analysis revealed a polycomb independent binding of RYBP at the *P3* promoter

In order to have a deeper understanding of our results generated from the luciferase reporter assays we analysed existing ChIP-seq binding datasets of RYBP and ncPRC1 core factor RNF2 in *Plagl1 locus* from both ES cells and cardiac progenitor cells (CPCs). ChIP-seq raw data were downloaded as BEDGRAPH files for RYBP (ES cells: GSM4052120; CPC: GSM1657391) and RNF2 (ES cells: GSM4052131; CPC: GSM1657390) and analysed for binding peaks at the *Plagl1* locus using IGV programme (methods 3.4.7). The files were

analysed to their respective experimental reference genome annotations i.e., mm10 for ES cells and mm9 for CPCs as disclosed by the authors. This analysis revealed binding of RYBP and RNF2 at *P1* and *P2* promoters and not at the *P3* promoter in ES cells (Figure 32). Since the PRC1s mediated regulatory functions have an affinity to bind to the CpG islands, the presence of CpG islands in *P1* and *P2* promoters (Figure 32) can explain the binding of both RYBP and RNF2 at these promoters as well as lack of binding at the *P3* promoter. *Plagl1* is not expressed in ES cells and hence the binding of RYBP and RNF2 at *P1* and *P2* promoters suggests a PRC mediated repression of these promoters. The *P3* promoter contains a ∽70 bp long TATA box (Figure 32) and surrounded with consensus sites for lineage specific transcription factors and hence might not express *Plagl1* in ES cells.

In CPCs both RYBP and RNF2 maintain strong binding at the *P2* promoter as the *P2* promoter is active only during abnormal conditions. At the *P1* promoter both RYBP and RNF2 are dispersedly bound suggesting a weak binding of these factors at this promoter. At the *P3* promoter RYBP showed adequate binding (indicated in red arrow) independent to RNF2 binding suggesting a direct activity of RYBP to regulate *Plagl1* expression via *P3*.

Since *Plagl1* expressed from the CPCs stage of cardiac differentiation (Figure 17) the direct binding of RYBP at the *P3* promoter in CPCs and not in ES cells together with the luciferase reporter assays (Figures 29, 30 and 31) gave a clear indication that RYBP indeed activates this promoter in a polycomb independent mechanism.

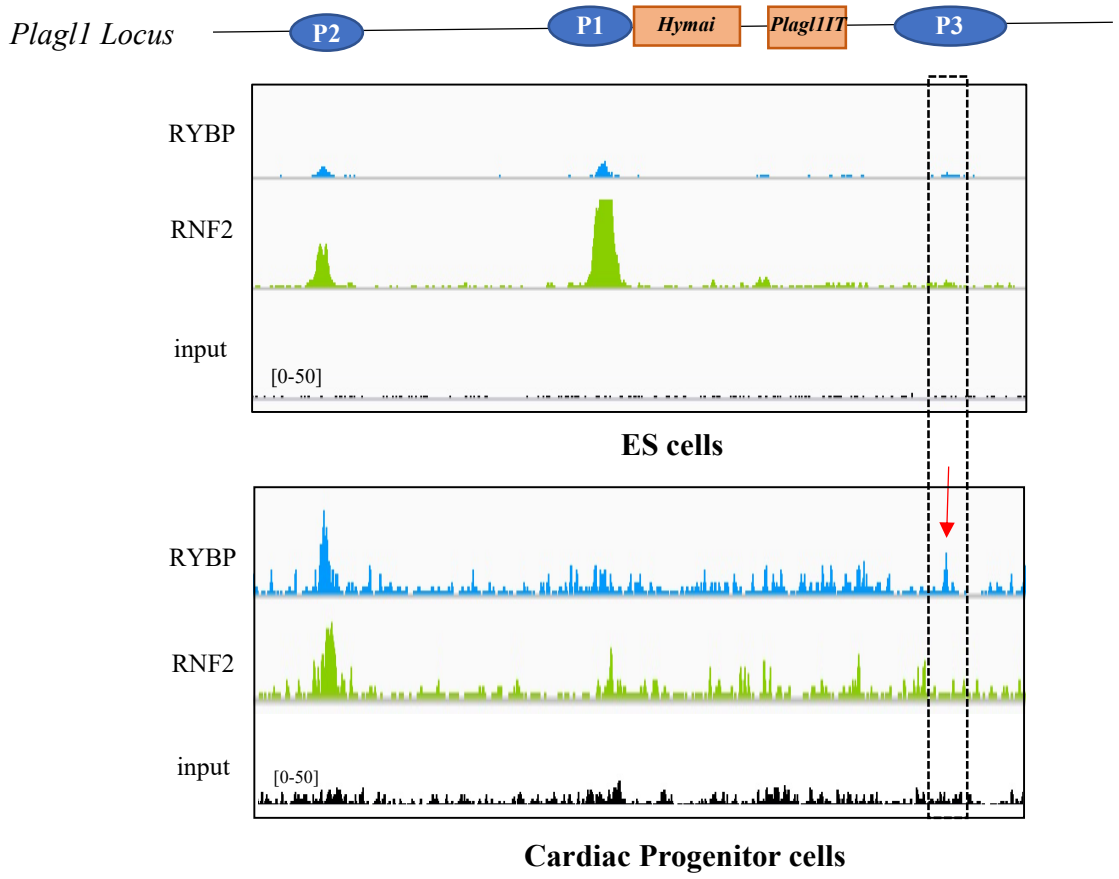


Figure 32: RYBP binds at the P3 promoter independent to RNF2 binding in CPCs and not in ES cells

Published ChIP-seq raw files of RYBP and RNF2 were downloaded from experiments generated from ES cells and CPCs. The downloaded files from Geo accession viewer (https://www.ncbi.nlm.nih.gov/geo/) under the IDs RYBP (ES cells: GSM4052120; CPC: GSM1657391) and RNF2 (ES cells: GSM4052131; CPC: GSM1657390) were processed in Integrative genome viewer (IGV). Binding peaks are displayed for RYBP (Blue) and RNF2 (Green) at 0-50 data range in both ES cells and CPCs at Plagl1 locus (displayed on top) with indicative promoter and ncRNA regions. The binding of RYBP at P3 in CPCs is indicated in red arrow highlighted with black dotted box.

4.16 RYBP did not activate the *P1* and *P3* promoters synergetic with E2F and YY1 transcription factors

After identifying that RYBP activates *Plagl1 P1* and *P3* promoters in a polycomb independent manner we were wondering the possible mechanism by which RYBP activates the two promoters. RYBP was previously identified to activate *Kdm2b* expression in ES cells by associating with pluripotency factor POU domain pluripotency factor 1 (POUF51, classically called as OCT4) (Li et al., 2017). Since we already reported that the expression of pluripotency genes including *Oct4* is not affected in the *Rybp* null mutant ES cells and *Plagl1* expression is only first seen after cardiac lineage commitment we ruled out the possibility of RYBP activation mechanism via OCT4. The other reported activation mechanism by RYBP was demonstrated to bridge E2F (E2F2 and E2F3) and YY1 transcription factors to activate *Cdc6*. In order to investigate if RYBP can activate *Plagl1* expression via E2F and YY1 we next performed luciferase reporter assay by co-transfecting HEK293T cells with the *P1*, *P2* and *P3* luciferase constructs in combination with RYBP, E2F2, E2F3 and YY1 overexpression. Luciferase assay was performed as mentioned earlier (methods 3.1.4). Our results determined that E2F2, E2F3 and YY1 could not elevate the activation levels of *P1* and *P3* promoters by RYBP (Figure 33A and C). Breaking down the results, in the case of the *P1* promoter interestingly E2F2 could induce high level of activation. In cells transfected with only E2F3, YY1 and in different combinations of RYBP, E2F2, E2F3 and YY1 overexpression the luciferase levels of the *P1* promoter did not exhibit any statistical differences (Figure 33A). In the *P2* promoter, as expected RYBP caused a decrease in luciferase levels when compared to the *P2* promoter control. Intriguingly single transfections and combinations of RYBP, E2F2, E2F3 and YY1 overexpression all resulted in the activation of the *P2* promoter (Figure 33B). In the *P3* promoter, single transfection with RYBP resulted in the highest activation of P3 and this activation did not get pronounced with the presence of E2F2, E2F3 or YY1 (Figure 33C). These results determined that RYBP is not activating *Plagl1 P1* and *P3* via interacting E2F and YY1. Interestingly RYBP could rather activate the *P2* promoter in combination with E2F and YY1 transcription factors revealing the unexpected role of RYBP in disease conditions and tumor formations.

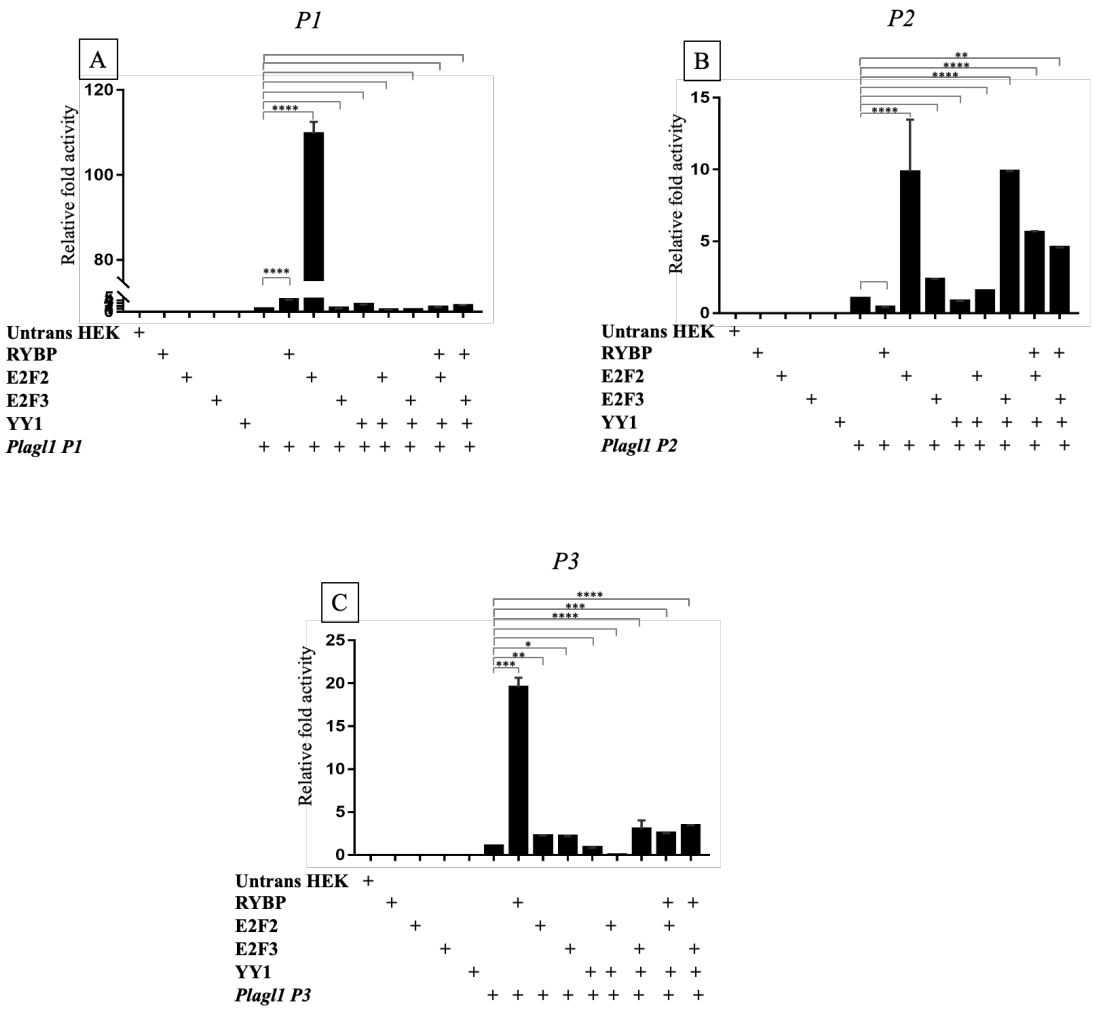


Figure 33: E2F and YY1 transcription factors cannot elevate the activation of P1 and P3 promoters in combination with RYBP

(A and C) *Luciferase reporter assay determined that E2F and YY1 cannot elevate the activation levels of P1 and P3 promoters by RYBP.* ***(B)*** *Luciferase reporter assay determined that E2F and YY1 increased the activation level of P2 promoter by RYBP. Values are expressed as fold changes of luciferase activity normalized to P1, P2 or P3 single transfected signals respectively for A, B and C respectively. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed from the value taken as 1 according to the statistical method one-way ANOVA (*p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001).*

4.17 *Hymai* and *Plagl1it* overexpression did not affect the activity of RYBP at the *Plagl1 P1* and *P3* promoters

Next, we further dissected the possible mechanism by which RYBP activates *Plagl1 P1* and *P3* promoters, considering the potential contribution of the ncRNAs located in the *Plagl1* genomic locus. NcRNAs such as *Xist*, *Meg3* and *Bvht* have been previously shown to transcriptionally regulate its target genes. The ncRNAs in the *Plagl1* genomic locus *Hymai* and *Plagl1it* ncRNAs express similar to *Plagl1* during cardiac differentiation. So we wondered if the two ncRNAs can synergistically act with RYBP. To test this hypothesis, we PCR amplified *Hymai* and *Plagl1it* from d14 differentiated wild type cDNA and cloned into pcDNA3.1 overexpression vector (methods 3.2.3). The cloned fragments were verified by sequencing and were transiently transfected to HEK293T cells with the *P1*, *P2* and *P3* luciferase constructs in combination with RYBP, *Hymai* and *Plagl1it* overexpression. Luciferase assay was performed as described earlier (methods 3.1.4). Our results determined that neither *Hymai* nor *Plagl1it* could synergistically act with RYBP to enhance the activation levels on the *Plagl1 P1* and *P3* promoters (Figure 34A, B and C). In brief, on the *P1* and *P3* promoters, *Hymai* (*P1*: 5.5-fold and *P3*: 11-fold) and *Plagl1it* (*P1*: 4.42-fold and *P3*: 3-fold) could exert activation compared to the activation of just RYBP (*P1*: 4-fold and *P3*: 3-fold) (Figure 34A and C). The activation levels of the *P1* and *P3* promoters induced by *Hymai* and *Plagl1it* were not increased in combination with RYBP when compared to the effects induced by just *Hymai* and *Plagl1it*. On the *P2* promoter, the two ncRNAs displayed no significant changes in combination with RYBP (Figure 34B). Our results also suggested that the two ncRNAs did not affect the *P2* promoter which lies 30 kb upstream to the TSS implying the target range of the two ncRNAs.

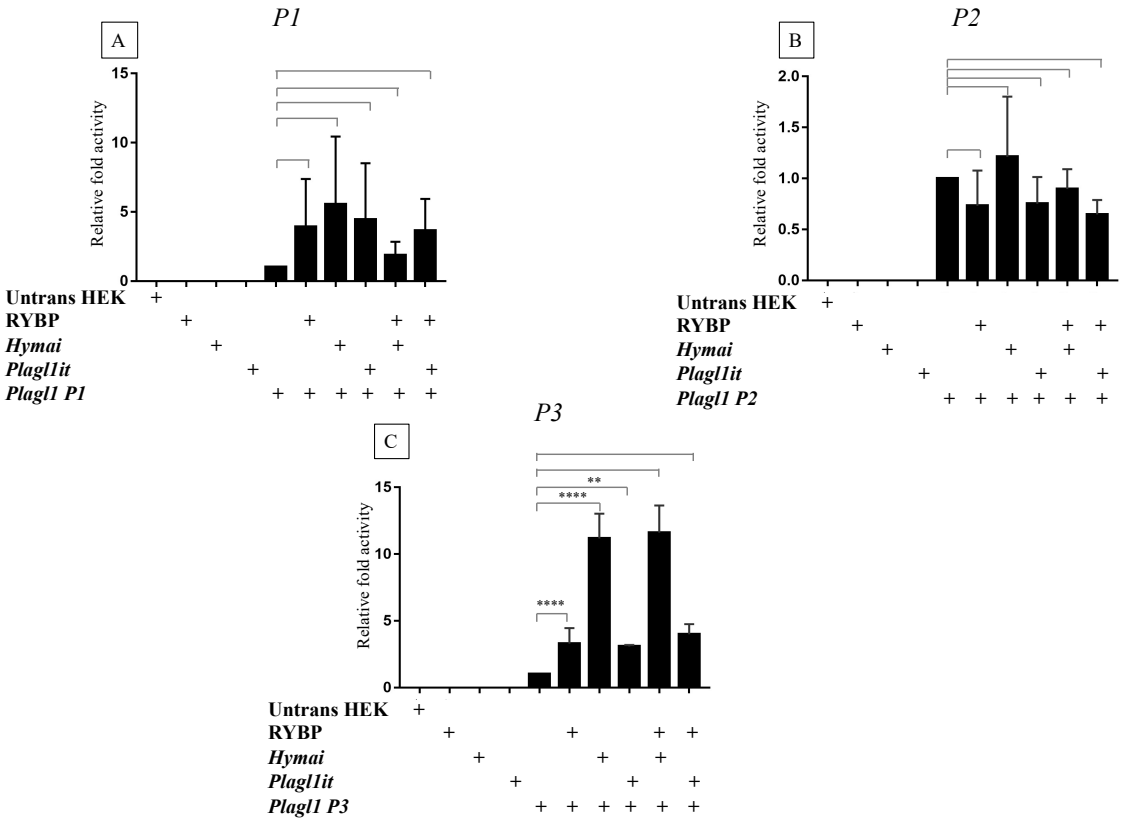


Figure 34: Hymai and Plagl1it did not enhance the activation levels of P1 and P3 promoter by RYBP

Luciferase reporter assay determined that Hymai and Plagl1it cannot elevate the activity of RYBP at the ***(A)*** *P1,* ***(B)*** *P2 and* ***(C)*** *P3 promoters. Values are expressed as fold changes of luciferase activity normalized to P1, P2 or P3 single transfected signals for A, B and C respectively. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed from the value taken as 1 according to the statistical method one-way ANOVA (**p < 0.01; ****p < 0.0001).*

4.18 RYBP activates the *P3* promoter via NKX2-5

In order to identify the mechanism by which RYBP could activate *Plagl1 P1* and *P3* promoter we next identified the minimum region required by RYBP to activate the promoter and combined with transcription factor binding sites (TFBS) analysis to detect the possible regulatory mechanism. Since the response to RYBP overexpression was higher in the *P3* promoter than the *P1* and due to the presence of enhancer elements such as a 67 bp long TATA box as opposed to the relatively smaller TATA box in the *P1* promoter, we utilized only the *P3* promoter for further analysis.

4.18.1 RYBP activates the 3' region of the *P3* promoter containing binding sites for NKX2-5 and MEF2C

Therefore, to detect the RYBP responsive regions in the *P3* promoter we made deletion mutant clones of this promoter by restriction digesting *P3* at several sites and re-cloned them back into luciferase reporter constructs (methods 3.2.3). These sub-clones were transfected in combination with RYBP cDNA and transfected in HEK293T cells for luciferase reporter assays. Our results demonstrated a surge in the activation of 3' half of the *P3* promoter when compared to the full length and 5' sub-clones of the *P3* promoter (Figure 35C). Sub-clones f, g and h of the *P3* promoter displayed high activation levels (∽ 20-folds) as opposed to the 5-fold activation by just RYBP. To gain insights into the RYBP response elements present in these sub clones we performed TFBS analysis using TRANSFAC (Figure 35A and B). We identified the presence of three *Nkx2-5* and one *Mef2c* binding sites, from which the first two *Nkx2-5* and the *Mef2c* sites were at close proximity to the TATA box. Also, from clone d and e which contained the first two *Nkx2-5* sites the activation levels were not as high as clone f, g and h suggesting that the third *Nkx2-5* site have more response to RYBP (Figure 35C). Comparing the activation signals of clone e, f, g and h, the *Mef2c* site did not seem to affect the activation by RYBP whilst the presence of *Nkx2-5* sites itself displayed higher activation signals (Figure 35C). These results implied that RYBP may regulate the region via *Nkx2-5* consensus.

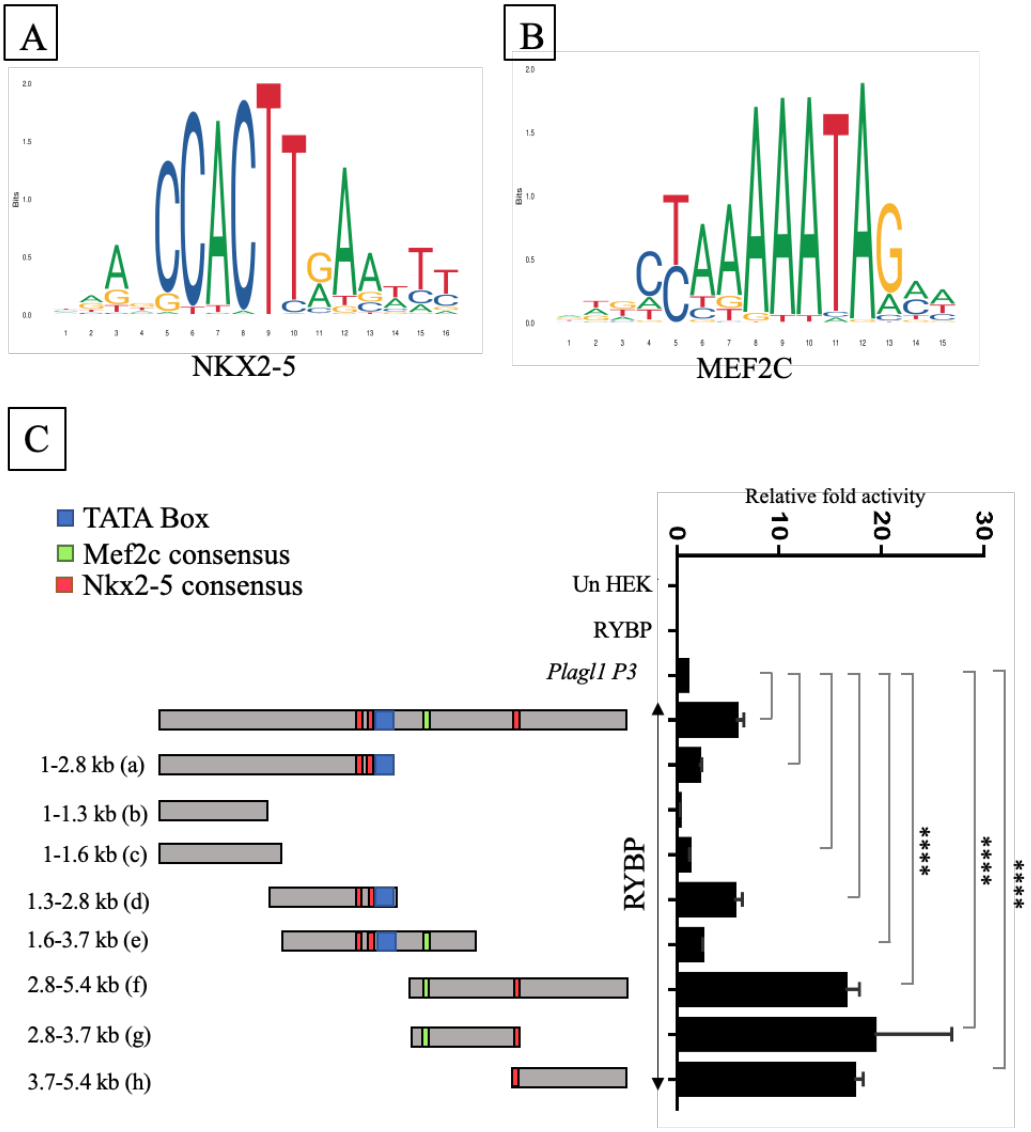


Figure 35: Deletion mutants of the P3 promoter resulted in the 3' fragments activated the highest by RYBP which contained NKX2-5 binding sites.

Consensus binding sites for ***(A)*** *NKX2-5 and* ***(B)*** *MEF2C procured from JASPAR database (http://jaspar.genereg.net).* ***(C)*** *Luciferase reporter assay using the various sub-clones of the P3 promoter labelled left to the schematic representation of the fragments. Clones containing regions 1-2.8 kb (a), 1-1.3 kb (b), 1-1.6 kb (c), 1.3-2.8 kb (d), 1.6-3.7 kb (e), 2.8-5.4 kb (f), 2.8-3.7 kb (g) and 3.7-5.4 kb (h) of the P3 promoter were transfected in HEK293T cells with RYBP. Values are expressed as fold changes of luciferase activity normalized to P3 single transfected signals. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed from the value taken as 1 according to the statistical method one-way ANOVA (****$p < 0.0001$).*

4.18.2 RYBP does not activate the *P3* promoter via MEF2C consensus sites

After identifying the presence of *Nkx2-5* and *Mef2c* binding sites at the RYBP responsive regions in the *P3* promoter, we further analysed the influence of these sites for activation by RYBP by performing site directed mutagenesis at these sites. Site directed mutagenesis was performed for the 3 *Nkx2-5* and 1 *Mef2c* sites by as altering the sites to form HindIII and BamHI sites (methods 3.2.4). Single mutants (clone 1, 2, 3 and 4) harbouring mutation for either *Nkx2-5* or *Mef2c* and progressive mutation of the sites harbouring more than one mutation of the *Nkx2-5* and *Mef2c* sites (clones 5, 6 and 7) were generated. The mutants were co-transfected with RYBP and luciferase reporter assay was carried out as mentioned earlier (methods 3.1.4). Our results showed the clones 1, 2 and 3 harbouring mutations for the 3 *Nkx2-5* binding sites displayed loss of activity implying the sites important for activation by RYBP (Figure 36). The activation signal did not get attenuated by the mutation of *Mef2c* site in clone 4 and the luciferase activity was over 20-fold implying that RYBP could activate this clone and that the *Mef2c* site was not important for the activation of the *P3* promoter by RYBP (Figure 36). The mutation of the *Nkx2-5* and *Mef2c* binding sites (clone 5, 6 and 7) also displayed attenuated luciferase levels suggesting that *Nkx2-5* site was indeed important for the activation of the P3 promoter by RYBP (Figure 36).

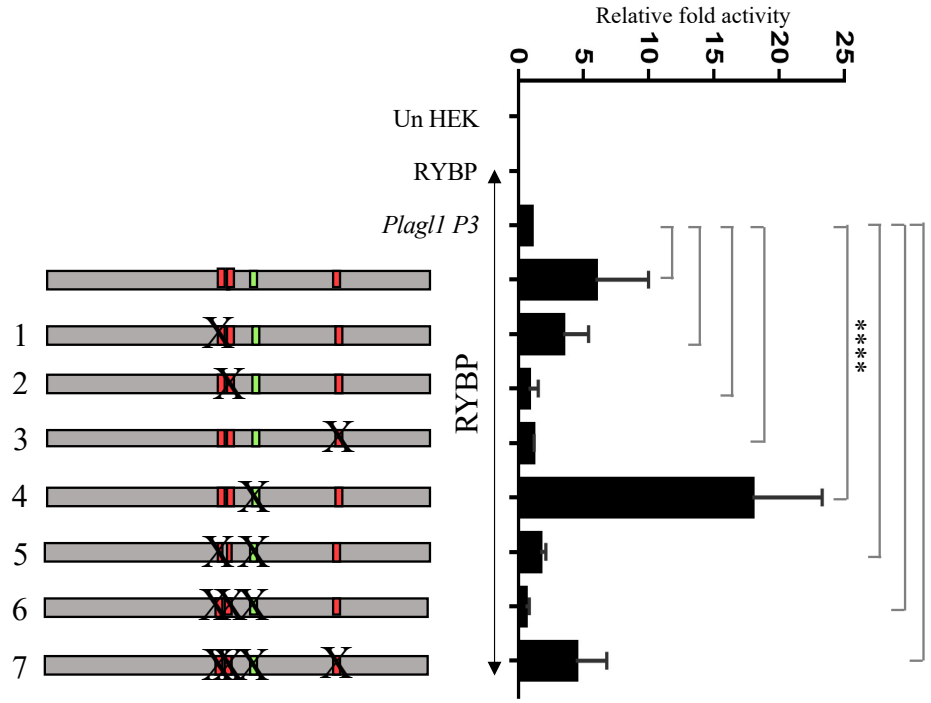


Figure 36: Binding site mutants of P3 promoter harbouring different mutations of NKX2-5 and MEF2C revealed that NKX2-5 is required for the activation of P3 by RYBP

*Luciferase reporter assay using the various mutants of the P3 promoter containing mutation for Nkx2-5 and Mef2c sites as indicated in the schematic representation. Clone 1 harbours mutation for the first Nkx2-5 site, clone 2 harbours mutation for the second NKX2-5 site, clone 3 harbours mutation for the third NKX2-5 site, clone 4 harbours mutation for the MEF2C site, clone 5 harbours mutation for the first Nkx2-5 and the Mef2c site, clone 6 harbours mutation for the first two Nkx2-5 sites and the Mef2c site and clone 7 harbours mutation for all the sites. The mutants were transfected in HEK293T cells with RYBP. Values are expressed as fold changes of luciferase activity normalized to P3 single transfected signals. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed from the value taken as 1 according to the statistical method one-way ANOVA (****$p < 0.0001$).*

4.18.3 RYBP and NKX2-5 can synergistically activate the *P3* promoter

Next in order to further clarify if RYBP and NKX2-5 can synergistically activate the *P3* promoter we performed luciferase reporter assay co-transfecting HEK293T cells with P3 promoter containing luciferase reporter and overexpression constructs for RYBP and NKX2-5. For this assay we generated a NKX2-5 overexpression construct by cloning NKX2-5 cDNA into pRK7FLAG vector. Luciferase reporter assay was performed as described earlier.

Our results showed that NKX2-5 can itself activate the *P3* promoter (10.5-fold) as opposed to the 3-fold activation by just RYBP (Figure 37). In combination with NKX2-5, the activation of the *P3* promoter by RYBP reached 80-folds implying that both RYBP and NKX2-5 can synergistically function in the activation of the *P3* promoter (Figure 37).

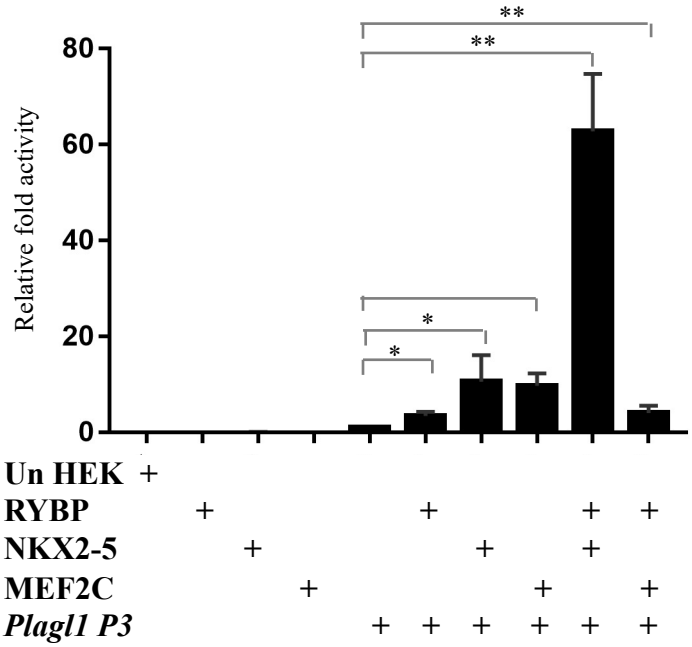


Figure 37: NKX2-5 overexpression elevated the activation of P3 promoter by RYBP

*Luciferase reporter assay using the P3 promoter in combination with RYBP and NKX2-5 revealed that the P3 promoter was activated extensively. Values are expressed as fold changes of luciferase activity normalized to P3 single transfected signals. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed from the value taken as 1 according to the statistical method one-way ANOVA (*p < 0.05; **p < 0.01).*

4.18.4 RYBP can interact with NKX2-5 protein

Since all previous results indicated that NKX2-5 might mediate the effects of RYBP in the activation of *Plagl1 P3*, we next examined if RYBP could interact with the NKX2-5 protein itself. To test this hypothesis HEK293T cells were co-transfected with RYBP in combination with either RING1-FLAG, FLAG-NKX2-5, FLAG-MEF2C and FLAG-PLAGL1 overexpression constructs. FLAG tagged RING1, an established interactor of RYBP (Garcia et al., 1999) in the ncPRC1s was used as the positive control in this experiment. Since the earlier experiments revealed that MEF2C might not affect the activation of *P3* promoter by

RYBP, we used FLAG tagged MEF2C as a potent negative control along with the PLAGL1. Co-immunoprecipitation (Co-IP) was performed with anti-RYBP tagged agarose beads (methods 3.3.3) and the immunoprecipitates were run in Western blot (methods 3.3.1) and hybridised with anti-FLAG antibody. Our results demonstrated that RYBP can interact with NKX2-5 protein. In brief, the cell lysate controls showed bands for RING1 (51 kDa), NKX2-5 (40 kDa), MEF2C (55 kDa) and PLAGL1 (81 kDa) upon hybridising with anti-FLAG antibody. In the immunoprecipitates, as expected RING1 was immunoprecipitated with RYBP (Figure 38). The presumed negative controls MEF2C and PLAGL1 were not immunoprecipitated by RYBP (Figure 38). These results demonstrated that RYBP interacted with NKX2-5 and this interaction is a novel finding which is vital for the activation of *Plagl1* expression.

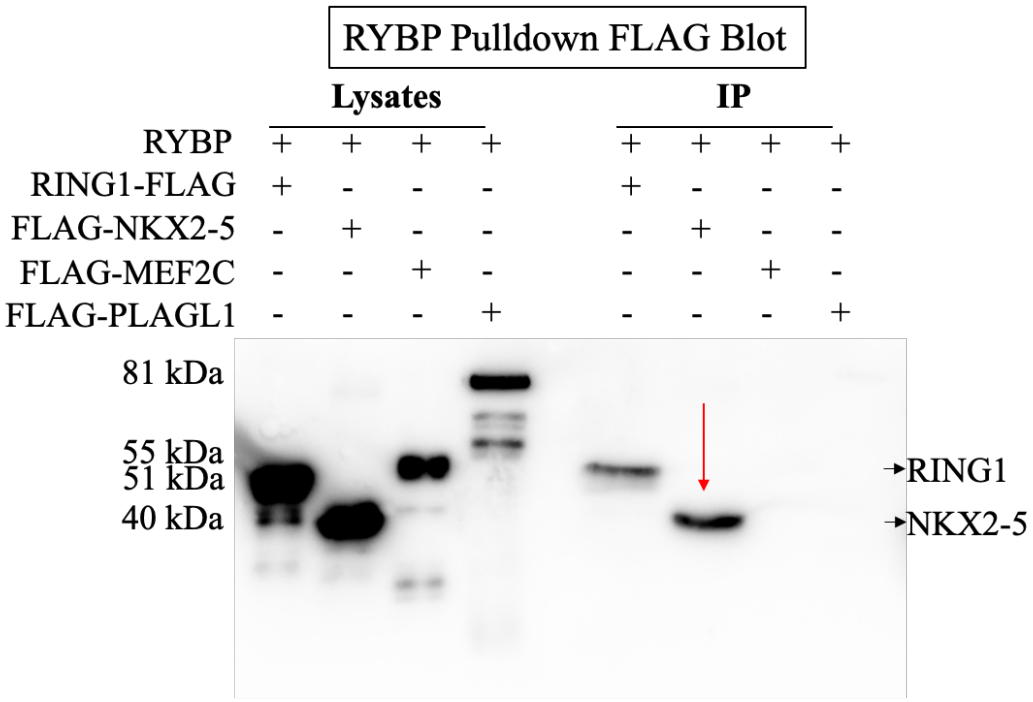


Figure 38: Co-immunoprecipitation revealed that RYBP interacts with NKX2-5 protein

Co-IP was performed from cell lysates derived from HEK293T cells co-transfected with RYBP with either RING1-FLAG, FLAG-NKX2-5, FLAG-MEF2C and FLAG-PLAGL1. The input lysates are in the left and the IPs are presented in the right with indicating labels at the top. The size of the bands is represented left to blot. RYBP interacted with the RING1 (51 kDa, positive control) and NKX2-5 (40 kDa) indicated with red arrow. RYBP did not interact with MEF2C and PLAGL1. ***Abbreviations:*** *IP: immunoprecipitates, kDa: kilo Daltons.*

4.18.5 RYBP bound at the NKX2-5 consensus at the *P1* and *P3* promoters in CMCs but not in ES cells

We next performed chromatin-immunoprecipitation (ChIP) coupled with qRT-PCR to identify if RYBP can directly regulate the *P3* promoter. ChIP was performed using EpiXplore ChIP kit (methods 3.2.2). The shearing of the chromatin was performed using a sonicator and the sheared chromatin were sustained at an average size of 200-800 bp (Figure 8, methods 3.2.2). The sheared chromatin was incubated overnight with anti-RYBP antibody and magnetic beads conjugated with mouse IgG (provided by the kit, details in methods 3.2.2). The bound chromatin was separated using a magnetic stand and eluted. The immunoprecipitated chromatin was eluted and used for qRT-PCR analysis using primers specific to the *Nkx2-5* and *Mef2c* consensus sites. This experiment was performed with chromatin isolated from wild type ES cells (d0) and d7 differentiated CMCs. Since *Plagl1* expression is not found in ES cells and only seen from the progenitor formation stage, d7 was used to identify the specific binding of RYBP at both the *P1* and *P3* promoter. 1% of the sheared DNA used for ChIP was used for the input control. The data was presented as input percentage by normalizing the Ct values of the ChIP to the respective adjusted input Ct values. Our results revealed that RYBP can bind at the *Nkx2-5* consensus sites in both *P1* and *P3* promoter in the d7 but not in d0 as expected (Figure 39). Since NKX2-5 is a cardiac transcription factor and not expressed in d0, the binding of RYBP at the *Nkx2-5* sites in d7 differentiating CMCs also implies the reason for *Plagl1* expression from cardiac progenitor formation stage.

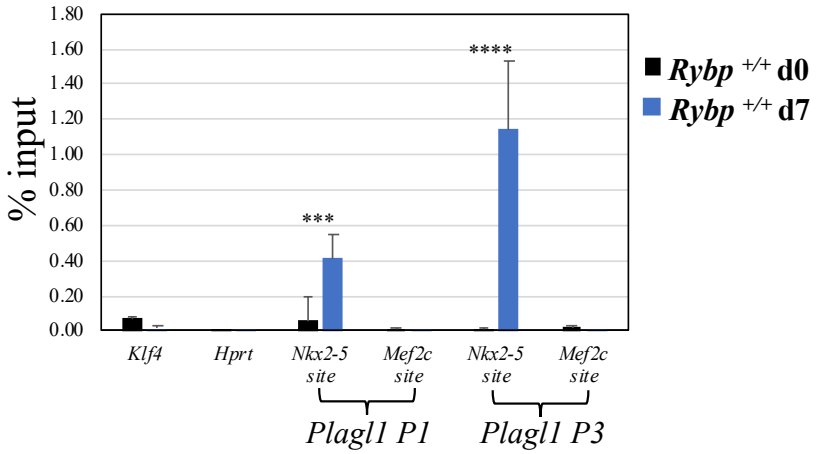


Figure 39: RYBP bound at the Nkx2-5 sites at P1 and P3 promoter in the wild type d7 differentiating CMCs and not in d0 ES cells.

ChIP-qRT-PCR was performed with sheared chromatin derived from d0 and d7 wild type cultures using primers specific to Nkx2-5 and Mef2c consensus. Primers specific to the previously described Klf4 promoter was used a positive control.

*RYBP bound discretely at the Nkx2-5 sites at both P1 and P3 promoter in the d7 differentiating CMCs and not in ES cells. Values are presented as input percentage normalized to adjusted input Ct values. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed in the $Rybp^{+/+}$ d7 compared to $Rybp^{+/+}$ d0 by the statistical method two-way ANOVA (****$p < 0.0001$).*

4.19 *Hymai* and *Plagl1it* synergistically functions with NKX2-5 to activate the *P3* promoter

As *Hymai* and *Plagl1it* ncRNAs could not affect the activation of the *P3* promoter by RYBP, we next wondered if the two ncRNAs could influence the activation of *P3* by NKX2-5. HEK293T cells were co-transfected with *P3* promoter containing luciferase reporter construct with either *Hymai*, *Plagl1it*, NKX2-5 overexpression or in combination with the ncRNA (*Hymai* or *Plagl1it*) along with NKX2-5 overexpression. Luciferase reporter assay was performed as mentioned earlier (methods 3.1.4). Our results showed that both *Hymai* and *Plagl1it* could act synergistic with NKX2-5 (Figure 40). In brief, both *Hymai* (8.5-fold) and *Plagl1it* (3-fold) could activate *P3* promoter as determined in the earlier experiments (Figure 40). NKX2-5 overexpression activated the *P3* promoter (5.6-fold) and this activation got enhanced with the presence of *Hymai* (72-fold) and *Plagl1it* (96-fold) (Figure 40). These

results suggested that both *Hymai* and *Plagl1it* ncRNAs synergistically activate the *P3* promoter with NKX2-5 implying that the ncRNAs also function in this activation mechanism (Figure 40).

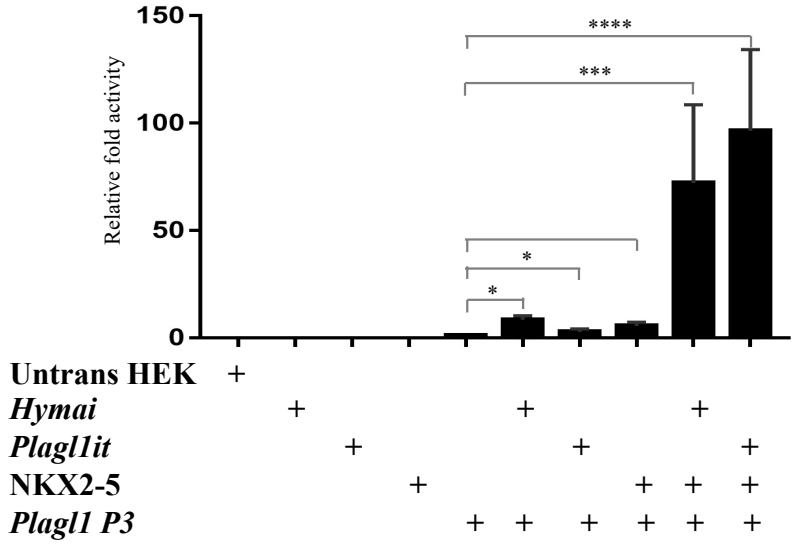


Figure 40: Hymai and Plagl1it ncRNA can enhance the activation of P3 promoter by NKX2-5.

*Luciferase reporter assay using the P3 promoter in combination with Hymai, Plagl1it and NKX2-5 revealed that the activation of the P3 promoter by NKX2-5 was enhanced significantly by both Hymai and Plagl1it. Values are expressed as fold changes of luciferase activity normalized to P3 single transfected signals. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed from the value taken as 1 according to the statistical method one-way ANOVA (*p < 0.05; ***p < 0.001; ****p < 0.0001).*

4.20 PLAGL1 is important for the formation of terminally differentiated CMCs and sarcomere organisation

Plagl1 is required for the proper formation of heart *in vivo* (Yuasa et al., 2010). *Rybp* null mutant cardiac differentiated cultures, lack *Plagl1* expression and could not form contracting CMCs. In order to characterize the functions of PLAGL1 during *in vitro* cardiac differentiation, we next checked if PLAGL1 affects the expression of sarcomeric genes which is crucial for contractility.

4.20.1 PLAGL1 co-expressed with CTNT in the differentiating CMCs

To elucidate if the expression of *Plagl1* admissible towards the formation of terminally differentiated CMCs, we performed ICC with wild type d7 and d14 differentiating CMCs co-stained for PLAGL1 and cardiomyocyte marker CTNT. CTNT is a classical cardiomyocyte marker and component of thin filament structure of the sarcomere. ICC was performed as mentioned earlier (methods 3.3.4). From our results, CTNT staining was visualized in the nuclei and dispersed along the cell of the d7 differentiating CMCs (Figure 41b). Nuclear staining of CTNT is previously reported in maturing muscle murine cells (Asumda & Chase, 2012). In d14 CMC cultures, CTNT staining was present in the cytoplasm (Figure 41f). PLAGL1 signals were strongly present in the cells that showed strong CTNT staining (indicated in white arrow, Figure 41d and h). These results determined that PLAGL1 and CTNT were co-expressed in the same differentiating CMCs and cells displaying strong CTNT staining also had strong PLAGL1 signals indicating the relevance of PLAGL1 expression towards the formation of terminally differentiated CMCs (Figure 41d and h).

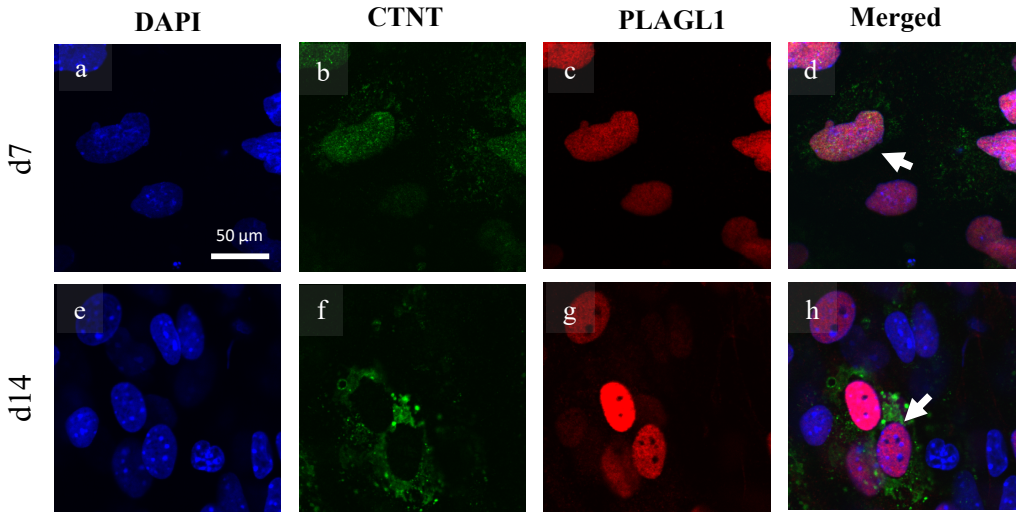


Figure 41: PLAGL1 and CTNT are co-expressed in the same differentiating CMCs

Immunocytochemical analysis of PLAGL1 (c, g) and CTNT (b, f) in day 7 and 14 samples of in vitro cardiac differentiated. DAPI (a, e) was used to stain the nuclei. Both PLAGL1 and CTNT co-stained the same differentiating CMCs (highlighted with white arrow). Immunostainings: blue: DAPI (nuclei), green: RYBP, red: PLAGL1. Olympus Confocal IX 81, Obj.: 60 x; Scale bar: a-h: 50 µm.

4.20.2 Sarcomere gene promoter regions contain consensus binding sites for PLAGL1

Since the expression of PLAGL1 protein had relevance to the formation of terminally differentiated CMCs, we were curious whether PLAGL1 could be involved in regulating sarcomeric gene expression. In order to achieve this, we first performed a consensus motif search for PLAGL1 binding at the sarcomeric gene promoters. Promoter sequence of sarcomere thin filament markers *Actc1*, *Tnnt2*, *Tnni3*, *Tpm1*, *Tpm4* and thick filament markers *Myh7*, *Myom1*, *Ttn* were downloaded from ENSEMBL (https://www.ensembl.org/index.html) in FASTA format based on the indicating promoter regions by the database. In case of more than 1 indicating promoter region, the promoter immediately upstream to the TSS was considered. The sequences were then analysed for the binding of PLAGL1, NKX2-5, MEF2C and TBX5 using JASPAR tool (http://jaspar.genereg.net). TFBS was performed by choosing consensus of these TFs from mouse (methods 3.4.5). The results were generated as binding scores with scores higher than 10 considered significant match to the experimentally identified consensus. This analysis resulted in several binding sites for PLAGL1, NKX2-5, MEF2C and TBX5 at each promoter region indicating that PLAGL1 could regulate these promoters parallel to NKX2-5, MEF2C and TBX5 which are already shown to affect the regulation of sarcomeric genes (Figure 42).

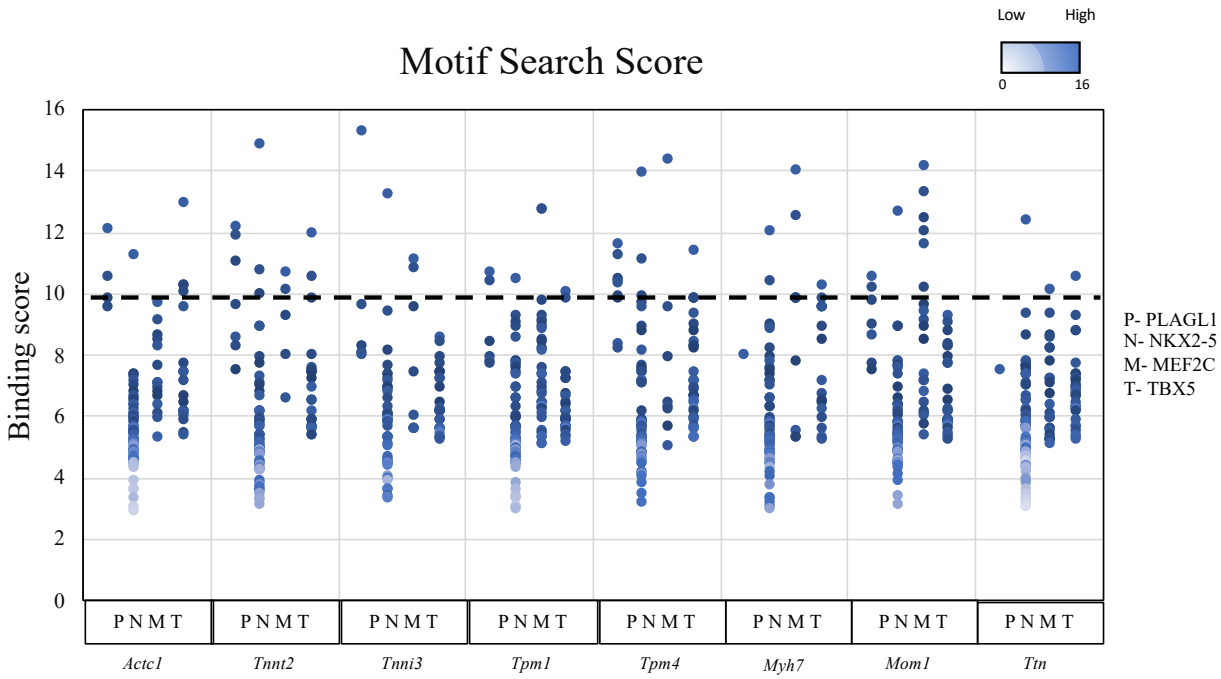


Figure 42: PLAGL1 can bind to the promoter of sarcomeric genes

Motif search for TFBS of PLAGL1, NKX2-5, MEF2C and TBX5 at the sarcomeric promoters were analysed from JASPAR. The resulting binding score was used to generate Manhattan plot. The gene names and the indicating TFs are presented below. The binding strength is presented at the left with binding score higher than 10 considered significant (indicated by a dotted line).

4.20.3 PLAGL1 can activate the *Tnnt2* promoter

Since our previous results indicated that PLAGL1 could potentially regulate the sarcomeric genes via their promoters, we tested this hypothesis by PCR amplifying and cloning *Tnnt2* promoter from genomic DNA derived from wild type ES cells (methods 3.2.3). The promoter was cloned into luciferase reporter constructs, verified by sequencing the fragment and co-transfected into HEK293T cells with either NKX2-5, MEF2C or PLAGL1. NKX2-5 and MEF2C were previously shown to affect the regulation of sarcomeric genes and were used as positive controls. Luciferase reporter assay was performed as indicated earlier (methods 3.1.4). Both NKX2-5 (16-fold) and MEF2C (4-fold) could activate the *Tnnt2* promoter as expected (Figure 43). PLAGL1 activated the *Tnnt2* promoter over 14-folds elucidating that PLAGL1 can activate the expression of sarcomeric genes via their promoters as analysed by *Tnnt2* (Figure 43). This novel finding also establishes the critical role of PLAGL1 during cardiac

differentiation and that the loss of *Plagl1* expression could affect differentiation by signalling sarcomere gene expression, which is vital for contractility.

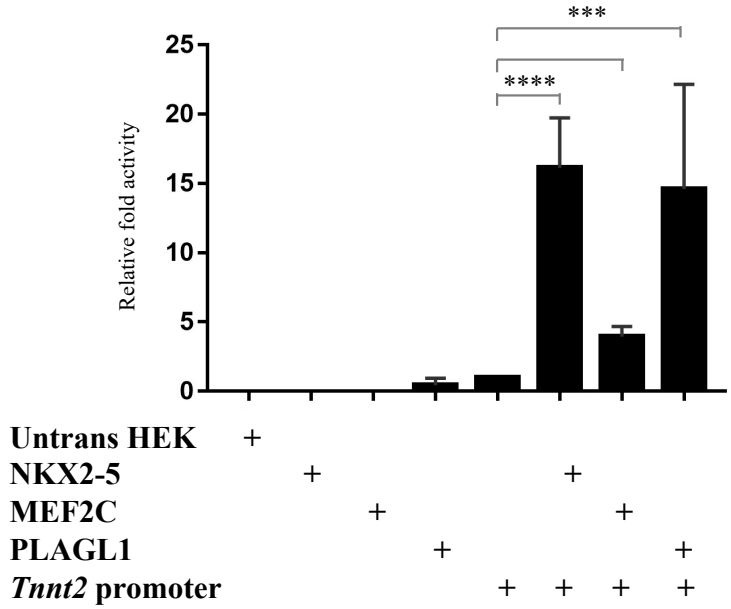


Figure 43: PLAGL1 activates sarcomeric thin filament marker Tnnt2 promoter

*Luciferase reporter assay using the Tnnt2 promoter in combination with NKX2-5, MEF2C and PLAGL1 revealed that the activation of the Tnnt2 promoter was enhanced significantly by PLAGL1. The activation of the Tnnt2 promoter by PLAGL1 was similar to the activity of the positive control NKX2-5 and higher than the activity of MEF2C. Values are expressed as fold changes of luciferase activity normalized to P3 single transfected signals. The presented values are averages of three independent experiments; error bars indicate standard deviation. Values indicated by asterisks significantly differed from the value taken as 1 according to the statistical method one-way ANOVA (***p < 0.001; ****p < 0.0001).*

5. DISCUSSION

RYBP is a member of the non-canonical polycomb repressive complex 1s (ncPRC1s), which are classically highlighted for their role as repressors (Garcia et al., 1999). As expected, genetic alterations of genes coding for polycomb proteins majorly resulted in the upregulation of genes in the mutant ES cells further confirming their role as epigenetic repressors. Recently, multiple studies have also revealed the downregulation of several genes in the polycomb mutant cells (Morey et al., 2013, 2015; Obier et al., 2015; Ujhelly et al., 2015; Zhao et al., 2020) For example, polycomb YY1 associated factor 2 (YAF2) is an ortholog of RYBP that shares over 65% homology to RYBP. Knock out of YAF2 in mouse ES cells resulted in 351 genes upregulated and 146 genes downregulated in the *Yaf2*$^{-/-}$ cells (Zhao et al., 2018). Furthermore, knockdown of Polycomb group ring finger 2 (PCGF2 also called as MEL18) using shRNA resulted in 720 upregulated genes and 148 downregulated genes in the shMel18 mutant ES cells. PCGF2/MEL18 is a distinctive member of the canonical PRC1.2 complex which functions in the maintenance of gene repression in ES cells and also plays essential roles during the early mesoderm precursor formation from mouse ES cells (Morey et al., 2015). To understand the mechanism of downregulation of target genes upon compromised expression of polycomb proteins has been of significant interest lately (Gao et al., 2014). Recently, the RYBP containing ncPRC1.3 and ncPRC1.5 complexes were identified to activate genes related to autism in the CNS by interacting with newly identified interacting partners AUTS2 and CK2 (Figure 5) (Gao et al., 2014).

Previously our group reported that compromised expression of RYBP leads to severe alterations in the expression of many genes essential for cardiac differentiation in the mutant cells (Figure 15) (Ujhelly et al., 2015). Several cardiac transcription factors such as *Isl1*, *Tbx5* and *Tnnt2* express at reduced mRNA levels in the *Rybp* null mutant CMCs (Figure 16). One of the most downregulated genes was *Plagl1*, which was of utmost interest to us since PLAGL1 is a key cardiac transcription factor (Yuasa et al., 2010). PLAGL1 can bind at critical cardiac gene loci, such as *Nkx2-5*, Tropomyosin 2 (*Tpm2*), Plectin (*Plec*) and Myosin heavy chain 7B (*Myh7b*) identified by ChIP-seq assay in N2A cells (Varrault et al., 2017).

In the current study, we provide evidence that *Plagl1* is not expressed in the *Rybp*$^{-/-}$ cells neither in mRNA nor at protein levels during any stages of *in vitro* cardiac differentiation. Based on our results we hypothesized that *Plagl1* might be a downstream target of RYBP. To examine this hypothesis, we first tried to understand the complex regulatory regions of *Plagl1*

expression. We dissected the *Plagl1* genomic locus, described all splice variants and the coded protein products in detail. The *Plagl1* genomic locus has had many schematic illustrations updated several times. The original description of the *Plagl1* locus showed that *Plagl1* had 8 exons with one promoter region (*P1* promoter). The *P1* promoter resides upstream to exon 4 containing a conserved CpG island which is also present in rat and human PLAGL1 (Figure 19) (Smith et al., 2002). Later Valleley and colleagues (Valleley et al., 2010) identified a second promoter region, which they termed as *P2* situated 30 kb upstream to the previously described *P1* promoter. The *P2* promoter has an unmethylated CpG island which can produce biallelic expression of *Plagl1* in leukocytes and pancreas during disease condition such as TNDM. Another independent study about the *Plagl1* genomic locus showed that *Plagl1* had 9 exons with a 3 kb promoter region upstream to exon 4 (*P1* promoter) that can be activated by cardiac transcription factor NKX2-5 (Yuasa et al., 2010). Next, demonstrating that MEF2C can regulate *Plagl1* expression in rat mesenchymal cells (Czubryt et al., 2010), identified a novel 5.4 kb alternate promoter region (*P3* promoter) containing a MEF2C consensus site immediately upstream to the exon 10 by *in silico* analysis. The latest illustration of the mouse *Plagl1* locus gave the most detailed description of the promoters and the presence of the two ncRNAs *Hymai* and *Plagl1it* (Platas et al., 2012). By performing 5' and 3' RACE along with EST analysis the position of the transcripts was mapped from the three different promoter regions which were described previously naming them as *P1, P2* and *P3* promoters (Platas et al., 2012). In our study, we further verified the positions of the exons, promoters, regulatory elements and the ncRNAs *Hymai* and *Plagl1it* ncRNAs. Our schematic illustration of the *Plagl1* locus is an updated depiction from all previously available data (Figure 19). We also provide novel information about the promoters potentially active during cardiac development as previous studies only aimed at characterizing the promoters and elucidating imprinting at the genomic locus. We performed gene expression analysis using exon specific primers *Plagl1 1/2*, *Plagl1 6/7*, and *Plagl1 10/11* to validate which promoter were producing protein coding transcripts during *in vitro* cardiac differentiation. Our results suggested that only the *P1* and *P3* promoters are active during cardiac differentiation (Figure 22). This data further strengthens the theory that the *Plagl1 P2* promoter is only active and produces biallelic expression of *Plagl1* in specific tissues such as leukocytes and pancreas during disease states such as TNDM.

PLAGL1 protein was previously identified to express abundantly in the embryonic heart in a chamber restricted pattern (Yuasa et al., 2010). Present study provides novel information characterizing the expression of PLAGL1 during *in vitro* differentiation of ES cells into cardiac

lineage. In the wild type cultures, *Plagl1* expression was not observed at day 0 and day 2 indicating that PLAGL1 may not have any role in maintaining the pluripotency of ES cells and in consequent germ layer formation (Figure 18). PLAGL1 expression was first detectable at day 4, which is a very early stage when the cardiac progenitors form indicating its role in early lineage commitments. Our gene expression analyses and co-expression of RYBP and PLAGL1 in the differentiating cells raised our hypothesis that *Plagl1* could be a downstream target of RYBP during cardiac morphogenesis (Figure 26).

By performing luciferase reporter assay we have examined the role of RYBP in regulating *Plagl1* at the promoter levels, and we determined that RYBP activates the *P1* and *P3* promoters. In order to clarify the mechanism by which RYBP activated the *Plagl1* promoters we demonstrated that RYBP activates both *P1* and *P3* promoter in a polycomb independent manner (Figure 29, 30 and 31). By performing similar promoter activity assays using RYBP in combination with E2F (E2F2 and E2F3) and YY1 transcription factors we also determined that the activation of the *Plagl1* promoters by RYBP differs from the previously identified activation mechanisms by RYBP (Figure 33).

Several ncRNAs have been identified to play vital roles in cellular processes including chromatin remodelling, DNA repair and translation (Wang & Chang, 2011). The recent development in the ncRNAs have increased the interest in identifying the interactions and functions of ncRNA. Several ncRNAs have been already identified in the polycomb mediated regulation of genes. Inactive X specific transcripts (*Xist*), Braveheart (*Bvht*), and Maternally expressed gene 3 (*Meg3*) are some of the many ncRNAs identified to interact with polycomb complex proteins (Bousard et al., 2019; Klattenhoff et al., 2013; Sunwoo et al., 2015; Wu et al., 2018). *Xist* mediates whole chromosome transcriptional silencing during the dose compensation process in mammals (Sahakyan et al., 2018). *Bvht* and *Meg3* are determined to likely induce cardiac lineage commitment and are expressed upstream to *Mesp1*, potentiate to regulate a core cardiac gene network (Klattenhoff et al., 2013; Wu et al., 2018).
Hymai and *Plagl1it* ncRNAs are both imprinted and express only from one allele from the *Plagl1* genomic locus (Benedetti et al., 2017). The expression of *Hymai* is partially connected to the expression of *Plagl1* since *Hymai* is also transcribed upon the regulation of *P1* promoter and altered expression of both PLAGL1 and *Hymai* are indicative of disease condition such as TNDM and tumors. *Plagl1it* ncRNA was identified from independent RACE experiments (Platas et al., 2012). Although few studies have described the functions of *Hymai*, not much is known about *Plagl1it*. From our expression studies by qRT-PCR we also determined the

altered expression pattern of the ncRNAs *Hymai* and *Plagl1it* in the *Rybp* null mutant cultures during *in vitro* cardiac differentiation (Figure 20). High expression of the two ncRNAs at day 14 of wild type cardiac differentiation CMCs shows that the ncRNAs might function during the terminal stages of CMC formation (Figure 20). Therefore, we were keen to investigate whether *Hymai* and *Plagl1it* affected the activation of *Plagl1* by RYBP. Our results determined that the two ncRNAs could activate the *Plagl1 P1* and *P3* promoters themselves. But the overexpression of either *Hymai* or *Plagl1it* did not alter the activity of the *Plagl1* promoters regulated by RYBP (Figure 34).

To identify the mechanism by which RYBP can activate *Plagl1* expression, we subcloned the *P3* promoter as various smaller fragments and determined their inducibility by RYBP. By combining the acquired results with TFBS we demonstrated that consensus binding sites for cardiac progenitor transcription factors *Nkx2-5* and *Mef2c* was required for the activation of the *P3* promoter by RYBP (Figure 35). Two previous studies have shown that *Nkx2-5* and *Mef2c* can activate the expression of *Plagl1*. (Czubryt et al., 2010; Yuasa et al., 2010). By performing site directed mutagenesis of the *Nkx2-5* and *Mef2c* consensus sites we confirmed that the *Nkx2-5* consensus site was indeed essential for the activation of the *P3* by RYBP (Figure 36).

By combining Co-IP and ChIP-qRT-PCR methods we also proved that RYBP interacts with NKX2-5 protein (Figure 37) and this interaction is vital in the regulation of *Plagl1* in the wild type CMCs. RYBP is bound at the *Nkx2-5* consensus sites in both *P1* and *P3* promoters in d7 CMCs and not in d0 ES cells where *Plagl1* is normally expressed (Figure 38). Since *Nkx2-5* only expressed after cardiac lineage commitment and vital for the progenitor formation, these results nicely recapitulate the expression kinetics of *Plagl1* during cardiac differentiation. Our results also indicate a potential interaction of *Hymai* and *Plagl1it* with NKX2-5 as they increased the fold activation levels of the *P3* promoter by NKX2-5 extensively (Figure 40).

Our results also suggest that although aberrant expression of *Nkx2-5*, *Hymai* and *Plagl1it* in the *Rybp* null mutant CMCs could itself impact *Plagl1* regulation, the consequence of the absence of *Rybp* might result in a more several decline in the activation of *Plagl1* (Figure 20). Since *Plagl1* expression was not detected neither at mRNA nor at protein levels in the *Rybp* null mutant cultures, the loss of *Plagl1* functions could at least partially contribute to the phenotype of the *Rybp* null mutant CMCS during cardiac differentiation. *Rybp* null mutant CMCs lack proper sarcomere formation and subsequent contractility (Figure 16) (Henry et al., 2020). Therefore, to examine whether *Plagl1* expression is important for sarcomere formation,

we determined the colocalization of PLAGL1 and a sarcomere marker CTNT by ICC. We found that PLAGL1 expression was profoundly present in the CTNT positive cells, suggesting that PLAGL1 preferentially expressed in cells differentiating towards terminal CMCs. Luciferase reporter assays showed that *Plagl1* can activate the expression of *Tnnt2* promoter. These data together provide vital understanding about the role of *Plagl1* in CMC formation and the molecular mechanism by which RYBP functions during cardiac morphogenesis (Figure 43 and 44).

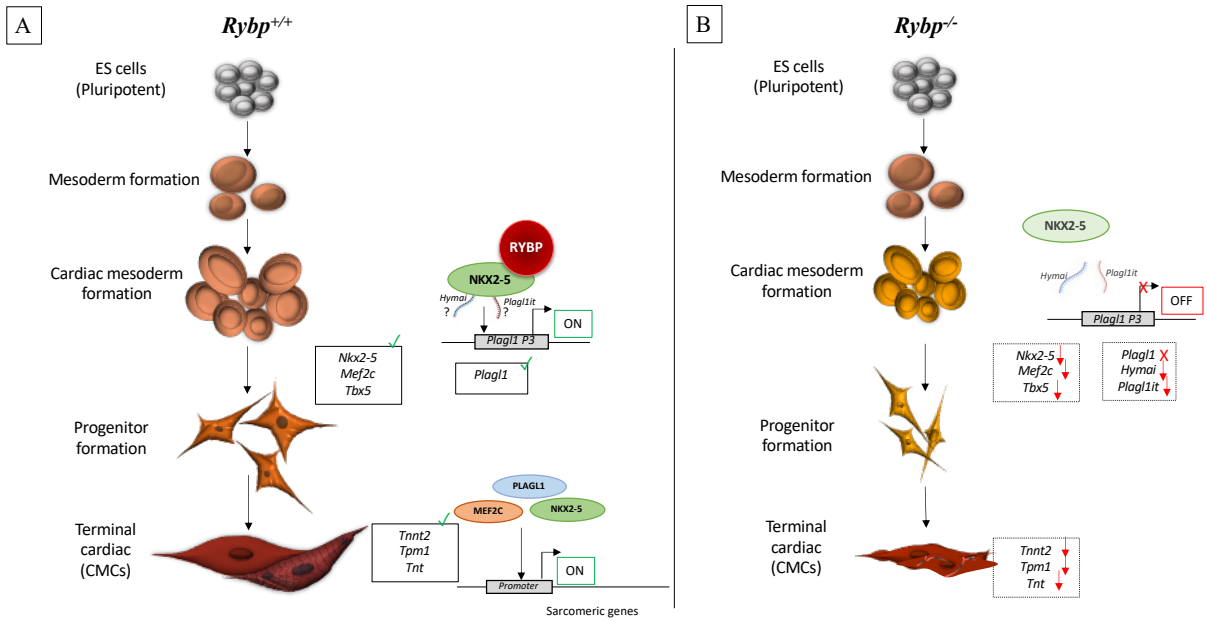


Figure 44: Model of in vitro cardiac differentiation in the presence and absence of Rybp

*(**A**) Differentiation of wild type ES cells towards CMCs in the presence of Rybp. The expression of key cardiac transcription factors such as NKX2-5, MEF2C and TBX5 guide cardiac lineage commitment and differentiation. In the wild type, RYBP interacts with NKX2-5 to activate Plagl1 expression associating with Hymai and Plagl1it. PLAGL1 activates Tnnt2 expression via its promoter and thereby affects sarcomere formation. (**B**) Differentiation of pluripotent ES cells towards CMCs is impaired in the absence of Rybp. During cardiac differentiation progenitor transcription factors (Nkx2-5, Mef2c and Tbx5) exhibit reduced expression levels in the Rybp null mutants. Moreover, in the absence of Rybp, Plagl1 expression is absent, which also affects sarcomere formation. As a result, the formation of terminally differentiated cardiomyocytes is impaired and contractility is compromised in the lack of functional Rybp.*

Taken together our results provide vital novel information about the regulatory functions of RYBP during cardiac development. We propose that RYBP acts by activating specific cardiac genes via *Plagl1* (Figure 44). The interaction between RYBP and NKX2-5 protein is a novel finding, and these results broadens the understanding about the alliance between polycomb proteins and lineage specific markers to regulate differentiation. Overall, these results also affirm the theory that in certain cases polycomb proteins, such as RYBP could also exert their roles as transcriptional activators during mammalian embryonic development.

6. ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor Dr. Melinda K. Pirity, for giving me the opportunity to carry out my PhD studies at her laboratory. Her expert advice, guidance and encouragement have been invaluable at all stages of my project.

I would like to thank all the current and alumni members of the Laboratory of Embryonic and Induced Pluripotent Stem Cells. Special thanks to Dr. Izabella Bajusz, Dr. Gergő Kovas, Viktória Szabó, Enikő Sutus and Krisztina Dudás who provided me an excellent environment to work and learn. Apart from the technical and intellectual support I also made some good friendship and fond memories with all of them.
I also extend my thanks to the new members in our laboratory Lilla Kókity, Lili Adorján and Katalin Kokavszky who showed immense enthusiasm for learning and scientific discussions in the past one year. I will also cherish our time playing "Unstable Unicorns" card games.

I express my gratitude to Dr. Miklós Erdélyi for giving me the opportunity to work in the Institute of Genetics, Bilogical Research Centre (Szeged, Hungary). I express my thanks to Dr. Lajos Haracska, Dr. Gabriella Endre and Dr. Eva Monostori for letting me use the instruments at their laboratories. I thank Dr. Ferhan Ayaydin for assisting with the laser confocal microscopy and Dr. Laszló Kozma Bognar for the technical assistance with Luciferase assays during my project work.

I also thank Dr. Krisztina Buzás and Dr. Tamás Csont for accepting to review my PhD dissertation for defense.

Special thanks to my dear wife Dr. Amanda Grace Vaz, my loving parents Mr. Henry Periyanayagam and Dr. Sarumathy Masilamani and other family members for their ever-ending love and support throughout my stay and studies in Hungary.

I would like to jointly thank the University Grant Commission New Delhi, India, and **Stipendium Hungaricum**, Tempus Public Foundation, Hungary for awarding me the Doctoral research fellowship. I also thank the BRC for awarding me the **Straub Fellowship** for PhD Research and the National Agency of Research, Development and Innovation for granting me the **Young Researcher Excellence Fellowship**.

This study was carried out with the support of **GINOP-2.3.2-15-2016-00001** and **GINOP-2.3.2-15-2016-00039.**

7. REFERENCES

Abdollahi, A., Bao, R., & Hamilton, T. C. (1999). LOT1 is a growth suppressor gene down-regulated by the epidermal growth factor receptor ligands and encodes a nuclear zinc-finger protein. *Oncogene*, *18*(47), 6477–6487. Https://doi.org/10.1038/sj.onc.1203067

Adamson, E. D., & Gardner, R. L. (1979). Control of early development. *British Medical Bulletin*, *35*(2), 113–119. Https://Doi.Org/10.1093/Oxfordjournals.Bmb.A071557

Alam, S., Zinyk, D., Ma, L., & Schuurmans, C. (2005). Members of the *Plag* gene family are expressed in complementary and overlapping regions in the developing murine nervous system. *Developmental Dynamics*, *234*(3), 772–782. Https://doi.org/10.1002/dvdy.20577

Aloia, L., Di Stefano, B., & Di Croce, L. (2013). Polycomb complexes in stem cells and embryonic development. In *Development (Cambridge)* (Vol. 140, Issue 12, pp. 2525–2534). Oxford University Press for The Company of Biologists Limited. Https://doi.org/10.1242/dev.091553

Aranda, S., Mas, G., & Di Croce, L. (2015). Regulation of gene transcription by Polycomb proteins. *Science Advances*, *1*(11). Https://doi.org/10.1126/sciadv.1500737

Arima, T., Drewell, R. A., Arney, K. L., Inoue, J., Makita, Y., Hata, A., Oshimura, M., Wake, N., & Surani, M. A. (2001). A conserved imprinting control region at the HYMAI/ZAC domain is implicated in transient neonatal diabetes mellitus. *Human Molecular Genetics*, *10*(14), 1475–1483. Https://doi.org/10.1093/hmg/10.14.1475

Arnold, S. J., Sugnaseelan, J., Groszer, M., Srinivas, S., & Robertson, E. J. (2009). Generation and analysis of a mouse line harboring GFP in the Eomes/Tbr2 locus. *Genesis*, *47*(11), 775–781. Https://doi.org/10.1002/dvg.20562

Arrigoni, R., Alam, S. L., Wamstad, J. A., Bardwell, V. J., Sundquist, W. I., & Schreiber-Agus, N. (2006). The Polycomb-associated protein Rybp is a ubiquitin binding protein. *FEBS Letters*, *580*(26), 6233–6241. Https://doi.org/10.1016/j.febslet.2006.10.027

Artyomov, M. N., Meissner, A., & Chakraborty, A. K. (2010). A Model for Genetic and Epigenetic Regulatory Networks Identifies Rare Pathways for Transcription Factor Induced Pluripotency. *Plos Computational Biology*, *6*(5), e1000785. Https://doi.org/10.1371/journal.pcbi.1000785

Asumda, F. Z., & Chase, P. B. (2012). Nuclear cardiac troponin and tropomyosin are expressed

early in cardiac differentiation of rat mesenchymal stem cells. *Differentiation; Research in Biological Diversity*, *83*(3), 106–115. Https://doi.org/10.1016/j.diff.2011.10.002

Bajusz, I., Kovács, G., & Pirity, M. (2018). From Flies to Mice: The Emerging Role of Non-Canonical PRC1 Members in Mammalian Development. *Epigenomes*, *2*(1), 4. Https://doi.org/10.3390/epigenomes2010004

Barbour, H., Daou, S., Hendzel, M., & Affar, E. B. (2020). Polycomb group-mediated histone H2A monoubiquitination in epigenome regulation and nuclear processes. In *Nature Communications* (Vol. 11, Issue 1, pp. 1–16). Nature Research. Https://doi.org/10.1038/s41467-020-19722-9

Batalov, I., & Feinberg, A. W. (2015). Differentiation of cardiomyocytes from human pluripotent stem cells using monolayer culture. *Biomarker Insights*, *10*(Suppl 1), 71–76. Https://doi.org/10.4137/BMI.S20050

Bilanges, B., Varrault, A., Basyuk, E., Rodriguez, C., Mazumdar, A., Pantaloni, C., Bockaert, S., Theillet, C., Spengler, D., & Journot, L. (1999). Loss of expression of the candidate tumor suppressor gene ZAC in breast cancer cell lines and primary tumors. *Oncogene*, *18*(27), 3979–3988. Https://doi.org/10.1038/sj.onc.1202933

Blackledge, N. P., Farcas, A. M., Kondo, T., King, H. W., mcgouran, J. F., Hanssen, L. L. P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., Ishikura, T., Long, H. K., Sheahan, T. W., Brockdorff, N., Kessler, B. M., Koseki, H., & Klose, R. J. (2014). Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell*, *157*(6), 1445–1459. Https://doi.org/10.1016/j.cell.2014.05.004

Bondue, A., Lapouge, G., Paulissen, C., Semeraro, C., Iacovino, M., Kyba, M., & Blanpain, C. (2008). Mesp1 acts as a master regulator of multipotent cardiovascular progenitor specification. *Cell Stem Cell*, *3*(1), 69–84. Https://doi.org/10.1016/j.stem.2008.06.009

Bousard, A., Raposo, A. C., Żylicz, J. J., Picard, C., Pires, V. B., Qi, Y., Gil, C., Syx, L., Chang, H. Y., Heard, E., & da Rocha, S. T. (2019). The role of Xist -mediated Polycomb recruitment in the initiation of X-chromosome inactivation . *EMBO Reports*, *20*(10). Https://doi.org/10.15252/embr.201948019

Brade, T., Männer, J., & Kühl, M. (2006). The role of Wnt signalling in cardiac development and tissue remodelling in the mature heart. In *Cardiovascular Research* (Vol. 72, Issue 2, pp. 198–209). Cardiovasc Res. Https://doi.org/10.1016/j.cardiores.2006.06.025

Bradford, M. M. (1976). A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. In *ANALYTICAL BIOCHEMISTRY* (Vol. 72).

Cai, C. L., Liang, X., Shi, Y., Chu, P. H., Pfaff, S. L., Chen, J., & Evans, S. (2003). Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Developmental Cell*, *5*(6), 877–889. Https://doi.org/10.1016/S1534-5807(03)00363-0

Calés, C., Pavón, L., Starowicz, K., Pérez, C., Bravo, M., Ikawa, T., Koseki, H., & Vidal, M. (2016). Role of Polycomb RYBP in Maintaining the B-1-to-B-2 B-Cell Lineage Switch in Adult Hematopoiesis. *Molecular and Cellular Biology*, *36*(6), 900–912.

Cao, Q., Wang, X., Zhao, M., Yang, R., Malik, R., Qiao, Y., Poliakov, A., Yocum, A. K., Li, Y., Chen, W., Cao, X., Jiang, X., Dahiya, A., Harris, C., Feng, F. Y., Kalantry, S., Qin, Z. S., Dhanasekaran, S. M., & Chinnaiyan, A. M. (2014). The central role of EED in the orchestration of polycomb group complexes. *Nature Communications*, *5*(1), 1–13. Https://doi.org/10.1038/ncomms4127

Cao, R., & Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Molecular Cell*, *15*(1), 57–67. Https://doi.org/10.1016/j.molcel.2004.06.020

Carter, M., & Shieh, J. (2015). Biochemical Assays and Intracellular Signaling. In *Guide to Research Techniques in Neuroscience* (pp. 311–343). Elsevier. Https://doi.org/10.1016/b978-0-12-800511-8.00015-0

Chen, D., Zhang, J., Li, M., Rayburn, E. R., Wang, H., & Zhang, R. (2009). RYBP stabilizes p53 by modulating MDM2. *EMBO Reports*, *10*(2), 166–172. Https://doi.org/10.1038/embor.2008.231

Chen, S., Jiao, L., Shubbar, M., Yang, X., & Liu, X. (2018). Unique Structural Platforms of Suz12 Dictate Distinct Classes of PRC2 for Chromatin Binding. *Molecular Cell*, *69*(5), 840-852.e5. Https://doi.org/10.1016/j.molcel.2018.01.039

Chesley, P. (1935). Development of the short-tailed mutant in the house mouse. *Journal of Experimental Zoology*, *70*(3), 429–459. Https://doi.org/10.1002/jez.1400700306

Cvetkovic, D., Pisarcik, D., Lee, C., Hamilton, T. C., & Abdollahi, A. (2004). Altered expression and loss of heterozygosity of the LOT1 gene in ovarian cancer. *Gynecologic*

Oncology, *95*(3), 449–455. Https://doi.org/10.1016/j.ygyno.2004.08.051

Czermin, B., Melfi, R., mccabe, D., Seitz, V., Imhof, A., & Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell*, *111*(2), 185–196. Https://doi.org/10.1016/S0092-8674(02)00975-3

Czubryt, M. P., Lamoureux, L., Ramjiawan, A., Abrenica, B., Jangamreddy, J., & Swan, K. (2010). Regulation of cardiomyocyte Glut4 expression by ZAC1. *Journal of Biological Chemistry*, *285*(22), 16942–16950. Https://doi.org/10.1074/jbc.M109.097246

David, R., Jarsch, V. B., Schwarz, F., Nathan, P., Gegg, M., Lickert, H., & Franz, W. M. (2011). Induction of mesp1 by Brachyury(T) generates the common multipotent cardiovascular stem cell. *Cardiovascular Research*, *92*(1), 115–122. Https://doi.org/10.1093/cvr/cvr158

Deruiter, M. C., Poelmann, R. E., vanderplas-de Vries, I., Mentink, M. M. T., & Gittenberger-de Groot, A. C. (1992). The development of the myocardium and endocardium in mouse embryos - Fusion of two heart tubes? *Anatomy and Embryology*, *185*(5), 461–473. Https://doi.org/10.1007/BF00174084

Dorafshan, E., Kahn, T. G., & Schwartz, Y. B. (2017). Hierarchical recruitment of Polycomb complexes revisited. *Nucleus*, *8*(5), 496–505. Https://doi.org/10.1080/19491034.2017.1363136

Dushnik-Levinson, M., & Benvenisty, N. (1995). Embryogenesis in vitro: Study of differentiation of embryonic stem cells. *Neonatology*, *67*(2), 77–83. Https://doi.org/10.1159/000244147

Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, *292*(5819), 154–156. Https://doi.org/10.1038/292154a0

Fan, F., & Wood, K. V. (2007). Bioluminescent assays for high-throughput screening. In *Assay and Drug Development Technologies* (Vol. 5, Issue 1, pp. 127–136). Assay Drug Dev Technol. Https://doi.org/10.1089/adt.2006.053

Farcas, A. M., Blackledge, N. P., Sudbery, I., Long, H. K., mcgouran, J. F., Rose, N. R., Lee, S., Sims, D., Cerase, A., Sheahan, T. W., Koseki, H., Brockdorff, N., Ponting, C. P., Kessler, B. M., & Klose, R. J. (2012). KDM2B links the polycomb repressive complex 1 (PRC1) to recognition of cpg islands. *Elife*, *2012*(1), 205. Https://doi.org/10.7554/elife.00205

Fijnvandraat, A. C., Van Ginneken, A. C. G., De Boer, P. A. J., Ruijter, J. M., Christoffels, V. M., Moorman, A. F. M., & Lekanne Deprez, R. H. (2003). Cardiomyocytes derived from embryonic stem cells resemble cardiomyocytes of the embryonic heart tube. *Cardiovascular Research*, *58*(2), 399–409. Https://doi.org/10.1016/S0008-6363(03)00282-7

Gao, Z., Lee, P., Stafford, J. M., Schimmelmann, M. Von, Schaefer, A., & Reinberg, D. (2014). An AUTS2-Polycomb complex activates gene expression in the CNS. *Nature*, *516*(7531), 349–354. Https://doi.org/10.1038/nature13921

Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y., & Reinberg, D. (2012). PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. *Molecular Cell*, *45*(3), 344–356. Https://doi.org/10.1016/j.molcel.2012.01.002

Garcia, E., Marcos-Gutiérrez, C., del Mar Lorente, M., Moreno, J. C., & Vidal, M. (1999). RYBP, a new repressor protein that interacts with components of the mammalian Polycomb complex, and with the transcription factor YY1. *The EMBO Journal*, *18*(12), 3404–3418. Https://doi.org/10.1093/emboj/18.12.3404

Gardner, R. L., & Brook, F. A. (1997). Reflections on the biology of embryonic stem (ES) cells. *International Journal of Developmental Biology*, *41*(2), 235–243. Https://doi.org/10.1387/ijdb.9184330

Geng, Z., & Gao, Z. (2020). Mammalian prc1 complexes: Compositional complexity and diverse molecular mechanisms. In *International Journal of Molecular Sciences* (Vol. 21, Issue 22, pp. 1–18). MDPI AG. Https://doi.org/10.3390/ijms21228594

Gessert, S., & Kühl, M. (2010). The multiple phases and faces of Wnt signaling during cardiac differentiation and development. In *Circulation Research* (Vol. 107, Issue 2, pp. 186–199). Circ Res. Https://doi.org/10.1161/CIRCRESAHA.110.221531

Gould, A. (1997). Functions of mammalian Polycomb group and trithorax group related genes. *Current Opinion in Genetics and Development*, *7*(4), 488–494. Https://doi.org/10.1016/S0959-437X(97)80075-5

Gracheva, E., Chitale, S., Wilhelm, T., Rapp, A., Byrne, J., Stadler, J., Medina, R., Cardoso, M. C., & Richly, H. (2016). ZRF1 mediates remodeling of E3 ligases at DNA lesion sites during nucleotide excision repair. *Journal of Cell Biology*, *213*(2), 185–200.

Https://doi.org/10.1083/jcb.201506099

Guo, X., Xu, Y., Wang, Z., Wu, Y., Chen, J., Wang, G., Lu, C., Jia, W., Xi, J., Zhu, S., Jiapaer, Z., Wan, X., Liu, Z., Gao, S., & Kang, J. (2018). A Linc1405/Eomes Complex Promotes Cardiac Mesoderm Specification and Cardiogenesis. *Cell Stem Cell*, *22*(6), 893-908.e6. Https://doi.org/10.1016/j.stem.2018.04.013

He, J., Shen, L., Wan, M., Taranova, O., Wu, H., & Zhang, Y. (2013). Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to cpg islands of developmental genes. *Nature Cell Biology*, *15*(4), 373–384. Https://doi.org/10.1038/ncb2702

Henry, S., Szabó, V., Sutus, E., & Pirity, M. K. (2020). RYBP is important for cardiac progenitor cell development and sarcomere formation. *Plos ONE*, *15*(7 July). Https://doi.org/10.1371/journal.pone.0235922

Hescheler, J., Fleischmann, B. K., Lentini, S., Maltsev, V. A., Rohwedel, J., Wobus, A. M., & Addicks, K. (1997). Embryonic stem cells: A model to study structural and functional properties in cardiomyogenesis. In *Cardiovascular Research* (Vol. 36, Issue 2, pp. 149–162). Oxford Academic. Https://doi.org/10.1016/S0008-6363(97)00193-4

Hoffmann, A. (2015). Role of ZAC1 in transient neonatal diabetes mellitus and glucose metabolism. *World Journal of Biological Chemistry*, *6*(3), 95. Https://doi.org/10.4331/wjbc.v6.i3.95

Iglesias-Platas, I., Court, F., Camprubi, C., Sparago, A., Guillaumet-Adkins, A., Martin-Trujillo, A., Riccio, A., Moore, G. E., & Monk, D. (2013). Imprinting at the PLAGL1 domain is contained within a 70-kb CTCF/cohesin-mediated non-Allelic chromatin loop. *Nucleic Acids Research*, *41*(4), 2171–2179. Https://doi.org/10.1093/nar/gks1355

Iglesias-Platas, I., Martin-Trujillo, A., Cirillo, D., Court, F., Guillaumet-Adkins, A., Camprubi, C., Bourc'his, D., Hata, K., Feil, R., Tartaglia, G., Arnaud, P., & Monk, D. (2012). Characterization of Novel Paternal ncrnas at the Plagl1 Locus, Including Hymai, Predicted to Interact with Regulators of Active Chromatin. *Plos ONE*, *7*(6), e38907. Https://doi.org/10.1371/journal.pone.0038907

Ismail, I. H., mcdonald, D., Strickfaden, H., Xu, Z., & Hendzel, M. J. (2013). A small molecule inhibitor of polycomb repressive complex 1 inhibits ubiquitin signaling at DNA double-strand breaks. *Journal of Biological Chemistry*, *288*(37), 26944–26954. Https://doi.org/10.1074/jbc.M113.461699

Ivanyuk, D., Budash, G., Zheng, Y., Gaspar, J. A., Chaudhari, U., Fatima, A., Bahmanpour, S., Grin, V. K., Popandopulo, A. G., Sachinidis, A., Hescheler, J., & Šarić, T. (2015). Ascorbic Acid-Induced Cardiac Differentiation of Murine Pluripotent Stem Cells: Transcriptional Profiling and Effect of a Small Molecule Synergist of Wnt/β-Catenin Signaling Pathway. *Cellular Physiology and Biochemistry*, *36*(2), 810–830. Https://doi.org/10.1159/000430140

Jacobs, D. I., Mao, Y., Fu, A., Kelly, W. K., & Zhu, Y. (2013). Dysregulated methylation at imprinted genes in prostate tumor tissue detected by methylation microarray. *BMC Urology*, *13*, 37. Https://doi.org/10.1186/1471-2490-13-37

Kamps, J. A. (2016). Micromanaging cardiac regeneration: Targeted delivery of micrornas for cardiac repair and regeneration. *World Journal of Cardiology*, *8*(2), 163. Https://doi.org/10.4330/wjc.v8.i2.163

Kaplun, A., Krull, M., Lakshman, K., Matys, V., Lewicki, B., & Hogan, J. D. (2016). Establishing and validating regulatory regions for variant annotation and expression analysis. *BMC Genomics*, *17*(2), 219–227. Https://doi.org/10.1186/s12864-016-2724-0

Kawai, T., Takahashi, T., Esaki, M., Ushikoshi, H., Nagano, S., Fujiwara, H., & Kosai, K. (2004). Efficient Cardiomyogenic Differentiation of Embryonic Stem Cell by Fibroblast Growth Factor 2 and Bone Morphogenetic Protein 2. *Circulation Journal*, *68*(7), 691–702. Https://doi.org/10.1253/circj.68.691

Keller, G.-A., Gould, S., Delucat, M., & Subramani, S. (1987). Firefly luciferase is targeted to peroxisomes in mammalian cells (catalase/peroxisomal targeting/protein sorting). In *Cell Biology* (Vol. 84).

Keller, G. M. (1995). In vitro differentiation of embryonic stem cells. *Current Opinion in Cell Biology*, *7*(6), 862–869. Https://doi.org/10.1016/0955-0674(95)80071-9

Ketel, C. S., Andersen, E. F., Vargas, M. L., Suh, J., Strome, S., & Simon, J. A. (2005). Subunit Contributions to Histone Methyltransferase Activities of Fly and Worm Polycomb Group Complexes. *Molecular and Cellular Biology*, *25*(16), 6857–6868. Https://doi.org/10.1128/mcb.25.16.6857-6868.2005

Kingston, R. E., Chen, C. A., & Rose, J. K. (2003). Calcium Phosphate Transfection. *Current Protocols in Molecular Biology*, *63*(1). Https://doi.org/10.1002/0471142727.mb0901s63

Klattenhoff, C. A., Scheuermann, J. C., Surface, L. E., Bradley, R. K., Fields, P. A.,

Steinhauser, M. L., Ding, H., Butty, V. L., Torrey, L., Haas, S., Abo, R., Tabebordbar, M., Lee, R. T., Burge, C. B., & Boyer, L. A. (2013). Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell*, *152*(3), 570–583. Https://doi.org/10.1016/j.cell.2013.01.003

Kokkinopoulos, I., Ishida, H., Saba, R., Coppen, S., Suzuki, K., & Yashiro, K. (2016). Cardiomyocyte differentiation from mouse embryonic stem cells using a simple and defined protocol. *Developmental Dynamics*, *245*(2), 157–165. Https://doi.org/10.1002/dvdy.24366

Kovacs, G., Szabo, V., & Pirity, M. K. (2016). Absence of Rybp Compromises Neural Differentiation of Embryonic Stem Cells. *Stem Cells International*, *2016*, 1–12. Https://doi.org/10.1155/2016/4034620

Kowalczyk, A. E., Krazinski, B. E., Godlewski, J., Kiewisz, J., Kwiatkowski, P., Sliwinska-Jewsiewicka, A., Kiezun, J., Wierzbicki, P. M., Bodek, G., Sulik, M., & Kmiec, Z. (2015). Altered expression of the PLAGL1 (ZAC1/LOT1) gene in colorectal cancer: Correlations to the clinicopathological parameters. *International Journal of Oncology*, *47*(3), 951–962. Https://doi.org/10.3892/ijo.2015.3067

Kuroda, M. I., Kang, H., De, S., & Kassis, J. A. (2020). Dynamic Competition of Polycomb and Trithorax in Transcriptional Programming. *Annual Review of Biochemistry*, *89*(1), 235–253. Https://doi.org/10.1146/annurev-biochem-120219-103641

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., & Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the enhancer of zeste protein. *Genes and Development*, *16*(22), 2893–2905. Https://doi.org/10.1101/gad.1035902

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, *227*(5259), 680–685. Https://doi.org/10.1038/227680a0

Laugesen, A., Højfeldt, J. W., & Helin, K. (2016). Role of the polycomb repressive complex 2 (PRC2) in transcriptional regulation and cancer. *Cold Spring Harbor Perspectives in Medicine*, *6*(9). Https://doi.org/10.1101/cshperspect.a026575

Lee, J., Sutani, A., Kaneko, R., Takeuchi, J., Sasano, T., Kohda, T., Ihara, K., Takahashi, K., Yamazoe, M., Morio, T., Furukawa, T., & Ishino, F. (2020). In vitro generation of functional murine heart organoids via FGF4 and extracellular matrix. *Nature*

Communications, *11*(1), 1–18. Https://doi.org/10.1038/s41467-020-18031-5

Lescroart, F., Wang, X., Lin, X., Swedlund, B., Gargouri, S., Sànchez-Dànes, A., Moignard, V., Dubois, C., Paulissen, C., Kinston, S., Göttgens, B., & Blanpain, C. (2018). Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. *Science*, *359*(6380), 1177–1181. Https://doi.org/10.1126/science.aao4174

Li, H., Lai, P., Jia, J., Song, Y., Xia, Q., Huang, K., He, N., Ping, W., Chen, J., Yang, Z., Li, J., Yao, M., Dong, X., Zhao, J., Hou, C., Esteban, M. A., Gao, S., Pei, D., Hutchins, A. P., & Yao, H. (2017). RNA Helicase DDX5 Inhibits Reprogramming to Pluripotency by mirna-Based Repression of RYBP and its PRC1-Dependent and -Independent Functions. *Cell Stem Cell*, *20*(4), 462-477.e6. Https://doi.org/10.1016/j.stem.2016.12.002

Li, Z., Ding, Y., Zhu, Y., Yin, M., Le, X., Wang, L., Yang, Y., & Zhang, Q. (2014). Both gene deletion and promoter hyper-methylation contribute to the down-regulation of ZAC/PLAGL1 gene in gastric adenocarcinomas: A case control study. *Clinics and Research in Hepatology and Gastroenterology*, *38*(6), 744–750. Https://doi.org/10.1016/j.clinre.2013.06.007

Lin, Q., Schwarz, J., Bucana, C., & Olson, E. N. (1997). Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science*, *276*(5317), 1404–1407. Https://doi.org/10.1126/science.276.5317.1404

Liu, P. Y., Chan, J. Y. H., Lin, H. C., Wang, S. L., Liu, S. T., Ho, C. L., Chang, L. C., & Huang, S. M. (2008). Modulation of the cyclin-dependent kinase inhibitor p21waf1/Cip1 gene by Zac1 through the antagonistic regulators p53 and histone deacetylase 1 in hela cells. *Molecular Cancer Research*, *6*(7), 1204–1214. Https://doi.org/10.1158/1541-7786.MCR-08-0123

Madani, M. M., & Golts, E. (2014). Cardiovascular Anatomy. In *Reference Module in Biomedical Sciences*. Elsevier. Https://doi.org/10.1016/b978-0-12-801238-3.00196-3

Magin, T. M., Mcwhir, J., & Melton, D. W. (1992). A new mouse embryonic stem cell line with good germ line contribution and gene targeting frequency. *Nucleic Acids Research*, *20*(14), 3795–3796. Https://doi.org/10.1093/nar/20.14.3795

Marikawa, Y., & Alarcón, V. B. (2009). Establishment of trophectoderm and inner cell mass lineages in the mouse embryo. In *Molecular Reproduction and Development* (Vol. 76, Issue 11, pp. 1019–1032). Mol Reprod Dev. Https://doi.org/10.1002/mrd.21057

Marques, S. M., & Esteves Da Silva, J. C. G. (2009). Firefly bioluminescence: A mechanistic approach of luciferase catalyzed reactions. In *IUBMB Life* (Vol. 61, Issue 1, pp. 6–17). IUBMB Life. Https://doi.org/10.1002/iub.134

Mcculley, D. J., & Black, B. L. (2012). Transcription Factor Pathways and Congenital Heart Disease. In *Current Topics in Developmental Biology* (Vol. 100, pp. 253–277). Academic Press Inc. Https://doi.org/10.1016/B978-0-12-387786-4.00008-7

Mckusick, V. A. (1964). A Genetical View of Cardiovascular Disease. *Circulation*, *30*(3), 326–357. Https://doi.org/10.1161/01.CIR.30.3.326

Miró, X., Zhou, X., Boretius, S., Michaelis, T., Kubisch, C., Alvarez-Bolado, G., & Gruss, P. (2009). Haploinsufficiency of the murine polycomb gene Suz12 results in diverse malformations of the brain and neural tube. *DMM Disease Models and Mechanisms*, *2*(7–8), 412–418. Https://doi.org/10.1242/dmm.001602

Molkentin, J. D., Lin, Q., Duncan, S. A., & Olson, E. N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes and Development*, *11*(8). Https://doi.org/10.1101/gad.11.8.1061

Moretti, A., Caron, L., Nakano, A., Lam, J. T., Bernshausen, A., Chen, Y., Qyang, Y., Bu, L., Sasaki, M., Martin-Puig, S., Sun, Y., Evans, S. M., Laugwitz, K. L., & Chien, K. R. (2006). Multipotent Embryonic Isl1+ Progenitor Cells Lead to Cardiac, Smooth Muscle, and Endothelial Cell Diversification. *Cell*, *127*(6), 1151–1165. Https://doi.org/10.1016/j.cell.2006.10.029

Moretti, A., Laugwitz, K. L., Dorn, T., Sinnecker, D., & Mummery, C. (2013). Pluripotent stem cell models of human heart disease. *Cold Spring Harbor Perspectives in Medicine*, *3*(11), a014027. Https://doi.org/10.1101/cshperspect.a014027

Morey, L., Aloia, L., Cozzuto, L., Benitah, S. A., & Croce, L. Di. (2013). RYBP and Cbx7 Define Specific Biological Functions of Polycomb Complexes in Mouse Embryonic Stem Cells. *Cell Reports*, *3*(1), 60–69. Https://doi.org/10.1016/j.celrep.2012.11.026

Morey, L., Santanach, A., Blanco, E., Aloia, L., Nora, E. P., Bruneau, B. G., & Di croce, L. (2015). Polycomb Regulates Mesoderm Cell Fate-Specification in Embryonic Stem Cells through Activation and Repression Mechanisms. *Cell Stem Cell*, *17*(3), 300–315. Https://doi.org/10.1016/j.stem.2015.08.009

Mori, A. D., Zhu, Y., Vahora, I., Nieman, B., Koshiba-Takeuchi, K., Davidson, L., Pizard, A.,

Seidman, J. G., Seidman, C. E., Chen, X. J., Henkelman, R. M., & Bruneau, B. G. (2006). Tbx5-dependent rheostatic control of cardiac gene expression and morphogenesis. *Developmental Biology*, *297*(2), 566–586. Https://doi.org/10.1016/j.ydbio.2006.05.023

Müller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E., & Simon, J. A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell*, *111*(2), 197–208. Https://doi.org/10.1016/S0092-8674(02)00976-5

Mummery, C. L., Zhang, J., Ng, E. S., Elliott, D. A., Elefanty, A. G., & Kamp, T. J. (2012). Differentiation of human embryonic stem cells and induced pluripotent stem cells to cardiomyocytes: A methods overview. In *Circulation Research* (Vol. 111, Issue 3, pp. 344–358). NIH Public Access. Https://doi.org/10.1161/CIRCRESAHA.110.227512

Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W., & Roder, J. C. (1993). Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(18), 8424–8428. Https://doi.org/10.1073/pnas.90.18.8424

Neira, J. L., Román-Trufero, M., Contreras, L. M., Prieto, J., Singh, G., Barrera, F. N., Renart, M. L., & Vidal, M. (2009). The transcriptional repressor RYBP is a natively unfolded protein which folds upon binding to DNA. *Biochemistry*, *48*(6), 1348–1360. Https://doi.org/10.1021/bi801933c

Nowotschin, S., Costello, I., Piliszek, A., Kwon, G. S., Mao, C. An, Klein, W. H., Robertson, E. J., & Hadjantonakis, A. K. (2013). The T-box transcription factor eomesodermin is essential for AVE induction in the mouse embryo. *Genes and Development*, *27*(9), 997–1002. Https://doi.org/10.1101/gad.215152.113

Obier, N., Lin, Q., Cauchy, P., Hornich, V., Zenke, M., Becker, M., & Müller, A. M. (2015). Polycomb Protein EED is Required for Silencing of Pluripotency Genes upon ESC Differentiation. *Stem Cell Reviews and Reports*, *11*(1), 50–61. Https://doi.org/10.1007/s12015-014-9550-z

Perino, M. G., Yamanaka, S., Riordon, D. R., Tarasova, Y., & Boheler, K. R. (2017). Ascorbic acid promotes cardiomyogenesis through SMAD1 signaling in differentiating mouse embryonic stem cells. *PLOS ONE*, *12*(12), e0188569. Https://doi.org/10.1371/journal.pone.0188569

Pirity, Melinda K, Locker, J., & Schreiber-Agus, N. (2005). Rybp/DEDAF Is Required for Early Postimplantation and for Central Nervous System Development. *Molecular and Cellular Biology*, *25*(16), 7193–7202. Https://doi.org/10.1128/MCB.25.16.7193-7202.2005

Pirity, Melinda K., Wang, W. L., Wolf, L. V., Tamm, E. R., Schreiber-Agus, N., & Cvekl, A. (2007). Rybp, a polycomb complex-associated protein, is required for mouse eye development. *BMC Developmental Biology*, *7*. Https://doi.org/10.1186/1471-213X-7-39

Pucéat, M. (2008). Protocols for cardiac differentiation of embryonic stem cells. *Methods*, *45*(2), 168–171. Https://doi.org/10.1016/j.ymeth.2008.03.003

Ribarska, T., Goering, W., Droop, J., Bastian, K. M., Ingenwerth, M., & Schulz, W. A. (2014). Deregulation of an imprinted gene network in prostate cancer. *Epigenetics*, *9*(5), 704–717. Https://doi.org/10.4161/epi.28006

Rose, N. R., King, H. W., Blackledge, N. P., Fursova, N. A., Ember, K. J. I., Fischer, R., Kessler, B. M., & Klose, R. J. (2016). RYBP stimulates PRC1 to shape chromatin-based communication between Polycomb repressive complexes. *Elife*, *5*. Https://doi.org/10.7554/elife.18591

Rungarunlert, S., Klincumhom, N., Tharasanit, T., Techakumphu, M., Pirity, M. K., & Dinnyes, A. (2013). Slow turning lateral vessel bioreactor improves embryoid body formation and cardiogenic differentiation of mouse embryonic stem cells. *Cellular Reprogramming*, *15*(5), 443–458. Https://doi.org/10.1089/cell.2012.0082

Russ, A. P., Wattler, S., Colledge, W. H., Aparicio, S. A. J. R., Carlton, M. B. L., Pearce, J. J., Barton, S. C., Azim Surani, M., Ryan, K., Nehls, M. C., Wilsons, V., & Evans, M. J. (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. *Nature*, *404*(6773), 95–99. Https://doi.org/10.1038/35003601

Saga, Y., Kitajima, S., & Miyagawa-Tomita, S. (2000). Mesp1 expression is the earliest sign of cardiovascular development. In *Trends in Cardiovascular Medicine* (Vol. 10, Issue 8, pp. 345–352). Trends Cardiovasc Med. Https://doi.org/10.1016/S1050-1738(01)00069-X

Saga, Y., Miyagawa-Tomita, S., Takagi, A., Kitajima, S., Miyazaki, J. I., & Inoue, T. (1999). Mesp1 is expressed in the heart precursor cells and required for the formation of a single heart tube. *Development*, *126*(15), 3437–3447.

Sahakyan, A., Yang, Y., & Plath, K. (2018). The Role of Xist in X-Chromosome Dosage

Compensation. In *Trends in Cell Biology* (Vol. 28, Issue 12, pp. 999–1013). Elsevier Ltd. Https://doi.org/10.1016/j.tcb.2018.05.005

Sandelin, A., Alkema, W., Engström, P., Wasserman, W. W., & Lenhard, B. (2004). JASPAR: An open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Research*, *32*(DATABASE ISS.). Https://doi.org/10.1093/nar/gkh012

Sasai, Y., Eiraku, M., & Suga, H. (2012). In vitro organogenesis in three dimensions: Self-organising stem cells. In *Development (Cambridge)* (Vol. 139, Issue 22, pp. 4111–4121). Development. Https://doi.org/10.1242/dev.079590

Savolainen, S. M., Foley, J. F., & Elmore, S. A. (2009). Histology atlas of the developing mouse heart with emphasis on E11.5 to E18.5. In *Toxicologic Pathology* (Vol. 37, Issue 4, pp. 395–414). Toxicol Pathol. Https://doi.org/10.1177/0192623309335060

Schlisio, S., Halperin, T., Vidal, M., & Nevins, J. R. (2002). Interaction of YY1 with E2Fs, mediated by RYBP, provides a mechanism for specificity of E2F function. *EMBO Journal*, *21*(21), 5775–5786. Https://doi.org/10.1093/emboj/cdf577

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., & Cavalli, G. (2007). Genome Regulation by Polycomb and Trithorax Proteins. In *Cell* (Vol. 128, Issue 4, pp. 735–745). Elsevier. Https://doi.org/10.1016/j.cell.2007.02.009

Shen, X., Liu, Y., Hsu, Y. J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G. C., & Orkin, S. H. (2008). EZH1 Mediates Methylation on Histone H3 Lysine 27 and Complements EZH2 in Maintaining Stem Cell Identity and Executing Pluripotency. *Molecular Cell*, *32*(4), 491–502. Https://doi.org/10.1016/j.molcel.2008.10.016

Simon, J. A., & Kingston, R. E. (2013). Occupying Chromatin: Polycomb Mechanisms for Getting to Genomic Targets, Stopping Transcriptional Traffic, and Staying Put. *Molecular Cell*, *49*(5), 808–824. Https://doi.org/10.1016/j.molcel.2013.02.013

Smith, R. J., Arnaud, P., Konfortova, G., Dean, W. L., Beechey, C. V., & Kelsey, G. (2002). The mouse Zac1 locus: Basis for imprinting and comparison with human ZAC. *Gene*, *292*(1–2), 101–112. Https://doi.org/10.1016/S0378-1119(02)00666-2

Soibam, B., Benham, A., Kim, J., Weng, K. C., Yang, L., Xu, X., Robertson, M., Azares, A., Cooney, A. J., Schwartz, R. J., & Liu, Y. (2015). Genome-wide identification of MESP1 targets demonstrates primary regulation over mesendoderm gene activity. *Stem Cells*, *33*(11), 3254–3265. Https://doi.org/10.1002/stem.2111

Spengler, D., Villalba, M., Hoffmann, A., Pantaloni, C., Houssami, S., Bockaert, J., & Journot, L. (1997). Regulation of apoptosis and cell cycle arrest by Zac1, a novel zinc finger protein expressed in the pituitary gland and the brain. *EMBO Journal*, *16*(10), 2814–2825. Https://doi.org/10.1093/emboj/16.10.2814

Stanton, S. E., Blanck, J. K., Locker, J., & Schreiber-Agus, N. (2007). Rybp interacts with Hippi and enhances Hippi-mediated apoptosis. *Apoptosis*, *12*(12), 2197–2206. Https://doi.org/10.1007/s10495-007-0131-3

Sunwoo, H., Wu, J. Y., & Lee, J. T. (2015). The Xist RNA-PRC2 complex at 20-nm resolution reveals a low Xist stoichiometry and suggests a hit-and-run mechanism in mouse cells. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(31), E4216–E4225. Https://doi.org/10.1073/pnas.1503690112

Sweeney, H. L., & Hammers, D. W. (2018). Muscle contraction. *Cold Spring Harbor Perspectives in Biology*, *10*(2). Https://doi.org/10.1101/cshperspect.a023200

Takahashi, T., Lord, B., Schulze, P. C., Fryer, R. M., Sarang, S. S., Gullans, S. R., & Lee, R. T. (2003). Ascorbic acid enhances differentiation of embryonic stem cells into cardiac myocytes. *Circulation*, *107*(14), 1912–1916. Https://doi.org/10.1161/01.CIR.0000064899.53876.A3

Tan, K., Zhang, X., Cong, X., Huang, B., Chen, H., & Chen, D. (2017). Tumor suppressor RYBP harbors three nuclear localization signals and its cytoplasm-located mutant exerts more potent anti-cancer activities than corresponding wild type. *Cellular Signalling*, *29*, 127–137. Https://doi.org/10.1016/j.cellsig.2016.10.011

Tian, Q., Guo, S. M., Xie, S. M., Yin, Y., & Zhou, L. Q. (2020). Rybp orchestrates spermatogenesis via regulating meiosis and sperm motility in mice. *Cell Cycle*, *19*(12), 1492–1501. Https://doi.org/10.1080/15384101.2020.1754585

Tosic, J., Kim, G. J., Pavlovic, M., Schröder, C. M., Mersiowsky, S. L., Barg, M., Hofherr, A., Probst, S., Köttgen, M., Hein, L., & Arnold, S. J. (2019). Eomes and Brachyury control pluripotency exit and germ-layer segregation by changing the chromatin state. *Nature Cell Biology*, *21*(12), 1518–1531. Https://doi.org/10.1038/s41556-019-0423-1

Tsuda, T., Markova, D., Wang, H., Evangelisti, L., Pan, T. C., & Chu, M. L. (2004). Zinc Finger Protein Zac1 Is Expressed in Chondrogenic Sites of the Mouse. *Developmental Dynamics*, *229*(2), 340–348. Https://doi.org/10.1002/dvdy.10439

Tyser, R. C. V., Miranda, A. M. A., Chen, C. M., Davidson, S. M., Srinivas, S., & Riley, P. R. (2016). Calcium handling precedes cardiac differentiation to initiate the first heartbeat. *Elife*, *5*(OCTOBER2016). Https://doi.org/10.7554/elife.17113

Ujhelly, O., Szabo, V., Kovacs, G., Vajda, F., Mallok, S., Prorok, J., Acsai, K., Hegedus, Z., Krebs, S., Dinnyes, A., & Pirity, M. K. (2015). Lack of Rybp in Mouse Embryonic Stem Cells Impairs Cardiac Differentiation. *Stem Cells and Development*, *24*(18), 2193–2205. Https://doi.org/10.1089/scd.2014.0569

Valente, Tony, & Auladell, C. (2001). Expression pattern of Zac1 mouse gene, a new zinc-finger protein that regulates apoptosis and cellular cycle arrest, in both adult brain and along development. *Mechanisms of Development*, *108*(1–2), 207–211. Https://doi.org/10.1016/S0925-4773(01)00492-0

Valente, Tony, Junyent, F., & Auladell, C. (2005). Zac1 is expressed in progenitor/stem cells of the neuroectoderm and mesoderm during embryogenesis: Differential phenotype of the Zac1-expressing cells during development. *Developmental Dynamics*, *233*(2), 667–679. Https://doi.org/10.1002/dvdy.20373

Valleley, E. M. A., Cordery, S. F., Carr, L. M., maclennan, K. A., & Bonthron', D. T. (2010). Loss of expression of ZACIPLAGLI in diffuse large B-cell lymphoma is independent of promoter hypermethylation. *Genes Chromosomes and Cancer*, *49*(5), 480–486. Https://doi.org/10.1002/gcc.20758

Valleley, E. M., Cordery, S. F., & Bonthron, D. T. (2007). Tissue-specific imprinting of the ZAC/PLAGL1 tumour suppressor gene results from variable utilization of monoallelic and biallelic promoters. *Human Molecular Genetics*, *16*(8), 972–981. Https://doi.org/10.1093/hmg/ddm041

Van Den Ameele, J., Tiberi, L., Bondue, A., Paulissen, C., Herpoel, A., Iacovino, M., Kyba, M., Blanpain, C., & Vanderhaeghen, P. (2012). Eomesodermin induces Mesp1 expression and cardiac differentiation from embryonic stem cells in the absence of Activin. *EMBO Reports*, *13*(4), 355–362. Https://doi.org/10.1038/embor.2012.23

Van Weerd, J. H., Koshiba-Takeuchi, K., Kwon, C., & Takeuchi, J. K. (2011). Epigenetic factors and cardiac development. In *Cardiovascular Research* (Vol. 91, Issue 2, pp. 203–211). Https://doi.org/10.1093/cvr/cvr138

Varrault, A., Bilanges, B., Mackay, D. J. G., Basyuk, E., Ahr, B., Fernandez, C., Robinson, D.

O., Bockaert, J., & Journot, L. (2001). Characterization of the Methylation-sensitive Promoter of the Imprinted ZAC Gene Supports Its Role in Transient Neonatal Diabetes Mellitus. *Journal of Biological Chemistry*, *276*(22), 18653–18656. Https://doi.org/10.1074/jbc.C100095200

Varrault, A., Dantec, C., Le Digarcher, A., Chotard, L., Bilanges, B., Parrinello, H., Dubois, E., Rialle, S., Severac, D., Bouschet, T., & Journot, L. (2017). Identification of Plagl1/Zac1 binding sites and target genes establishes its role in the regulation of extracellular matrix genes and the imprinted gene network. *Nucleic Acids Research*, *45*(18), 10466–10480. Https://doi.org/10.1093/nar/gkx672

Vega-Benedetti, A. F., Saucedo, C., Zavattari, P., Vanni, R., Zugaza, J. L., & Parada, L. A. (2017). PLAGL1: an important player in diverse pathological processes. In *Journal of Applied Genetics* (Vol. 58, Issue 1, pp. 71–78). Springer Verlag. Https://doi.org/10.1007/s13353-016-0355-4

Vidal, M. (2009). Role of polycomb proteins Ring1A and Ring1B in the epigenetic regulation of gene expression. In *International Journal of Developmental Biology* (Vol. 53, Issues 2–3, pp. 355–370). Int J Dev Biol. Https://doi.org/10.1387/ijdb.082690mv

Voruganti, S., Xu, F., Qin, J.-J., Guo, Y., Sarkar, S., Gao, M., Zheng, Z., Wang, M.-H., Zhou, J., Qian, B., Zhang, R., & Wang, W. (2015). RYBP predicts survival of patients with non-small cell lung cancer and regulates tumor cell growth and the response to chemotherapy. *Cancer Letters*, *369*(2), 386–395. Https://doi.org/10.1016/j.canlet.2015.09.003

Wang, C., Xu, H., Lin, S., Deng, W., Zhou, J., Zhang, Y., Shi, Y., Peng, D., & Xue, Y. (2020). GPS 5.0: An Update on the Prediction of Kinase-specific Phosphorylation Sites in Proteins. *Genomics, Proteomics and Bioinformatics*, *18*(1), 72–80. Https://doi.org/10.1016/j.gpb.2020.01.001

Wang, K. C., & Chang, H. Y. (2011). Molecular Mechanisms of Long Noncoding rnas. In *Molecular Cell* (Vol. 43, Issue 6, pp. 904–914). Mol Cell. Https://doi.org/10.1016/j.molcel.2011.08.018

Wang, X., & Yang, P. (2008). In vitro differentiation of mouse embryonic stem (mes) cells using the hanging drop method. *Journal of Visualized Experiments*, *17*. Https://doi.org/10.3791/825

Watabe, T., & Miyazono, K. (2009). Roles of TGF-β family signaling in stem cell renewal and

differentiation. In *Cell Research* (Vol. 19, Issue 1, pp. 103–115). Cell Res. Https://doi.org/10.1038/cr.2008.323

Wilkinson, D. G., Bhatt, S., & Herrmann, B. G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature*, *343*(6259), 657–659. Https://doi.org/10.1038/343657a0

Wingender, E., Dietze, P., Karas, H., & Knüppel, R. (1996). TRANSFAC: A database on transcription factors and their DNA binding sites. In *Nucleic Acids Research* (Vol. 24, Issue 1, pp. 238–241). Oxford University Press. Https://doi.org/10.1093/nar/24.1.238

Wingender, Edgar. (2008). The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. *Briefings in Bioinformatics*, *9*(4), 326–332. Https://doi.org/10.1093/bib/bbn016

Wu, H., Zhao, Z. A., Liu, J., Hao, K., Yu, Y., Han, X., Li, J., Wang, Y., Lei, W., Dong, N., Shen, Z., & Hu, S. (2018). Long noncoding RNA Meg3 regulates cardiomyocyte apoptosis in myocardial infarction. *Gene Therapy*, *25*(8), 511–523. Https://doi.org/10.1038/s41434-018-0045-4

Yanagisawa, K. O., Fujimoto, H., & Urushihara, H. (1981). Effects of the Brachyury (T) mutation on morphogenetic movement in the mouse embryo. *Developmental Biology*, *87*(2), 242–248. Https://doi.org/10.1016/0012-1606(81)90147-0

Yanni, J., Tellez, J. O., MącZewski, M., mackiewicz, U., Beresewicz, A., Billeter, R., Dobrzynski, H., & Boyett, M. R. (2011). Changes in Ion channel gene expression underlying heart failure-induced sinoatrial node dysfunction. *Circulation: Heart Failure*, *4*(4), 496–508. Https://doi.org/10.1161/CIRCHEARTFAILURE.110.957647

Yuasa, S., Onizuka, T., Shimoji, K., Ohno, Y., Kageyama, T., Yoon, S. H., Egashira, T., Seki, T., Hashimoto, H., Nishiyama, T., Kaneda, R., Murata, M., Hattori, F., Makino, S., Sano, M., Ogawa, S., Prall, O. W. J., Harvey, R. P., & Fukuda, K. (2010). Zac1 is an essential transcription factor for cardiac morphogenesis. *Circulation Research*, *106*(6), 1083–1091. Https://doi.org/10.1161/CIRCRESAHA.109.214130

Yutzey, K. E., Gannon, M., & Bader, D. (1995). Diversification of Cardiomyogenic Cell Lineages in Vitro. *Developmental Biology*, *170*(2), 531–541. Https://doi.org/10.1006/dbio.1995.1234

Zakrzewski, W., Dobrzyński, M., Szymonowicz, M., & Rybak, Z. (2019). Stem cells: Past,

present, and future. In *Stem Cell Research and Therapy* (Vol. 10, Issue 1). Biomed Central Ltd. Https://doi.org/10.1186/s13287-019-1165-5

Zepeda-Martinez, J. A., Pribitzer, C., Wang, J., Bsteh, D., Golumbeanu, S., Zhao, Q., Burkard, T. R., Reichholf, B., Rhie, S. K., Jude, J., Moussa, H. F., Zuber, J., & Bell, O. (2020). Parallel PRC2/cprc1 and vprc1 pathways silence lineage-specific genes and maintain self-renewal in mouse embryonic stem cells. *Science Advances*, *6*(14), 5692–5693. Https://doi.org/10.1126/sciadv.aax5692

Zerbino, D. R., Wilder, S. P., Johnson, N., Juettemann, T., & Flicek, P. R. (2015). The Ensembl Regulatory Build. *Genome Biology*, *16*(1), 56. Https://doi.org/10.1186/s13059-015-0621-5

Zhang, J., Klos, M., Wilson, G. F., Herman, A. M., Lian, X., Raval, K. K., Barron, M. R., Hou, L., Soerens, A. G., Yu, J., Palecek, S. P., Lyons, G. E., Thomson, J. A., Herron, T. J., Jalife, J., & Kamp, T. J. (2012). Extracellular matrix promotes highly efficient cardiac differentiation of human pluripotent stem cells: The matrix sandwich method. *Circulation Research*, *111*(9), 1125–1136. Https://doi.org/10.1161/CIRCRESAHA.112.273144

Zhao, J., Wang, M., Chang, L., Yu, J., Song, A., Liu, C., Huang, W., Zhang, T., Wu, X., Shen, X., Zhu, B., & Li, G. (2020). RYBP/YAF2-PRC1 complexes and histone H1-dependent chromatin compaction mediate propagation of H2AK119ub1 during cell division. *Nature Cell Biology*, *22*(4), 439–452. Https://doi.org/10.1038/s41556-020-0484-1

Zhao, W., Huang, Y., Zhang, J., Liu, M., Ji, H., Wang, C., Cao, N., Li, C., Xia, Y., Jiang, Q., & Qin, J. (2017). Polycomb group RING finger proteins 3/5 activate transcription via an interaction with the pluripotency factor Tex10 in embryonic stem cells. *Journal of Biological Chemistry*, *292*(52), 21527–21537. Https://doi.org/10.1074/jbc.M117.804054

Zhao, W., Liu, M., Ji, H., Zhu, Y., Wang, C., Huang, Y., Ma, X., Xing, G., Xia, Y., Jiang, Q., & Qin, J. (2018). The polycomb group protein Yaf2 regulates the pluripotency of embryonic stem cells in a phosphorylation-dependent manner. *Journal of Biological Chemistry*, *293*(33), 12793–12804. Https://doi.org/10.1074/jbc.RA118.003299

Zhu, X., Yan, M., Luo, W., Liu, W., Ren, Y., Bei, C., Tang, G., Chen, R., & Tan, S. (2017). Expression and clinical significance of pcg-associated protein RYBP in hepatocellular carcinoma. *Oncology Letters*, *13*(1), 141–150. Https://doi.org/10.3892/ol.2016.5380

8. SUMMARY OF THE PHD THESIS

INTRODUCTION

In my PhD study, I have focused on the functions of a PcG protein RYBP, during the early onset of cardiac development using mouse ES cells based *in vitro* differentiation system. We have earlier shown that RYBP is expressed in the mouse developing heart *in vivo* and that ES cells lacking RYBP could not form functionally contracting CMCs *in vitro*. In my PhD thesis, I have further dissected the underlying molecular mechanisms. Genome-wide transcriptomics using wild-type and *Rybp* null mutant CMCs revealed a list of genes with significantly altered expression in the *Rybp* null mutant CMCs including alterations in key mechanisms such as ion homeostasis, cell adhesion, cardiac progenitor formation and sarcomere organisation. One of the most downregulated gene in the *Rybp* null mutant CMCs was Pleiomorphic adenoma gene like 1 (*Plagl1),* a zinc finger protein with transactivation and consensus specific DNA-binding activities. *Plagl1* is an essential factor for cardiac morphogenesis and is highly expressed in mouse hearts from E8.5 of embryonic development to adulthood in a chamber-restricted pattern. Importantly, ablation of *Plagl1* in mouse embryos caused atrial and ventricular septum defects which often lead to cardiomyopathies.

AIM

To identify the regulatory mechanism that cause the downregulation of *Plagl1* expression in the *Rybp* null mutant ES cells and CMCs and investigate how *Plagl1* deficiency could contribute to the non-contractility phenotype of the *Rybp* null mutant CMCs.

METHODS

To achieve these aims I used wild-type and *Rybp* null mutant mouse ES cells and differentiated them to form cardiac cultures up to 21 days and have collected samples from day 0 (pluripotent), day 2 (embryoid body), day 4 (early cardiac progenitor stage), day 7 and day 10 (early CMCs), 14 and 21 (matured CMCs) for further assays. This included: (1) Gene expression studies using quantitative real-time polymerase chain reaction; (2) Western blot analyses to define protein content and kinetics; (3) Immunocytochemistry to determine the subcellular localization of PLAGL1 and its co-localization with key cardiac transcription factors such as cardiac CTNT.

RESULTS AND CONCLUSIONS

Our results showed that PLAGL1 was first detected at the early progenitor formation stages (day 4) and the expression was the highest in the matured CMCs stage (day 14) of differentiation indicating its role in CMC formation. Since in the *Rybp* null mutant cultures, *Plagl1* is not expressed at any time point of *in vitro* cardiac differentiation our current hypothesis is that cells lacking RYBP are impaired in their CMC development at least partially due to the impaired inducibility of *Plagl1*.

Besides this, *in silico* analyses of the *Plagl1* genomic locus has revealed that *Plagl1* has a complex genomic structure containing 11 exons, 3 promoter regions (*P1*, *P2* and *P3*) and 2 ncRNAs *Hymai* and *Plagl1it.* From gene expression analyses using primers specific to the various splice variants of *Plagl1*, I determined that gene products from only *Plagl1 P1* and *P3* promoters are transcribed during *in vitro* cardiac differentiation. Furthermore, the expression of the two ncRNAs *Hymai* and *Plagl1it* were also greatly altered in the *Rybp* mutant cells suggesting that every gene product from the entire *Plagl1* locus is affected by the lack of *Rybp* in the mutant CMCs. Since the ncRNAs are transcribed from the same promoters as *Plagl1* and were induced only at low level in the *Rybp* null mutant CMCs, suggested that RYBP may regulate the promoters of *Plagl1*. Therefore, I cloned the *Plagl1 P1*, *P2* and *P3* promoters into luciferase reporter constructs and performed reporter assays to investigate whether RYBP can regulate any of the *Plagl1* promoters. My results revealed that RYBP can activate *Plagl1* via the *P1* and *P3* promoters. By using truncated mutants of the *P3* promoter, my experiments revealed that RYBP can activate the 3' part of the *P3* promoter that consisted of the TATA box and included consensus sites for cardiac transcription factors NKX2-5 and MEF2C. Site directed mutagenesis of the consensus sites of NKX2-5 and MEF2C and luciferase assays with RYBP disclosed that the NKX2-5 sites were essential for the activation of the *P3* promoter by RYBP. ChIP-qRT-PCR analysis to determine if RYBP bound at the *Plagl1* promoters revealed that RYBP bound at the NKX2-5 consensus in both the *P1* and *P3* promoter. Our results also determined that PLAGL1 co-expressed with CTNT in the wild type differentiating CMCs and PLAGL1 can activate *Tnnt2* promoter. Since these transcription factors are required for the development of multiple cardiac cells types our results suggest that *Plagl1* at least partially, mediates the effects of *Rybp* during cardiac differentiation.

Taken together, my results provide novel interpretation on how the absence of RYBP can cause impairment in cardiac differentiation, which might also be relevant to disease conditions such as cardiomyopathies or arrhythmias.

9. ÖSSZEFOGLALÓ

BEVEZETÉS

A PhD tanulmányom középpontjában a polikomb csoportba tartozó RYBP fehérje szívizom irányú differenciációban betölött szerepének elemzése állt, melyet egér embrionális őssejteken alapuló *in vitro* differenciációs rendszerben vizsgáltunk. Vizsgálataink alapját azon korábbi megfigyelésünk képezte, hogy az RYBP termelődése egér szívben az embrionális fejlődés során kifejezett, továbbá hogy RYBP hiányában *in vitro* kultúrákban az őssejtek nem képesek funkcionálisan összehúzódó kardiomiocitákat (CMC) kialakítani. A doktori disszertációm során tovább elemeztem a molekuláris mechanizmusok hátterét. Először, *Rybp null* mutáns és vad típusú kardiális minták teljes genomi transzkriptom elemzését felhasználva kimutattunk egy sor olyan gént, amelynek expressziója szignifikánsan eltér az *Rybp null* mutáns szívizomsejtekben és olyan fő kardiális mechanizmusok résztvevői, mint pl. az ion homeosztázis, sejt adhézió, kardiális progenitor kialakítás és szarkomer organizáció. Az *Rybp null* mutáns szívizomsejtekben az egyik legkiugróbb változást egy DNS-kötő aktivitással rendelkező cink-ujj fehérje, a Pleiomorphic adenoma gene like 1 (*Plagl1*) mutatta. Ez a gén *Rybp* hiányában sem őssejtekben, sem pedig a CMC populációkban nem volt kimutatható.
A *Plagl1* a kardiális morfogenézis egyik nélkülözhetetlen tagja és a 8.5-ik naptól nagy mennyiségben termelődik a fejlődő egér szívkamrákban. Egér embrióban kimutatták, hogy *Plagl1* hiányában a pitvari és kamrai válaszfal hibásan fejlődik ki, ami gyakran szívbetegségek kialakulásához vezet.

CÉLKITŰZÉSEK

Azonosítani a *Plagl1* csökkent termelődésében szerepet játszó szabályozó mechanizmusokokat, *Rybp null* mutáns őssejtekben és kardiomiocitákban. Úgyszintén, megvizsgálni, hogy a hibás *Plagl1* termelődés vajon miként játszhat szerepet az összehúzódásra képtelen *Rybp null* mutáns szívizomsejtek fenotípus kialakításában.

MÓDSZEREK

Céljaink megvalósításához, *Rybp null* mutáns és vad típusú egér őssejteket differenciáltattuk *in vitro* kardiális irányba. Az kolóniákat 21 napig növesztettük és az analízishez szükséges mintákat 0, 2, 4, 7, 10, 14, és 21 napokon gyűjtöttük be, ahol a nulladik nap a pluripotens, 2.

nap az embrionális test, 4. nap a korai kardiális progenitor, a 7. és 10. nap a korai kardiomiocita a 14. és 21. nap pedig az érett kardiomiocita állapotot reprezentálja. Mintáink vizsgálatához a következő módszereket alkalmaztuk: (1) Kvantitatív valós idejű polimeráz lánc reakcióval (qRT-PCR) gén kifejeződést vizsgáltunk; (2) A fehérjék mennyiségének és kinetikájának kimutatásához Western-blot analízis alkalmaztunk; (3) Immunocitokémiai módszerekkel megvizsgáltuk a PLAGL1 sejteken belüli lokalizációját ill. hogy milyen más kulcsfontosságú kardiális transzkripciós faktorokkal mint pl. a kardiális troponinT-vel lokalizálódik együtt. (4) Luciferáz riporter próbával és helyspecifikus mutációs analízissel megvizsgáltuk az RYBP fehérje *Plagl1 P1*, *P2* és *P3* promóter szakaszára ill. a PLAGL1 fehérje kardiális troponinT promóterére kifejtett hatását, tehát hogy vajon transzkripcionálisan képes-e az RYBP a *Plag1* gén szabályozására.

EREDMÉNYEK (ÉS KÖVETKEZTETÉSEK)

Az eredmények rávilágítottak, hogy a PLAGL1 fehérje a korai kardiális progenitor kialakulásának stádiumában kezd el termelődni és a legnagyobb mértékű kifejeződést a 14. napon az érett kardiomiocita állapotban éri el, jelezve a PLAGL1 fontos szerepét a kardiomiociták kialakulásában. Az *Rybp* null mutáns szívizomtelepekben a *Plagl1* egyik vizsgált időpontban sem mutatott expressziót ezért úgy véljük, hogy a *Plagl1* hibás indukciója legalább részben felelőssé tehető a mutáns kolóniák rendellenesen fejlődő szívizomsejtjeinek kialakításában. Emellett *in silico* feltártuk a *Plagl1* összetett genomi szerkezetét, ami 12 exonból, 3 promóter régióból (*P1*, *P2* és *P3*) és 2 nem-kódoló RNS-ből (*Hymai* és *Plagl1it*) áll. A különböző *Plagl1* splice variánsok kimutatásához specifikus primereket terveztünk és gén expressziós analízissel kimutattuk, hogy *in vitro* kardiális differenciáció során csak a *P1* és *P3 Plagl1* promóterek gén terméke íródik át. Továbbá a *Hymai* és *Plagl1it* nem kódoló RNS-ek termelődése szintén változást mutatott az *Rybp* mutáns sejtekben, ez arra utal, hogy az *Rybp* hiánya a mutáns kardiomiocitákban hatással van a *Plagl1* lokusz összes géntermékre. A nem kódoló RNS-eket ugyanazon promóter régiók szabályozzák, mint a *Plagl1*-et és csak kis mértékben indukálódtak az *Rybp* mutáns szívizomsejtekben, mindez arra engedtek következtetni, hogy az RYBP nagy valószínűséggel közvetlenül szabályozhatja a *Plagl1* promóterét. Ahhoz, hogy ezt a feltevésünket kísérletileg is alátámasszuk a *P1*, *P2* és *P3 Plagl1* promótereket luciferáz riporter konstrukcióba klónoztuk. A riporter-assay eredményei feltárták, hogy az RYBP a *P1* és *P3* promótereken keresztül képes kifejteni aktiváló hatását. Ezután mutáns *P3* promótereket hoztunk létre és segítségükkel kimutattuk, hogy az RYBP a

P3 promóter, TATA-boxokból valamint NKX2-5 és MEF2C kötő szekvenciákból álló, 3' végét aktiválja. A konszenzus NKX2-5 és MEF2C szakaszok célzott mutagenezisével és RYBP luciferáz-assay segítségével megállapítottuk, hogy az *Nkx2-5* kötő helyek megléte nélkülözhetetlen a *P3* promóter aktiválásához. ChIP-qRT-PCR segítségével kimutattuk, hogy az RYBP képes kötődni az *Nkx2-5* kötő konszenzus szekvenciához mindkét, *P1* és *P3*, promóter esetében. Vad típusú kardiomiocitákon végzett kísérleteinkből az is kiderült, hogy a PLAGL1 ugyanazon sejtekben található mint a kariális troponinT és a PLAGL1 képes aktiválni a kardiális troponinT promóterét. Az előbb említett transzkripciós faktorok jelenléte több kardiális sejttípus kialakulásához is szükséges ezért az eredményeink arra engednek következtetni, hogy a kardiális differenciáció során a PLAGL1 legalább részben közvetíti az RYBP hatását. Összességében az eredményeink az RYBP hiányában fellépő hibás kardiális differenciáció új oldalait világítják meg, melyek a különböző kardiális betegségekben, mint pl. a veleszületett szívbetegségek, aritmiák kialakulásában is relevánsak lehetnek

10. APPENDICES

Appendix A: List of Publications

Appendix B: Declaration

APPENDIX A

LIST OF PUBLICATIONS

The thesis is based on the following publications

1) **Surya Henry**, Viktória Szabó, Enikő Sutus, Melinda Katalin Pirity., RYBP is important for cardiac progenitor cell development and sarcomere formation. *Plos One (2020)* doi: /10.1371/journal.pone.0235922

(Q1, I.F: 3.24)

2) Izabella Bajusz, **Surya Henry**, Enikő Sutus, Gergő Kovács, Melinda Pirity., Evolving role of RING1 and YY1 binding protein in germ cell specific transcription regulation. *MDPI Genes* (2019)
doi: /10.3390/genes10110941

(Q1, I.F: 3.75)

Cumulative impact factor: 6.99

MTMT identifier: 10053037

APPENDIX B

DECLARATION

I declare that the contribution of Surya Henry was significant in the listed publications and the doctoral process is based on the publications listed. The results reported in the PhD dissertation and the publications have not been used to acquire any PhD degree previously and will not be used in the future either.

Szeged,

31.08.2021

……..

Dr. Melinda K. Pirity