CRISPR/Cas9-mediated Engineering of Subtype-specific Cell Lines for iPSC-derived Cardiomyocyte Phenotyping

A Thesis Submitted for the Degree of Doctor of Philosophy

Trinity Term, 2022

Naeramit Sontayananon

Oriel College

Division of Cardiovascular Medicine

Radcliffe Department of Medicine

University of Oxford

Supervisors:

Professor Katja Gehmlich, Professor Benjamin Davies and Professor Charles Redwood

COVID-19 impact statement

The COVID-19 pandemic has had a considerable effect on my PhD studies. In compliance with the national lockdown announcement, the University research buildings were only made accessible to essential workers, meaning that I could not perform any laboratory experiments between March and July 2022. In addition, the access was strictly limited during August-September 2022, when a specific capacity was determined for working at a given space according to the social distancing policies. Thus, I could only work on-site at allocated time slots, which was challenging, particularly when conducting key laborious experiments involving stem cells. These together delayed the overall progress of my research plan.

The final characterisation of the fluorescent *MYL7-mClover* and *HCN4-mScarlet* reporter lines could not be accomplished as it unexpectedly required optimisation for cell preparations and sorting conditions for FACS-aided live cell isolation for which there was insufficient time (Chapter 3). The exploration for applications of the *MYL7-mClover* line was also incomplete (Chapter 6) due to the unfinished reporter characterisation. Repeating CRISPR/Cas9 engineering of the *MYL2-mClover* and *PITX2-mClover* reporter lines was also constrained by time (Chapter 2).

Naeramit Sontayananon

September 2022

CRISPR/Cas9-mediated Engineering of Subtype-specific Cell Lines for iPSC-derived Cardiomyocyte Phenotyping

Naeramit Sontayananon, Oriel College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy, Trinity 2022

Abstract

Induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) represent a useful tool for cardiovascular research. However, the mixed subtypes, including atrial-like, ventricular-like and nodal-like, that result from differentiation protocols present a challenge for selective studies. This thesis exploited the emerging genome engineering technology "CRISPR/Cas9" to generate subtype-specific fluorescent reporter systems to label specific CM subtypes, facilitating subsequent isolation for use in subtype-specific disease modelling.

Atrial-specific myosin light chain 7 (*MYL7*)-*mClover* and pituitary homeobox 2 (*PITX2*)-*mClover*, ventricular-specific myosin light chain 2 (*MYL2*)-*mClover* and nodal-specific hyperpolarisation activated cyclic nucleotide-gated potassium channel 4 (*HCN4*)-*mScarlet* iPSCs were established by targeted insertion of the fluorophores at the stop codons of the target genes. Each fluorophore was preceded by a P2A linker, allowing bicistronic expression. Unfortunately, the *PITX2-mClover* and *MYL2-mClover* lines failed quality control and could not be used for fluorescent reporter iPSC-CM derivation. The *MYL7-mClover* and *HCN4-mScarlet* lines, however, exhibited faithful reporter systems where the expression of the fluorophores mirrored that of their respective target genes.

Heterozygous *TTN* c.59926+1 G>A, a truncating mutation of the titin gene, is associated with atrial fibrillation and dilated cardiomyopathy. This mutation was introduced into parental iPSCs for preliminary characterisations. Subsequently, the mutation was engineered in the *MYL7-mClover* iPSCs for exploring the application of subtype-specific reporter systems in disease studies. The fluorescent sorting of the cells could not be optimised in the available time, however, the engineered mutant *MYL7-mClover* line was differentiated towards the atrial and the ventricular lineages.

No prominent consequences in heart failure-associated foetal gene reprogramming, overall *TTN* expression, myofibrillar protein isoform switching, sarcomere organisation, sarcomeric protein localisation were identified in the heterozygous mutant model. Interestingly, the Ca^{2+} transients were altered specifically in the atrial-like heterozygous model. This alteration may underlie irregular, fast atrial rhythm in atrial fibrillation. The atrial-specific phenotype of *TTN* c.59926+1 G>A may be further clarified when using purer atrial-like iPSC-CMs obtainable by fluorescence-activated cell sorting of the differentiated *MYL7-mClover* iPSC-CMs.

Declaration of own work and attribution

Except where otherwise stated, this DPhil thesis is entirely my own work. No portion of the document has been submitted for another degree of this University or a qualification at any other institution. Also, as clearly attributed, the following were performed by collaborators:

All karyotyping work (Chapter 4, 5 & 6) was performed by the Chromosome Dynamics Core Facility at the Wellcome Centre for Human Genetics.

The *R-GECO*-carrying adenoviruses used for the Ca^{2+} transient studies (Chapter 4 & 6) were generated by Dr Alexander Sparrow.

The *mScarlet-ACTN2*-harbouring adenoviruses used for the contractility studies (Chapter 4 & 6) were generated by Dr Violetta Steeples. The contractility analysis (SarcTrack) was conducted with Dr Yiangos Psaras and Professor Chris Toepfer.

The gRNAs used for targeting *TTN* c.11952 and *TTN* c.59926+1 (Chapter 5 & 6) were co-designed with Mr Max Cumberland (University of Birmingham).

Part of Chapter 1 (section 1.6) was previously published as “Fluorescent PSC-derived cardiomyocyte reporter lines: generation approaches and their applications in cardiovascular medicine”
Work was presented previously: Understanding how heart muscle cells function and developing new strategies for repairing the damaged heart are important challenges for tackling heart disease. Human cells from the heart are required in the laboratory for these investigations, but are difficult to obtain in large numbers from patients. As an alternative, stem cells can be used which can be cultured indefinitely in the laboratory and then turned into heart muscle cells. The current methods lead to a mixture of cell types, belonging to the different regions of the heart, creating difficulties when trying to understand the biology of a particular area of the heart, or for future applications when these stem cell-derived cell types are used to repair the heart. Genetic modification of stem cells provides a solution to this problem, as these techniques allow fluorescent markers, so-called reporters, to be inserted into key genes that are active in the different cell types. The different coloured reporters thus allow the identification and purification of specific cell types. In this review, we discuss the various methods that can be used to establish these reporter systems and highlight their applications in different aspects of cardiovascular medicine.

Naeramit Sontayananon

September 2022

Acknowledgments

The work presented in the thesis could have not been accomplished without contributions from many other people. Firstly, I would like to express my gratitude toward my supervisors, Professor Katja Gehmlich, Professor Benjamin Davies and Professor Charles Redwood, for providing access to an outstanding research network. Their patient supervision and academic rigour have had a substantial influence on the development of this thesis.

Secondly, I am grateful for all members, both former and present, within Professor Davies', Professor Gehmlich's and Professor Redwood's research groups, of whom only a limited number are directly described in this document. Of particular importance are Dr Phalguni Ruth and Dr Violetta Steeples, who trained me in numerous laboratory techniques and guided daily experiments. I would also like to thank Dr Charlotte Hooper for molecular analysis advice, Dr Alexander Sparrow for sharing his expertise in Ca^{2+} transient measurement and analysis, Dr Amine Bouchareb for frequent constructive discussion, Mrs Polinka Hernandez-Pliego for her guidance in molecular cloning and Mr Daniel Biggs for his flawless laboratory management skill. Also, I would like to thank Professor Chris Toepfer and Dr Yiangos Psaras for their help in contractility analysis.

Thirdly, my special thanks go to Professor Carolyn Carr and Dr Phillip Hublitz for assessment during my Transfer and Confirmation of Status. Their invaluable suggestions had a solid contribution onto the direction of the thesis.

Fourthly, I am thankful for having been financially supported by the Development and Promotion of Science and Technology Talents Project. Without this funding, my PhD study could have not been initiated.

Fifthly, I wish to dedicate this work to my parents and older sisters, for their immeasurable contributions to my educational development and their unquestionable love.

Table of contents

COVID-19 impact statement **i**

Abstract **ii**

Declaration of own work and attribution **iii**

Acknowledgements **iv**

Table of contents **v**

List of tables **xiii**

List of figures **xiv**

Non-standard abbreviations and acronyms **xx**

Chapter 1 – Introduction

1.1 The heart ***1***

1.1.1 Macroscopic physiology of the heart 1

1.1.2 Histology of the heart 2

1.1.3 Electrical conduction within the heart 3

1.1.4 Autonomic regulation of the heart 5

1.1.5 Heart failure 6

1.1.6 Inherited cardiomyopathies 7

1.2 Cardiomyocytes **(CMs)** ***8***

1.2.1 Typical CM structure 9

1.2.2 Cardiac action potential 10

1.2.1 Atrial and ventricular-type action potential 10

1.2.2 Nodal-type action potential 12

1.2.3 Ca^{2+} signalling in CMs 14

1.2.3.1 Atrial and ventricular-type Ca^{2+} transients 15

1.2.3.2 Nodal-type Ca^{2+} transient 16

1.2.4 Regulation of sarcomere contraction 17

1.2.4.1 Thin filament 18

1.2.4.2 Thick filament 19

1.2.4.2.1 Actomyosin cross-bridge cycling 20

1.2.4.2.2 Atrial and ventricular myosins 20

1.2.4.2.3 Cardiac myosin binding protein C 21
1.2.4.3 Titin: a gigantic and complex muscle protein 22
1.2.4.3.1 The diversity of titin isoform 23
1.2.4.3.2 Complex titin interactome 24
1.2.4.3.2.1 Z-disc interactome 24
1.2.4.3.2.2 I-band interactome 25
1.2.4.3.2.3 A-band interactome 26
1.2.4.3.2.4 M-line interactome 26
1.2.5 β-adrenergic stimulation of contractile activity 26
1.3 Titin in DCM ***27***
1.3.1 Influence of mutational position on *TTNtv* pathology 28
1.3.2 Proposed pathomechanisms 28
1.3.2.1 Haploinsufficiency 28
1.3.2.2 Poison peptide 30
1.4 TTNtv-associated atrial fibrillation ***31***
1.5 Induced pluripotent stem cell-derived CMs (iPSC-CMs) ***32***
1.5.1 iPSC-CM differentiation 33
1.5.2 Limitations of iPSC-CMs 34
1.5.2.1 Immaturity 34
1.5.2.1.1 Electrophysiology and Ca^{2+} signalling 34
1.5.2.1.2 Metabolism 35
1.5.2.1.3 Structure and morphology 36
1.5.2.2 Population heterogeneity and CM purification methods 36
1.5.2.2.1 Percoll gradient 37
1.5.2.2.2 Microfluidic approaches 37
1.5.2.2.3 Metabolic selection 37
1.5.2.2.4 Mitochondrial dyes 37
1.5.2.2.5 Molecular beacons 38

1.5.2.2.6 Cell surface markers 38

1.5.2.2.7 Genetic approaches 38

1.6 Establishment of fluorescent reporter line ***39***

1.6.1 Transient reporter line 39

1.6.2 Random integration of fluorescent reporter transgene 41

1.6.3 Targeted integration of fluorescent reporter transgene 41

1.6.4 Genetic engineering technologies 42

1.6.5 CRISPR/Cas9 43

1.7 Fluorescence-guided subtype-specific iPSC-CM purification ***45***

1.7.1 Ventricular markers 46

1.7.2 Atrial markers 46

1.7.3 Nodal markers 47

1.8 Hypotheses and aims ***47***

1.8.1 Hypotheses 47

1.8.2 Specific aims 48

Chapter 2 – Materials and methods

2.1 Materials ***50***

2.1.1 Bacteria 50

2.1.2 iPSC line 50

2.1.3 Antibodies 50

2.1.4 Taqman assay probes 51

2.2 Methods ***52***

2.2.1 Molecular cloning 52

2.2.1.1 Gibson Assembly 52

2.2.1.2 Ligation-dependent cloning 53

2.2.1.3 Golden Gate Assembly 53

2.2.1.4 Bacterial transformation 53

2.2.1.5 Bacterial electroporation 54

2.2.1.6 Colony PCR 54
2.2.2 Long single-stranded DNA (lssDNA) purification 54
2.2.3 *In vitro* gRNA synthesis 57
2.2.4 Genomic DNA (gDNA) extraction 58
2.2.5 PCR and DNA gel electrophoresis 58
2.2.6 cDNA synthesis and quantitative PCR (qPCR) 58
2.2.7 Protein extraction and Western blot (WB) 59
2.2.8 iPSC culture 60
2.2.9 iPSC quality control 60
2.2.9.1 FACS analysis for iPSC pluripotency 60
2.2.9.2 iPSC karyotyping 61
2.2.10 iPSC transfection and clonal establishment 61
2.2.10.1 iPSC transfection 61
2.2.10.1.1 Ribonucleoprotein (RNP) electroporation 61
2.2.10.1.2 Plasmid reverse transfection 61
2.2.10.2 iPSC low density plating and colony picking 62
2.2.11 iPSC-CM differentiation 62
2.2.12 Live iPSC-CM imaging 63
2.2.13 iPSC-CM Immunofluorescence (IF) 64
2.2.14 FACS analysis for iPSC-CMs 64
2.2.15 Contractility measurement for iPSC-CMs 65
2.2.16 Ca^{2+} transient measurement for iPSC-CMs 65
2.2.17 Statistical analysis 66

Chapter 3 – CRISPR/Cas9-mediated generation of CM subtype-specific fluorescent iPSCs

3.1 Introduction ***67***

3.2 gRNA and targeting donor design ***68***

3.2.1 *MYL7-mClover* line 69
3.2.1.1 MYL7 gRNA efficiency test 69

3.2.1.2 *MYL7-mClover* targeting vector cloning 71
3.2.2 *HCN4-mScarlet* line 74
3.2.2.1 HCN4 gRNA efficiency test 74
3.2.2.2 *HCN4-mScarlet* targeting vector cloning 76
3.2.3 *PITX2-mClover* line 80
3.2.3.1 PITX2 gRNA efficiency test 80
3.2.3.2 *PITX2-mClover* targeting vector cloning 82
3.2.4 *MYL2-mClover* line 87
3.2.4.1 MYL2 gRNA efficiency test 87
3.2.4.2 MYL2 gRNA vector cloning 89
3.2.4.3 *MYL2-mClover* lssDNA synthesis 90
3.3 CRISPR/Cas9-mediated generation of fluorescent reporter iPSC lines ***93***
3.3.1 Generation of *MYL7-mClover* line 93
3.3.2 Generation of *HCN4-mScarlet* line 97
3.3.3 Generation of *PITX2-mClover* line 100
3.3.4 Generation of *MYL2-mClover* line 105
3.4 iPSC quality checks ***108***
3.4.1 Quality checks for *MYL7-mClover* line 108
3.4.2 Quality checks for *HCN4-mScarlet* line 110
3.4.3 Quality checks for *PITX2-mClover* line 112
3.4.4 Quality checks for *MYL2-mClover* line 114
3.5 Discussion ***114***
3.5.1 Modes of CRISPR/Cas9 delivery 114
3.5.2 DNA donor template 115
3.5.3 Summary 117
Chapter 4 – Characterisation of engineered fluorescent reporter iPSC-CM lines
4.1 Introduction ***118***
4.2 MYL7-mClover line ***119***

4.2.1 Optimisation of retinoic acid (RA)-induced atrial differentiation 119

4.2.1.1 Transcriptional expression of atrial and ventricular markers in response to RA 119

4.2.1.2 Protein expression of atrial and ventricular myosin light chains in response to RA 121

4.2.1.3 Effect of RA on cell size 122

4.2.1.4 Effect of RA on Ca^{2+} transient dynamics 122

4.2.1.5 Effect of RA on contractility 124

4.2.2 Correlation between *mClover* and *MYL7* expression 126

4.2.3 Response of *mClover* transcriptional and protein expression to atrial differentiation 126

4.2.4 Investigation of mClover fluorescence expression by microscopy 128

4.2.5 Investigation of mClover fluorescence expression by FACS 131

4.3 HCN4-mScarlet line ***137***

4.3.1 Correlation between *mScarlet* and *HCN4* expression 137

4.3.2 Investigation of mScarlet fluorescence expression by microscopy 138

4.3.3 Investigation of mScarlet fluorescence expression by FACS 139

4.4 PITX2-mClover line ***142***

4.5 MYL2-mClover line ***142***

4.6 Discussion ***143***

4.6.1 RA-induced atrial cell specification 148

4.6.2 *MYL7* expression during *in vitro* atrial differentiation 151

4.6.3 *HCN4* expression during iPSC-CM maturation 152

4.6.4 Purification of iPSC-derived nodal-like cells 154

4.6.5 Summary 155

Chapter 5 – *TTNtv* modelling in KOLF2-C1 iPSC-CMs

5.1 Introduction ***158***

5.2 CRISPR/Cas9-assisted generation of TTNtv iPSC line ***159***

5.2.1 CRISPR design 159

5.2.2 *TTN* c.59926+1 G>A donor design 161

5.2.3 CRISPR/Cas9-assisted generation of *TTNtv* iPSC line 161

5.2.4 iPSC quality control 164

5.3 Characterisation of TTNtv iPSC line ***166***

5.3.1 Molecular characterisation of *TTNtv* 166

5.3.1.1 Effect of *TTNtv* on cardiac gene expression 166

5.3.1.2 Effect of *TTNtv* on *TTN* exon splicing 172
5.3.2 Functional and structural consequences of *TTNtv* 175
5.4 Discussion ***177***
5.4.1 *TTNtv* pathogenesis 178
5.4.2 Titin isoforms and diseases 180
5.4.3 *TTNtv* and CM development 180
5.4.4 *TTNtv* disrupts sarcomerogenesis 181
5.3.5 Concluding remarks 183
Chapter 6 – Subtype-specific characterisation of *TTNtv* in *MYL7-mClover* iPSC-CMs
6.1 Introduction ***184***
6.2 CRISPR/Cas9-assisted generation of TTNtv in MYL7-mClover iPSC line ***185***
6.2.1 CRISPR and *TTN* c.59926+1 G>A donor design 185
6.2.2 CRISPR/Cas9-assisted generation of *TTNtv* iPSC line 185
6.2.3 iPSC quality checks 188
6.3 Characterisation of TTNtv in ventricular-like MYL7-mClover iPSC-CMs ***190***
6.3.1 Molecular analysis 190
6.3.1.1 Effect of *TTNtv* on transcriptional expression of cardiac genes 190
6.3.1.2 Effect of *TTNtv* on *TTN* splicing 194
6.3.2 Functional analysis 197
6.3.2.1 Effect of *TTNtv* on spontaneous contraction 197
6.3.2.2 Effect of *TTNtv* on contractility 199
6.3.2.3 Effect of *TTNtv* on Ca^{2+} transient dynamics 201
6.3.3 Effect of *TTNtv* on sarcomeric protein localisation 203
6.4 Characterisation of TTNtv in atrial-like MYL7-mClover iPSC-CMs ***211***
6.4.1 Molecular analysis 211
6.4.1.1 Effect of *TTNtv* on transcriptional expression of cardiac genes 211
6.3.1.2 Effect of *TTNtv* on *TTN* splicing 215
6.4.2 Functional analysis 218
6.4.2.1 Effect of *TTNtv* on spontaneous contraction 218
6.4.2.2 Effect of *TTNtv* on Ca^{2+} transient dynamics 218
6.4.3 Effect of *TTNtv* on sarcomeric protein localisation 220
6.5 Discussion ***225***
6.5.1 *TTNtv* delays sarcomere maturation 228

6.5.2 *TTNtv* and dysregulated protein quality control (PQC) 229
6.5.3 Length-dependent Ca^{2+} sensitivity of tension/activation 231
6.5.4 Arrhythmogenic potential of RA-induced atrial-like iPSC-CMs 232
6.5.5 Concluding remarks 233

Chapter 7 – Discussion

7.1 Generation of subtype-specific fluorescent reporter lines ***235***
7.2 Faithful reporter systems: do they enable specific iPSC-CM subtype identification? ***236***
7.3 Subtype-specific TTNtv modelling ***238***
7.4 Future directions ***240***
7.4.1 Validation of subtype-specific fluorescent reporter lines 240
7.4.2 Dissecting *TTNtv* pathogenesis 240
7.5 Concluding remarks ***241***
References **243**
Appendix **262**

List of tables

Chapter 1 – Introduction

Table 1-1 Approximate ion concentration across the sarcolemma at the resting stage 11

Chapter 2 – Materials and methods

Table 2-2 List of antibodies 50

Table 2-2 List of Taqman assays 51

Appendix

Table S1 Recipes for buffers and solutions 262

Table S2 List of commercial kits and other commercial products 262

Table S3 List of instruments 263

Table S4 List of primers and DNA oligos 254

List of figures

Chapter 1 – Introduction

Figure 1-1 Unidirectional blood flow within the heart 2

Figure 1-2 Layers of the heart 3

Figure 1-3 Simplified illustration of an intercalated disc 4

Figure 1-4 Cardiac conduction system 5

Figure 1-5 Common inherited cardiomyopathies 7

Figure 1-6 Microscopic anatomy of CMs 9

Figure 1-7 Simplified illustration of morphology of CM subtypes 10

Figure 1-8 Atrial and ventricular-type action potentials and their underlying ion currents 12

Figure 1-9 Nodal-type action potential and its underlying ion currents 13

Figure 1-10 Ca^{2+}-induced Ca^{2+}-release mechanism and excitation-contraction coupling 14

Figure 1-11 Simplified illustration of Ca^{2+} transient profile 15

Figure 1-12 Atrial and ventricular Ca^{2+} transient profile 15

Figure 1-13 Local Ca^{2+} release mechanism of the SR of nodal cardiomyocyte 17

Figure 1-14 Simplified illustration of nodal-type Ca^{2+} transient profile 17

Figure 1-15 Schematic illustration of sarcomere structure 18

Figure 1-16 Schematic of thin filament configuration during diastole and systole 19

Figure 1-17 Schematic illustration of thin filament structure 20

Figure 1-18 Actomyosin cross-bridge cycle 21

Figure 1-19 Simplified depiction of a single titin protein organisation relative to the sarcomere regions 22

Figure 1-20 Human *TTN* isoforms 23

Figure 1-21 Titin binding partners 25

Figure 1-22 β-adrenergic response in CMs 27

Figure 1-23 Simplified illustration of NMD mechanism 29

Figure 1-24 Electrocardiogram of atrial fibrillation 31

Figure 1-25 Schematic illustration of embryonic heart development with respect to the embryonic stages 32

Figure 1-26 Immaturity of iPSC-CMs 35

Figure 1-27 Fluorescent reporter line generation methods 40

Figure 1-28 Naturally occurring and engineered CRISPR/Cas9 44

Figure 1-29 *S. pyogenes* CRISPR/Cas9 machinery 44

Figure 1-30 Mechanisms of CRISPR/Cas9-assisted genome editing 45

Chapter 2 – Materials and methods

Figure 2-1 Simplified directional Gibson Assembly diagram 52

Figure 2-2 Design for lssDNA purification 56

Figure 2-3 Simplified illustration of lssDNA purification 57

Figure 2-4 A conventional, directed monolayered iPSC-CM differentiation protocol 61

Chapter 3 – CRISPR/Cas9-mediated generation of CM subtype-specific fluorescent iPSCs

Figure 3-1 Proposed mechanism of self-cleavage P2A peptide 68

Figure 3-2 Targeting constructs for genomic insertion of a promoterless fluorescent reporter at a stop codon 69

Figure 3-3 TIDE analysis for CRISPRs targeting the *MYL7* locus 70

Figure 3-4 MYL7 CRISPR-C location relative to the *MYL7* locus 71

Figure 3-5 *MYL7-mClover* targeting vector cloning 74

Figure 3-6 TIDE analysis for CRISPRs targeting the *HCN4* locus 75

Figure 3-7 HCN4 CRISPR-A location relative to the *HCN4* locus 76

Figure 3-8 *HCN4-mScarlet* targeting vector cloning 79

Figure 3-9 TIDE analysis for CRISPRs targeting the *PITX2* locus 81

Figure 3-10 PITX2 CRISPR-A location relative to the *PITX2* locus 82

Figure 3-11 *PITX2-mScarlet* targeting vector cloning 86

Figure 3-12 TIDE analysis for CRISPRs targeting the *MYL2* locus 88

Figure 3-13 MYL2 CRISPR-B location relative to the *MYL2* locus 89

Figure 3-14 MYL2 CRISPR-B cloning and *MYL2-mClover* donor purification 92

Figure 3-15 Genotypic screening for *MYL7-mClover* line 95

Figure 3-16 Conformational PCR of isolated *MYL7-mClover* iPSCs after clonal expansion 96

Figure 3-17 Genotypic screening for *HCN4-mScarlet* line 98

Figure 3-18 Conformational PCR of isolated *HCN4-mScarlet* iPSCs after clonal expansion 100

Figure 3-19 Genotypic screening for *PITX2-mClover* line 102

Figure 3-20 Conformational PCR of isolated *PITX2-mClover* iPSCs after clonal expansion 104

Figure 3-21 Genotypic screening for *MYL2-mClover* line 106

Figure 3-22 Conformational PCR of isolated *MYL2-mClover* iPSCs after clonal expansion 107

Figure 3-23 *MYL7-mClover* iPSC line's quality control 109

Figure 3-24 *HCN4-mScarlet* iPSC line's quality control 111

Figure 3-25 *PITX2-mClover* iPSC line's quality control 113

Chapter 4 – Characterisation of engineered fluorescent reporter iPSC-CM lines

Figure 4-1 Gene expression profile for atrial differentiation in day 30 *MYL7-mClover* iPSC-CMs 120

Figure 4-2 Protein expression analysis of atrial differentiation in day 30 *MYL7-mClover* iPSC-CMs 121

Figure 4-3 Quantification of cell area of day 30 *MYL7-mClover* iPSC-CMs upon atrial induction 122

Figure 4-4 Ca^{2+} transient analysis of day 35 *MYL7-mClover* iPSC-CMs differentiated with or without RA treatment. 124

Figure 4-5 Contractility analysis of day 35 *MYL7-mClover* iPSC-CMs differentiated with or without RA treatment 125

Figure 4-6 Relationship between *MYL7* and *MYL7*-driven *mClover* expression patterns 126

Figure 4-7 *mClover* transcriptional expression in day 30 *MYL7-mClover* iPSC-CMs upon RA-mediated atrial differentiation 127

Figure 4-8 mClover protein expression in day 30 *MYL7-mClover* iPSC-CMs upon RA-mediated atrial differentiation 127

Figure 4-9 Live cell imaging of day 30 *MYL7-mClover* iPSC-CMs differentiated by the atrial and ventricular protocols 129

Figure 4-10 Immunofluorescent images of day 30 *MYL7-mClover* iPSC-CMs differentiated by the atrial and ventricular protocols 130

Figure 4-11 FACS analysis for MYL7 expression of day 25 fixed *MYL7-mClover* iPSC-CMs differentiated by the atrial and ventricular protocols 132

Figure 4-12 FACS analysis for mClover expression of day 25 fixed *MYL7-mClover* iPSC-CMs differentiated by the atrial and ventricular protocols 133

Figure 4-13 FACS analysis for ACTN2 and mClover expression of day 25 fixed *MYL7-mClover* iPSC-CMs differentiated by the atrial and ventricular protocols 134

Figure 4-14 FACS analysis for mClover expression of day 20 live *MYL7-mClover* iPSC-CMs differentiated by the ventricular protocol 136

Figure 4-15 Relationship between *HCN4* and *HCN4*-driven *mScarlet* expression patterns 137

Figure 4-16 Live cell imaging of day 23 *HCN4-mScarlet* iPSC-CMs 138

Figure 4-17 FACS analysis of day 25 fixed *HCN4-mScarlet* iPSC-CMs differentiated by the ventricular protocol 140

Figure 4-18 FACS analysis for mScarlet expression of day 20 live *HCN4-mScarlet* iPSC-CMs differentiated by the ventricular protocol 141

Figure 4-19 RNA expression of day 25 *MYL2-mClover* iPSC-CMs differentiated by the ventricular protocol 142

Figure 4-20 WB analysis of day 25 *MYL2-mClover* iPSC-CMs differentiated by the ventricular protocol 143

Figure 4-21 RA signalling and its regulatory role in atrial lineage specification 149

Figure 4-22 Proposed time course *MYL7* expression in iPSC-CMs differentiated with and without RA 152

Chapter 5 – *TTNtv* modelling in KOLF2-C1 iPSC-CMs

Figure 5-1 Schematic illustration of *TTN* c.11952 C>A and c.59926+1 G>A locations with respect to the *TTN* locus 159

Figure 5-2 TTN CRISPR locations 159

Figure 5-3 TIDE analysis for gRNAs containing TTN c.11952 and c.59926+1 CRISPRs 160

Figure 5-4 *TTN* c.59926+1 G>A ssODN donor template design 161

Figure 5-5 Genotypic screening for *TTN* c.59926+1 G>A change 162

Figure 5-6 Genotypic confirmation for *TTN* c.59926+1 G>A editing 163

Figure 5-7 *TTN* c.59926+1 G>A iPSC line's quality control 165

Figure 5-8 Transcriptional expression profile of day 21 *TTN* c.59926+1 G>A KOLF2-C1 iPSC-CMs 167

Figure 5-9 Transcriptional expression profile of day 31 *TTN* c.59926+1 G>A KOLF2-C1 iPSC-CMs 168

Figure 5-10 *TTN* transcriptional levels of day 21 *TTN* c.59926+1 G>A KOLF2-C1 iPSC-CMs 170

Figure 5-11 *TTN* transcriptional levels of day 31 *TTN* c.59926+1 G>A KOLF2-C1 iPSC-CMs 171

Figure 5-12 *TTN* c.59926+1 G>A caused changed *TTN* transcript species in KOLF2-C1 iPSC-CMs 173

Figure 5-13 *TTN* c.59926+1 G>A caused intronic retention in KOLF2-C1 iPSC-CMs 174

Figure 5-14 *TTN* c.59926+1 G>A impaired the sarcomere formation in KOLF2-C1 iPSC-CMs 176

Figure 5-15 Simplified illustration of sarcomerogenesis in CMs 182

Chapter 6 – Subtype-specific characterisation of *TTNtv* in *MYL7-mClover* iPSC-CMs

Figure 6-1 Genotypic screening for *TTN* c.59926+1 G>A in *MYL7-mClover* iPSCs 185

Figure 6-2 Genotypic confirmation for *TTN* c.59926+1 G>A in *MYL7-mClover* iPSCs 187

Figure 6-3 *MYL7-mClover, TTN* c.59926+1 G>A iPSC line's quality control 189

Figure 6-4 *TTN* transcriptional levels of day 31 ventricular-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A 191

Figure 6-5 Transcriptional expression profile of day 31 ventricular-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A 193

Figure 6-6 *TTN* c.59926+1 G>A caused changed *TTN* transcript species in ventricular-like *MYL7-mClover* iPSC-CMs 195

Figure 6-7 *TTN* c.59926+1 G>A caused intronic retention in ventricular-like *MYL7-mClover* iPSC-CMs 196

Figure 6-8 Live cell imaging of day 30 ventricular-like *MYL7-mClover* iPSC-CMs bearing *TTN* c.59926+1 G>A 198

Figure 6-9 Contractility analysis of day 35 ventricular-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A 200

Figure 6-10 Ca^{2+} transient analysis of day 35 ventricular-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A 203

Figure 6-11 *TTN* c.59926+1 G>A impaired sarcomere formation in ventricular-like *MYL7-mClover* iPSC-CMs 205

Figure 6-12 N-terminus titin was expressed in ventricular-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A 206

Figure 6-13 *TTN* c.59926+1 G>A generated truncated titins lacking the M-line region in ventricular-like *MYL7-mClover* iPSC-CMs 207

Figure 6-14 *TTN* c.59926+1 G>A impeded myofibrillogenesis in ventricular-like *MYL7-mClover* iPSC-CMs 208

Figure 6-15 Truncated titin may interact with ACTN2 in ventricular-like *MYL7-mClover* iPSC-CMs carrying homozygous *TTN* c.59926+1 G>A 209

Figure 6-16 Normal Z-disc and M-line protein localisation in ventricular-like *MYL7-mClover* iPSC-CMs harbouring heterozygous *TTN* c.59926+1 G>A 210
Figure 6-17 *TTN* transcriptional levels of day 31 atrial-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A 212
Figure 6-18 Transcriptional expression profile of day 31 atrial-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A 215
Figure 6-19 *TTN* c.59926+1 G>A caused changed *TTN* transcript species in atrial-like *MYL7-mClover* iPSC-CMs 216
Figure 6-20 *TTN* c.59926+1 G>A caused intronic retention in atrial-like *MYL7-mClover* iPSC-CMs 217
Figure 6-21 Ca^{2+} transient analysis of day 35 atrial-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A 219
Figure 6-22 *TTN* c.59926+1 G>A abolished sarcomerogenesis in atrial-like *MYL7-mClover* iPSC-CMs 221
Figure 6-23 *TTN* c.59926+1 G>A truncated titin in atrial-like *MYL7-mClover* iPSC-CMs 222
Figure 6-24 *TTN* c.59926+1 G>A impeded myofibrillogenesis in atrial-like *MYL7-mClover* iPSC-CMs carrying *TTNtv* c.59926+1 G>A 223
Figure 6-25 Normal Z-disc and M-line protein localisation in atrial-like *MYL7-mClover* iPSC-CMs harbouring *TTN* c.59926+1 G>A 224
Figure 6-26 UPS mechanism 230

Appendix

Figure S1 Plasmid map for V294-neo-dTA 267
Figure S2 Plasmid map for PL452 268
Figure S3 Plasmid map for Litmus28 269
Figure S4 Plasmid map for pX330-puro 270
Figure S5 Plasmid map for 18ACH20P_2368293 271
Figure S6 Plasmid map for pJET1.2/blunt 272
Figure S7 MYL2-P2A-mClover gBlock template 273

Non-standard abbreviations and acronyms

3'HA	3' homology arm
5'HA	5' homology arm
A-band	Anisotropic band
ACTN2	Alpha actinin 2
BMP4	Bone morphogenetic protein 4
CD	Ca^{2+} transient duration
CMs	Cardiomyocytes
DCM	Dilated cardiomyopathy
dsDNA	Double stranded DNA
dTA	Diphtheria toxin A
E-C coupling	Excitation-contraction coupling
ESCs	Embryonic stem cells
ESC-CMs	Embryonic stem cell-derived cardiomyocytes
E8	Essential 8
FACS	Fluorescence-activated cell sorting
F-actin	Filamentous actin
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
gDNA	Genomic DNA
gRNA	Single guide RNA
HA	Homology arm
HCM	Hypertrophic cardiomyopathy
HCN4	Hyperpolarisation activated cyclic nucleotide-gated potassium channel 4
HDR	Homology-directed repair
HEY2	Hairy/enhancer-of-split related with YRPW motif protein 2
I-band	Isotropic band
IF	Immunofluorescence
iPSCs	Induced pluripotent stem cells
iPSC-CMs	Induced pluripotent stem cell-derived cardiomyocytes
Indel	Insertion/deletion
IRX4	Iroquois homeobox protein 4
KCNJ3	K^+ channel subfamily J Member 3

KO	Knock-out
L28	Litmus28
lssDNA	Long single stranded DNA
LTCC	L-type Ca^{2+} channel
MESP1	Mesoderm posterior basic helix-loop-helix transcription factor 1
M-line	Middle line
MURF1	Muscle ring finger protein 1
MURF2	Muscle ring finger protein 2
MYBPC3	Myosin binding protein C 3
MYH6	Myosin heavy chain 6
MYH7	Myosin heavy chain 7
MYL2	Myosin (regulatory) light chain 2
MYL7	Myosin (regulatory) light chain 7
MYOM1	Myomesin 1
NCX	Na^{+}/Ca^{2+} exchanger
Neo	Neomycin resistant gene
NKX2.5	NK2 homeobox 5
NMD	Nonsense mediated decay
NPPA	Atrial natriuretic peptide
NPPB	Brain natriuretic peptide
NR2F2	Nuclear receptor subfamily 2 group F member 2
P2A	Porcine teschovirus-1 2A
PAM	Protospacer adjacent motif
PBS	Phosphate buffered saline
PITX2	Pituitary homeobox 2
PLN	Phospholamban
PQC	Protein quality control
qPCR	Quantitative PCR
RA	Retinoic acid
RBM20	RNA binding motif protein 20
RGECO	Red genetically encoded calcium indicator
ROCKi	Rho-kinase inhibitor
RNP	Ribonucleoprotein
RYR	Ryanodine receptor

SERCA	Sarco/endoplasmic reticulum Ca^{2+}-ATPase
SHOX2	Short stature homeobox 2
SLN	Sarcolipin
ssDNA	Single-stranded DNA
ssODN	Single-stranded oligo deoxynucleotide
SR	Sarcoplasmic reticulum
T-CAP	Titin-cap
TIDE	Tracking of indel by decomposition
TTN	Titin
TTNtv	Titin truncating variants
T-tubule	Transvers tubule
UPS	Ubiquitin-proteasome system
WB	Western blot
WNT	Wingless integrated
WT	Wild type
Z-disc	Zwischenscheibe disc

Chapter 1

Introduction

1.1 The heart

The heart is a vital organ whose role is to continuously supply blood throughout the body via the circulatory system. Its function is tightly regulated by electrical activation and propagation, ensuring rhythmic cycles of cardiac muscle contraction (systole) and relaxation (diastole) associated with the blood pumping. Aberration in the cardiac function can ultimately lead to heart failure.

1.1.1 Macroscopic physiology of the heart

The human heart is comprised of 4 contractile chambers partitioned by the septum and valves (Figure 1-1) (Guyton and Hall, 2000), tightly controlling a unidirectional blood flow. The superior vena cava and inferior vena cava collect deoxygenated blood from the upper body and the lower body, respectively, and pass it into the right atrium. During diastole, the tricuspid valve remains open, allowing the deoxygenated blood to travel from the right atrium into the right ventricle. During systole, the pressure within the right ventricle increases, forcing the tricuspid valve shut resulting in the blood entering the pulmonary artery via the pulmonary valve. Within the pulmonary circulation system, CO_2 in the blood is exchanged for O_2. Then, the oxygenated blood returns to the heart at the left atrium.

Similar to the situation from the right side, the oxygenated blood passes into the left ventricle via the mitral valve, which remains open during diastole. The valve closes during systole as a consequence of the increased pressure within the left ventricle, enabling the oxygenated blood to exit the heart via the aortic valve. Then, the oxygenated blood travels via the aorta and the subsequent arterial system to supply nutrients, gas and signalling molecules for all tissues of the body. Finally, the depleted blood returns to the heart via the venous system to complete the cycle.

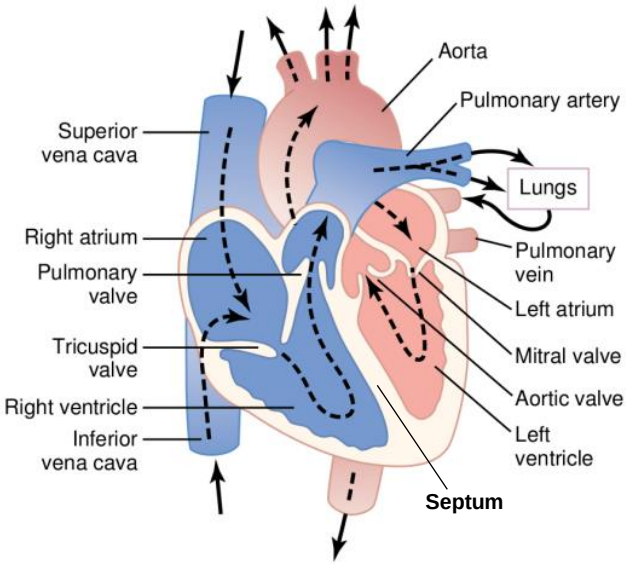


Figure 1-1 Unidirectional blood flow within the heart. The deoxygenated blood (blue) returns to the right atrium via the superior vena cava and the inferior vena cava and then is passed into the right ventricle via the tricuspid valve. The deoxygenated blood enters the pulmonary circulatory system via the pulmonary valve for gas exchange. The oxygenated blood (red) fills the left atrium and then travels into the left ventricle through the mitral valve. The blood leaves the heart via the aortic valve to provide nutrients, gas and signalling molecules all over the body. Modified from (Guyton and Hall, 2000).

1.1.2 Histology of the heart

The majority of the heart tissue is of mesoderm origin, comprised of 4 main layers (Tran et al., 2020) (Figure 1-2). Firstly, the pericardium is a fibrous sac surrounding the heart from all sides except the roots of the great vessels. It positions the heart in the mediastinum, limits overfilling of the heart and provides a mechanical barrier against infection. The sac is made of connective tissues and layered with mesothelium at the innermost area. Secondly, the epicardium (or the visceral pericardium) is composed of connective tissues, where cardiac ganglia develop, nerves travel and adipose cells accumulate. Multipotent epicardial cells can migrate into the myocardium and the endocardium, giving rise to valve cells, fibroblasts (responsible for extracellular matrix deposition and signalling molecular secretion), smooth muscles and components of the coronary vasculature. Thirdly, the myocardium is primarily composed of cardiac muscle cells or cardiomyocytes (CMs), which are responsible for the contractile activity of the heart. This layer is supplied with a complex system of the coronary vasculature, lymphatics and autonomic innervation. Fourthly, the endocardium is made of a monolayer of endothelial cells and subendothelial connective tissue. It provides the lining of the heart

chambers and valves and a separation between the blood stream and the myocardium. It also houses the cardiac conduction system (the atrioventricular node, bundle of His, bundle branches and Purkinje's fibres (Sánchez-Quintana, 2003, Nagarajan et al., 2019)), which regulates heart rhythm.

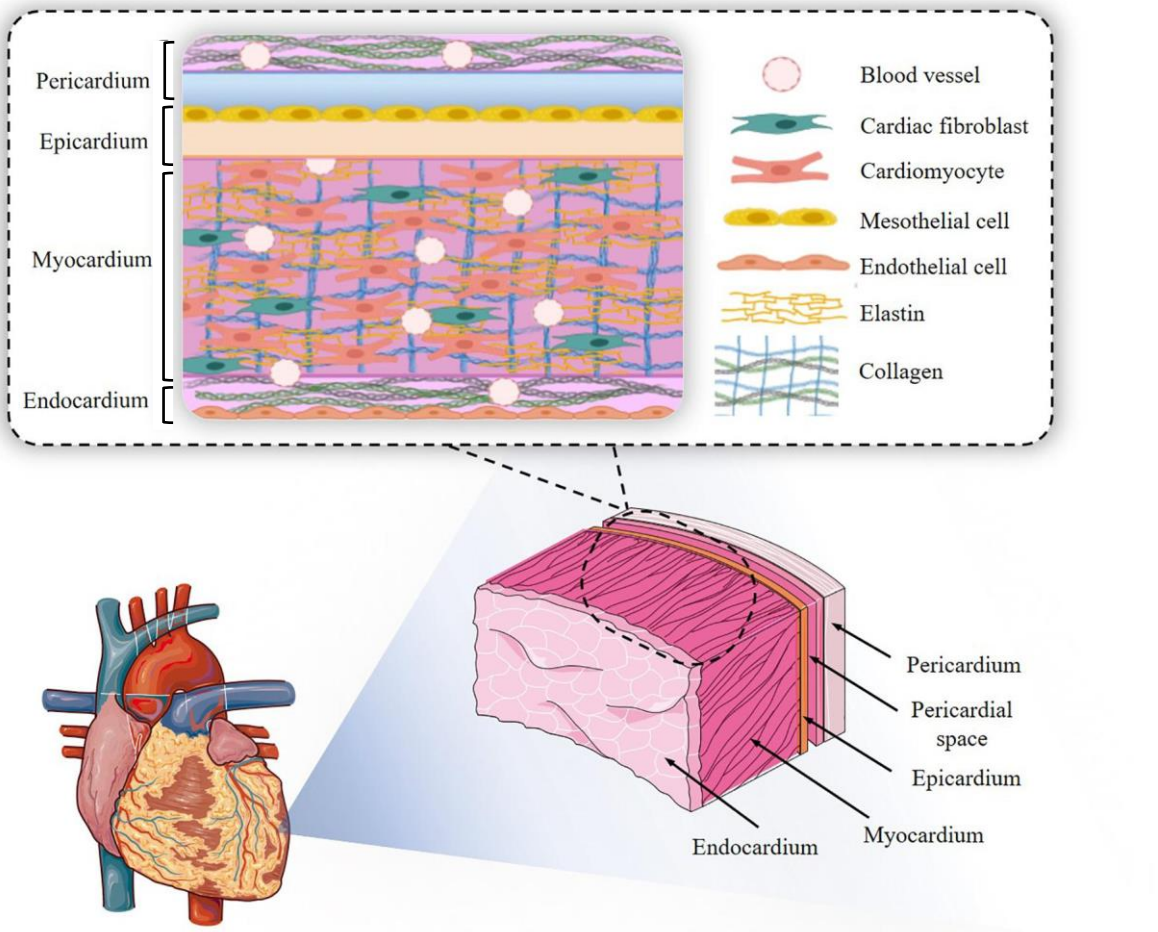


Figure 1-2 Layers of the heart. The human heart is comprised of 4 main layers: the epicardium, the pericardium, the myocardium and the endocardium. The epicardial sac holds the heart at the middle of the thoracic cavity. The epicardium functions as an outer protective layer of the heart. The myocardium is primarily made of cardiac muscles and thus, has a role in the cardiac contractility. The endothelium provides a mechanical barrier between the myocardium and the blood stream. Modified from (Montero et al., 2020).

1.1.3 Electrical conduction within the heart

Coordination of contraction within the atria and ventricles is partly mediated by a syncytial structure called an "intercalated disc" (Figure 1-3), allowing action potential propagation between connecting CMs. It is composed of a number of ion channels, fascia adherens, desmosome and gap junction (Kleber and Saffitz, 2014). However, this mechanism alone is

not efficient enough to meet the precise timing of chamber filling and ejection required for proper cardiac function.

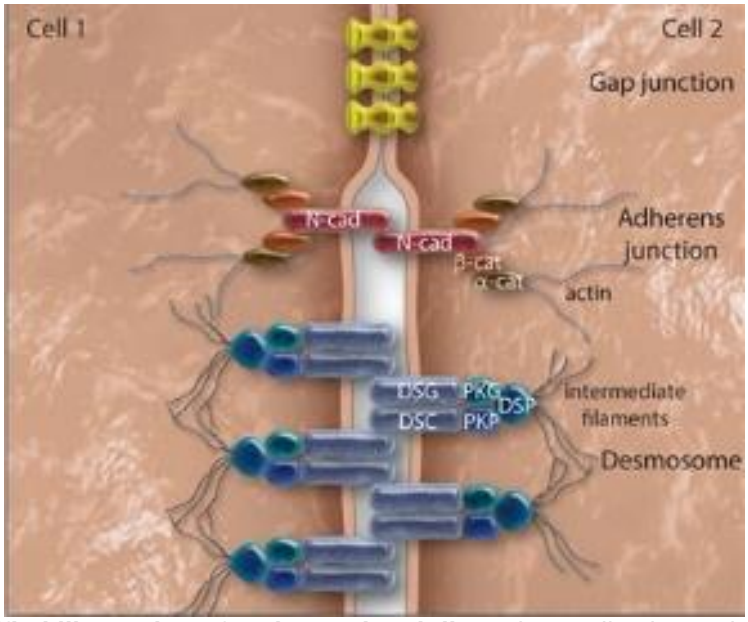


Figure 1-3 Simplified illustration of an intercalated disc. The cardiac intercalated disc provides an electromechanical connection between adjacent CMs. It is consisting of gap junctions, adherens junctions and desmosomes. Gap junctions allow specific ion exchange between connecting cells. Fascia adherens junctions (or adherens junctions), which are comprised of N-cadherin (N-cad), α-catenin and β-catenin, form a mechanical link between the cells via the actin network. Desmosomes (or macula adherens junctions) are intermediate filament-linked complexes of plakophilin (PKP), plakoglobin (PKG) desmoglein (DSG), desmocollin (DSC) and desmoplakin (DSP), specialising in cell-cell adhesion. Taken from (Vermij et al., 2017).

The heart develops its own conduction system comprised of specialised cells which rhythmically initiate and propagate electrical signals throughout the heart (Figure 1-4A) (Monteiro et al., 2017). Localising within the upper region of the right atrium, the sinoatrial node spontaneously generates and propagates the stimulatory signal simultaneously throughout the right atrium and left atrium via the intermodal pathway and Bachmann's bundle, respectively, stimulating atrial depolarisation/contraction and ejection of the blood from the atria into the ventricles. The atrial activity correlates with the P-wave in an electrocardiogram (Figure 1-4B).

Subsequently, the electrical signal is concentrated at the atrioventricular node before accelerating down the bundle of His and into the left and right branches. The atrio-ventricular transition is relatively delayed (~0.1 sec) to allow a complete atrial contraction and efficient ventricular filling prior to the ventricular contraction. The signal continues travelling through

the Purkinje fibres innervating the heart's apex, causing the ventricular depolarisation/contraction beginning at the bottom of the ventricles towards the pulmonary and aortic valve. This results in propelling the blood into the pulmonary and systemic circulations. The ventricular contraction appears as the QRS complex on an electrocardiogram.

Finally, as the impulse dissipates, the ventricles proceed to repolarisation/relaxation. This can be detected as the T-wave on an electrocardiogram.

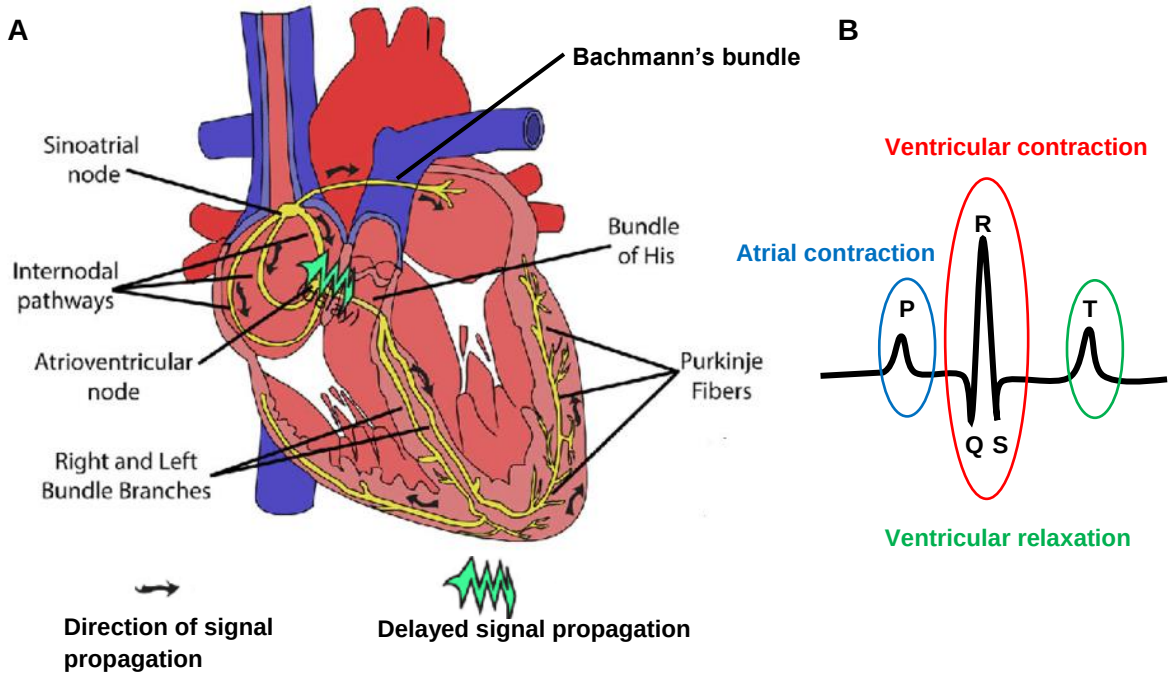


Figure 1-4 Cardiac conduction system. A) Simplified demonstration of the electric conduction pathway of the heart. B) Typical electrocardiogram morphology corresponding to the cardiac activity. The P-wave represents the atrial activation as a consequence of the electric propagation within the atrial chambers. The QRS complex reflects the ventricular contraction which corresponds to when the impulse distributes within the ventricles. The T-wave represents the ventricular relaxation phase, where the electrical signal disappears from the ventricles. Modified from (Monteiro et al., 2017).

1.1.4 Autonomic regulation of the heart

In addition to the regulation by the intrinsic conduction system, the heart rhythm can be tuned by the autonomic nervous system, including the sympathetic and parasympathetic systems, to meet the demand for the blood contents under different conditions (Gordan et al., 2015).

Activation of the sympathetic system by e.g. stress, exercise, emotional excitement and heart failure leads to the release of the excitatory hormones epinephrine and norepinephrine. These chemicals interact with the β-adrenergic receptors of nodal, atrial and ventricular CMs, resulting in increased electrical conductance (dromotropy), cardiac contractility (inotropy) and heart rate (chronotropy).

Meanwhile, during rest, sleep, emotional tranquillity or post-sympathetic activation, the parasympathetic system is activated. Acetylcholine secreted from the vagal nerve of this system exerts negative effects for dromotropy, chronotropy and inotropy by stimulating the muscarinic receptors of CMs.

1.1.5 Heart failure

Despite the delicate regulatory mechanisms enabling adaptation to different physiological needs, cardiac function can be undermined under disease states. Persistence of the pathogenesis without a proper intervention results in heart failure. Heart failure is defined as a cardiac condition which fails to provide sufficient cardiac output (the amount of blood ejected over a given interval) to meet the requirements of the body tissues (McMurray et al., 2012). Its aetiology is diverse e.g. abnormalities in valve function, heart rhythm (arrhythmia), the conduction system and cardiac muscle (cardiomyopathy) and infection.

Heart failure can be classified into 2 major groups. Firstly, systolic heart failure is described as a condition of impaired cardiac contractility, resulting in increased end diastolic volume and reduced left ventricular ejection fraction (<35%)(McMurray et al., 2012). Thus, this condition is also known as heart failure with reduced ejection fraction. Secondly, diastolic heart failure involves reduced ventricular compliance independent from changes in contractility, leading to depleted ventricular filling and stroke volume (the amount of blood ejected during each contractile cycle). Since the ejection fraction is not affected (≥50%)(McMurray et al., 2012), this condition is also known to as heart failure with preserved ejection fraction.

1.1.6 Inherited cardiomyopathies

Cardiomyopathy is one the factors contributing to heart failure. It can be either acquired or inherited. 2 common types of cardiomyopathies are hypertrophic cardiomyopathy (HCM) and dilated cardiomyopathy (DCM), both of which are usually inherited (Watkins et al., 2011) (Figure 1-5).

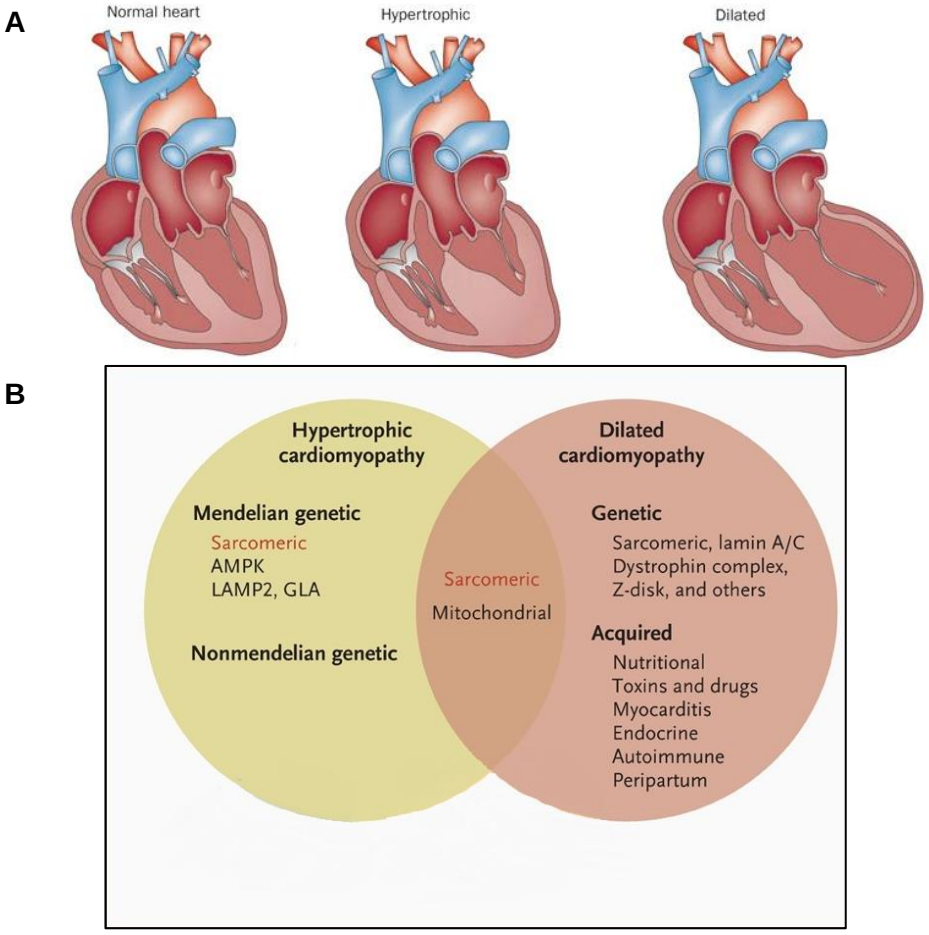


Figure 1-5 Common inherited cardiomyopathies. A) Morphology of hypertrophic cardiomyopathy (HCM) with thickened left ventricular wall and dilated cardiomyopathy (DCM) with left ventricular dilation. Modified from (Hershberger et al., 2013). B) Categories of HCM and DCM. *AMPK*, AMP-activated protein kinase; *LAMP2*, lysosomal-associated membrane protein 2; *GLA*, α-galactosidase. Modified from (Watkins et al., 2011).

HCM and DCM share some genetic causes, especially sarcomeric genes (Figure 1-5B). HCM is characterised primarily by thickening of the left ventricular wall with CM enlargement (hypertrophy) (Figure 1-5A) and secondarily by CM disarray and myocardial fibrosis (Watkins et al., 2011). Its prevalence is estimated as high as 1:500 in young adults (Maron et al., 1995). Its most common genetic markers are mutations (autosomal dominant) in components of the

CM's contractile unit "sarcomere" – myosin heavy chain 7 (*MYH7*) and cardiac myosin binding protein C (encoded by myosin binding protein C 3 (*MYBPC3*)), each accounting for approximately 25-35%. The diseases-causing mutations in *MYH7* are generally missense, encoding malfunctioning proteins capable of being incorporated into the sarcomere. Meanwhile, truncating mutations, which lead to insufficiency of proteins from the wild-type (WT) allele to compensate functionally for the non-functioning products from the mutant allele, are more common in the case of *MYBPC3*.

DCM is a disease of left ventricular dilation (weakened ventricular wall, Figure 1-5A) with systolic dysfunction (reduced left ventricular ejection fraction). The estimated DCM prevalence is elusive, ranging from 1:2700 (Maron et al., 1995) up to 1:250-400 (Hershberger et al., 2013). Compared to HCM, DCM aetiology is complex and heterogeneous. Over 30 genes have been reported to be associated with DCM, mostly missense mutations, including those involving sarcomere constituents (e.g. titin (*TTN*) and *MYH7*), force transduction mechanisms (e.g. desmin and dystrophin), nuclear envelope components (e.g. Lamin A/C), Ca^{2+} handling apparatus (e.g. phospholamban (*PLN*)), ion channels (e.g. sodium channel type 5 subunit alpha) etc. (Hershberger et al., 2013). Of these, *TTN* is the most common contributor (25%). While the other disease loci are associated with missense mutations, most of the pathogenic *TTN* mutations cause truncation of the protein. However, how these trigger the DCM phenotype is not well understood. Most DCM are inherited in an autosomal dominant fashion, though X-linked and recessive inheritances have also been identified (Watkins et al., 2011).

1.2 Cardiomyocytes (CMs)

CMs play a pivotal role in myocardial contraction and the cardiac pump at both physiological and pathological conditions. Further understanding of how CMs transduce the electrical signal produced by pace-making nodal cells to activate the Ca^{2+} transient pathway and subsequent CM contraction, collectively known as excitation-contraction (E-C) coupling, would aid in the dissection of the patho-mechanism of cardiomyopathy.

1.2.1 Typical CM structure

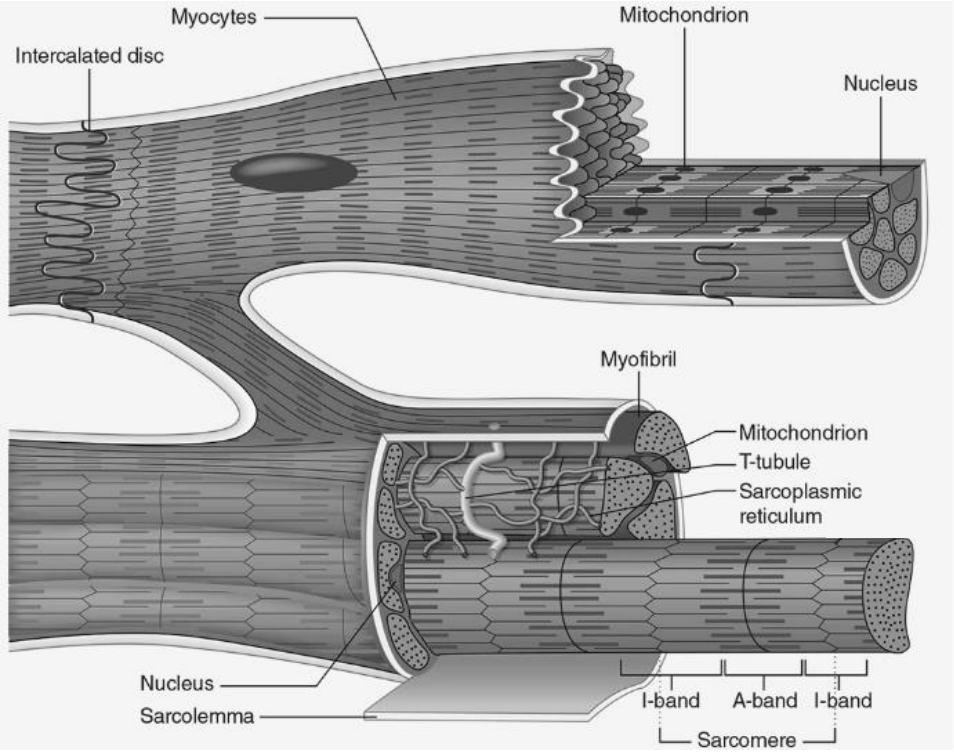


Figure 1-6 Microscopic anatomy of CMs. Taken from (Smith and Fernhall, 2022).

Typical human adult CMs possess a branching, cylindrical morphology, each connecting via an intercalated disc (Figure 1-6) (Smith and Fernhall, 2022). As highlighted in 1.1.3, intercalated discs are important for synchronisation of myocardial contraction. CMs develop a specialised cell membrane structure called the transvers tubule (T-tubule), which invaginates into the cell interior and connects with the sarcoplasmic reticulum (SR). Signalling structures mediating sarcomere activation are concentrated on T-tubules. Under a light microscope, sarcomere structure can be visualised as a striation of alternating isotropic bands (I-bands) and anisotropic bands (A-bands). A series of repeating sarcomere units is known as a myofibril.

CMs can be categorised into 3 major subtypes according to their position and function within the heart. Atrial and ventricular CMs perform contractile activity upon electrical stimulation for the atrial and ventricular chambers, respectively. Compared to ventricular CMs, atrial CMs are smaller in cell area with a more spindle-like morphology (Figure 1-7) (Manfra et al., 2017).

Nodal CMs are pace-making cells that form the sinoatrial and atrioventricular nodes and thus function in the spontaneous generation of stimulatory impulses and regulation of heart rhythm. Nodal cell size is smaller than atrial CMs with an elongated, crescent-like morphology and a few, sporadic T-tubules (Petkova and Dobrzynski, 2021) (Figure 1-7).

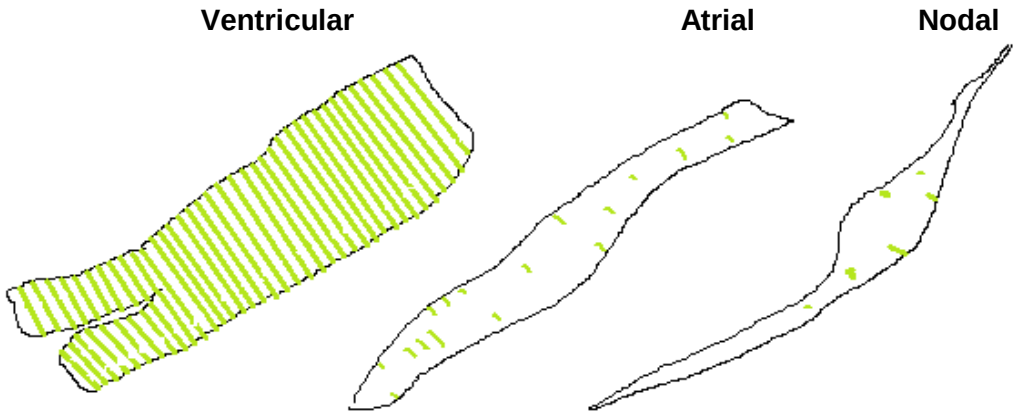


Figure 1-7 Simplified illustration of morphology of CM subtypes. Green colour represents T-tubule distribution.

1.2.2 Cardiac action potential

Coordination of CM's contraction is achieved by the transmission of the electrical action potential. The action potential is defined as a rapid oscillation of CM's membrane potential, the difference in electrical potential between the cell interior and the extracellular space.

1.2.2.1 Atrial and ventricular-type action potential

Typical action potential morphology for contractile CMs can be divided into 5 distinct phases (Figure 1-8) (Jost, 2009). In an unstimulated condition, CMs are maintained at the stable resting membrane potential (~-85 to -90 mV) (Kussauer et al., 2019, Grant, 2009) (Phase 4). The resting membrane potential is determined by a gradient of key ions such as Na^+, K^+ and Ca^{2+} (Table 1-1), which are largely membrane-impermeable (Klabunde, 2011).

Ion	Extracellular concentration (mM)	Intracellular concentration (mM)
Na^+	145	20
K^+	4	150
Ca^{2+}	2.5	0.0001

Table 1-1 Approximate ion concentration across the sarcolemma at the resting stage. (Klabunde, 2011)

Upon electrical activation (Phase 0), fast-type Na^+ channels, are opened, allowing a rapid Na^+ influx and an increase of the membrane potential beyond the threshold (~-70 mV) (Jost, 2009, Ikonnikov et al., 2013). The membrane potential continues depolarising till reaching slightly above 0 mV, followed by a closure of the Na^+ channels. Some of these Na^+ ions can move to nearby CMs via gap junctions, thereby triggering depolarisation of the connecting cells. Long lasting (L-type) Ca^{2+} channels (LTCCs) are also opened at a membrane potential above -40 mV, leading to a small Ca^{2+} influx. These Ca^{2+} ions are important for the E-C coupling mechanism.

In Phase 1, transient outward K^+ channels are briefly opened, leading to a quick and small K^+ efflux. Here, the membrane potential is slightly repolarised (~0 mV) (Ikonnikov et al., 2013).

The LTCCs remain open during Phase 2. The Ca^{2+} efflux is counter-balanced by the activity of both slow, delayed rectifying K^+ channels and rapid, delayed rectifying K^+ channels, resulting in a plateau of the membrane potential slightly below 0 mV (Ikonnikov et al., 2013).

In Phase 3, the LTCCs gradually close while both the slow and rapid, delayed rectifying K^+ channels are still active, depressing the membrane potential. At the late stage of this phase, additional K^+ channels are activated and further manipulate the K^+ efflux, resulting in the regeneration of the resting membrane potential. In addition, the normal ion concentration gradients (Table 1-1) are restored via the activity of Na^+/Ca^{2+} exchanger (NCX), ATP-dependent Ca^{2+} pump and ATP-dependent Na^+/K^+ pump (Jost, 2009, Ikonnikov et al., 2013).

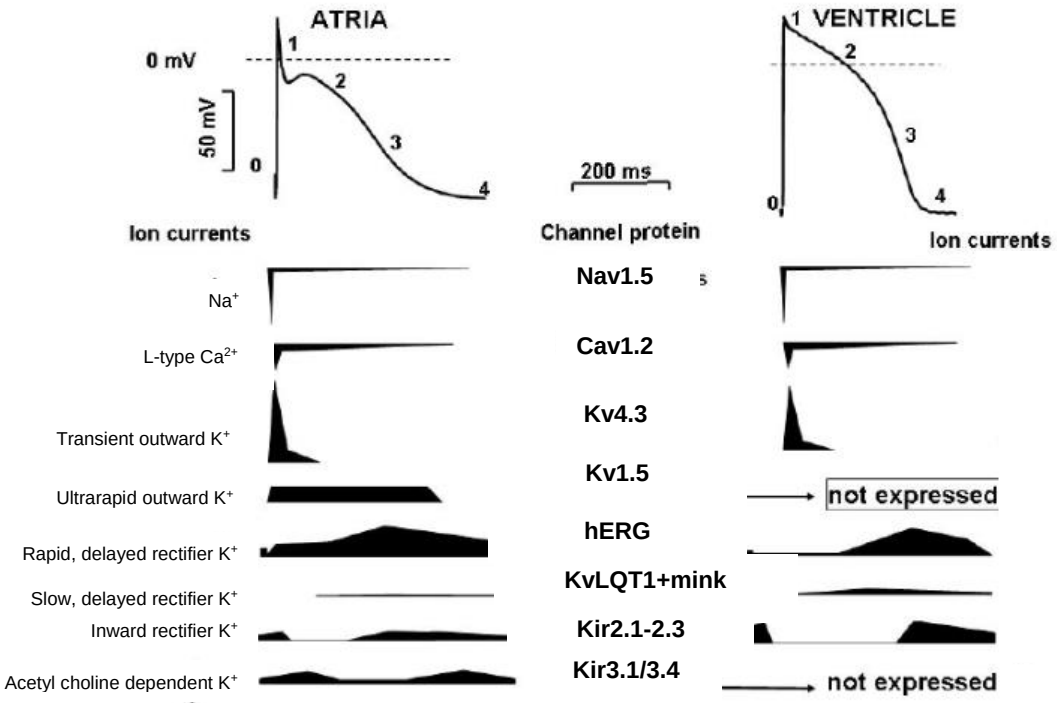


Figure 1-8 Atrial and ventricular-type action potentials and their underlying ion currents. The numbers refer to phases of the action potential. For each ion current diagram, the horizontal line indicates the 0 current level where the area above the line shows an outward current and the area below the line shows an inward current. Nav1.5; voltage-gated sodium channel 1.5; Cav1.2, voltage-gated calcium channel 1.2; Kv4.3, voltage-gated potassium channel 4.3; Kv1.5, voltage-gated potassium channel 1.5; hERG, human ether-a-go-go related gene; KvLQT1, long QT 1 syndrome-associated voltage-gated potassium channel; mink, minimal potassium subunit; Kir2.1-2.3, inwardly rectifying potassium channels 2.1/2.2/2.3; Kir3.1/3.4, inwardly rectifying potassium channels 3.1/ 3.4. Modified from (Jost, 2009).

All the ion channels presented in ventricular CMs are also identified in atrial CMs (Figure 1-8). However, atrial CMs express 3 additional channels: voltage-gated K^+ channel 1.5, inwardly rectifying K^+ channels 3.1 and inwardly rectifying K^+ channel 3.4. These channels enhance the membrane repolarisation, leading to the less prominent Phase 2 and decreased action potential duration.

1.2.2.2 Nodal-type action potential

Unlike the contractile subtypes, nodal CMs display no stable resting membrane potential but spontaneously produce action potentials (Figure 1-9) (Klabunde, 2011). Nodal-type action potential has only 3 phases, lacking Phase 1 and 2.

Phase 4 is spontaneous depolarisation. At approximately -60 mV, hyperpolarisation activated cyclic nucleotide-gated potassium channel 4 (HCN4) is activated and mediates a Na^+ influx, slowly increasing the membrane potential. At approximately -50 mV, transient (T-type) Ca^{2+} channels are active and mediate a transient Ca^{2+} influx. The Na^+ and Ca^{2+} influxes together drive the membrane potential toward the threshold level (~-40 to -30 mV).

Phase 0 is mainly regulated by Ca^{2+} movement. At approximately -40 mV, the LTCCs become active, continuously increasing the membrane potential until approaching slightly above 0 mV. Because nodal CMs express no fast-type Na^+ channels, the depolarisation is developed at a slower pace when compared to atrial and ventricular CMs.

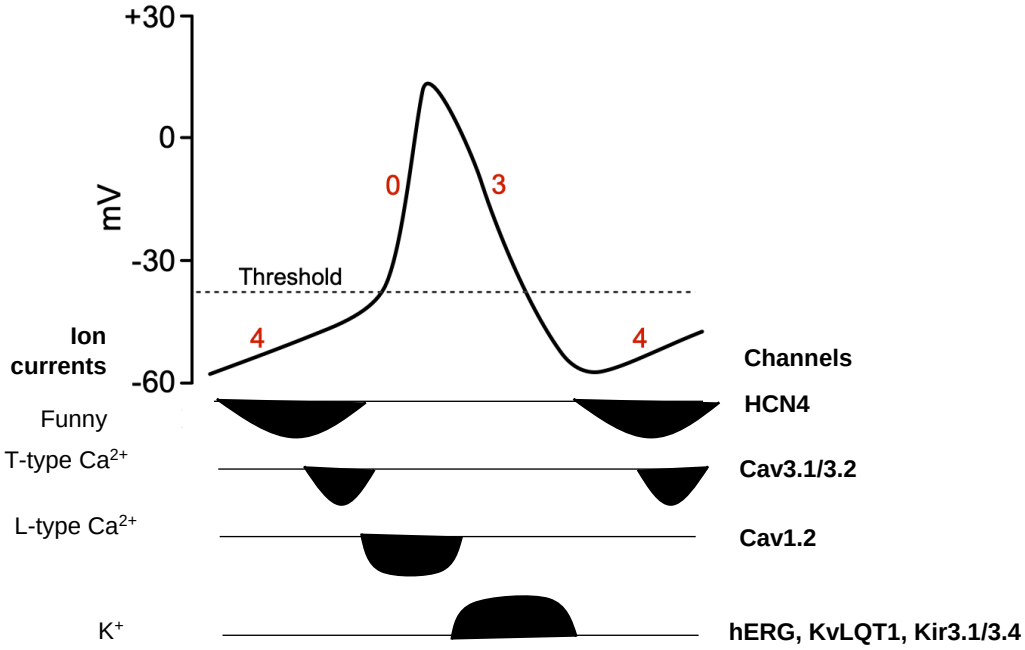


Figure 1-9 Nodal-type action potential and its underlying ion currents. The numbers refer to phases of the action potential. For each ion current diagram, the horizontal line indicates the 0 current level where the area above the line shows an outward current and the area below the line shows an inward current. HCN4, hyperpolarisation activated cyclic nucleotide-gated potassium channel 4; Cav3.1/3.2, voltage-gated calcium channel 3.1/3.2. Cav1.2, voltage-gated calcium channel 1.2; hERG, human ether-a-go-go related gene; KvLQT1, long QT 1 syndrome -associated voltage-gated potassium channel; Kir3.1/3.4, inwardly rectifying potassium channels 3.1/ 3.4. Modified from (Klabunde, 2011).

In phase 3, K^+ channels are active, causing membrane repolarisation. As a consequence, the membrane potential reaches -60 mV, where it is able to activate HCN4 for the subsequent

spontaneous depolarisation. The mechanism by which HCN4 regularly and rhythmically triggers cycles of action potential generation is known as the membrane clock.

1.2.3 Ca^{2+} signalling in CMs

LTCC activation increases the intracellular Ca^{2+} concentration level. However, it is not sufficient to effectively initiate sarcomere contraction. Rather, this initial Ca^{2+} pool enables additional Ca^{2+} release from the SR, via ryanodine receptors (RYRs) (Figure 1-10). This mechanism is generally known as Ca^{2+}-induced Ca^{2+}-release. To reduce the cytosolic Ca^{2+} concentration during diastole, Ca^{2+} ions inside CMs are sequestered into the SR via sarco/endoplasmic reticulum Ca^{2+}-ATPase (SERCA). Additional Ca^{2+} ions are pumped out to the extracellular space by NCX. The temporal rise and fall of the intracellular Ca^{2+} level is called a Ca^{2+} transient (Figure 1-11).

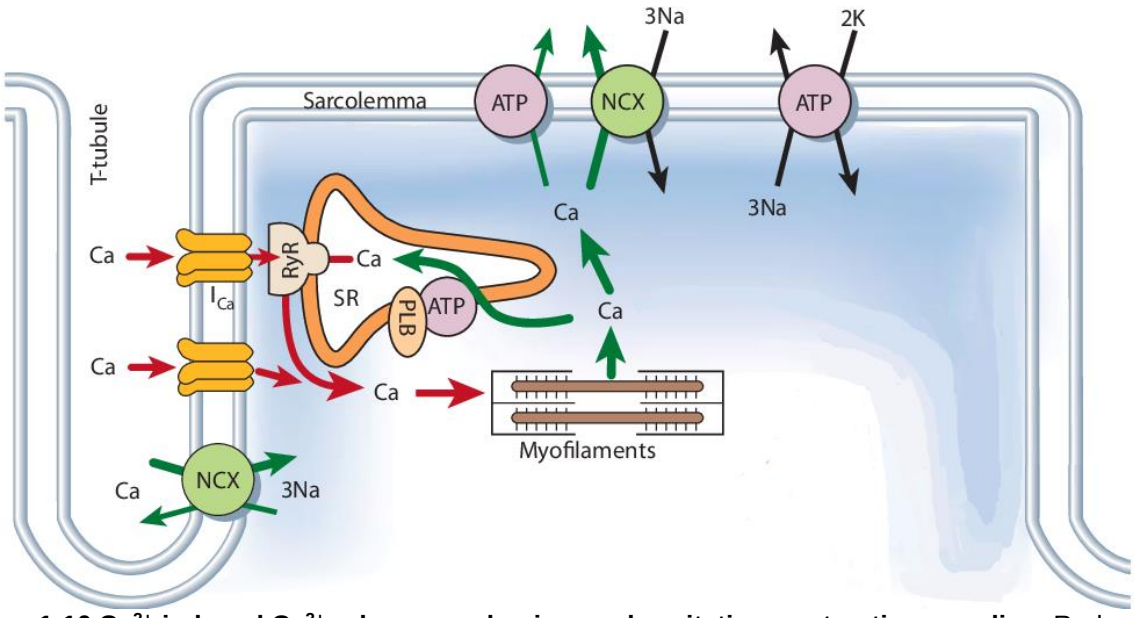


Figure 1-10 Ca^{2+}-induced Ca^{2+}-release mechanism and excitation-contraction coupling. Red and green arrows indicate pathways regulating intracellular Ca^{2+} increase and decrease, respectively. I_{Ca}, Ca^{2+} current component; NCX, Na^+/Ca^{2+} exchanger; RyR, ryanodine receptor; SR, sarco/endoplasmic reticulum; PLB, phospholamban. Modified from (Bers, 2002).

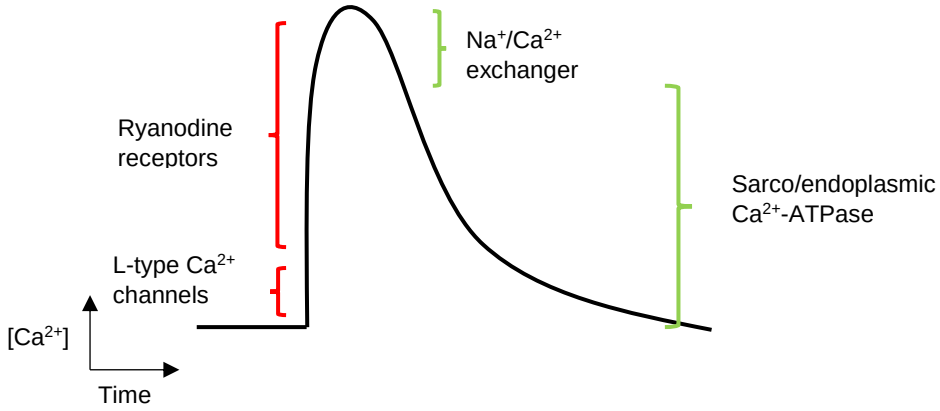


Figure 1-11 Simplified illustration of a Ca^{2+} transient profile. Red, Ca^{2+} rise; green, Ca^{2+} fall.

1.2.3.1 Atrial and ventricular-type Ca^{2+} transients

Atrial and ventricular CMs share similarity in most of the Ca^{2+} handling components. However, the *NCX* expression is lower in atrial CMs (Wang et al., 1996) and thus, the Ca^{2+} reuptake depends primarily on SERCA. The SERCA activity is modulated by PLN and sarcolipin (SLN), the latter being selectively expressed in the atrium (Ng et al., 2010). Binding with either PLN or SLN reduces the apparent Ca^{2+} affinity and transport velocity of SERCA (Asahi et al., 2002). PLN is less expressed in atrial CMs and coupled to higher SERCA levels (e.g. 4-5 times lower PLN:SERCA ratio in the rat and mouse atrium compared to the ventricle (Koss et al., 1998)). Overall, the Ca^{2+} decay kinetic is faster and Ca^{2+} transient duration is shorter in atrial CMs (Figure 1-12) (Walklate et al., 2021).

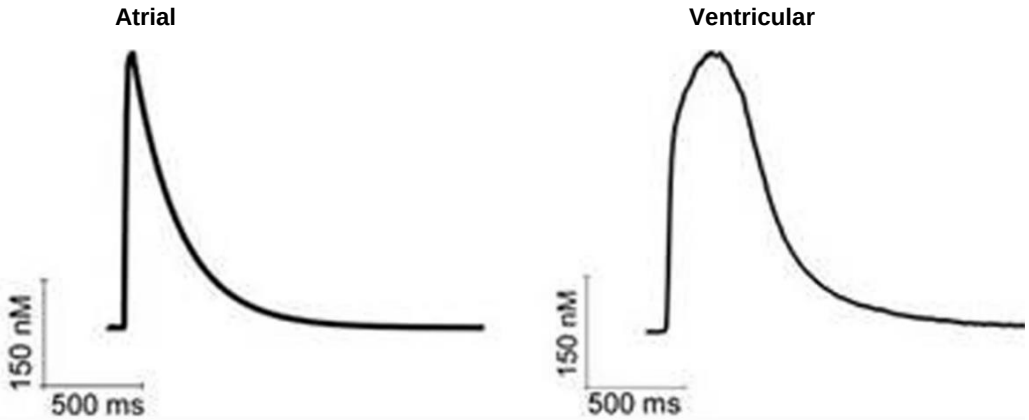


Figure 1-12 Atrial and ventricular Ca^{2+} transient profile. Adapted from (Walklate et al., 2021).

The presence of T-tubules facilitates the E-C coupling via concentrating associated ion channels, especially LTCCs, in close contact to the SR (Figure 1-10). In ventricular CMs, this structural network runs extensively at the cell periphery and deeply into the cell interior, allowing a global/homogenous Ca^{2+} transient activation (Bootman et al., 2006). By contrast, atrial CMs possess a less extensive T-tubule structure, which initiates Ca^{2+} sparks, i.e. rapid increase of intracellular Ca^{2+} level, only at the cell periphery followed by a propagation of the signal into the cell interior.

1.2.3.2 Nodal-type Ca^{2+} transient

Similar to the action potential characteristics, nodal CMs display a spontaneous Ca^{2+} transient feature. In addition to the extrinsic, Ca^{2+}-induced Ca^{2+}-release mechanism governing the intracellular Ca^{2+} rise, nodal CMs possess a local Ca^{2+} release pathway (Bogdanov et al., 2006). The RYR function is negatively regulated by triadin (Singh et al., 2013). Interacting with the Ca^{2+}-sensitive calsequestrin, triadin inhibits the RYR's Ca^{2+} transport activity (Figure 1-13A). Following the SR refilling by SERCA, the increased luminal Ca^{2+} concentration causes detachment of calsequestrin from triadin, thereby relieving its inhibitory effect and allowing a spontaneous Ca^{2+} release (Figure 1-13B). The localised Ca^{2+} release occurs concomitantly with the Ca^{2+} influx associated with the T-type Ca^{2+} channels and exerts a stimulatory effect on NCX, increasing the Na^{+} influx during the late stage of diastolic depolarisation (Bogdanov et al., 2006). Once the luminal Ca^{2+} level is depleted, especially by the Ca^{2+}-induced Ca^{2+}-release mechanism, calsequestrin is able to bind to triadin, preventing Ca^{2+} leaking from the SR. This intrinsic Ca^{2+} regulation is known as the calcium clock. The membrane clock and calcium clock are interdependent and their coupled mechanism contributes to "automaticity" of nodal CMs (Lakatta et al., 2010).

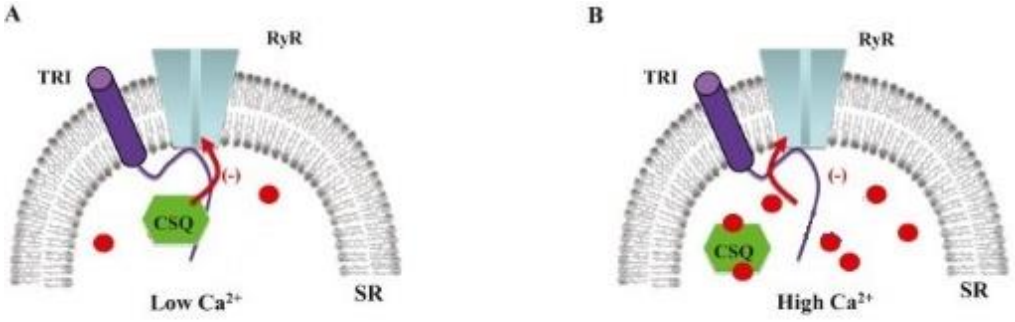


Figure 1-13 Local Ca^{2+} release mechanism of the SR of nodal cardiomyocytes. A) At low concentration of luminal Ca^{2+}, triadin (TRI)/calsequestrin (CSQ) complex blocks Ca^{2+} release by ryanodine receptor (RyR). B) Increased luminal Ca^{2+} level relieves the TRI/CSQ interaction, allowing spontaneous Ca^{2+} release. Modified from (Singh et al., 2013).

Due to the additional T-type Ca^{2+} channels and local Ca^{2+} release contributions, the kinetics of Ca^{2+} upstroke is slow at its early stage in nodal CMs as compared to atrial and ventricular CMs (Figure 1-14) (Lakatta et al., 2010).

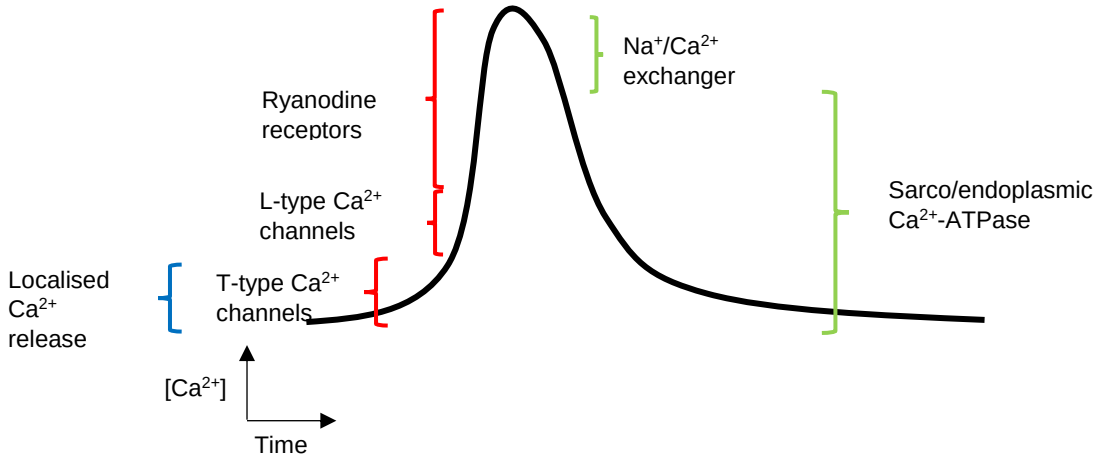


Figure 1-14 Simplified illustration of nodal-type Ca^{2+} transient profile. Red Ca^{2+} rise via Ca^{2+}-induced Ca^{2+}-release mechanism; blue, Ca^{2+} rise via local Ca^{2+} release mechanism; green, Ca^{2+} fall.

1.2.4 Regulation of sarcomere contraction

The Ca^{2+} signalling regulated by the mechanism described in 1.2.2.1 is required for sarcomere contraction. The sarcomere is composed primarily of interlocking thick, thin and titin filaments, projecting from the zwischenscheibe (Z)-disc or Z-line and the middle (M)-line (Figure1-15) (Bootman et al., 2006). The thin filament is primarily comprised of filamentous actin (F-actin), which provides binding sites for the contractile motor protein myosin that makes up the thick

filament. It is the interaction between actin and myosin (actomyosin cross-bridge) that generates force during the sarcomere contraction.

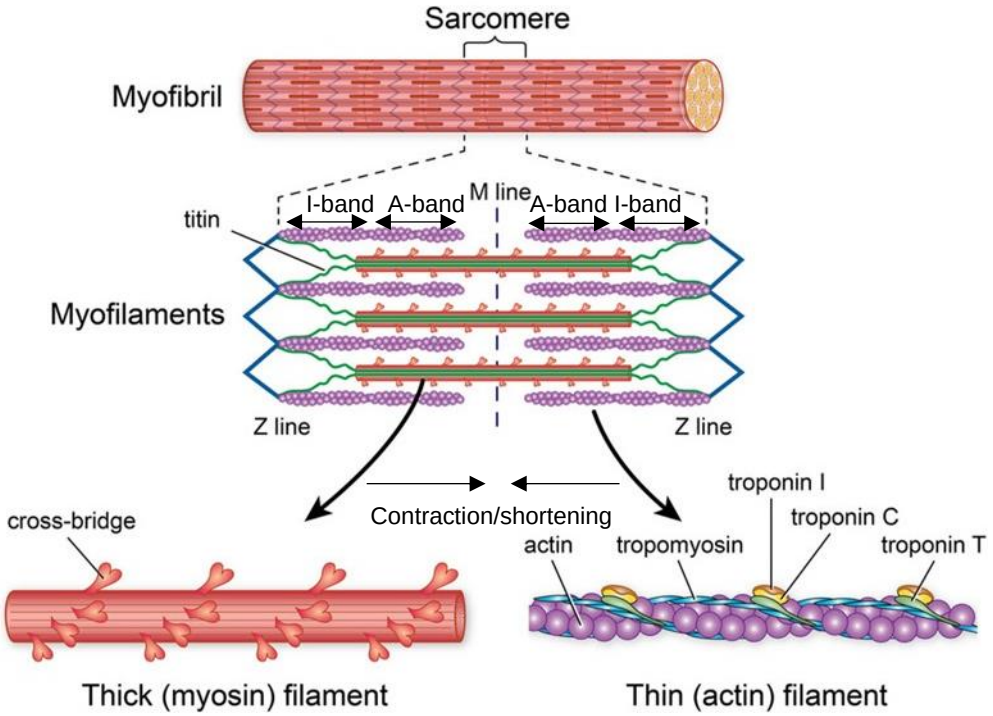


Figure 1-15 Schematic illustration of sarcomere structure. A single sarcomere can be divided into 4 main regions: Z-disc/line, I-band, A-band and M-line. The movement of sarcomere from the Z-disc toward the M-line is defined as sarcomere contraction and vice versa for sarcomere relaxation. The sarcomere contraction and relaxation are regulated by the molecular interaction between thick and thin filaments. Modified from (Bootman et al., 2006).

1.2.4.1 Thin filament

Thin filament is composed of actin, tropomyosin and troponin complex (troponin I, C and T) (Figure 1-15 & 1-16) in a 7:1:1 ratio (Shave et al., 2010). Troponin T stabilises troponin/tropomyosin/F-actin complex, where troponin I precludes the actin's myosin binding site in the absence of Ca^{2+} i.e. during diastole. Upon Ca^{2+} binding to troponin C, the inhibitory effect of troponin I is relieved, allowing tropomyosin conformational change, actomytosin cross-bridging and subsequent sarcomere contraction (Figure 1-16).

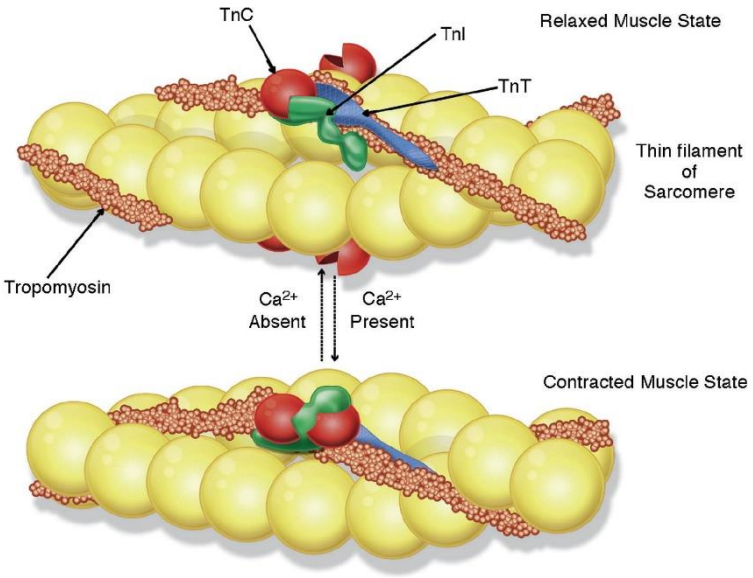


Figure 1-16 Schematic of thin filament configuration during diastole and systole. TnC, troponin C; TnI, troponin I; TnT, troponin T. Each yellow sphere represents an actin monomer. Taken from (Shave et al., 2010).

1.2.4.2 Thick filament

Cardiac thick filament is composed of myosin II proteins stabilised by MYBPC3. The myosin II consists of 2 heavy chains, 2 regulatory light chains and 2 essential light chains (Figure 1-17) (Sheikh et al., 2015). Each heavy chain contains a catalytic head (or sub-fragment 1), a lever arm, a sub-fragment 2 domain and a coiled-coil tail domain. The sub-fragment 1 is an essential domain that forms a cross-bridge with actin. The lever arm, where myosin light chains localise, is associated with the myosin's power stroke mechanism during sarcomere contraction. Myosin light chains stabilise the myosin heavy chain's lever arm and plays roles in actomyosin cross-bridge kinetics and force development required for physiological needs (England and Loughna, 2013).

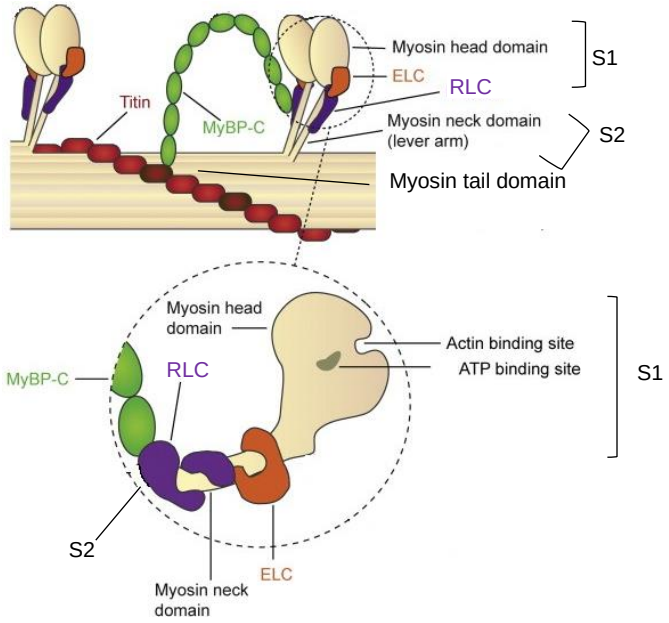


Figure 1-17 Schematic illustration of thin filament structure. Thin filament composition includes myosin and myosin binding protein C (MyBP-C), both of which are stabilised within the sarcomere by titin. Myosin consists of two heavy chains, each of which interacts with one essential light chain (ELC) and one regulatory light chain (RLC) at its lever arm region. The myosin's head domain (S1) harbours sites for ATP and actin binding. MyBP-C controls the myosin's mobility during actomyosin cross-bridging. S1, sub-fragment 1; S2, sub-fragment 2. Modified from (Sheikh et al., 2015).

1.2.4.2.1 Actomyosin cross-bridge cycling

The actomyosin cross-bridge cycle can be divided into 4 major states (Figure 1-18). Firstly, at the resting state, the sub-fragment 1's catalytic core is occupied by ATP (Figure 1-18A). Secondly, ATP hydrolysis allows the sub-fragment1-ADP-Pi complex to approach the thin filament (Figure 1-18B). Thirdly, upon the Ca^{2+}-induced troponin configuration change, an actomyosin cross-bridge develops (Figure 1-18C). Fourthly, upon the release of ADP-Pi, the myosin undergoes a power stroke i.e. sliding of the actin filament toward the M-line, resulting in sarcomere contraction (Figure 1-18D).

1.2.4.2.2 Atrial and ventricular myosin

In the human heart, myosin heavy chains exist in 2 isoforms including myosin heavy chain 6 (MYH6) and MYH7, which are predominantly expressed in the atria and ventricles, respectively (Ng et al., 2010). MYH6 displays ~3 times faster cross-bridge cycling kinetics

compared to MYH7, which partly explains the faster contraction kinetics in the atria (Walklate et al., 2021).

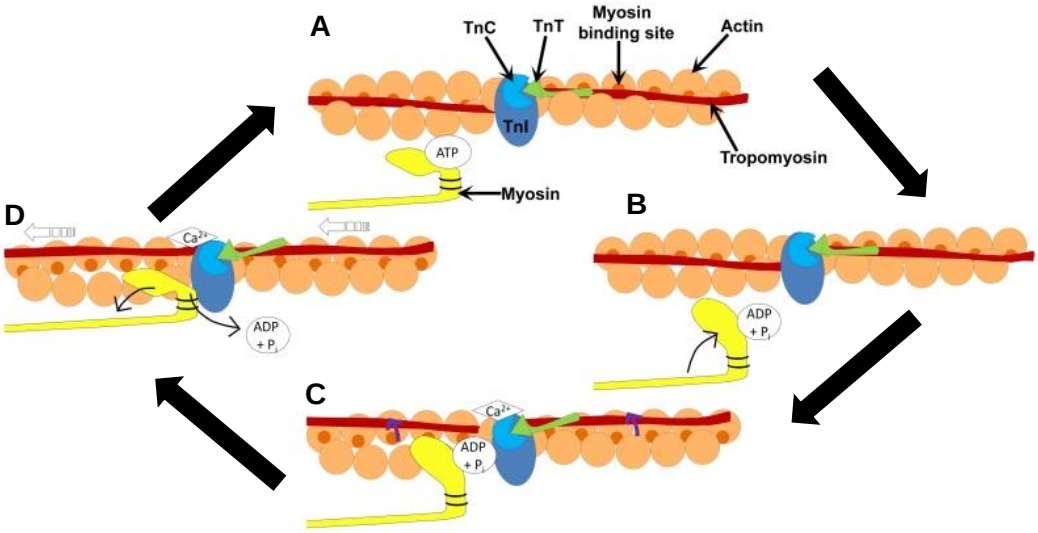


Figure 1-18 Actomyosin cross-bridge cycle. A) When bound with ATP, the myosin filament is structurally unable to approach the F-actin. B) Upon ATP hydrolysis, the myosin moves toward the F-actin. C) The presence of Ca^{2+} causes a conformational change of the troponin/tropomyosin/actin complex, permitting actomyosin cross-bridging. D) Release of the ADP:Pi enables an acto-myosin sliding. TnC, troponin C; TnT, troponin T; TnI, troponin I. Modified from (England and Loughna, 2013).

Each myosin heavy chain isoform has its own myosin light chain partners. Atrial-specific MYH6 is coupled with myosin (essential) light chain 4 and myosin (regulatory) light chain 7 (MYL7). Meanwhile, ventricular-specific MYH7 interacts with myosin (essential) light chain 3 and myosin (regulatory) light chain 2 (MYL2). MYL2 is classified as a slow isoform according to its effect on cross-bridge prolongation, which allows a sustained actin activation and enhanced force production (Sheikh et al., 2015). This property further explains the slow contraction kinetics in the ventricles.

1.2.4.2.3 Cardiac myosin binding protein C

The A-band localisation of MYBPC3 is mediated by its C-terminus interaction with myosin and titin (Figure 1-17). Meanwhile, interacting with the myosin's sub-fragment 2 (Gruen and Gautel, 1999), the N-terminus functions as a molecular brake for actomyosin cross-bridge by modulating the myosin's ATPase activity (Yang et al., 2001). Phosphorylation of the C-

terminus by β-adrenergic stimulation removes its inhibitory effect, resulting in hypercontractility (Yang et al., 2001).

1.2.4.3 Titin: a gigantic and complex muscle protein

Titin is another component of the sarcomere. It is the largest protein known to the human body (with a total 364 exons, 363 of which are included in the *TTN* meta-transcript devoid of exon 48), whose molecular weight of the full-length isoforms are in the range of 2.9-3.3 MDa (in cardiac muscles) or 3.7 MDa (in skeletal muscles) (Freiburg et al., 2000). The *TTN* meta-transcript represents an inferred complete model that includes all possible in-frame coding exons (Bang et al., 2001). In CMs, a single titin molecule constitutes half of the sarcomere, spanning from the Z-disc to the M-line. Titin can be divided into 4 regions (Figure1-19). The N-terminal titin anchors titin to the Z-line of the sarcomere. The extensible I-band provides an elasticity for titin and the sarcomere. This domain consists of 3 different spring-like elements: tandem immunoglobulin-like, proline-glutamic acid-valine-lysine-rich and N2B elements. The inextensible A-band contains immunoglobulin-like motifs alternating with fibronectin motifs and serves as a stable platform for myosin attachment during CM contraction. The C-terminal titin harbours a serine/threonine kinase domain and stabilises myosin at the M-line/band of the sarcomere.

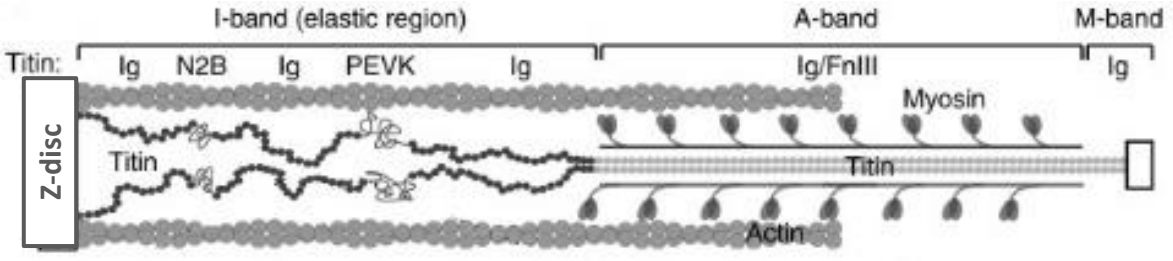


Figure 1-19 Simplified depiction of a single titin protein organisation relative to the sarcomere regions. A full-length titin spans half of the sarcomere from the Z-disc to the M-band. Its structure comprises of 4 domains: i) N-terminal Z-disc-binding domain, ii) extensible I-band domain containing elastic immunoglobulin-like (Ig), proline-glutamic acide-valine-lysine (PEVK)-rich and N2B elements, iii) inextensible A-band domain containing Ig/Fibronectin III (FnIII) elements anchoring titin with myosin and iv) M-band/M-line-binding domain. Taken from (Giganti et al., 2018).

1.2.4.3.1 The diversity of titin isoforms

The diversity of titin isoforms, which differ in length, is primarily associated with I-band exon usage through alternative splicing while exons encoding the other domains are conserved (Figure 1-20). N2-B (2.9 MDa) and N2-BA (3.3 MDa) are the titin isoforms exclusively expressed in cardiac muscles whereas the N2-A isoform is generated specifically in skeletal muscles (3.4 MDa in psoas muscles and 3.7 MDa in soleus muscles) (Freiburg et al., 2000). The length of titin determines the passive tension of the sarcomere. As the short cardiac N2-B has fewer elastic elements, it holds a greater passive stiffness than the longer N2-BA isoform (Freiburg et al., 2000, Trombitás et al., 2000). While healthy adult CMs predominantly express the N2-B isoform, an upregulation of the compliant N2-BA is frequently observed in the compromised heart (Makarenko et al., 2004, Nagueh et al., 2004, Neagoe et al., 2002). RNA binding motif protein 20 (RBM20) is an RNA splicing factor expressed in cardiac and skeletal muscles. It is responsible for a diversity of isoforms of cardiac transcripts including *TTN*, where its mutations are linked with cardiomyopathy. Knocking out *Rbm20* induced the DCM phenotype with an increased *Ttn* N2-BA:N2-B ratio in a rat model (Guo et al., 2012).

The splicing mechanism also produces the less abundant *TTN* isoforms Novex 1,2 and 3 which incorporate unique exons novex1, novex2 and novex3, respectively, with an unclear function. The Novex1 and 2 (both 3 MDa) are very similar in exon composition to the N2-B isoform while the Novex 3 (~700 kDa) has an alternative C-terminus which lacks the majority of the A-band and M-line domains (Bang et al., 2001).

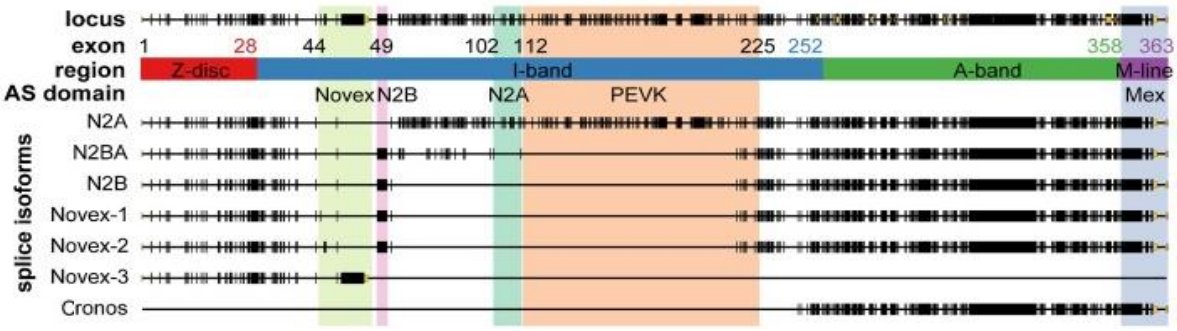


Figure 1-20 Human *TTN* isoforms. Modified from (Giganti et al., 2018). The exons presented here is numbered according to the meta-transcript excluding novex3 exon (exon48) in the total exon annotation https://www.cardiodb.org/titin/.

Apart from the RNA splicing pathway, an alternative titin isoform is generated using an internal promoter (Zou et al., 2015). This unusual promoter activity is identified just upstream of human *TTN* exon 241 (according to the total exon annotation) or within the A/I junction, and is responsible for the Cronos titin expression. Though the Cronos isoform lacks the whole Z-disc and I-band regions, its exclusive expression did not completely disrupt sarcomere formation in CMs (Zaunbrecher et al., 2019). Contractility was also preserved in this cell model although depressed contractile force was detected. This isoform was believed to play a role in the functional rescue of titin from proximal mutations that disrupt the generation of full-length titin molecules.

1.2.4.3.2 Complex titin interactome

Titin not only exhibits a diversity in the structure but also interacts with a wide range of proteins along its different regions (Figure 1-21) (Guo et al., 2018). Mutations of these binding partners are linked with cardiomyopathy.

1.2.4.3.2.1 Z-disc interactome

The titin Z-disc binds to alpha actinin, which in CMs, is encoded by alpha actinin 2 (*ACTN2*). ACTN2 crosslinks F-actin within the Z-disc of adjoining sarcomeres. The *ACTN2* Q9R mutation is associated with DCM (Mohapatra et al., 2003). This mutation abolished the ACTN2 binding ability to myogenesis-regulating muscle limb protein and decreased ACTN2 nuclear localisation in myoblasts (Mohapatra et al., 2003), leading to impaired myogenic differentiation potential.

Two adjacent titin molecules are glued together by a single titin-CAP (T-CAP) at their N-termini. T-CAP is known to interact with minimal K^+ subunit β (Furukawa et al., 2001) and fast-type Na^+ channel (Mazzone et al., 2008) on T-tubules. Hence, T-CAP may play a role in the E-C coupling (Ibrahim et al., 2013). *TCAP* homozygous missense mutations are associated with HCM (T137I and R153H) and DCM (E132Q) by unknown mechanisms (Hayashi et al., 2004).

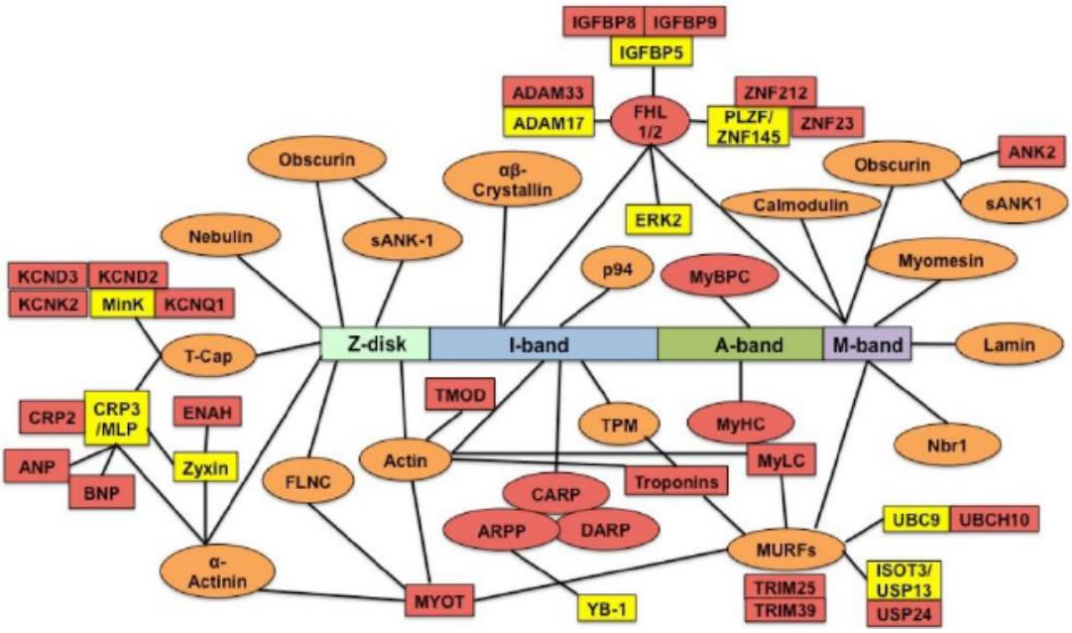


Figure 1-21 Titin binding partners. Titin is presented in the long rectangle with its different domains. Titin's direct partners are oval (both in orange and red) and indirect partners are rectangle (both in red and yellow). Red oval and rectangle highlight targets whether their expression is associated with RNA binding motif protein 20. sANK1, small ankyrin 1; p94, calpain 3; FHL1/2, four and a half LIM domain protein; MyBPC, myosin binding protein C; Nbr1, neighbour of BRCA1 gene 1; MURF, muscle ring finger protein; MyHC, myosin heavy chain; MyLC, myosin light chain; TPM, tropomyosin; CARP, cardiac ankyrin repeat protein; DARP, diabetes related ankyrin repeat protein; ARPP, ankyrin repeat domain protein 2; TMOD, tropomodulin; MYOT, myotilin; FLNC, filamin C; T-cap, titin-cap; minK, minimal potassium channel subunit ; CRP3/MLP, muscle LIM protein/cysteine-rich protein 3); CRP2, cysteine-rich protein 2); ENAH, enabled homolog; ANP, atrial natriuretic peptide; BNP, brain natriuretic peptide; KCND2/3, potassium voltage-gated channel subfamily D member 2/3; KCNK2, potassium channel subfamily K member 2; KCNQ1, potassium voltage-gated channel subfamily Q member 1; ADAM17, a disintegrin and metalloprotease 17; ADAM33, a disintegrin and metalloprotease-33 protein; IGFBP, insulin growth factor binding protein; ZNF, zinc finger protein; ERK2, extracellular signal regulated kinase 2; ANK2, ankyrin 2; YB-1, Y-box binding protein 1; ISOT-3/USP13, isopeptidase T-3; USP24, ubiquitin specific peptidase 24; UBC9, ubiquitin-conjugating enzyme 9; UBCH10, ubiquitin-conjugating enzyme H10. Taken from (Guo et al., 2018).

1.2.4.3.2.2 I-band interactome

The titin I-band is responsible for titin's passive tension, which is important for the diastolic function. The I-band region also harbours a binding site for a small heat shock protein called alpha crystallin B. This protein plays a role in a stress response (Kötter et al., 2014). Under stress, the titin's N2-B elements preferentially unfold and form aggregates, resulting in increased stiffness. Translocalisation of alpha crystallin B from the cytosol to the sarcomere suppresses this stiffening effect (Kötter et al., 2014). The R157H mutation in the gene

encoding αlpha crystallin B was suggested to impair the protein interaction with titin, leading to the DCM phenotype (Inagaki et al., 2006).

1.2.4.3.2.3 A-band interactome

As described in 1.2.4.2.3, the A-band titin provides a scaffold for myosin and MYBPC3. Also, as highlighted in 1.1.6 both missense mutations in *MYH7* and truncating mutations in *MYBPC3* are the most common genetic factor predisposing HCM.

1.2.4.3.2.4 M-line interactome

Myomesin crosslinks adjacent myosin filaments and titin at the M-line. Like titin, myomesin, encoded by the myomesin 1 gene (*MYOM1*), contains spring-like elements: immunoglobulin-like and fibronectin III domains. This protein functions as a mechanical stress sensor, in addition to titin, where it can be stretched upon myosin sheer stress introduced by contractile activity (Agarkova and Perriard, 2005). Little is known about its contribution in biology and disease. Knocking out *MYOM1* impaired sarcomere organisation, contractile function and Ca^{2+} homeostasis in a CM model (Hang et al., 2021). A missense *MYOM1* mutation (V1490I) is associated with HCM but its underlying mechanism is unknown (Siegert et al., 2011).

The titin's M-line region also contains a binding site for muscle ring finger proteins 1/2 (MURF1/2), which regulate myofibrillar protein turnover. A dysregulated MURF activity can modify the sarcomere content and the cardiac mass, which in turn compromise cardiac function. Missense mutations of MURF1 (S5L) and MURF2 (C50Y) were suggested to be mediators of HCM pathogenesis through protein degradation pathways (Su et al., 2014). Knocking out the gene encoding MURF1 in mice did not interfere with cardiac development but exerted an anti-hypertrophic effect (Willis et al., 2007).

1.2.5 β-adrenergic stimulation of contractile activity

As described in 1.1.4, epinephrine or norepinephrine targets the cardiac β-receptors in CMs. This stimulation leads to increased cellular cyclic AMP level by adenylyl cyclases, which in

turn activates protein kinase A. Protein kinase A catalyses phosphorylation of numerous targets associated with the E-C coupling mechanism (Figure 1-22) (El-Armouche and Eschenhagen, 2009). For example, phosphorylation of LTCC (Kamp and Hell, 2000), RYR (serine 2030 (Xiao et al., 2006)) and PLN (serine 16 (Chu et al., 2000)) together cause overall cellular Ca^{2+} load associated with the Ca^{2+} transient generation, resulting in increased force production during systole. Phosphorylation of troponin I (serine 23/24 (Kobayashi and Solaro, 2005)) and MYBPC3 (serine 282 (Sadayappan et al., 2005)) decrease myofilament Ca^{2+} sensitivity, promoting sarcomere relaxation. The overall effect of the β-adrenergic response is increased cardiac inotropy.

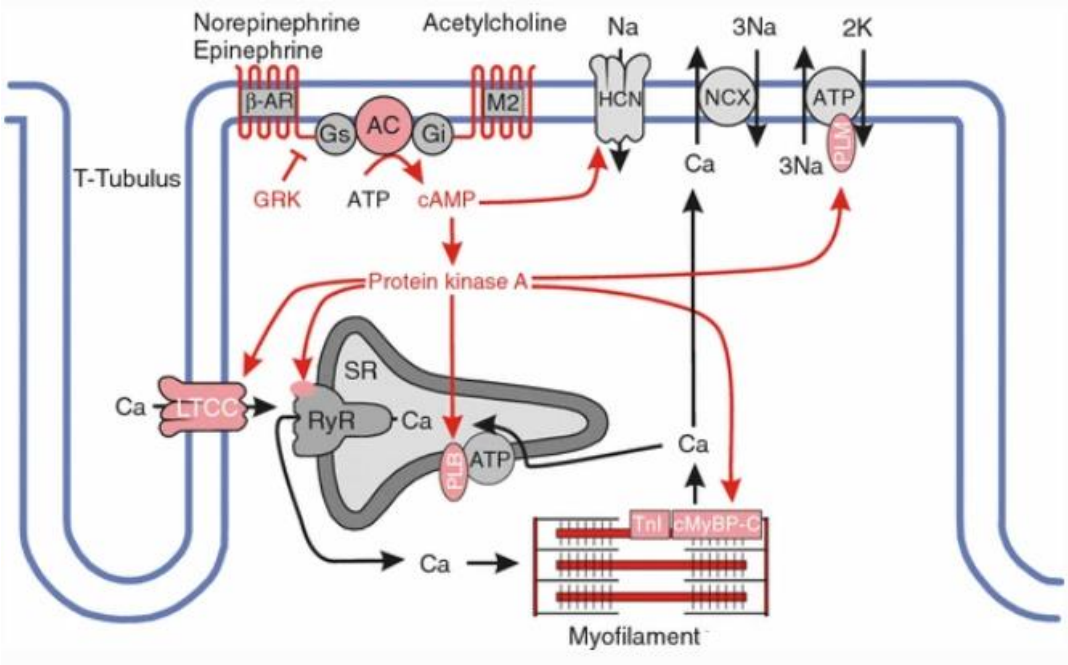


Figure 1-22 β-adrenergic response in CMs. Upon norepinephrine/epinephrine stimulation, the CM's β-receptor (β-AR) liberates its coupled stimulatory G-protein (Gs). Binding of Gs with adenylyl cyclase (AC) promotes ATP conversion into cyclic AMP (cAMP). This catalytic activity can be suppressed upon binding to inhibitory G-protein (Gi) associated with the acetylcholine-activated muscarinic receptor (M2) response. cAMP activates protein kinase A, which in turn phosphorylates several proteins modulating the E-C coupling mechanism including, L-type Ca^{2+} channel (LTCC), ryanodine receptor (RYR), phospholamban (PLB), phospholemman (PLM), cardiac myosin binding protein C (cMyBP-C) and troponin I (TnI). cAMP itself also triggers the Na^{+} transport activity by hyperpolarisation-activated cyclic nucleotide-gated channel (HCN). Overall, β-adrenergic stimulation exerts a positive inotropic effect. Taken from (El-Armouche and Eschenhagen, 2009).

1.3 Titin in DCM

Mutations in the *TTN* gene are largely associated with DCM with *TTN* truncating variants (*TTNtv*) being the most common contributor, accounting for ~25% of familial DCM (Herman et

al., 2012). However, not all *TTNtv* mutations are pathogenic since they can be found in about 1-3% of the general population (Golbus et al., 2012).

1.3.1 Influence of mutational position on *TTNtv* pathology

The discovery of asymptomatic *TTNtv* carriers has brought about efforts to elucidate how the mutational positions contribute to the DCM phenotype. As demonstrated previously, several titin isoforms are produced by alternative RNA splicing, where I-band exons are most differentially expressed. RNA sequencing suggested that the percentage of a given exon being included in the expressed *TTN* transcript repertoire (also known as percentage spliced in) is a useful metric for predicting the *TTNtv* pathogenicity (Roberts et al., 2015). *TTNtv* originating from consecutive exons (high percentage spliced in) is more likely be pathogenic whereas *TTNtv* located in exons with a low percentage spliced in index is better tolerated by e.g. transcriptional bypass. Besides, variants encoded in exons spliced in the N2-BA but not in the N2-B isoform have a weak association with the disease (Roberts et al., 2015). Thus, this notion partly explains healthy individuals carrying non-consecutive I-band *TTNtv* (Schafer et al., 2017, Hinson et al., 2015).

1.3.2 Proposed pathomechanisms

The mechanisms by which *TTNtv* induces DCM pathogenesis is still elusive, largely due to the titin's mammoth size and isoform diversity. Two molecular models are proposed based on whether truncated titin molecules can be stably transcribed and translated.

1.3.2.1 Haploinsufficiency

Happloinsufficiency describes a situation where a diploid organism has only one active allele while the other is inviable due to e.g. a loss-of-function mutation. The resultant protein insufficiency compromises the protein's normal function and the organism's fitness (Deutschbauer et al., 2005). *TTNt*v may also be pathogenic due to non-sense mediated decay (NMD) (Brogna and Wen, 2009).

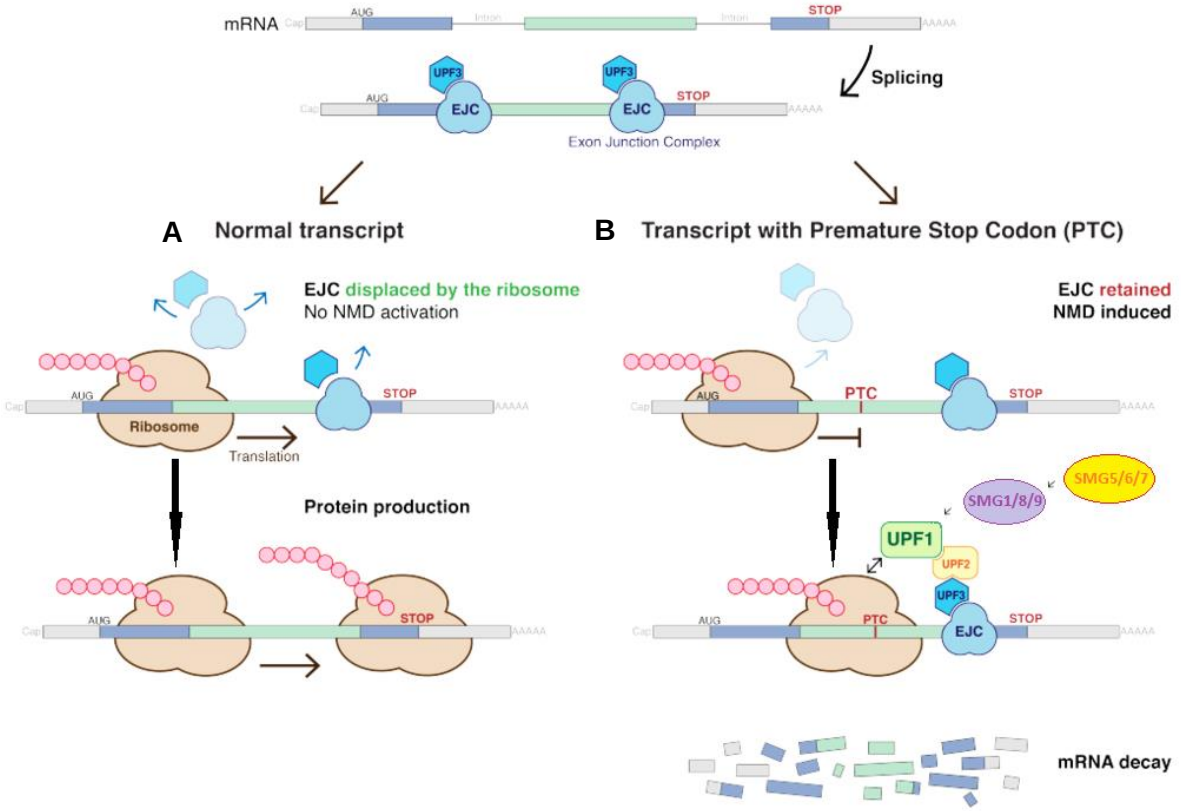


Figure 1-23 Simplified illustration of NMD mechanism. A) In normal translation, an exon-exon junction complex (EJC), which binds to a spliced mRNA, is removed by the ribosome undergoing translation toward a stop codon. B) However, a premature stop (PTC) precludes the complete translation, allowing a persistent EJC localisation. The EJC recruits upstream frameshift proteins 1/2/3 (UPF1/2/3) – at this stage, the ribosome may dissociate. The UPF complex interacts with the suppressor of morphogenetic effect on the genitalia proteins (SMG) 1/8/9 complex. Phosphorylation of UPF1 by SMG1 further recruits SMG5, SMG6 and SMG7. mRNA degradation mediated by the SMG6's endonuclease activity results in reduced transcriptional and/or protein levels from a mutant allele harbouring a PTC. Modified from Wikimedia Commons, retrieved on September 7, 2022.

It is conjectured that the presence of a premature stop codon in *TTNtv* could provide a substrate for activating the NMD machinery, which consequently subjects the mRNA to degradation, lowering the titin transcriptional and/or protein levels. In normal translation, exon-exon junction complexes bound to mRNA after splicing are displaced by the ribosome approaching a stop codon (Lykke-Andersen and Jensen, 2015) (Figure 1-23A). In the presence of a premature stop, an early termination of translation prevents the ribosome acquiring the actual stop codon and thus, the exon-exon junction complexes remain bound to the mRNA (Figure 1-23B). The exon-exon junction complex recruits upstream frameshift proteins, which further recruit a series of suppressor of morphogenetic effect on genitalia proteins. Of these, the suppressor of morphogenetic effect on genitalia protein 6 has a

nuclease activity which mediates degradation of the mutant mRNA, leading to reduced transcriptional and/or protein levels. However, in some studies, a loss of titin protein is reported independently from *TTN* allelic imbalance (Hinson et al., 2015, Fomin et al., 2021), possibly suggesting that NMD may not be the pathway underlying titin haploinsufficiency.

Reduction of titin protein levels may lead to metabolic and signalling remodelling as seen in Z-line and A-band *TTNtv*-harbouring rat models, switching metabolism from the fatty acid toward the glycolic pathway (Schafer et al., 2017), recapitulating heart failure (Stanley et al., 2005, Neubauer, 2007).

1.3.2.2 Poison peptide

This alternative model is proposed on the basis of the presence of truncated titin proteins which interferes with cardiac function. In order to support this notion, proof of stable truncated titin proteins and/or their incorporation into the sarcomere are required. Recent publications have reported stable truncated titin proteins in *TTNtv* DCM biopsies by Western blot analysis, showing their size corresponding to the respective mutational positions (McAfee et al., 2021, Fomin et al., 2021). However, how the truncated titin proteins contribute to disease pathology is still not completely understood.

(Fomin et al., 2021) demonstrated that the truncated titin protiens were unable to integrate into the sarcomere but rather presented as cytosolic inclusion bodies. This study also suggested that the accumulation of these aggregates compromised protein homeostasis by aggravating the ubiquitin-proteasome system (UPS), which is responsible for myofibrillar protein turnover. UPS dysfunction is a common characteristic in cardiomyopathies (Gilda and Gomes, 2017, Predmore et al., 2010), which potentially hinders CMs' capacity to cope with stress.

(McAfee et al., 2021) and (Fomin et al., 2021) agreed that a combination of *TTNtv*-associated-titin happloinsufficiency, independent of a transcriptional change, and stable truncated titin

proteins, possibly through dysregulated protein quality control (PQC), underlies the DCM pathogenesis.

1.4 TTNtv-associated atrial fibrillation

Atrial fibrillation is the most common arrhythmia defined by irregular and rapid heart rhythm, which is associated with heart failure, stroke, and premature death. Atrial fibrillation is classified as an electrical heart disease where the electrocardiogram displays an aberrant P wave, possibly repeatable or invisible, and an irregular, high QRS frequency (Figure 1-24). As described in 1.1.3, the P wave and the QRS complex represent an electrical propagation within the atria and the ventricles, respectively.

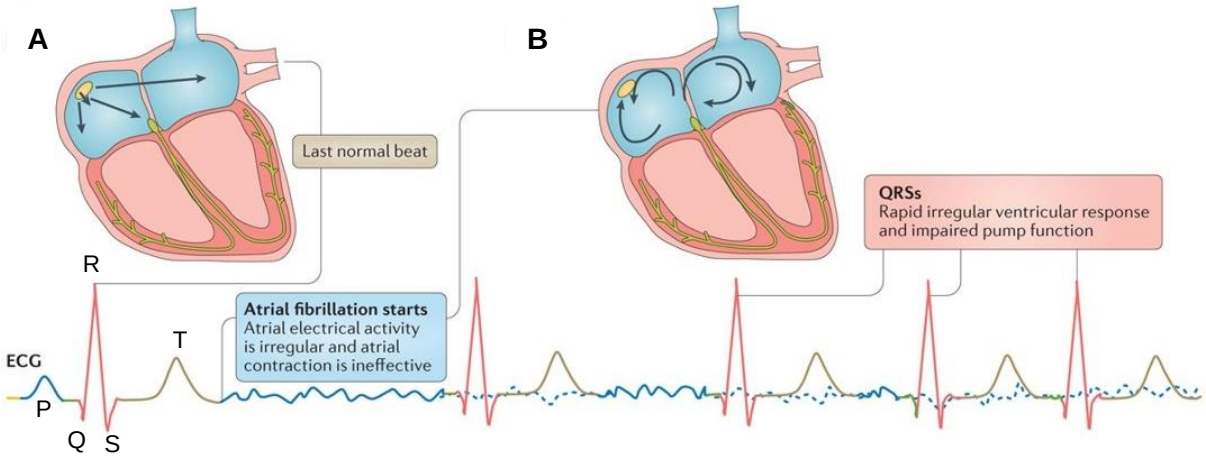


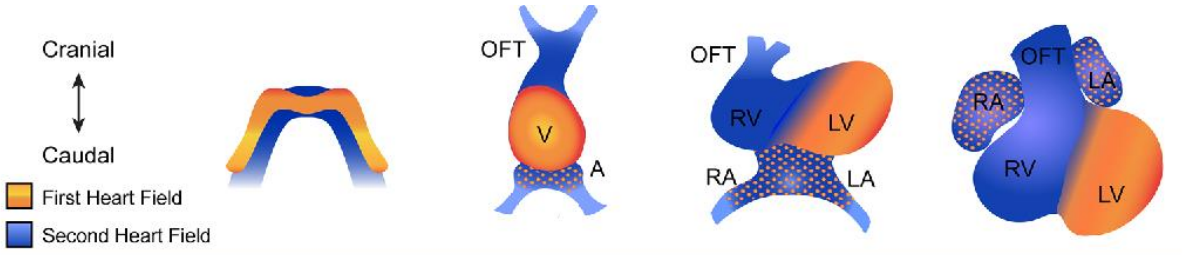
Figure 1-24 Electrocardiogram of atrial fibrillation. A) The last normal beat prior to the onset of atrial fibrillation with a normal electrocardiogram (ECG). B) Atrial fibrillation is characterised by a fast, uncoordinated atrial electrical activity in which the ECG displays an aberrant P wave (blue). The ventricles also respond to this irregularity which leads to impaired contraction (represented by an abnormal, high QRS frequency (red)). Modified from (Lip et al., 2016).

Atrial fibrillation is caused by an abnormal electrical conduction within the atria but its aetiology is not fully understood. Interestingly, genome wide association studies have identified *TTN* as being one of the most common genetic risk factors for atrial fibrillation, where the *TTN*-atrial fibrillation co-segregation was independent from a DCM background (Nielsen et al., 2018). Similarly, large-scale whole genome sequencing uncovered enrichment of rare *TTNtv* in individuals developing an early onset atrial fibrillation (Choi et al., 2018). A rare missense *TTN* variant (rs2042995) in human was found to be implicated in a P wave prolongation (Weng et

al., 2020). In agreement, a heterozygous A-band *Ttntv* zebrafish model recapitulated the atrial fibrillation's phenotype with an increased atrial fibrosis and a prolongation of the PR interval (Ahlberg et al., 2018). The PR interval correlates with the descending impulse from the atria to the ventricles. Hence, aberration of this parameter indicate impairment in the atrial electrical activity.

1.5 Induced pluripotent stem cell derived-CMs (iPSC-CMs)

Advances in pluripotent stem cells, including embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs), have provided an alternative source of cells for cardiovascular research (Gao and Pu, 2021). Human ESCs are derived from preimplantation embryos, generally donated from fertility clinics, which require ethical approval for use (Robertson, 2001). Meanwhile, human iPSCs can be directly reprogrammed from somatic cells (Takahashi and Yamanaka, 2006) and thus have gained more popularity. iPSCs can be differentiated into many cell linages including CMs by modulating key signalling pathways identified during embryonic development.



Stage	Cardiac Crescent	Linear Heart Tube	Looping Heart	Chamber Specialization/ Septation
Mouse Embryo Stage	E7.5	E8.0	E9-12	E10+
Human Embryo Stage	Day 15	Day 20	Day 28-30	Day 32+

Figure 1-25 Schematic illustration of embryonic heart development with respect to the embryonic stages. The cardiac crescent consisting of the first and second heart fields. The cells within the cardiac crescent undergo proliferation and differentiation, leading to formation of the linear heart tube. The cells from the first heart field contribute to the left ventricle (LV) and a portion of the atria (A) while the cells derived from the second heart field form the right ventricle (RV), a portion of the atria (A), inflow tract and outflow tract (OFT). The right forward looping of the linear heart tube results in the atria being positioned cranially relative to the ventricles. Maturation of the embryonic heart includes formation of the valves, the septum and the 4 chambers of the heart. Taken from (Paige et al., 2015).

1.5.1 iPSC-CM differentiation

The cardiac lineage is derived from epiblast-originated mesodermal cells during the primitive streak formation, a structure that is promoted upon a gradient of NODAL and wingless integrated (WNT) signals (Perea-Gomez et al., 2002). Eomesodermin-expressing mesoderm contributes to the cardiovascular system. Eomesodermin activates mesoderm posterior basic helix-loop-helix transcription factor 1 (MESP1), which is identified as an early cardiac mesoderm marker and plays important roles in cardiac progenitor cell specification (Bondue et al., 2008). However, the fate of $MESP1^{+}$ cells may be already determined and the cardiac progenitors of the so-called first and second heart fields (Figure 1-25) may present at or before gastrulation (Devine et al., 2014, Meilhac et al., 2004, Lescroart et al., 2014). The first heart field develops from cells that express NK2 homeobox 5 (NKX2.5*)* and T-BOX 5, giving rise to the left ventricle and a portion of the atria whereas the second heart field is a group of cells expressing NKX2.5 and ISLET1 that are mainly involved in formation of the right ventricle, atria, inflow tract and outflow tract (Evans et al., 2010).

iPSC-derived CMs (iPSC-CMs) differentiation can be achieved by modulating transforming growth factor β family proteins (including WNT3a, bone morphometric protein 4 (BMP4), activin A and NODAL) and WNT. BMP4 regulates expression of GATA binding protein 4, serum response factor and myocyte enhancer 2C, committing iPSCs into the mesodermal stage (Klaus et al., 2012). A combination of BMP4 and activin A activates expression of platelet-derived growth factor receptor alpha in the mesodermal subpopulation which are positive for kinase domain receptor (Kattman et al., 2011). These double positive cells represent the cardiac mesoderm marked by expression of MESP1, which can be differentiated further into the cardiac progenitor cells.

WNT has a biphasic role in iPSC-CMs differentiation. WNT promotes iPSC transition into the mesodermal stage. However, activation of this pathway at the slightly later stage blocks the cardiac differentiation. Inhibition of WNT in the mesodermal cells allow their transition into the cardiac mesoderm and the cardiac progenitor cells (Willems et al., 2011). Therefore, a WNT

agonist e.g. CHIR-99021 is required at the start of the differentiation followed by a WNT antagonist e.g. IWR1 or IWP-2 (Lian et al., 2012) at the subsequent stage.

1.5.2 Limitations of iPSC-CMs

In contrast to primary adult CMs, the self-renewal capacity of iPSCs enables the production of an infinite numbers of iPSC-CMs for studying cardiac biology and disease. However, iPSC-CMs possess unique characteristics which may limit their utilities. Firstly, they express embryonic-like, immature phenotypes, molecularly, structurally and functionally (Figure1-26A), which may not fully recapitulate the adult human heart. Secondly, current differentiation protocols generally yield mixed types of cells and mixed subtypes of CMs including atrial-like, ventricular-like and nodal-like cells.

1.5.2.1 Immaturity

1.5.2.1.1 Electrophysiology and Ca^{2+} signalling

iPSC-CMs possess automaticity due to the expression of HCN4 (Goversen et al., 2018). In contrast, primary ventricular and atrial CMs only contract upon acquiring electrical stimulation. iPSC-CMs have a higher resting membrane potential (~-50 to -70 mV) compared to adult CMs (~ -85 to 90 mV) (Hoekstra et al., 2012, Kussauer et al., 2019, Grant, 2009). The less negative membrane potential allows the HCN4 activity to spontaneously activate action potential generation (Figure 1-26C). Spontaneous Ca^{2+} release from the SR is also seen in iPSC-CMs (Koivumäki et al., 2018). This mechanism underlies the spontaneous Ca^{2+} transient generation and further enhances the spontaneous depolarisation via the activated NCX channel. These inherent iPSC-CMs' membrane and Ca^{2+} clocks resemble the phenotype of adult nodal cells.

In addition, iPSC-CMs display a poorly-developed SR, no (or few) t-tubules, low expression of Ca^{2+} handling proteins e.g. SERCA and RYR (Figure 1-26D), which together lead to the slow E-C coupling when compared to adult CMs (Kane et al., 2015).

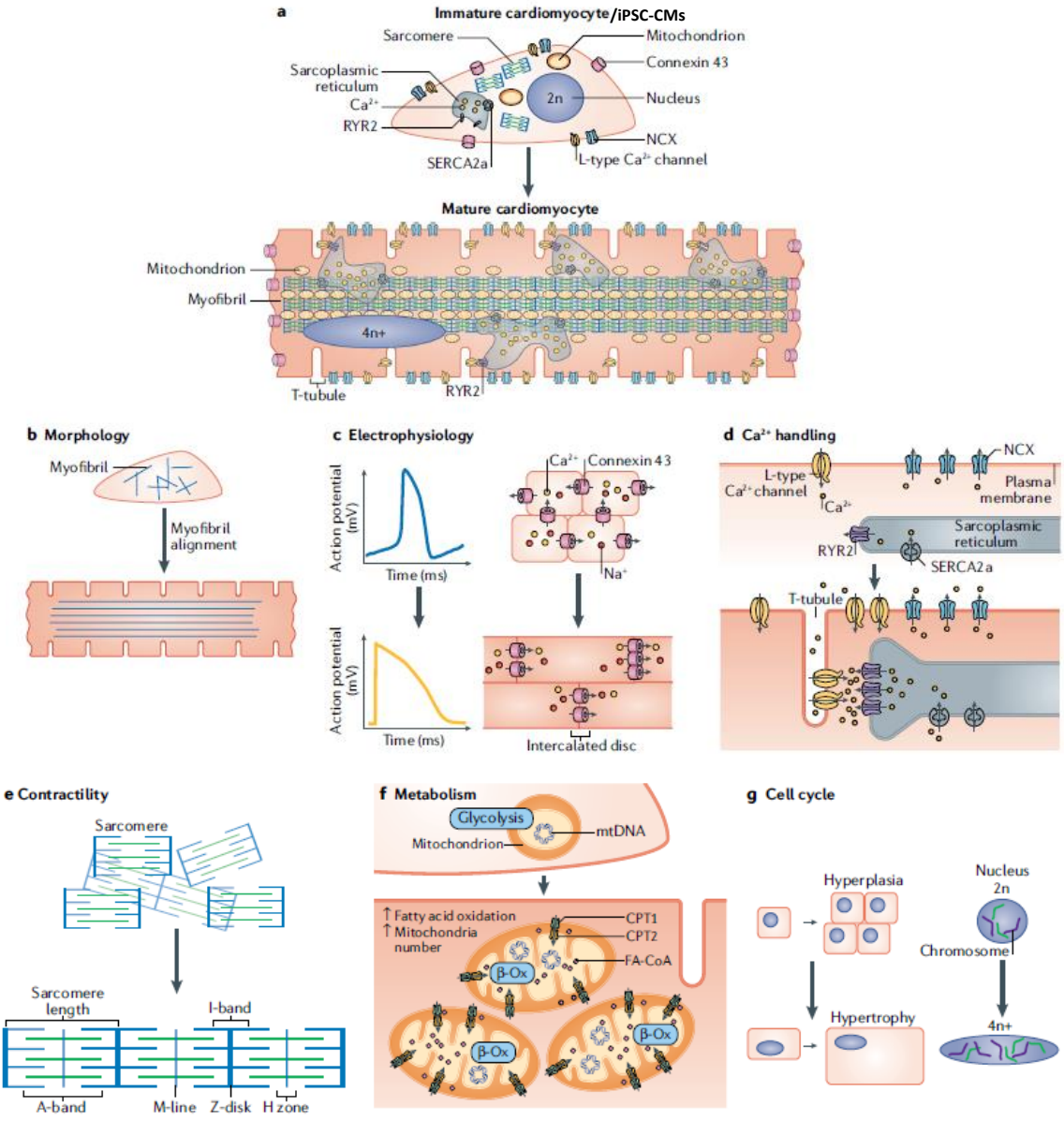


Figure 1-26 Immaturity of iPSC-CMs. A) iPSC-CMs display several immature characteristics with respect to structure and function, which are different from adult or mature CMs. Compared to mature CMs, iPSC-CMs have B) smaller cell size, poorly alinged myofilments, C) circumferential distribution of gap junction connexin 43 (this junctional protein is polarised to the intercalated disc in mature CMs), lower expression of Ca^{2+} handling proteins, no development of T-tubules, E) lower contractile force (due to the lower sarcomere content, less organised myofilament and differential myofilament protein expression), F) fewer mitochondria number, more preference on using glucose to generate energy by glycolysis (while mature CMs can additionally oxidise fatty acids by β-oxidation for ATP production), G) higher prolifertion potential (instead, mature CMs can expand their volume) and more nuclear DNA content. CPT, carnitine O-palmitoyltransferase; FA-CoA, fatty acyl-CoA ester; mtDNA, mitochondrial DNA; n, haploid content of chromosomes; NCX, sodium–calcium exchanger; RYR2, ryanodine receptor 2; SERCA2a, sarcoplasmic/endoplasmic reticulum calcium ATPase 2a. Modifiied from (Karbassi et al., 2020).

1.5.2.1.2 Metabolism

In eukaryotic cells, oxidative phosphorylation, an aerobic process of producing ATP by oxidising NADH or $FADH_2$, takes place on cristae within mitochondria. CMs are capable of

using several classes of energy substrates including carbohydrates (by glycolysis) and lipids (by β oxidation) to generate NADH/$FADH_2$. iPSC-CMs have a limited number of mitochondria with less developed cristae and thus, have a low oxidative capacity (Figure 1-26F) (Horikoshi et al., 2019, Feyen et al., 2020). Approaching the mature stage, CMs undergo metabolic reprogramming. Immature iPSC-CMs predominantly use glucose for energy production whereas mature CMs generate ATP via oxidation of fatty acid and glucose (Yang et al., 2019, Feyen et al., 2020).

1.5.2.1.3 Structure and morphology

iPSC-CMs exhibit various immature morphometric features; they are small and rather rounded (Figure 1-26B) with mononucleation (Figure 1-26C) and a circumferential gap junction arrangement (Figure 1-26E). Meanwhile, adult CMs have a higher frequency of binucleation and a cylindrical morphology presenting well aligned gap junctions within intercalated discs (Yang et al., 2014).

Sarcomeric organisation is the other structural immaturity found in iPSC-CMs (Figure 1-26E). Compared to the adult counterpart, iPSC-CMs possess poorly-aligned sarcomeres with shorter sarcomere length (2.2 versus 1.6 μm) (Denning et al., 2016). During human development, there are myofibrillar protein isoform switching events including MYH6→MYH7, skeleton troponin I→ cardiac troponin I, MYL7→MYL2 and titin N-2BA→N-2B. iPSC-CMs predominantly express the immature isoforms of these elements (MYH6, skeleton troponin I and titin N2-BA) (Yang et al., 2014, Ahmed et al., 2020). These differences have consequences in CM function. For example, the titin N2-BA isoform has a greater molecular size than the N2-B isoform, conferring lower passive stiffness (Makarenko et al., 2004).

1.5.2.2 Population heterogeneity and CM purification methods

Despite continuing improvements, the *in vitro* CM differentiation still yields mixed populations, which may preclude scalable CM production, complicate phenotypic analyses and cause adverse effects in post-CM transplantation applications. The heterogeneity comes from

undifferentiated cells, cell types other than CMs (e.g. fibroblasts, endothelial, epicardial and smooth muscle cells (Zhang et al., 2019b, Christoforou et al., 2013, Zhang et al., 2019a) and mixed CM subtypes including atrial-like, ventricular-like and nodal-like cells (Poulin et al., 2021, Scheel et al., 2014). Various methods have been developed to purify CM populations.

1.5.2.2.1 Percoll gradient

Percoll is a colloidal suspension of polyvinyl pyrrolidone-coated silica that facilitates a separation of different non-cell and cell components based on their ability to travel differentially upon centrifugation force. Although this method is simple to apply, it provides only a limited fractionation which yields rather highly variable purity (Laflamme et al., 2007).

1.5.2.2.2 Microfluidic approaches

Microfluidic devices have been used to separate cells based on size (Zhang et al., 2012) and electrophysiology (Myers et al., 2013). However, this strategy would need substantial validation if used for iPSC-CM purification due to their phenotypic variability and immaturity.

1.5.2.2.3 Metabolic selection

CMs can generate energy from substrates other than glucose, whereas other cell types need glucose for survival, growing iPSC-CMs in a glucose-deprived condition could eliminate non-CM populations (Tohyama et al., 2013) However, the efficiency of this method depends on the stage of CM maturity. Fatty acid supplementation has been shown to promote iPSC-CM maturation; therefore, a glucose-free, fatty acid-augmented medium effectively enriches iPSC-CM populations (Lopez et al., 2021, Yang et al., 2019, Feyen et al., 2020, Lin et al., 2017). The free fatty acid addition is complicated by its cytotoxicity effects on CMs (de Vries et al., 1997) but these can be ameliorated by complexing with bovine serum albumin.

1.5.2.2.4 Mitochondrial dyes

The heart is one of the organs that demands a high level of energy from oxidative phosphorylation and thus, CMs harbour numerous mitochondria. Fluorescent mitochondria

dyes e.g. tetramethylrhodamine methyl ester perchlorate permit the labelling of CMs (Hattori et al., 2010) which could be further purified by fluorescence-activated cell sorting (FACS). Nonetheless, the efficiency of this approach depends on the stage of cell maturity, which probably excludes the majority of immature CM populations. Moreover, non-CMs and undifferentiated cells also have been found to take up a considerable quantity of the dye (Elliott et al., 2011)

1.5.2.2.5 Molecular beacons

Molecular beacons are synthetic nucleotide probes carrying a fluorescent reporter at one end and a quencher at the other end. Molecular beacons form a hairpin structure in which the quencher comes into the vicinity of the fluorophore, hence is kept fluorescently inactive. The beacons can be designed to specifically recognise mRNAs of interest. The mRNA/beacon hybridisation distances the fluorophore from its quencher, allowing fluorescently labelling of the target mRNA. FACS assay can be used to detect and isolate iPSC-CMs labelled with CM-specific beacons (Myers et al., 2013). Molecular beacons represent a powerful CM purification tool (>97% purity (Myers et al., 2013)). The only concern is how beacons can be efficiently delivered into the cells, which may need optimisation for each iPSC line.

1.5.2.2.6 Cell surface markers

Unlike blood cells, CMs lack specific surface markers. Although signal-regulatory protein alpha (SIRPA) and vascular cell adhesion molecule 1 (VCAM1) were suggested as being CM markers in ESC-derived CMs (ESC-CMs) (Elliott et al., 2011), they have been not widely implemented for CM purification. Contamination of NKX2.5 negative populations in the sorted fractions expressing these two protein questions their utility. In addition, these markers are known to be expressed in non-CMs (Dubois et al., 2011, Osborn et al., 1989).

1.5.2.2.7 Genetic approaches

iPSC-CMs can be engineered to express selectable markers under the control of CM-specific promoters. These cells are collectively known as reporter cell lines (Sontayananon et al.,

2020). iPSC-CM purification methods vary by the means of selection e.g. antibiotic treatment when using drug-resistant reporters or FACS when using fluorescent reporters. Although antibiotic selection offers a robust and efficient iPSC-CM refinement, a prolonged drug supplementation can alter CMs' Ca^{2+} signalling (Belus and White, 2001) and electrophysiology (Hannes et al., 2015). The additive advantage of fluorescent reporters is enabling a real time tracking of the cell differentiation and maturation. The complication of all genetic strategies is that they involve genetic manipulation which requires genetic expertise and may cause unforeseeable functional consequences.

1.6 Establishment of fluorescent reporter line

A transgenic construct carrying fluorophore is a prerequisite for generating fluorescent reporter lines. The DNA sequence encoding the fluorophore needs to be inserted downstream of the promoter of interest. Fluorescent reporter lines can be engineered as transient lines, where the transgene presents as a free episome, or as stable lines, in which the transgene becomes stably incorporated into the host genome, either randomly or site-specifically.

1.6.1 Transient fluorescent reporter line

Transient reporter lines are established by transfection of non-integrating plasmids or transduction of adenoviral vectors (Nayerossadat et al., 2012) into host cells (Figure 1-27A). These reporter systems are only active for a period of time as long as the episomal DNA is still viably intact. Uncontrollable copy number is the main disadvantage of the system, which depends on the initial transfection efficiency and the loss of the plasmid over time, leading to cell-cell variation of fluorescent levels. Moreover, the applicability in developmental studies is relatively limited. A transgene needs to be introduced at the specific time frame where the ectopic promoter driving the fluorophore expression is readily active. These difficulties can be overcome using the following alternative approaches.

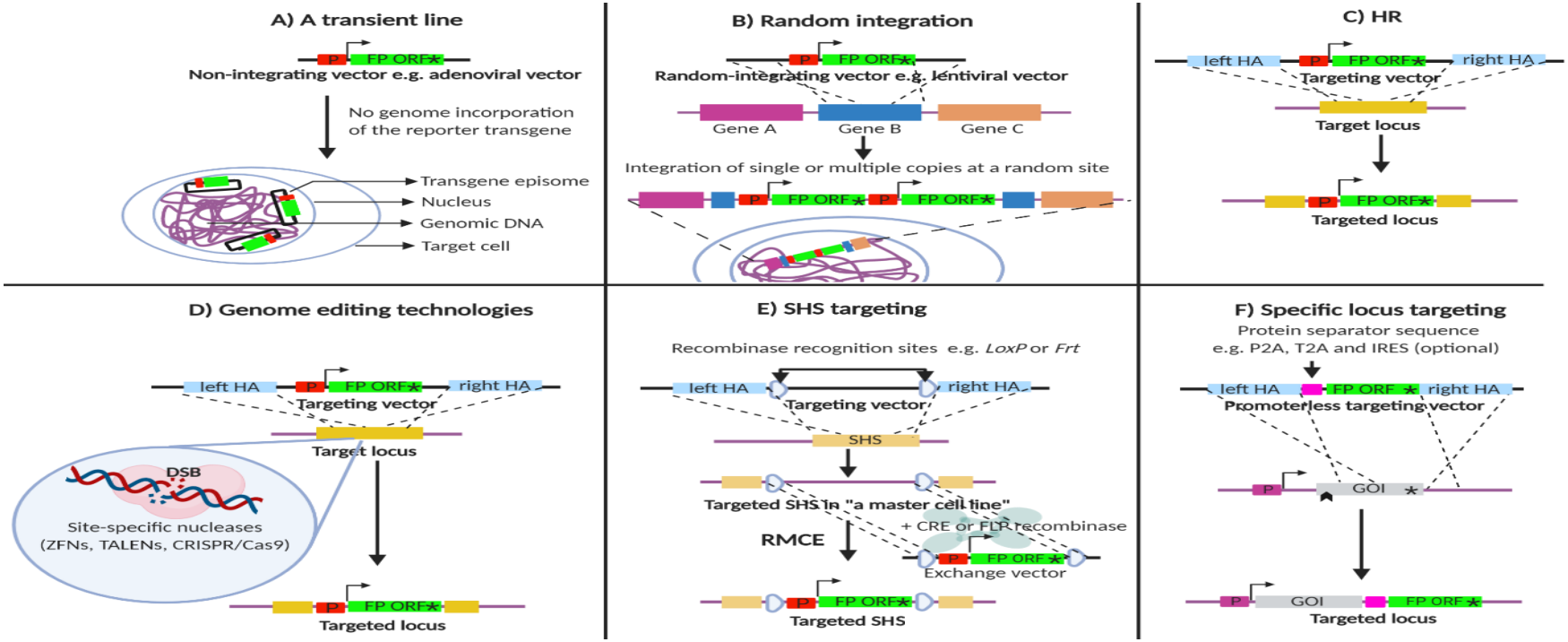


Figure 1-27 Fluorescent reporter line generation methods. A common fluorescent reporter transgene construct consists of an open reading frame (ORF) of fluorophore (FP) following an ectopic promoter sequence (A and B). A) A transient line is generated by introducing a fluorescent reporter transgene using a non-integrating vector e.g. adenovirus. B) An integrating vector e.g. lentivirus can be used to insert the transgene into host's genome at random sites. These two approaches (A and B) cannot control a copy number of transgenes in each single cells, where the random integration (B) can also lead to a series of the transgene localisation at a random locus. C) To site-specifically target the genome by homologous recombination (HR), the transgene construct requires flanking homology arms (HA). D) Using genome editing tools e.g. zinc finger nucleases (ZNFs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9), the efficiency of targeted insertion of the transgene is tremendously enhanced. These programmable nucleases introduce DNA double strand break (DSB) at a precise location which stimulates the HR pathway. A fluorescent reporter can be targeted into E) a safe harbour site (SHS) or F) a gene of interest (GOI). E.) A master cell line is engineered by introducing a pair of recombination sites e.g. *LoxP* or *Frt* into SHS. Various fluorescent reporter lines can be derived from a master cell line by transfecting an exchange vector harbouring a fluorescent transgene flanked by equivalent recombination sites and relevant recombinase enzyme e.g. Cre or Flp. This strategy is known as recombinase-mediated cassette exchange (RMCE). F) Instead of depending on an ectopic promoter activity, an endogenous promoter can be utilised to drive expression of FP. A carefully designed promoterless targeting vector is a prerequisite for HR-mediated incorporating FP's ORF into e.g. an endogenous stop codon. To preserve normal expression of GOI, a linker sequence e.g. a porcine teschovirus-1 2A peptide (P2A), a thosea asigna 2A peptide (T2A) or an internal ribosomal entry site (IRES) may be included in the targeting construct. Taken from (Sontayananon et al., 2020).

1.6.2 Random integration of fluorescent reporter transgene

This strategy relies on a reporter construct being cloned, packaged and delivered by genome-integrating viral systems e.g. lentivirus (Nayerossadat et al., 2012) or as a linear or circular naked DNA (Figure 1-27B). In the latter situation, a plasmid or bacterial artificial chromosome is transfected into host cells and its stable genomic integration is established by antibiotic selection.

Despite the simplicity, the random behaviour of genome integration may be causative of both technical and biological risks (Den Hartogh and Passier, 2016). Firstly, the copy number of integrating constructs is unpredictable and a degree of transgene expression loss after a few cell passages (Fiedorowicz et al., 2020) raises a concern over the reliability of the gene reporter system. Secondly, since genomic insertion position is unknown, there is a possibility that the fluorescent expression may be influenced, either negatively or positively, by the flanking proximal regulatory sequences and/or vice versa (Wilson et al., 1990). Thirdly, transgenes may be marked by epigenetic factors and/or chromosomal architecture alteration specifically during iPSC-CM differentiation, possibly silencing the fluorescent expression (Klatt et al., 2020, Calero-Nieto et al., 2010).

1.6.3 Targeted integration of fluorescent reporter transgene

By exploiting the endogenous repair mechanism “homologous recombination” (Sung and Klein, 2006), a fluorescent reporter transgene can be placed at a desired genomic position (Figure 1-29C), thereby avoiding deleterious effects of random integration. A transgene is carefully designed to harbour flanking arms homologous to a target region, allowing recombination between the vector and the genome and thus, targeted fluorophore integration. Despite its specificity, without the application of site-specific nucleases (see 1.6.4), the targeting efficiency by this strategy is extremely low, making such experiments challenging (Leavitt and Hamlett, 2011).

1.6.4 Genome engineering technologies

The advance in genome engineering technologies has revolutionised the way scientists are manipulating the genome. Site-specific endonucleases, e.g. clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR associated protein 9 (Cas9), can be programmed to target and cleave sequences of interest (Kim and Kim, 2014) (Figure 1-27D). The resultant targeted double strand break (DSB) of DNA provides a substrate that triggers the homologous recombination pathway, hence enhancing the targeting efficiency (Rouet et al., 1994).

Endonuclease-assisted production of fluorescent reporter lines most commonly employ one of two strategies. In the first strategy, a fluorescent reporter transgene is introduced into so-called safe harbour sites (Figure 1-27E). For example, Adeno-associated virus integration site 1 (Oceguera-Yanez et al., 2016) is considered as a suitable genomic region for integration and transgene expression in many cell types (Papapetrou and Schambach, 2016). Moreover, ectopic DNA insertion into these loci does not exhibit any major harmful consequences. Safe harbour site targeting design requires the transgene to be flanked with arms homologous against the target site. Simplifying engineering at these well-used sites, systems have been developed to facilitate sequence exchange by enzymatic recombination (Figure 1-27E). In this scenario, recombination sites such as heterotypic *LoxP* are firstly introduced into the safe harbour site of host's cells. This resulting "master cell line" serves as a substrate for producing multiple reporter lines using recombination-mediated cassette exchange mechanism (Irion et al., 2007, Pei et al., 2015). This mechanism involves an exchange vector carrying a fluorescent reporter transgene flanked by a matching pair of recombination sites being transfected into the master cell line, which permits sequence exchange through the action of a co-delivered recombinase (Oceguera-Yanez et al., 2016).

The disadvantage of all previous techniques is that the expression of fluorescent reporter is under the control of an ectopic promoter. In this regard, the fluorescent reporter may not faithfully represent the actual biological situation of the target gene. Avoiding this concern, the second strategy of nuclease-aided fluorescent reporter line establishment takes advantage of

endogenous promoter activity (Figure 1-27F). Open reading frames containing coding sequences of fluorophores can be incorporated just downstream of the gene of interest or alternately, directly upstream of the gene's coding sequence. Here, the fluorophore expression is regulated by the endogenous promoter. To allow a bicistronic expression, additional peptide sequences e.g. porcine teschovirus-1 2A (P2A) (Liu et al., 2017) are required to be interspaced between the gene of interest and the fluorophore.

1.6.5 CRISPR/Cas9

CRISPR/Cas9 is an RNA-programmable endonuclease system evolved in prokaryotes as an adaptive defence mechanism against foreign DNA (Barrangou et al., 2007). The system incorporates invading DNA sequences between the tandem repeats within the CRISPR array (Sander and Joung, 2014) (Figure 1-28A). Transcripts from this array are called CRISPR RNAs, containing the product of the acquired foreign sequences known as spacers. Each CRISPR RNA hybridises with a so-called transactivating RNA prior to complexing with Cas9 protein. The spacer-encoded sequence directs Cas9 to cleave complementary DNA sequences, only if they are present adjacent to sequences known as protospacer adjacent motifs (PAMs).

CRISPR/Cas9 from *Streptococcus pyogenes* has been widely used in genome engineering (Sander and Joung, 2014). To simplify this system, the two RNA components are fused into a single guide RNA (gRNA) (Figure 1-28B). The 5' end of gRNA harbours 20-nt sequence corresponding to the protospacer, determining its target specificity. Recognition by *S. pyogenes* Cas9 requires an NGG PAM lying immediately downstream of target sequences, where N is any nucleotide (Figure 1-29).

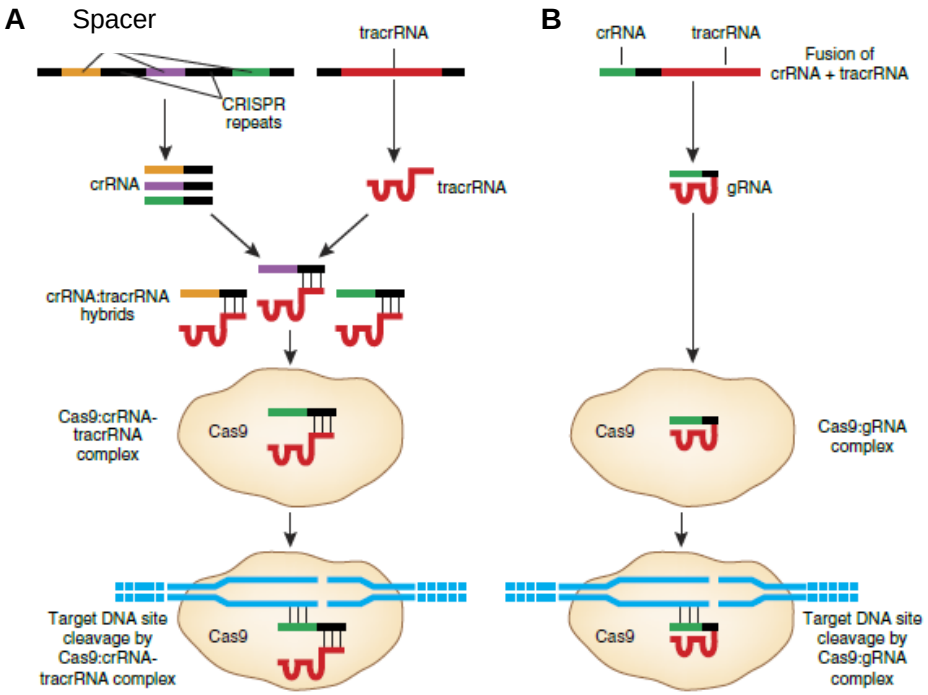


Figure 1-28 Naturally occurring and engineered CRISPR/Cas9. A) Targeted cleavage of foreign DNA by the naturally occurring CRISPR/Cas9 system requires complexing between Cas9 protein and a hybridised product of CRISPR RNA (cr-RNA) and trans-activating RNA (tr-RNA). cr-RNA is transcribed from the CRISPR array containing tandem repeat motifs interspaced by spacers acquired from the former invading DNA. B) The two RNA elements can be fused in to a single guide RNA (gRNA) for use in genome engineering application. Modified from (Sander and Joung, 2014).

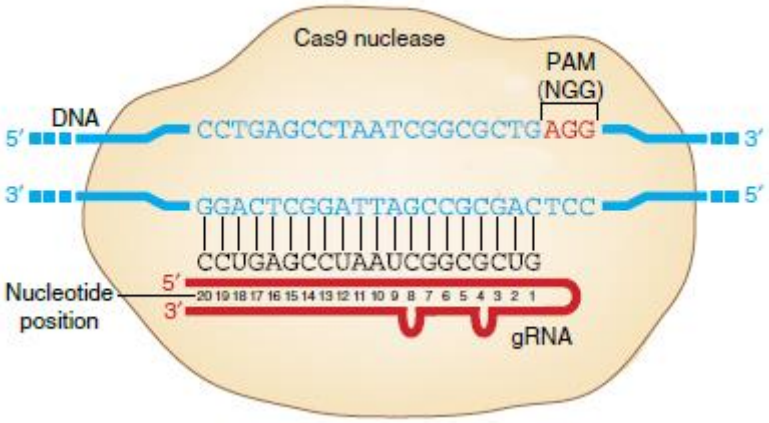


Figure 1-29 *S. pyogenes* CRISPR/Cas9 machinery. The *S. pyogenes* Cas9 specificity relies on the complementarity between the first 20-nt of gRNA and the target sequence containing an NGG PAM. Taken from (Sander and Joung, 2014).

Targeting by CRISPR/Cas9 activate endogenous repair pathways (Sander and Joung, 2014, Jiang and Doudna, 2017) (Figure 1-31). The non-homologous end joining repair pathway can seal the break with some errors, resulting in introduction of insertion/deletion (indel) mutations. This pathway can be exploited to knock-out (KO) a target gene. Alternatively, through the

homology-directed repair pathway (HDR) in the presence of a DNA donor, a targeted genomic insertion/replacement can be achieved. Modifications can be made at the level of a single nucleotide change or insertion of a whole gene sequence such as a fluorophore.

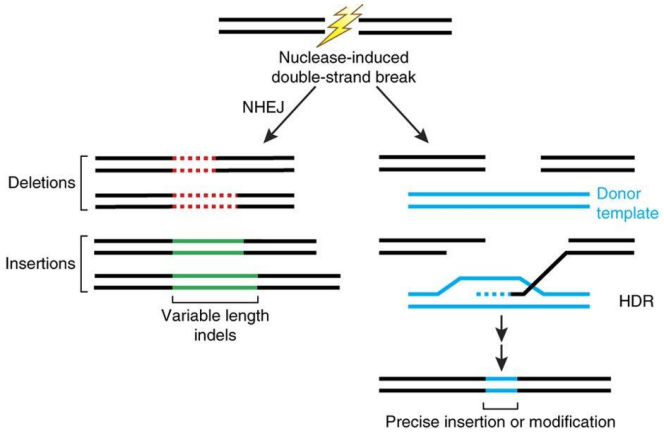


Figure 1-30 Mechanisms of CRISPR/Cas9-assisted genome editing. A DSB generated by Cas9 can be repaired by the non-homologous end joining (NHEJ) or homology-directed repair (HDR) mechanisms. The NHEJ occasionally introduces insertion/deletion mutations at the break site. Depending on a DNA donor sequence, the HDR mediates sequence replacement/insertion in a precise manner. Taken from (Sander and Joung, 2014).

1.7 Fluorescence-guided subtype-specific iPSC-CM purification

Depending on the choice of promoter used, purification of fluorescent iPSC-CM reporter lines yields different population compositions. Although using pan-cardiac fluorescent reporters including *ACTN2* (Fiedorowicz et al., 2020), *TTN* (Sharma et al., 2018) and *NCX* (Fu et al., 2010) allow a highly efficient removal of non-cardiac cells, CM subtype identification is not feasible. Heterogeneous CMs would be undesirable for most areas of cardiovascular research. Firstly, the contamination of other subtypes could induce post-transplantation arrhythmogenic effects, raising a safety concern for iPSC-CMs in cardiac tissue repair (Soma et al., 2021). Secondly, since each CM subtype responds differentially to drugs (Gunawan et al., 2021), a pharmacological screening in mixed CM cultures may be biased against specific CM subtype(s), hindering an assessment of the actual chemical potency. Thirdly, pooled iPSC-CMs may not accurately predict pathogenic mechanisms that are restricted within specific heart regions. In contrast, subtype-specific fluorescent reporter systems enable isolation of single CM subtypes, which could be specifically tailored to meet the needs of the downstream applications.

1.7.1 Ventricular markers

Among all CM subtypes, ventricular-like cells represent the greatest population (up to 84%) in ESC-CMs and iPSC-CMs (Zhang et al., 2009, Moore et al., 2008, Churko et al., 2018).

MLY2 is exclusively expressed in ventricular CMs throughout their development (Chuva de Sousa Lopes et al., 2006). Thus, *MLY2* is widely accepted as a ventricular-specific marker. Several fluorescent *MYL2* iPSC-CM reporter lines have been shown to allow successful FACS-based enrichment of cells displaying ventricular-like electrophysiology and molecular phenotype (Huber et al., 2007, Bizy et al., 2013, Chirikian et al., 2021). The general disadvantage of *MYL2*-based ventricular-like cell purification is that the *MYL2* expression level is low at the early stage of differentiation but increases with time in culture, corresponding to the cell maturity (Piccini et al., 2015).

1.7.2 Atrial markers

Atrial-like iPSC-CM purification is in high demand due to this subtype being relatively a minor population (Churko et al., 2018). Generation of atrial-specific reporter lines is also challenging because of a lack of a reliable genetic marker.

SLN is expressed predominantly in the atria (Minamisawa et al., 2003). Fluorescent *SLN* iPSC-CM reporter lines have been shown to enable purification of cells with atrial-like electrophysiology and Ca^{2+} signalling (Josowitz et al., 2014, Chirikian et al., 2021). Moreover, a dual *SLN/MYL2* fluorescent reporter allows simultaneous isolation of atrial-like and ventricular-like cells (Chirikian et al., 2021). The possible drawback of *SLN* is that its expression becomes downregulated when iPSC-CMs mature (Josowitz et al., 2014).

Nuclear receptor subfamily 2 group F member 2 (*NR2F2*) has been shown to mediate atrial specification in human iPSC-CMs (Devalla et al., 2015) and mice (Wu et al., 2013). Relying on this gene activity, atrial-like cells have been successfully purified from ESC-CMs (Schwach et al., 2017). The limitation of *NR2F2* is its expression in non-CMs such as vascular smooth muscle cells, endocardial cells, and epicardial cells (Lin et al., 2012).

MYL7 is expressed throughout the linear heart tube and only becomes restricted to the atria in the later stages (Kubalak et al., 1994, Franco et al., 1998). Considering this expression pattern, *MYL7* may be a good candidate for an atrial reporter. The concern of using this marker is that: though iPSC-CMs primarily express *MYL7* at the early differentiation days, its level progressively declines in the later stages (Bedada et al., 2014). Currently, there is no stable iPSC-CM reporter line for *MLY7* produced.

Another potential atrial marker is pituitary homebox 2 (*PITX2*), which is exclusively expressed in the adult left atrium (Clauss and Kääb, 2011). PITX2 inhibits sinoatrial node development by suppressing the expression of nodal-regulating transcription factor called short stature homeobox 2 (*SHOX2*). Currently, there is no fluorescent *PITX2* iPSC-CM line reported.

1.7.3 Nodal markers

Nodal-like cells are the rarest subtype resulting from the *in vitro* differentiation (Churko et al., 2018) and there are not many nodal-specific fluorescent reporters published.

HCN4 is a nodal-specific channel. However, there is no *HCN4* reporter iPSC-CM line published. *HCN4* is transcriptionally under the control SHOX2 (Clauss and Kääb, 2011). *SHOX2* overexpression is able to activate the nodal gene programme and steer the ESC-CM differentiation towards the nodal pathway (Ionta et al., 2015). The utility of *SHOX2* as a nodal-marker has been also confirmed by fluorescent *SHOX2* iPSC-CMs (Chen et al., 2016).

1.8 Hypotheses and specific aims

1.8.1 Hypotheses

According to the expression pattern seen in the adult heart, I hypothesise that:

I. *MYL7* and *PITX2* are atrial-specific markers in iPSC-CMs. Reporters driven by these genes may enable atrial-like CM identification and purification.

II. *MYL2* is a ventricular-specific marker in iPSC-CMs. A reporter driven by this gene may enable ventricular-like CM identification and purification.

III. *HCN4* is a nodal-specific marker in iPSC-CMs. A reporter driven by this gene may enable nodal-like CM identification and purification.

Titin is the key sarcomere component which provides a scaffold for thick and thin filaments. Therefore, I hypothesise that:

IV. *TTNtv* disrupts sarcomere formation, leading to sarcomere insufficiency and impaired contractility in DCM.

Though titin is expressed similarly in both atrial and ventricular CMs, these two CM subtypes display differences in e.g. Ca^{2+} handling apparatus and myofibrillar proteins. Thus, I hypothesise that:

V. *TTNtv* impairs Ca^{2+} transient dynamics and contractility differentially in atrial and ventricular CMs.

1.8.2 Specific aims

Aim 1 – In accordance with Hypotheses I-III, I aim to generate fluorescent *MYL7*, *PITX2*, *MYL2* and *HCN4* reporter iPSC-CM lines using CRISPR/Cas9 (Chapter 3).

Aim 2 – Following Aim 1, I aim to explore the utility of the engineered fluorescent reporter lines in purification of CM subtypes (Chapter 4). It includes characterisation of correlation between the fluorescent reporter and target gene expression and use of the fluorescent reporters in facilitating FACS sorting of their respective CM subtype.

Aim 3 – In accordance with Hypothesis IV, I aim to generate a *TTNtv* iPSC-CM model using CRISPR/Cas9 and characterise *TTNtv* consequences in sarcomere organisation, myofibrillar remodelling and expression of CM failure markers (Chapter 5).

Aim 4 – I aim to explore the utility of the engineered fluorescent reporters in the study of *TTNtv* disease (Chapter 6). In accordance with Hypothesis V, the aim includes introduction of a common *TTNtv* into the fluorescent reporter line(s) and characterisation of *TTNtv*

consequences in the reporter lines by measuring Ca^{2+} transient dynamics, contractility and sarcomere remodelling in atrial and ventricular-like iPSC-CMs.

Chapter 2

Materials and methods

Note: Recipes for all buffers and solutions are included in Appendix: Table S1. Commercial kits and other commercial products are summarised in Appendix: Table S2. All instruments are summarised in Appendix: Table S3. All primers, SYBR green probes and other DNA oligos are summarised in Appendix: Table S4. Maps of plasmids and a gBlock template are included in Appendix: Figure S1-S7.

2.1 Materials

2.1.1 Bacteria

In this study, 5 sources of *Escherichia coli* were used for molecular cloning purposes. Chemically competent DH5α, NEB5α (NEB) and ElectroSHOX were for heat shock transformation. Electro-competent ElectroSHOX (Bioline) or NEB® 5-alpha (NEB) was used for transformation by electroporation.

2.1.2 iPSC line

KOLF2-C1 iPSC line (WTSIi018-B-1) was a gift from the Wellcome Sanger Institute. This stem cell line was derived from healthy male's skin fibroblasts.

2.1.3 Antibodies

All antibodies used in this study are listed in Table 2-1. The concentrations are stated for immunofluorescence (IF), Western blot (WB) or FACS.

Table 2-1: List of antibodies

Primary antibodies				
Target specificity	**Host**	**Clone/Cat. No.**	**Working concentration**	**Source**
ACTN2	Rabbit	Ab68167	1:200 (IF), 1:2000 (WB), 1:100 (FACS)	Abcam
ACTN2	Mouse	EA53/ab9465	1:200 (IF)	Abcam
F-actin (Alexa Fluor 633 conjugated phalloidin)	-	A22284	1:400 (IF)	Invitrogen
GAPDH	Rabbit	ABS16	1:1000 (WB),	Millipore

GFP	Rabbit	SAB4301138	1:1000 (WB), 1:100 (FACS)	Sigma
MYL2	Mouse	60229-1-Ig	1:1000 (WB) 1:100 (FACS)	Proteintech
MYL7	Mouse	311011	1:2000 (WB) 1:100 (FACS)	Synaptic system
MYOM1	Mouse	B4	1:200 (IF)	DSHB
T-CAP	Rabbit	Ab133646	1:30 (IF)	Abcam
TNNT2	Rabbit	Ab45932	1:200 (IF)	Abcam
TTN M8	Rabbit	M8	1:50 (IF)	(Obermann et al., 1996)
TTN T12	Mouse	T12	1:10 (IF)	(Fürst et al., 1988)
Secondary antibodies				
Specificity	**Host**	**Clone/Cat. No.**	**Working concentration**	**Source**
Alexa Fluor 488-conjugated anti rabbit IgG	Goat	A11034	1:1000 (IF)	Invitrogen
Alexa Fluor 568-conjugated anti rabbit IgG	Goat	A11036	1:1000 (IF)(FACS)	Invitrogen
Alexa Fluor 647-conjugated anti mouse IgG	Donkey	A31571	1:1000 (IF) (FACS)	Invitrogen
Alexa Fluor 647-conjugated anti Rabbit IgG	Donkey	A31573	1:1000 (FACS)	Invitrogen
HRP-conjugated anti-mouse IgG	Sheep	NA931	1:10000 (WB)	Sigma
HRP-conjugated anti-rabbit IgG	Donkey	NA934	1:10000 (WB)	Sigma

2.1.4 Taqman assay probes

Taqman assay probes (ThermoFisher) are summarised in Table 2-2.

Table 2-2 List of Taqman assays

Transcript	**Assay ID**
ACTN2	Hs00153809_m1
GAPDH (VIC-MGB labelled)	4626317E
HCN4	Hs00975492_m1
HEY2	Hs01012057_m1
IRX4	Hs00212560_m1
KCNJ3	Hs04334861_s1
MYH6	Hs01101425_m1
MYH7	Hs01110632_m1
MYL2	Hs00166405_m1
MYL7	Hs1085598_g1
NKX2.5	Hs00231763_m1
NPPA	Hs00383230_g1
NPPB	Hs00173590_m1
NR2F2	Hs00819630_m1
PITX2	Hs04234069_mH
T-CAP	Hs00366220_m1
TNNT2	Hs00943911_m1

2.2 Methods

2.2.1 Molecular Cloning

2.2.1.1 Gibson Assembly

Gibson Assembly is a molecular cloning method which facilitates joining of multiple double-stranded DNA (dsDNA) fragments in a single, isothermal reaction (Figure 2-1) (Gibson et al., 2009). The joining DNA fragments were designed to contain directional homology ends (25-nt). 5' exonuclease partially digests these sequences, facilitating the annealing of the fragments. DNA polymerase fills the gap between the joining fragments and then the reaction is completed by sealing the fragments with DNA ligase.

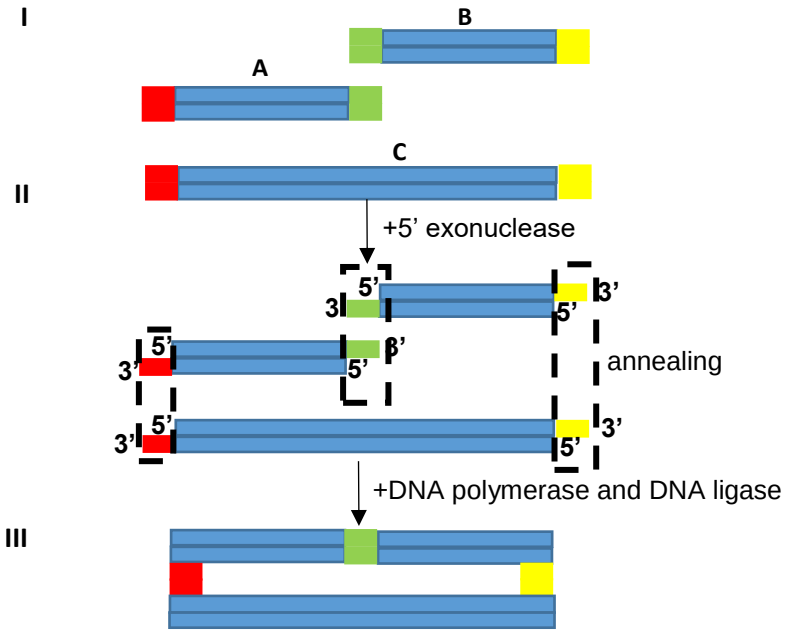


Figure 2-1 Simplified directional Gibson Assembly diagram. Gibson Assembly is a cloning approach for joining multiple DNA fragments in a single, isothermal reaction. I) For a directional Gibson Assembly, each pair of joining DNA fragments is designed to share short homology sequences at their joining ends (highlighted by the same colour). For example, there is only one way to complete a 3-way combination of fragment A, B and C:- that is A-B: B-C: C-A. II) 5' exonuclease in Gibson Assembly reaction resects the 5' end of each fragment, creating 3' DNA overhangs. The partial single –stranded DNA regions then anneal to their compatible sequences. III) DNA polymerase fills the gapped sequence and DNA ligase finally seals the joining fragments together.

All Gibson Assembly reactions were performed using NEBuilder® HIFI DNA Assembly Kit (NEB) as per the manufacturer's recommendation with equimolarity (0.1 pmol) for each joining fragment.

2.2.1.2 Ligation-dependent cloning

In this thesis, a ligation-dependent cloning involved two or three fragments. 100 ng of phosphatase (NEB)-treated linearised vector was ligated with the respective linearised insert(s) at 1:3 and 1:3:3 molarity ratio for 1-insert cloning and 2-insert cloning, respectively (or with 10 µM annealed oligos). Ligation was performed using T4 DNA ligase (NEB) as recommended by the manufacturer.

2.2.1.3 Golden Gate Assembly

This seamless cloning strategy was used for the cloning of the CRISPR target sequence into a sgRNA scaffold. 2 µl of 100 µM oligo was mixed with equal concentration of its paired oligo in 16 µl Tris-EDTA buffer (Invitrogen). The oligo mix was heated at 100°C for 5 min and cooled down slowly to room temperature to allow annealing. 1 µl of the annealed oligos was mixed with 150 ng of a CRISPR cloning plasmid (pX330-Puro, Figure S4), 1 µl *Bbs*I, 1 µl T4 DNA ligase, and 1 µl 10x T4 DNA ligase buffer in a total 20 µl reaction. The mixture was then subjected to thermocycling using the following parameters: 10 cycles of 5 min 37°C and 10 min 16°C, 1 cycle of 5 min 50°C and 5 min 80°C. Subsequently, 1 µl of 1 mM ATP and 1 µl of ATP-dependent Plasmid Safe Nuclease (Epicentre) were added into the CRISPR reaction and incubated at 37°C for 1 hr to remove non-circular DNAs.

2.2.1.4 Bacterial transformation

5 µl of Gibson Assembly, Golden Gate Assembly or ligation mixture was mixed with 100 µl ice-thawed chemically competent *E. coli* and incubated on ice for 20 min. The cells were transferred to 42°C for exactly 90 sec and immediately returned to ice for 5 min. If ampicillin was used, the whole culture was spread evenly onto an LB agar supplemented with 100 µg/ml ampicillin, and the plate was incubated at 37°C overnight. If kanamycin was the selection (in case of any plasmids containing a neomycin resistant cassette(*Neo)*), the cells were recovered in an antibiotic-free LB liquid for 45 min on a 225-rpm rocking platform at 37°C. The cells were pelleted (3000 rpm, 2 min) and re-suspended in 100 µl LB and plated onto an LB agar

containing 50 µg/ml kanamycin. Single colonies were picked and cultured in LB containing appropriate antibiotic at 37°C on a 225-rpm shaking platform overnight. Plasmids were purified using GeneJET Plasmid Miniprep Kit (ThermoFisher), or GeneJET Plasmid Midiprep Kit (ThermoFisher) if iPSC transfection was their downstream application, and verified by restriction digestion and Sanger DNA sequencing.

2.2.1.5 Bacterial electroporation

2 µl of Gibson Assembly mixture was mixed with electro-competent *E. coli* and incubated on ice for 5 min. The cell/Gibson Assembly reaction mixture was transferred to a pre-chilled electroporation cuvette. Electroporation was performed using Gene Pulser Xcell (Biorad) by following parameters: 1350 V, 10 µF and 600 Ω. The bacteria were recovered in LB and incubated at 37°C on a 255-rpm rocking platform for 1 hr before spread on an LB agar containing an appropriate antibiotic.

2.2.1.6 Colony PCR

Colony PCR is a primary screening method for successful cloning based on PCR amplification of the insert. Single bacterial colonies grown on an agar plate were carefully picked using a clean micropipette tip. The tip was swirled gently into a PCR mixture containing screening primers prior to being resuspended into LB supplemented with an appropriate antibiotic. After the PCR confirmation, positive clones were sub-cultured for further analysis e.g. plasmid extraction and DNA sequencing.

2.2.2 Long single stranded DNA (lssDNA) purification

dsDNA molecules representing a HDR donor template were cloned into the pJET1.2/blunt plasmid (Figure 2-2A & S6) using CloneJET PCR Cloning Kit (ThermoFisher) by following the manufacture's protocol. The dsDNA inserts were amplified by PCR from the cloned plasmid using the biotinylated (Btn) pJET1.2 forward primer and an insert-specific reverse primer (Figure 2-2B). The PCR product was purified using Monarch® PCR & DNA Cleanbup Kit (NEB) as per the manufacturer's instruction.

Long single stranded DNA (lssDNA) purification was performed in a 1.5 ml Eppendorf tube on a magnetic rack. Dynabeads M280 Streptavidin (ThermoFisher) were washed twice with 600 µl chilled Tween Buffer. The beads were then saturated twice with 600 µl 2X Binding Buffer. 100 µl 2X Binding buffer and 100 µl of 10 µg biotinylated PCR product was added to the beads and incubated at room temperature for 1 hr with rolling and the unbound fraction was collected for analysis (Figure 2-3A). The beads were washed 3 times with 1X Binding Buffer (2X Binding Buffer diluted with distilled water) and the washes were collected. 100 µl of 20 mM NaOH was added to the bead mix and then incubated at 22°C for 10 min with 600 rpm shaking. On a magnetic rack, the supernatant was transferred to 50 µl of Neutralising Solution (this fraction is the eluted bottom strand lssDNA, Figure 2-3B). The second/top strand lssDNA was still non-covalently bound with the beads at this stage. The elution was repeated two more times. The fourth elution was in a stronger alkaline condition (150 mM NaOH) with equal concentration of neutralising buffer.

The beads were saturated again twice with 600 µl 2X Binding Buffer. The beads were incubated at room temperature for 5 min with 50 µl Hybridisation Buffer prior to adding 1 µl *Xho*I. The hybridisation primer was designed to create a small region of dsDNA sequence surrounding the polylinker *Xho*I site lying immediately downstream of the biotinylated pJET1.2 forward primer sequence (Figure 2-2B & 2-3C). The reconstitution of dsDNA allows the *XhoI* digestion to occur and release the top strand lssDNA. The restriction digestion was performed at 37°C with 600 rpm shaking for 1 hr. 50 µl Monarch® DNA Elution Buffer (NEB) was added and a total 100 µl supernatant was collected as the top strand lssDNA product (Figure 2-3D). The lssDNAs were concentrated using Monarch® PCR & DNA Cleanup Kit (NEB).

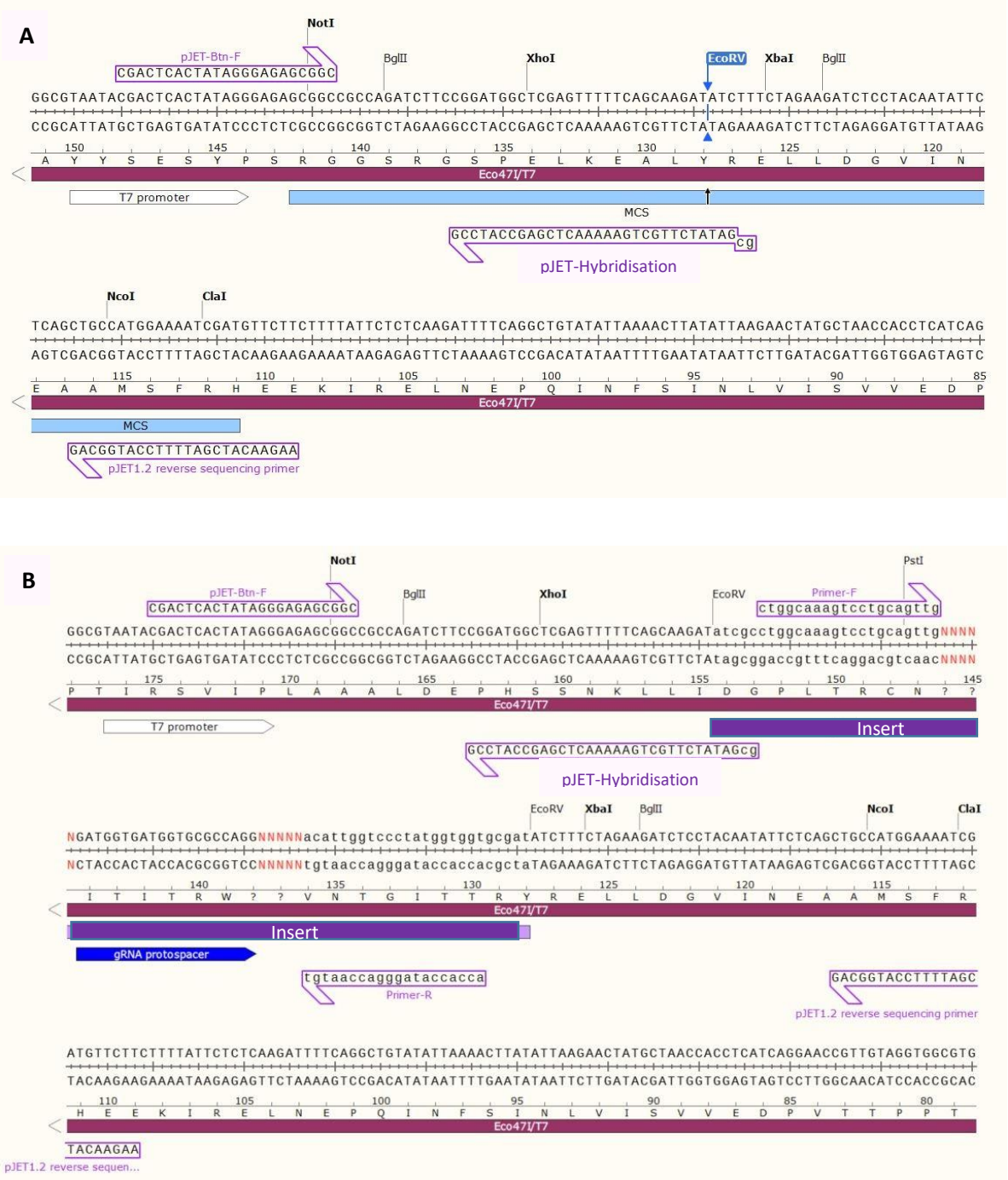


Figure 2-2 Design for lssDNA purification. A) A blunt end PCR product of interest can be cloned into pJET1.2/blunt vector at the broken *Eco*RV site and B) the modified target sequence is amplified by PCR using the biotinylated forward primer (pJET-Btn-F) and a target-specific reverse primer. After adhering to streptavidin dynabeads, alkali treatment to denature the DNA and several steps of eluting the non-biotinylated strand (Figure 2-3A and B), the dsDNA portion can be generated by annealing with the hybridisation primer, allowing *Xho*I site cleavage (Figure 2-3C). This restriction site was used for eluting the single strand (top strand) of the target DNA (Figure 2-3D).

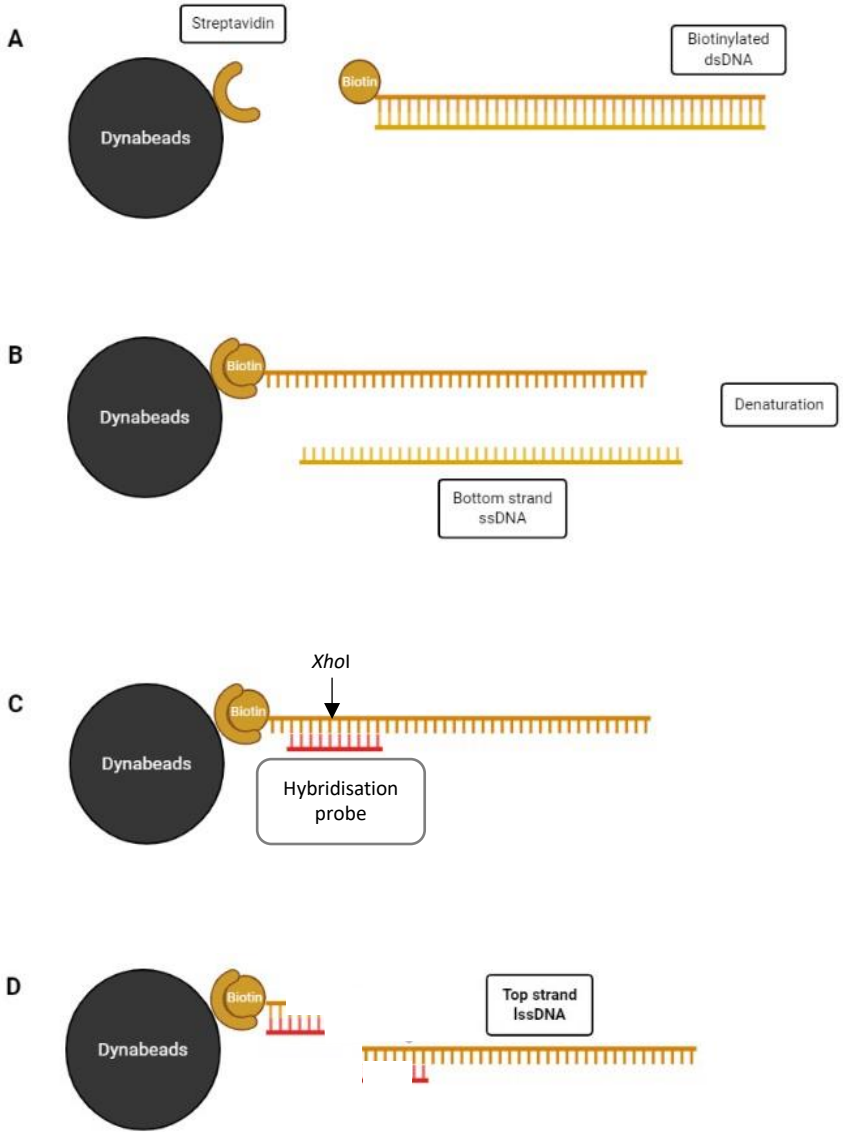


Figure 2-3 Simplified illustration of lssDNA purification. A) A biotinylated dsDNA is captured by a streptavidin-coated dynabead. B) The dsDNA is denatured on the bead, allowing non-biotinylated bottom lssDNA elution. C) *Xho*I site is regenerated upon hybridisation primer annealing where D) the *Xho*I catalysis releases the top lssDNA from the bead.

2.2.3 *In vitro* gRNA synthesis

CRISPR target sequences (20 nt) were designed *in silico* using http://crispor.tefor.net/. A T7 gRNA transcription template (Appendix, Table S4) was designed using EnGen® sgRNA Synthesis Kit, *S. pyogenes* protocol (NEB) and a gRNA was synthesised by following the manufacture's recommendation.

2.2.4 Genomic DNA (gDNA) extraction

Cells were lysed in the lysis buffer (Table S1) at 37°C overnight. 1:10 volume 8 M LiCl and 1 volume absolute isopropanol were mixed with the cell lysate and centrifuged at 13,000 rpm at room temperature for 15 min. The white DNA pellet was then washed with 70% ethanol and centrifuged at 13,000 rpm at room temperature for 15 min. The DNA pellet was let dry at 37°C for at least 10 min. gDNA was re-suspended with nuclease-free water.

2.2.5 PCR and DNA gel electrophoresis

Prior to electrophoresis, PCR samples were mixed with the Orange-G DNA loading buffer and loaded into an agarose gel (Sigma)/TAE buffer (Severn Biotech). Electrophoresis was performed at constant 130V in a TAE buffer. DNA bands were imaged under ultraviolet light on ChemiDoc™ MP Imaging System (Biorad). Where necessary, DNA band was excised from an agarose gel and purified using GeneJet Gel Extraction Kit (ThermoFisher).

2.2.6 cDNA synthesis and quantitative PCR (qPCR)

RNA extraction was performed using TRIzol™ (Invitrogen) as per the manufacturer's recommendation. RNA was converted to cDNA using qPCRBIO cDNA Synthesis Kit (PCR Biosystems) by following the manufacturer's instruction. cDNA was diluted to 20 ng/µl when quantitative PCR (qPCR) was the downstream application. qPCR was performed using PowerUP™ SYBR™ Green Master Mix (Applied Biosystems) or TaqMan™ Fast Advanced Master Mix (Applied Biosystems), and qPCR parameters were quantified by CFX96 Touch Real-Time PCR Detection System (BioRad). Unless stated otherwise, expression fold change was calculated by means of $\Delta\Delta$Ct using glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) as a normalisation gene. All Taqman assay probes and SYBR green probes (mClover qF/qR, mScarlet qF/qR) are presented in Table 2-2 and Table S4, respectively.

2.2.7 Protein extraction and Western blot (WB)

Cells were lysed in RIPA buffer (ThermoFisher) supplemented with 1X cOmplete™ EDTA-free protease inhibitor cocktail (Roche) on ice for 20 min and centrifuged at 10,000 rpm for 10 min. The clear supernatant was carefully transferred into a new Eppendorf tube. Protein concentration was measured using Pierce™ BCA Protein Assay Kit (ThermoFisher) as per the manufacturer's guidance.

30-40 µg of protein was mixed with 1X Laemmli loading buffer (Biorad) and boiled at 100°C for 5 min. The protein sample was cooled on ice for at least 5 min and loaded into AnykD™ Mini-PROTEAN®-TGX™ Stain-free Precast Protein Gel (Biorad) submerged in 1X TGS running buffer (Biorad). Electrophoresis was performed at constant 100V for approximately 1 hr.

The protein polyacrylamide gel was combined with Trans-Blot Turbo Mini 0.2 µM PVDF transfer paper (Biorad) and transfer apparatus. Protein transfer was performed by Trans-Blot Turbo transfer system (Biorad) using mixed molecular weight program.

The transfer membrane was blocked with 5% skimmed milk in 0.1% Tween-20 (VWR) in phosphate buffered saline (PBS) (Gibco) and incubated at room temperature for 1 hr on a rocking platform. The membrane was incubated with primary antibodies and incubated at 4°C overnight. The membrane was washed 3 times with 0.1% Tween-20/PBS, 5 min each on a rocking platform. The membrane was incubated with HRP-conjugated secondary antibodies in 0.1% Tween-20/PBS at room temperature on a rocking platform for 1 hr. The membrane was washed 3 times with 0.1% Tween-20/PBS, 5 min each on a rocking platform. The membrane was then incubated with ECL substrate (Biorad) away from light for 5 min. Protein bands were visualised on ChemiDoc™ MP Imaging System (Biorad) and quantified by ImageJsoftware (version 1.53c).

2.2.8 iPSC culture

iPSCs were maintained in Essential 8™ (E8) medium (Gibco) without antibiotic supplement on a vitronectin (Gibco) pre-coated culture vessel at 37°C with 5% CO_2 and the medium was changed daily. At approximately 65-70% density, the cells were dissociated using ReLeaSR™ (STEMCELL) as recommended by the manufacturer and re-plated at 1:5 to 1:6 dilution. To obtain single cells, the cells were disaggregated using TryPLE™ select enzyme (Gibco) by following the manufacturer's protocol. When culturing single cells, an additional 10 μM Rho-kinase inhibitor Y-27632 (ROCKi) (STEMCELL) was required. ROCKi was removed 24 hr after treatment.

For cryopreservation, approximately 10^6 cells were dissociated using ReLeSR™ and re-suspended in 500 μl E8. The cell suspension was transferred into a 2-ml cryovial and added in a dropwise manner with an equal volume of 2X freezing medium (20% DMSO in foetal bovine serum (Gibco)). The cells were then moved into a defined cooling unit "Mr Frosty" and stored at -80 °C overnight. Subsequently, the cells were preserved in liquid nitrogen, vapour phase.

For thawing, frozen cells were partially defrosted at 37 °C for 5 min or until a small ice crystal remained. The cell suspension was gently added to 5-ml Dulbecco's Modified Eagle Medium/F12 (Gibco) and centrifuged at 100 g for 5 min. The cell pellet was re-suspended with ROCKi-supplemented E8 and plated on a vitronectin-coated tissue-culture vessel.

2.2.9 iPSC quality control

2.2.9.1 FACS analysis for iPSC pluripotency

Approximately 10^6 iPSCs were dissociated and centrifuged at 100 g for 5 min. The cell pellet was washed once with PBS and centrifuged at 100 g for 5 min. The pellet was fixed with Cytofix (Human and Mouse Pluripotent Stem Cell Analysis Kit, BD Biosciences) or equivalent 4% paraformaldehyde at room temperature for 20 min. The subsequent staining processes were performed by following the manufacturer's recommendation. The cells were acquired on

LSRFORTESSA™ X20 CELL ANALYZER (BD Biosciences) and analysed by FlowJo software (version 10.6.2).

2.2.9.2 iPSC karyotyping

Approximately 10^5 iPSCs were plated onto a 30-mm rectangular coverslip for 24 hr. The cells were karyotyped by the Chromosome Dynamics Core Facility at the Wellcome Centre for Human Genetics using DAPI staining of metaphase spreads.

2.2.10 iPSC transfection and clonal establishment

2.2.10.1 iPSC transfection

2.2.10.1.1 Ribonucleoprotein (RNP) electroporation

0.6 µl of 20 µM EnGen® Spy Cas9 NLS (NEB) or Alt-R® S.p. HiFi Cas9 Nuclease (IDT) and 1 µg gRNA were diluted in 4.4 µl Buffer R provided with Neon™ Transfection System 10 µl kit (Invitrogen) and incubated at room temperature for 15 min. In the case of experiments using HDR templates, 200 ng of plasmid template or 1 µl of 100 µM single-stranded oligo deoxynucleotide (ssODN) was added into the RNP mix just prior to electroporation. 10^5 cells were re-suspended in 7 µl Buffer R and added into the RNP mix. Using a P10 tip, the cells were electroporated by Neon™ Electroporation System (Invitrogen) using the following programme: 1250 V, 30ms, 2 pulses. Immediately, the cells were transferred into a 48-well plate containing ROCKi-supplemented E8.

72 hr after transfection, gDNA was harvested for subsequent PCR analysis. In the case of HDR experiments, the remaining cells were seeded at low density. If the HDR template harboured a *Neo* cassette, the cells were treated with 350 µg/ml G418 for 7 days followed by 200 ug/ml G418 for another 7 days before low density plating.

2.2.10.1.2 Plasmid reverse transfection

5 µl FuGENE6® (Promega) was diluted in 20 µl Opti-MEM™ (Gibco) and incubated at room temperature for 5 min. In the meantime, 550-1000 ng plasmid in a final volume of 25 µl Opti-

MEM™ was incubated at room temperature for 5 min in a separate tube. The two diluted reagents were mixed thoroughly and incubated at room temperature for another 10 min.

2.5 x 10^5 iPSCs were seeded onto a 12-well plate just prior to transfection. The transfection mix was dropwise added into the cell culture and swirled to mix. After 24 hr, the transfection reagent was withdrawn and replaced with E8 medium. If plasmids contained a puromycin resistant cassette, the cells were selected in a medium containing 350 ng/ml puromycin 48 hr post-transfection for 48 hr before low density seeding.

2.2.10.2 iPSC low density plating and colony picking

10^3 cells were seeded evenly onto a 6-cm dish in E8 supplemented with 1X CloneR™ (STEMCELL) or 1X RevitaCell™ (Gibco). These cloning supplements were maintained in the culture medium for 2 days. Approximately 7 days after seeding, single colonies showing a clear and defined boundary and a regular rounded morphology were picked using a sterile micropipette tip under a stereoscope and transferred into a vitronectin-coated 96-well plate containing E8. Approximately 7 days later, the colonies were split into two 96-well plates, one for maintenance or freezing and the other for gDNA extraction.

2.2.11 iPSC-CM differentiation

iPSC-CMs were differentiated as described by (Lian et al., 2012) with some modifications (Figure 2-4).

iPSCs were seeded onto a reduced growth factor, phenol red-free Matrigel (Corning) coated vessel and maintained in E8 medium. When the cells reached approximately 90-95% confluency (day 0) (Figure 2-4), the medium was changed to RPMI-1640 (Gibco) supplemented with B-27™ minus insulin (Gibco) and 10 μM CHIR-99021 (Selleck). On day 1, RPMI-1640 supplemented with B-27™ minus insulin was added to halve the CHIR-99021 concentration without medium change. On day 2, the medium was replaced with RPMI-1640 supplemented with B-27™ minus insulin. On day 3, the medium was changed to RPMI-1640 supplemented with B-27™ minus insulin and 5 μM IWR-1 (Selleck). The medium was replaced

with plain RPMI-1640 supplemented with B-27™ minus insulin on day 5. From day 7 onward, the cells were cultured in RPMI-1640 supplemented with B-27™ and the medium was changed every other days. The cells started beating from day 9 to d11. Approximately at day 13-day 15, the cells were cultured in a glucose-free RPMI-1640 medium (Gibco) supplemented with B27™ for a consecutive 3-day period without medium change to enrich iPSC-CMs population. iPSC-CMs from day 30 onward were considered mature CMs and were harvested for phenotypic observation.

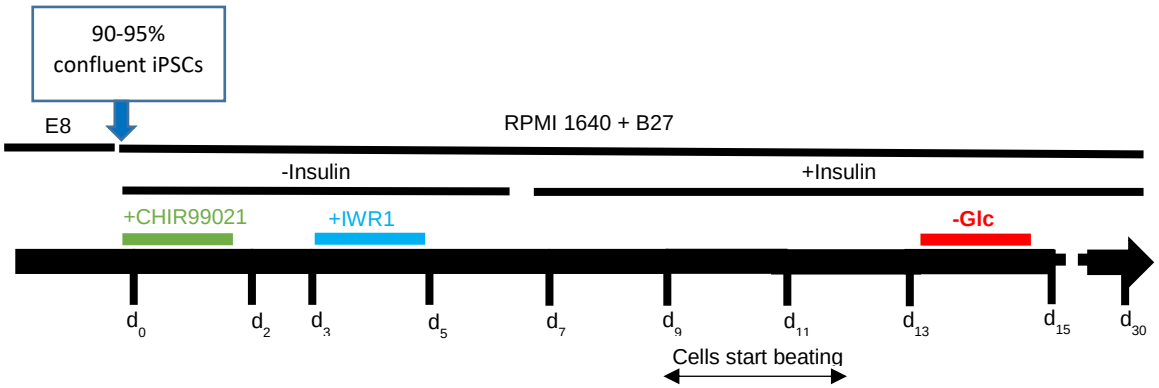


Figure 2-4 A conventional, directed monolayered iPSC-CM differentiation protocol. –Glc, glucose-deprived medium.

When required, iPSC-CMs were disaggregated using TryPLE or 0.25% Trypsin-EDTA (Gibco). The dissociated cells were re-suspended in a plating medium (RPMI-1640 supplemented with B-27, 10% knock-out serum replacement (Gibco) and 2 μM Thiazovivin (Selleck)). The plating medium was replaced with plain RPMI-1640 supplemented with B-27 after 24 hr.

2.2.12 Live iPSC-CM imaging

Live iPSC-CM imagining was performed by SP8 LIGHTNING confocal microscope (Leica) using a 20X dry objective lens at 37 °C. mClover fluorescent signals were acquired using a fluorescein filter set with 7.5X zoom factor at 512*512 pixel^2 resolution. mScarlet fluorescence signals were acquired using a mCherry filter set with 1X zoom factor at 512*512 pixel^2 resolution. Nuclei were stained with 1 μM siR-DNA (Universal Biologicals) 5 min prior to acquiring fluorescent signals using a Cy5 filter set.

2.2.13 iPSC-CM immunofluorescence (IF)

Approximately 5 x 10^4 iPSC-CMs were seeded onto a 12-mm round glass coverslip and cultured for at least 48 hr. The cells were washed once with PBS and fixed with 4% paraformaldehyde for 15 min at room temperature. The cells were washed 3 times with PBS before permeabilised with 0.2% Triton X-100 for 10 min at room temperature. The cells were blocked with 5% fish gelatin (Sigma) and 0.3 M glycine in PBS for 1 hr at room temperature. Subsequently, the cells were probed with primary antibodies in PBS at 4°C overnight. The cells were washed 3 times with PBS and stained with fluorescence-labelled secondary antibodies in PBS at room temperature for 1 hr away from light. The cells were washed 3 times with PBS and then the coverslip was mounted on a clean glass slide with DAPI-containing fluoroshield (Sigma). The coverslip was sealed with colourless nail polish and allowed to dry. Images were acquired by SP8 LIGHTNING confocal microscope using a 63X oil-immersion objective lens at 1024*1024 $pixel^2$ resolution with 7.5X zoom factor. mClover fluorescent signals were acquired using a fluorescein filter set. Targets labelled with Alexa Fluor-488, Alexa Fluor-568 and Alexa Fluor-647-conjugated antibodies were imaged using a Texas Red, a Fluorescein and a Cy5 filter set, respectively.

2.2.14 FACS analysis for iPSC-CMs

Approximately 10^5 iPSC-CMs were dissociated and centrifuged at 100 g for 5 min. The pellet was washed once with PBS and centrifuged at 100 g for 5 min. The pellet was fixed with Cytofix (BD Biosciences) or equivalent 4% paraformaldehyde at room temperature for 20 min. The cells were permeabilised with BD perm/wash buffer (BD Biosciences) for 10 minutes at room temperature. The cells were probed with primary antibodies in BD perm/wash buffer at 4°C overnight. The cells were washed twice with BD perm/wash buffer and incubated with fluorophore-conjugated secondary antibodies in BD perm/wash buffer at room temperature for 1 hour. The cells were washed twice with BD perm/wash and then resuspended in PBS. The cells were acquired on LSRFORTESSA™ X20 CELL ANALYZER and analysed by FlowJo software. Targets labelled with Alexa Fluor-488 and Alexa Fluor-647-conjugated antibodies

were analysed using a Texas Red, Fluorescein and Cy5 filter set, respectively. mScarlet fluorescent signals were acquired using a mCherry filter set.

For live cell analysis, after cell dissociation and washes, the cells were re-suspended in PBS supplemented with a chilled plating medium without phenol red. The cells were acquired on FX500 cell sorter (Sony) and analysed by FlowJo software.

2.2.15 Contractility measurement for iPSC-CMs

Approximately 1 x 10^5 iPSC-CMs were seeded onto a 3.5-cm glass bottom culture dish and cultured for at least 7 days prior to measurement. The cells were transduced with 1:10,000 diluted (from the stock) (unknown multiplicity of infection, unpublished) adenoviral vectors harbouring a *mScarlet-ACTN2* cassette for 1 hr and the medium was changed to plain RPMI-1640 supplemented with B27™. The *mScarlet-ACTN2* adeno viruses were a gift from Dr Violetta Steeples. 48 hr post treatment, the medium was changed to Tyrode's-HEPES solution. Sarcomere contraction behaviour videos were acquired by IX81 microscope (Olympus) using a 100x oil-immersion lens and a RFP filter set at 512*288 pixel^2 resolution at 37°C at 50 frames per second. Sarcomere contractility was analysed using SarcTrack scripts (Toepfer et al., 2019) in collaboration with Professor Chris Toepfer and Dr Yiangos Psaras.

2.2.16 Ca^{2+} transient measurement for iPSC-CMs

Approximately 1 x 10^5 iPSC-CMs were seeded onto a 3.5 glass bottom culture dish and cultured for at least 7 days prior to measurement. The cells were transduced with adenoviruses carrying a red genetically encoded calcium indicator (*RGECO*) (multiplicity of infection = 50) overnight and the medium was changed to plain RPMI-1640 supplemented with B27™. The *RGECO* adeno viruses were a gift from Dr Alex Sparrow (Sparrow et al., 2019). 48 hr post treatment, the medium was changed to Tyrode's-HEPES solution. Ca^{2+} transient videos were acquired by IX81 microscope using a 100x oil-immersion lens and an RFP filter set at 512*288 pixel^2 resolution at 37°C at 50 frames per second. Ca^{2+} transient

parameters were analysed using CalTrack scripts (Psaras et al., 2021) run on MatLab (version R2021b)

2.2.17 Statistical analysis

All statistical analyses were performed using GraphPad Prism software (version 9.3.0). For a two-group comparison, the Student t-test was used on normal distributed data, otherwise, the Mann-Whitney U-test was used. For testing more than two treatments modelled by normal distribution, ordinary one-way ANOVA was used with Tukey's post-hoc analysis. The non-parametric Kruskal-Wallis test was used with post-hoc Dunn correction only when data were not satisfied by one-way ANOVA assumptions.

Chapter 3

CRISPR/Cas9-mediated generation of CM subtype-specific fluorescent iPSCs

3.1 Introduction

The population heterogeneity in iPSC-CMs poses a limitation in cardiovascular research. Fluorescent reporter systems made specific to each iPSC-CM subtype, including atrial-like, ventricular-like and nodal-like cells may be useful for detection and purification of these subtypes, providing a cell source for a subtype-specific disease study. In order to produce CM subtype-specific fluorescent reporter lines, a fluorophore construct can be inserted downstream of CM subtype-specific promoter. Specifically, the work described in this chapter aims to replace the stop codon of target genes with a fluorophore preceded by a P2A self-cleaving peptide.

The self-cleavage peptide P2A is commonly used to promote bicistronic expression, especially to establish fluorescent reporter lines. In this chapter, a GSG sequence was added at the N-terminus of P2A to maximise the cleavage efficiency (Wang et al., 2015). The P2A mechanism is proposed as ribosomal skipping where the ribosome pauses and releases a nascent peptide when reaching the last proline codon of P2A (Kim et al., 2011, Liu et al., 2017) (Figure 3-1). Subsequently, the translation recommences at this proline site toward the following gene/open reading frame. This results in addition of the majority of the P2A oligo peptide at the C-terminus of the upstream open reading frame and a single proline residue at the N-terminus of the downstream open reading frame.

As described in Chapter 1, *MYL7* and *PITX2* are specific markers for adult atrial CMs. In addition, *MYL2* and *HCN4* are specific markers for adult ventricular and nodal CMs, respectively. This chapter aims to develop fluorescent reporter iPSC lines specific for *MYL7*, *PITX2*, *MYL2* and *HCN4* using CRISPR/Cas9 technology. Each of these reporter lines were generated using different targeting strategies and different CRISPR/Cas9 delivery methods

depending on specific CRISPR/Cas9 targeting efficiency and complexity of the structure of the target genes.

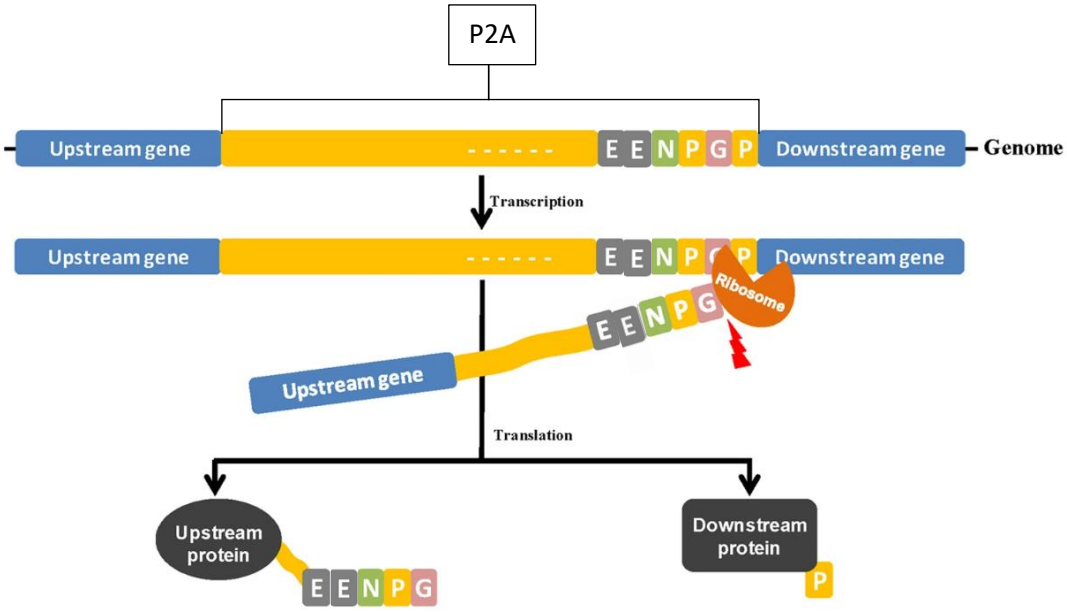


Figure 3-1 Proposed mechanism of self-cleavage P2A peptide. When the translation machinery reach the last proline codon of P2A (cleavage site), it halts and frees the nascent peptide containing the majority of the P2A sequence. The translation resumes and continues to translate the downstream gene containing the last proline residue of P2A. Modified from (Wang et al., 2015)

3.2 gRNA and targeting donor design

CRISPR-targetable sequences were screened *in silico* near the stop codon of each target gene. Target sequences were selected based on the proximity to the stop codon and the predicted maximum specificity with minimum off-target scores. The respective gRNAs containing the selected CRISPR target sequence were *in vitro* transcribed and purified before being reconstituted with EnGen® Spy Cas9 NLS, generating a RNP preparation. Subsequently, the CRISPR/Cas9 RNP was electroporated into KOLF2-C1 iPSCs and 48 hr post transfection, the cells were harvested for gDNA extraction. Approximately 500-bp PCR amplicons covering the CRISPR/Cas9 target regions were amplified and sequenced, allowing an assessment of the CRISPR-related indel rate by http://shinyapps.datacurators.nl/tide/ (Brinkman et al., 2014). The indel or tracking of indel by decomposition (TIDE) score determined how efficient the respective gRNAs cut the target genomic sequence.

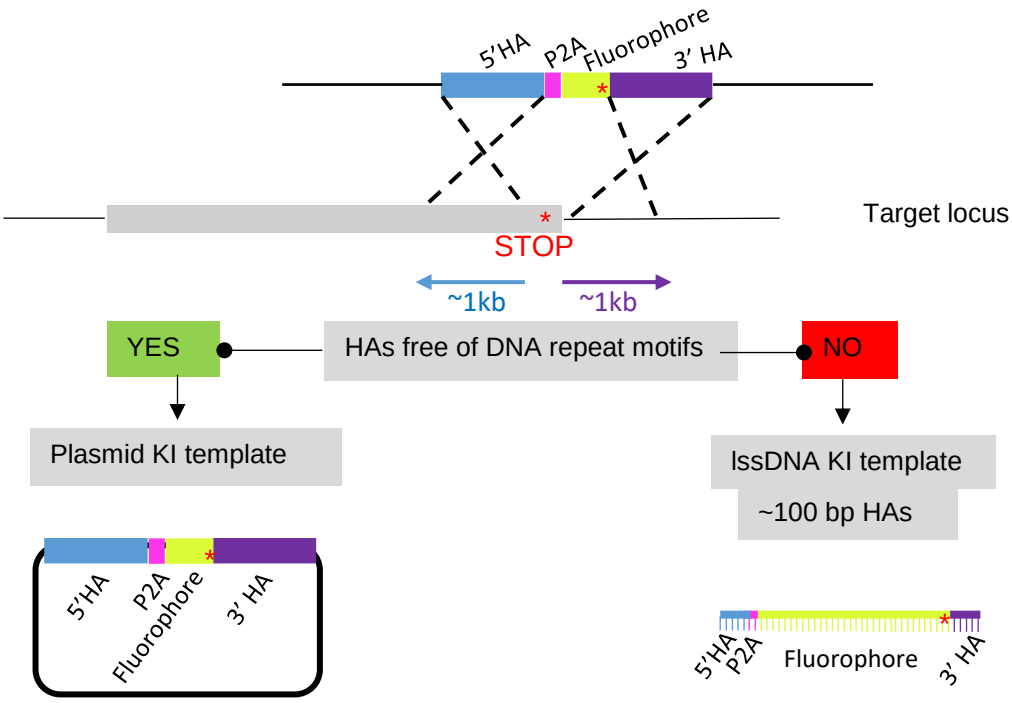


Figure 3-2 Targeting constructs for genomic insertion of a promoterless fluorescent reporter at a stop codon. For those targets showing a low degree of repetitive sequences up to 1kb either upstream or downstream of a stop codon (referred to as homology arms (HA) in a targeted knock-in (KI) context), a targeting construct is designed in a form of plasmid. For a locus where unique HAs cannot be extended to 1kb, a lssDNA template with approximately 100 bp HAs is preferred.

For conventional targeted knock-in of a large insert e.g. fluorophore, a plasmid with homology arms flanking the desired change serves as a targeting construct. With CRISPR/Cas9-assisted gene targeting, homology arms can be relatively modest in size; a typical design involves the sequence to be inserted being flanked by repetitive motif-free sequence up to 1 kb (Figure 3-2). As an alternative, homology arms can be shortened to approximately 100 bp when using lssDNA templates (Miura et al., 2018).

3.2.1 *MYL7-mClover* line

3.2.1.1 MYL7 gRNA efficiency test

3 different MYL7 gRNAs, which harboured 20-nt CRISPR sequences complementary to the *MYL7* locus, were synthesised and tested for their efficiency. CRISPR-A (2.2% TIDE) and CRISPR-B (1.3 % TIDE) showed minimal activity at the *MYL7* locus. Only CRISPR-C showed a modest TIDE score (14.6%) (Figure 3-3C).

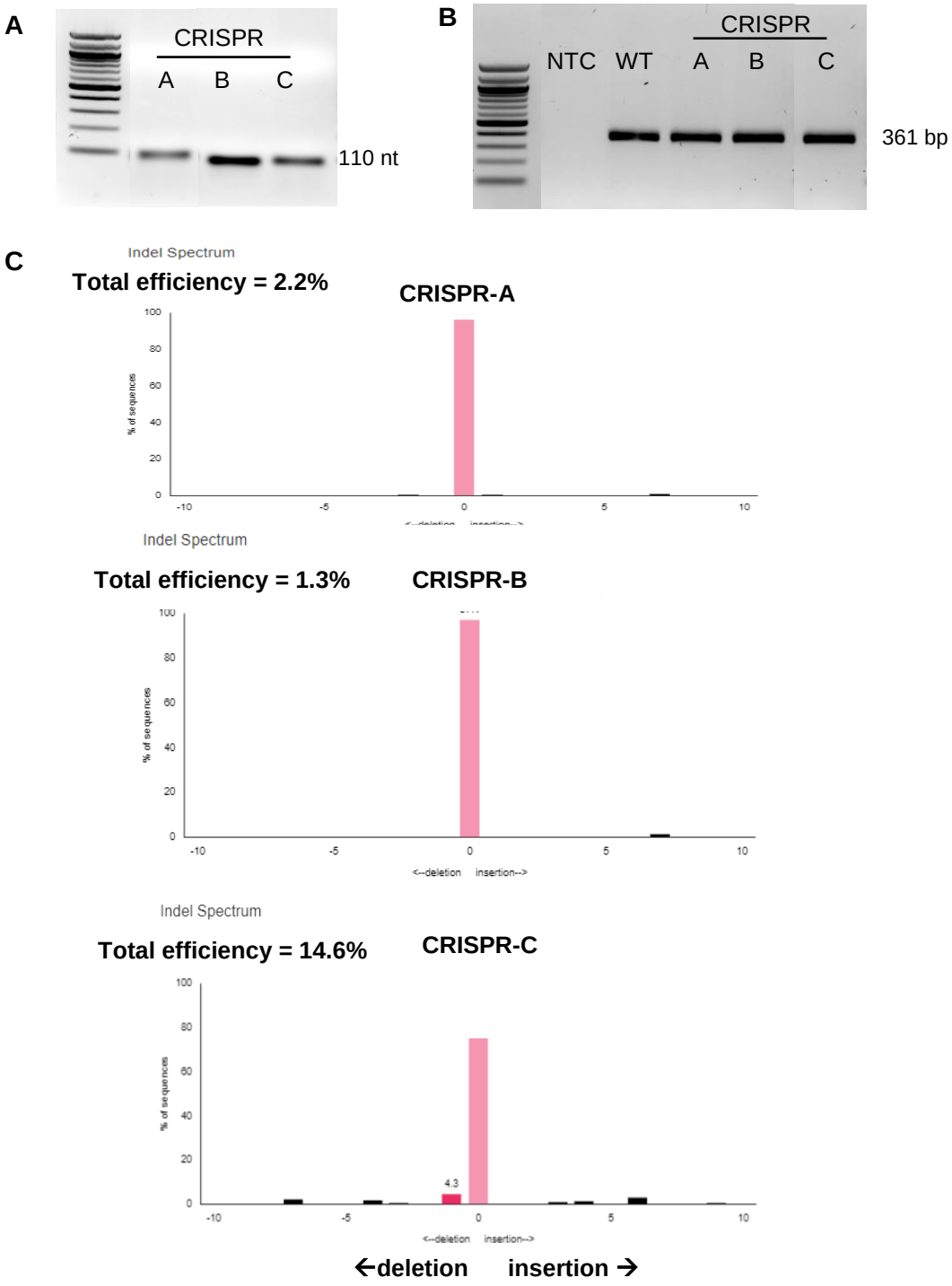


Figure 3-3 TIDE analysis for CRISPRs targeting the *MYL7* locus. A) *In vitro* transcribed gRNAs in a heat-denatured form containing the CRISPR A, B and C sequences. B) PCR products 48 hr after CRISPR/Cas9 transfection into KOLF2-C1 iPSCs using the primers MYL7-F2/R2 which amplified the sequence flanking the target site. These PCR products were sequenced and analysed for CRISPR activity by TIDE as compared to the non-transfected WT control. C) TIDE score for each CRISPR design. NTC, no template control.

Since CRISPR-C presented the highest TIDE score, it was selected for the targeted insertion of the fluorescent reporter into the *MYL7* locus. The *MYL7* locus displays only a few repeat motifs surrounding the stop codon (Figure 3-4), so the donor template was designed in a form of plasmid.

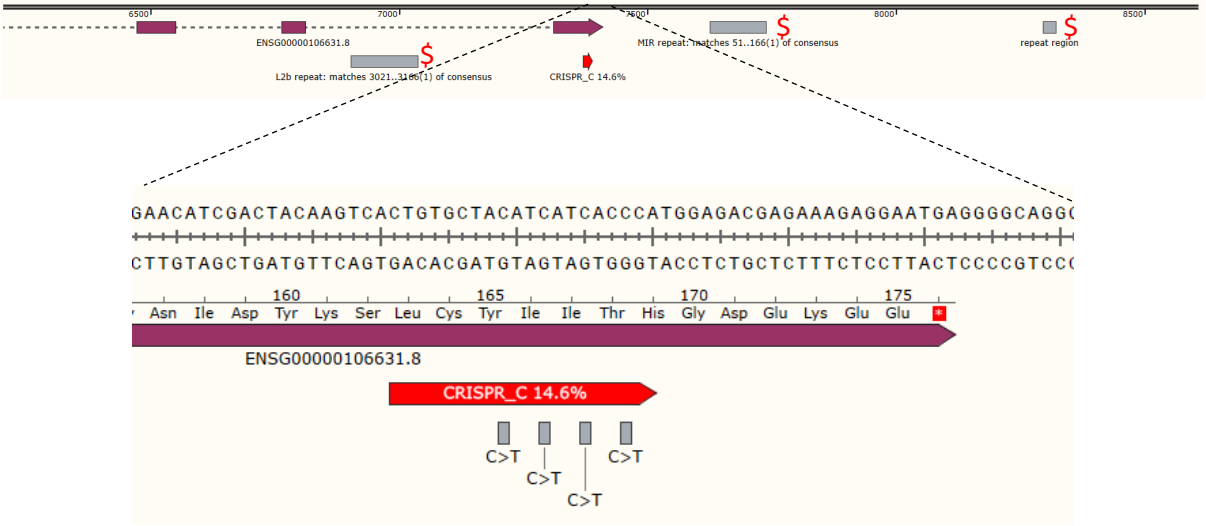


Figure 3-4 MYL7 CRISPR-C location relative to the *MYL7* locus. The features (C>T) within the CRISPR-C region hightlight the intentional silent mutations introduced into the 5' HA of the targeting vector required for the *MYL7-mClover* line generation. Approximately 1000 bp upstream and 1000 bp downstream of the stop codon contain 2 relatively large and 1 relatively small repeative motifs($).

3.2.1.2 *MYL7-mClover* targeting vector cloning

The targeting plasmid (Figure 3-5A) was initially synthesised by Gibson Assembly from 4 DNA fragments. The first fragment (3111 bp) (Figure 3-5B & C) was generated by linearising the destination plasmid V924-neo-dTA (Figure S1) via *Xho*I and *Hind*III digestion. The dTA or diphtheria toxin A is a negative selection marker which helps exclude random genomic integration events. The second fragment, the 5' Homology Arm (5' HA; 1130 bp), and the third fragment, the 3' Homology Arm (3' HA; 1338 bp) (Figure 3-5B & C) were amplified by PCR from KOLF2-C1 gDNA using the primers MYL7-5HR-F/R and MYL7-3HR-F/R2, respectively. The 5' HA fragment contained both the CRISPR recognition site and the intact PAM site. In order to prevent recleavage of the successfully targeted allele by the CRISPR/Cas9, either the target's PAM sequence or the CRISPR recognition sequence itself needs modification. As no silent mutations could be made at the PAM site, 4 silent mutations (all C>T) were introduced

onto the CRISPR region (Figure 3-4 & 3-5A). The fourth fragment P2A-mClover (846 bp) (Figure 3-5B & C) was amplified by PCR from the synthesized *MYL2*-P2A-*mClover* gBlock template (Figure S7) using the primers MYL7-mClover-F/R. The successful Gibson Assembly products were screened by *PstI* digestion (4317 and 2007 bp) (Figure 3-5D) before being verified by DNA sequencing. Since the successful clone N3 showed no undesired mutations, it was selected for the further cloning step.

Given the relatively low activity of the selected CRISPR/Cas9 at the *MYL7* locus, to ensure stable genomic integration, a selection cassestte *Neo* was introduced into the plasmid clone N3. This modification was achieved in 2 steps. Firstly, the annealed oligo linker (NS-TV-link1/2) containing *Bam*HI-*Bsr*GI-*Apa*I-*Xho*I restriction sites (26 bp) was inserted into the *Xho*I-linearised plasmid clone N3 (6324 bp) (Figure 3-5E). After *Xho*I and *Nco*I digestion (the correct pattern: 2065, 1371, 1241, 1019 and 651 bp) (Figure 3-5F) and sequence verification, the resultant plasmid clone L3 was selected for the next step. Secondly, the floxed *Neo* cassette fragment (1911 bp) (Figure 3-5G, right) obtained from the plasmid PL452 (Figure S2) previouly digested with *BamH*I and *Sal*I was cloned into the linearised plasmid clone L3 (6330 bp) (Figure 3-5G, left) which was prepared by *Bam*HI and *Xho*I digestion. The final targeting plasmid (8241 bp) should have 7 products (3299, 2014, 2007, 525, 261, 113 and 22 bp) when analysed by *Pst*I digestion. With the correct digestion pattern and verified sequence, the plasmid clone S3 (Figure 3-5H), which was later called the *MYL7-mClover* targeting vector, was selected for generating the *MYL7-mClover* reporter line.

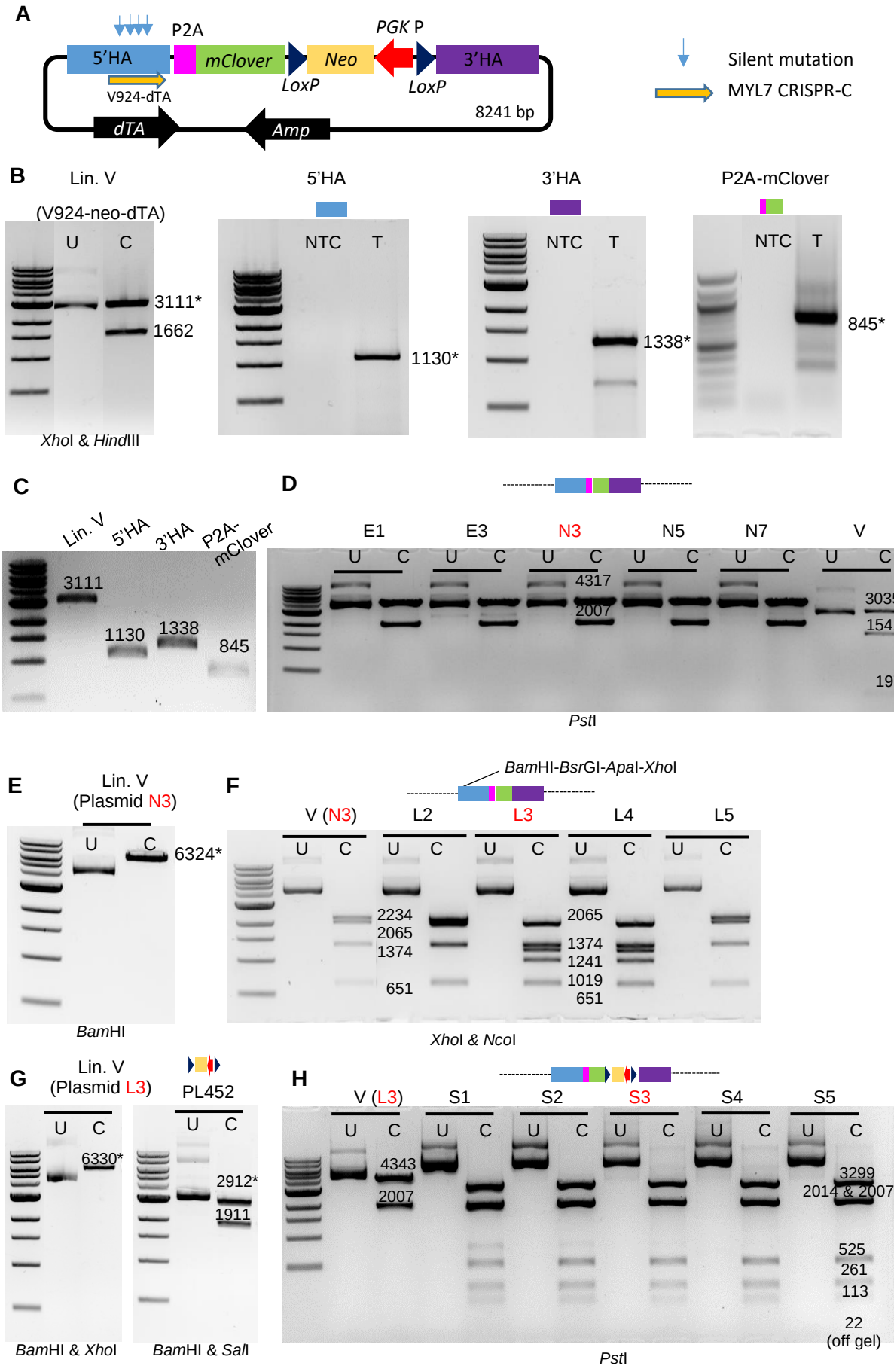

A
P2A
5'HA
mClover
Neo
PGK P
3'HA
LoxP
LoxP
V924-dTA
8241 bp
dTA
Amp
Silent mutation
MYL7 CRISPR-C
B
Lin. V
(V924-neo-dTA)
5'HA
3'HA
P2A-mClover
U
C
NTC
T
3111*
1662
1130*
1338*
845*
XhoI & HindIII
C
Lin. V
5'HA
3'HA
P2A-mClover
3111
1130
1338
845
D
E1
E3
N3
N5
N7
V
4317
2007
3035
1541
197
PstI
E
Lin. V
(Plasmid N3)
6324*
BamHI
F
BamHI-BsrGI-ApaI-XhoI
V (N3)
L2
L3
L4
L5
2234
2065
1374
651
2065
1374
1241
1019
651
XhoI & NcoI
G
Lin. V
(Plasmid L3)
PL452
6330*
2912*
1911
BamHI & XhoI
BamHI & SalI
H
V (L3)
S1
S2
S3
S4
S5
4343
2007
3299
2014 & 2007
525
261
113
22
(off gel)
PstI

Figure 3-5 *MYL7-mClover* targeting vector cloning. A) A simplified illustration of the *MYL7-mClover* plasmid construct labelled with 4 silent mutations, which were introduced to prevent recleavage of the targeted allele. The targeting plasmid was initially generated by Gibson Assembly to combine 4 DNA fragments (B-D). B) Preparation of Gibson Assembly fragments. C) Purified DNA fragments from B. D). Diagnostic screening for successful Gibson Assembly (N3). The oligo linker (*Bam*HI-*Bsr*GI-*Apa*I-*Xho*I) was introduced into E) the *Bam*HI-linearised fragment from clone N3 F) Diagnostic screening for successful linker cloning (L3). G) A floxed *Neo* fragment and linearised L3 fragment preparation. H) Diagnostic screening for successful combination of the fragments in G (S3, final form of the targeting vector). *Amp*, ampicillin resistant gene; *PGK* P, phosphoglycerate kinase promoter; U, uncut; C, cut; NTC, no template control; T, with template; V, vector; Lin. V, linearised vector; *, fragments used for relevant cloning; red, relevant clone.

3.2.2 *HCN4-mScarlet* line

3.2.2.1 HCN4 gRNA efficiency test

2 different gRNAs, which harboured 20-nt CRISPR sequences complementary to the *HCN4* locus, were synthesised and tested for their efficiency. CRISPR-A (Figure 3-6C) was selected for generating the *HCN4-mScarlet* reporter line due to its higher TIDE score (38.21%); CRISPR-B yielded 4.7% TIDE.

Approximately 1000 bp upstream and 1000 bp downstream of the stop codon contain only a few repeat motifs (Figure 3-7), and subsequently the donor template was designed in a form of a plasmid with these lengths of homology.

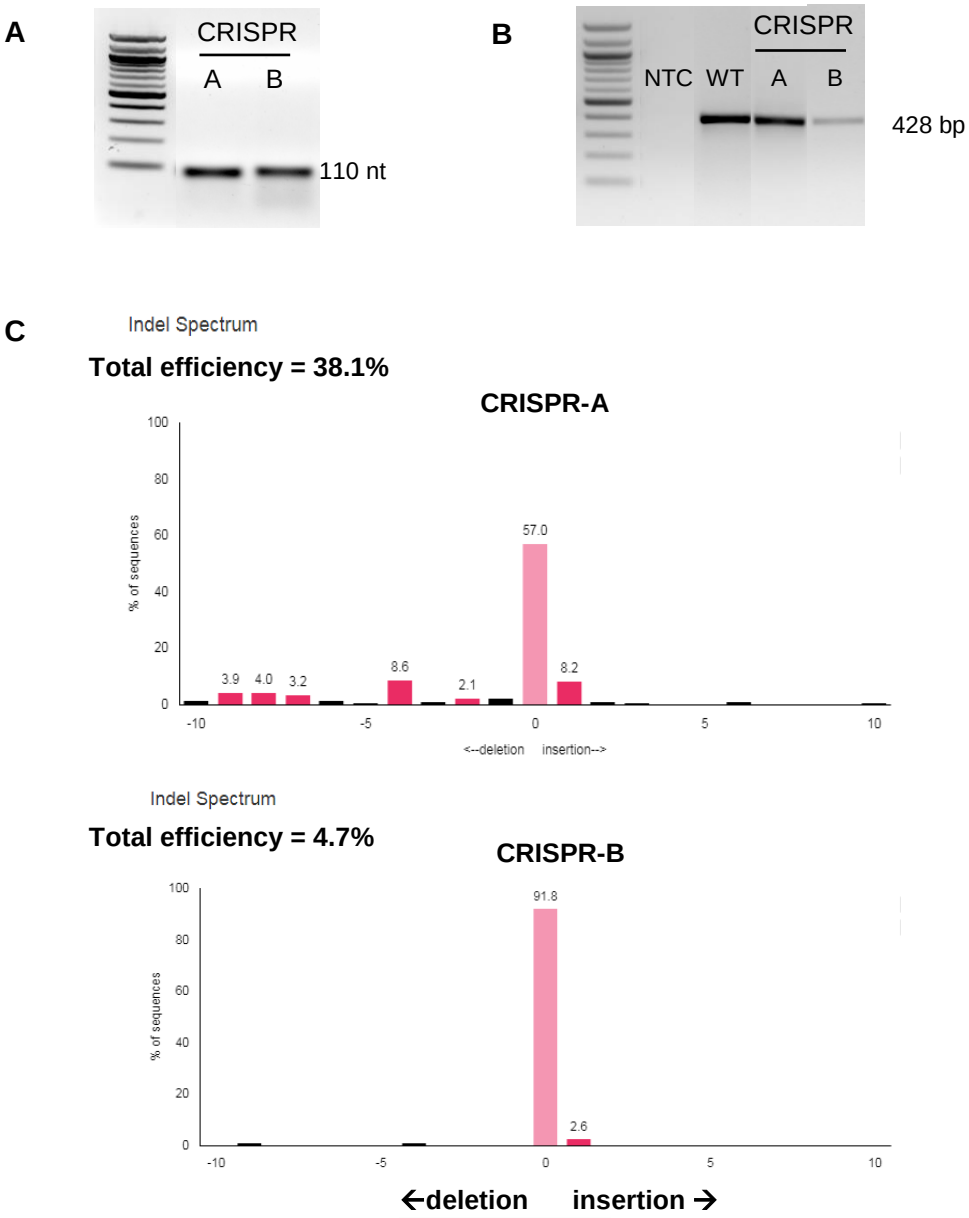


Figure 3-6 TIDE analysis for CRISPRs targeting the *HCN4* locus. A) *In vitro* transcribed gRNAs in a heat-denatured form containing the CRISPR A and B sequences. B) PCR products 48 hr after CRISPR/Cas9 transfection into KOLF2-C1 iPSCs using the primers HCN4/R which amplified the sequence flanking the stop codon. These PCR products were sequenced and analysed for CRISPR activity by TIDE as compared to the non-transfected WT control. C) TIDE score for each CRISPR design. NTC, no template control.

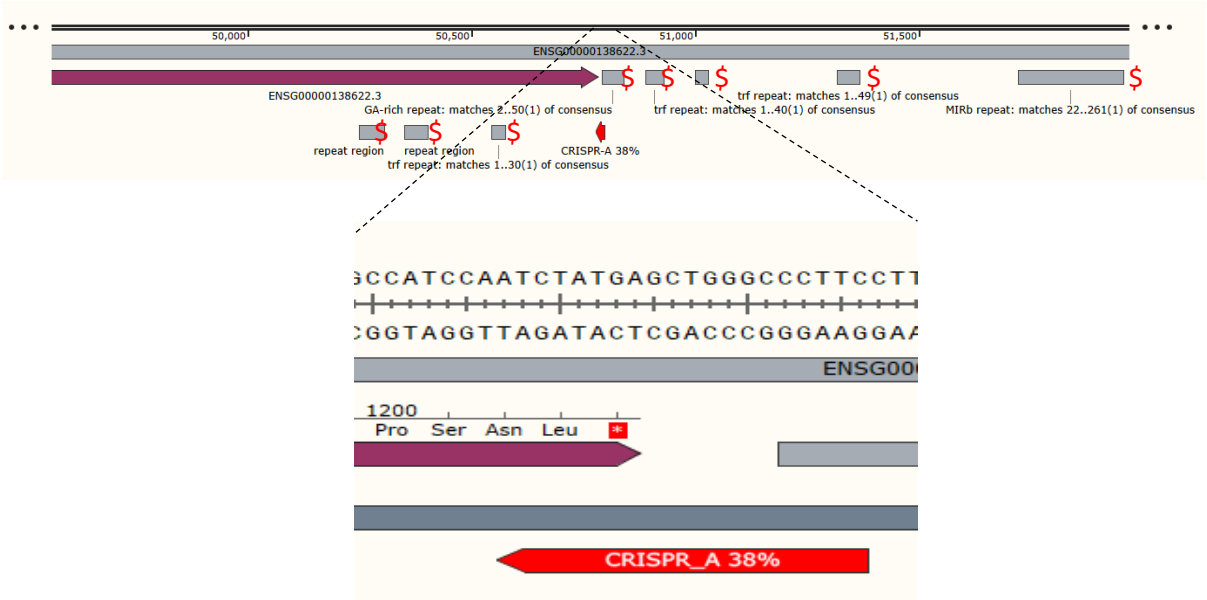


Figure 3-7 HCN4 CRISPR-A location relative to the *HCN4* locus. Approximately 1000 bp upstream and 1000 bp downstream of the stop codon contain only 7 relatively small and 1 relatively large repetitive motifs ($).

3.2.2.2 *HCN4-mScarlet* targeting vector cloning

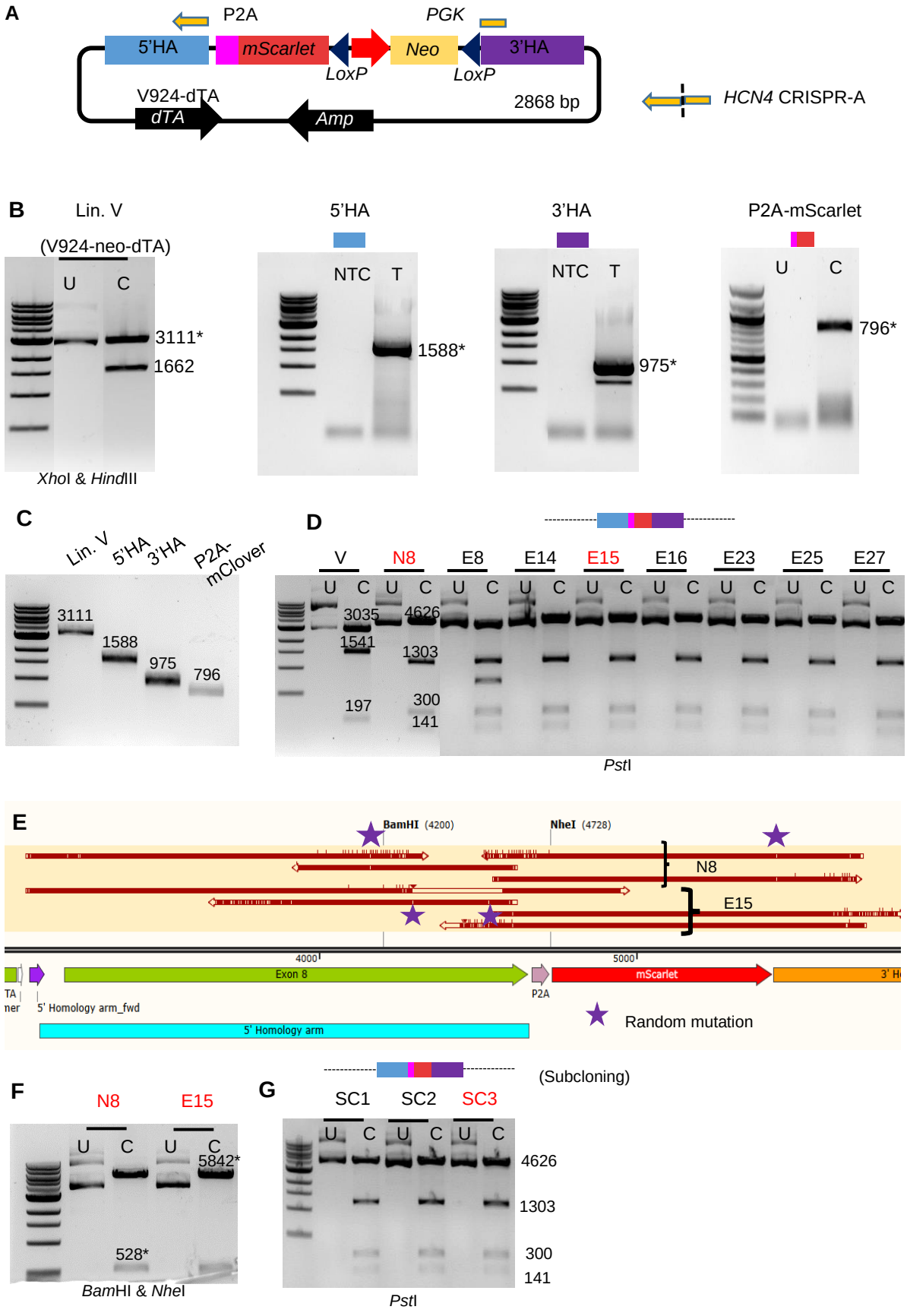
The targeting plasmid (Figure 3-8A) was initially synthesised by Gibson Assembly from 4 DNA fragments. The first fragment (3111 bp) (Figure 3-8B & C) was generated by linearising a destination plasmid V924-neo-dTA via *Xho*I and *Hind*III digestion. The second fragment 5' HA (1588 bp) and the third fragment 3' HA (975 bp) (Figure 3-8B & C) were amplified by PCR from KOLF2-C1 gDNA using the primers HCN4-5HR-F2/R2 and HCN4-3HR-F2/R2, respectively. Since the HCN4 CRISPR-A recognition site was interrupted by the fluorophore insertion in the final targeting vector, no sequence modifications on the targeting construct were required to prevent recleavage of the targeted allele (Figure 3-8A). The fourth fragment P2A-mScarlet (796 bp) (Figure 3-8B & C) was amplified by PCR from the plasmid 18ACH20P_2368293 (Figure S5) using the primers HCN4-mScarlet-F2/R2. The successful Gibson Assembly products were screened by *PstI* digestion with the correct products being 4626, 1303, 300 and 141 bp (Figure 3-8D).

Despite successful assembly, DNA sequencing showed random mutations in all of the clones generated by Gibson Assembly. Interestingly, the clones N8 and E15 had mutations in

different locations which then allowed a subcloning strategy to be developed to correct the mutations (Figure 3-8E). The clone E15 contained no mutations except the two within the *Nhe*I and *Bam*HI flanked region whereas the same region was free of undesired mutations in the clone N8. The small fragment (528 bp) made by *Nhe*I and *Bam*HI digestion of the clone N8 was therefore subcloned into the large fragment (5842 bp) from the clone E15 digested with the same restriction enzymes (Figure 3-8F). The resultant plasmid clone SC3 with the correct *Pst*I digetion pattern (4626, 1303, 300 and 141 bp) (Figure 3-8G) and verified sequence was slected for the next cloning steps.

To ensure a stable genomic integration, a selection cassestte *Neo* was introduced into the plasmid clone SC3. This modification can be divided into 3 steps. Firstly, since there was no unique restriction site which could be used to linearlise the clone SC3 between the *mScarlet* and the 3'HA region, the 3'HA was initially removed by *Bsr*GI *and Hind*III digestion (Figure 3-8H). The 3'HA-free fragment (5426 bp) was then used as a vector for cloning the annealed oligo linker (NS-TV-link3/4) containing *Mfe*I-*Spe*I-*Psp*OMI-*Apa*I restriction sites (26 bp). The resultant clone Q1 with the correct *Mef*I and *Xho*I digestion pattern (3135 and 2322 bp) (Figure 3-8I) and verified sequence was selected for the next step.

Secondly, the floxed *Neo* cassette frament A (1543 bp) (Figure 3-8K, left) and fragment B (352 bp) (Figure 3-8K, right) obtained by A) *Eco*RI and *Sac*I and B) *Sac*I and *Spe*I digestion of the plasmid PL452, respectively were cloned into the *Spe*I and *Mfe*I linealised fragment of the plasmid clone Q1 (5449 bp) (Figure 3-8J). The resultant clone R9 with *Ase*I digestion confirmation (3490, 2057 and 1752 bp) (Figure 3-8L) and verified sequence was selected for the next step.


A
P2A
PGK
5'HA
mScarlet
Neo
3'HA
LoxP
LoxP
V924-dTA
dTA
Amp
2868 bp
HCN4 CRISPR-A
B
Lin. V
(V924-neo-dTA)
U
C
3111*
1662
XhoI & HindIII
5'HA
NTC
T
1588*
3'HA
NTC
T
975*
P2A-mScarlet
U
C
796*
C
Lin. V
5'HA
3'HA
P2A-mClover
3111
1588
975
796
D
V
N8
E8
E14
E15
E16
E23
E25
E27
U C U C U C U C U C U C U C U C U C U C
3035
4626
1541
1303
197
300
141
PstI
E
BamHI (4200)
NheI (4728)
N8
E15
4000
5000
Exon 8
mScarlet
P2A
5' Homology arm_fwd
5' Homology arm
Random mutation
F
N8
E15
U C U C
5842*
528*
BamHI & NheI
G
(Subcloning)
SC1
SC2
SC3
U C U C U C
4626
1303
300
141
PstI

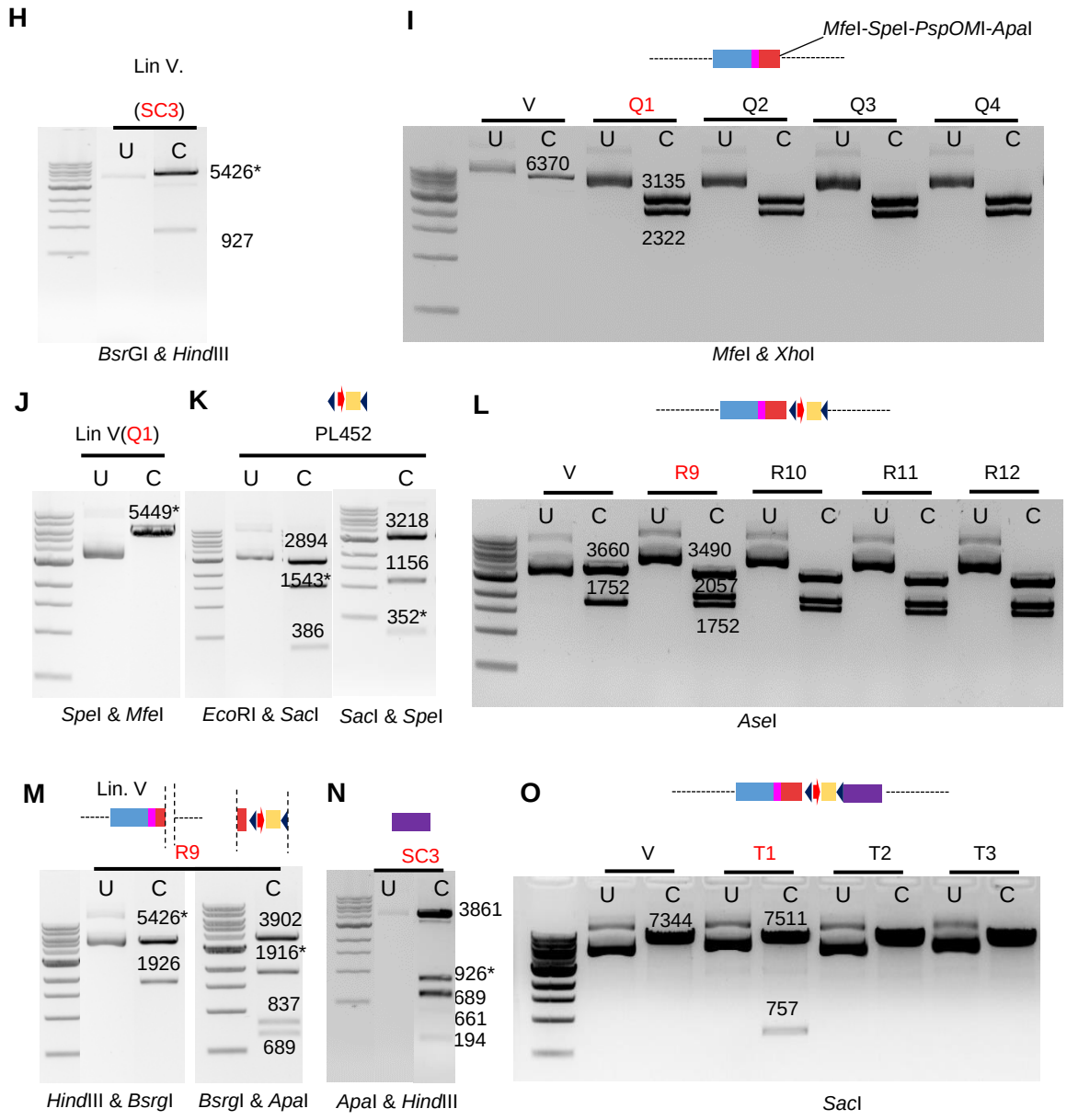


Figure 3-8 *HCN4-mScarlet* targeting vector cloning. A) A simplified illustration of the *HCN4-mScarlet* plasmid construct showing the CRISPR-A recognition site split between the 5'HA and the 3'HA. The targeting plasmid was initially generated by Gibson Assembly to combine 4 DNA fragments (B-D). B) Preparation of Gibson Assembly fragments. C) Purified DNA fragments from B. D). Diagnostic screening for successful Gibson Assembly. E) Products of Gibson Assembly in the clones N8 and clone E15 contained different mutational sites, allowing (F-G) subcloning to correct the mutations. F) Preparation of fragment for subcloning. G) Diagnostic screening for successful subcloning (SC3). H) Preparation of the clone SC3 for *Mfe*I-*Spe*I-*PspOM*I-*Apa*I linker cloning. I) Diagnostic screening for successful linker cloning (Q1). J-K) Preparation of fragments for *Neo* cloning. L) Diagnostic screening for successful *Neo* cloning (R9 containing no 3' HA). M-N) Preparation of fragments for reintroduction of the 3'HA.O) Diagnostic screening for the final form of the targeting vector (T1). *Amp*, ampicillin resistant gene; *PGK* P, phosphoglycerate kinase promoter; U, uncut; C, cut; NTC, no template control; T, with template; V, vector; Lin. V, linearised vector; *, fragments used for relevant cloning(s); purple stars, undesired mutations.

Thirdly, the fragment C (5426 bp) (Figure 3-8M, left) and the fragment D (1916 bp) (Figure 3-8M, right) obtained by C) *Hind*III and *Bsrg*I and D) *Bsrg*I and *Apa*I digestion of the plasmid clone R9, respectively were integrated with the 926-bp fragment generated by *Apa*I and *Hind*III digestion of the plasmid clone SC3. The final targeting vector (8268 bp) should have 2 products (7511 and 757 bp) when analysed by *Sac*I digestion. With the correct digstion pattern and verified sequence, the plasmid clone T1 (Figure 3-8O), which was later called the *HCN4-mScarlet* targeting vector, was selected for generating the *HCN4-mScarlet* reporter line.

3.2.3 *PITX2-mClover* reporter line

3.2.3.1 PITX2 gRNA efficiency test

2 different gRNAs, which harboured 20-nt CRISPR sequences complementary to the *PITX2* locus, were synthesised and tested for their efficiency. CRISPR-A yielded 14.3% TIDE (Figure 3-9C) and CRISPR-B had minimal activity (4.3% TIDE). Due to the higher efficiency, the CRISPR-A was selected for generating the *PITX2-mClover* reporter line.

Approximately 1000 bp upstream and 1000 bp downstream of the stop codon contain only a few, small repeat motifs (Figure 3-10), subsequently the donor template was designed in a form of a plasmid with these homology regions.

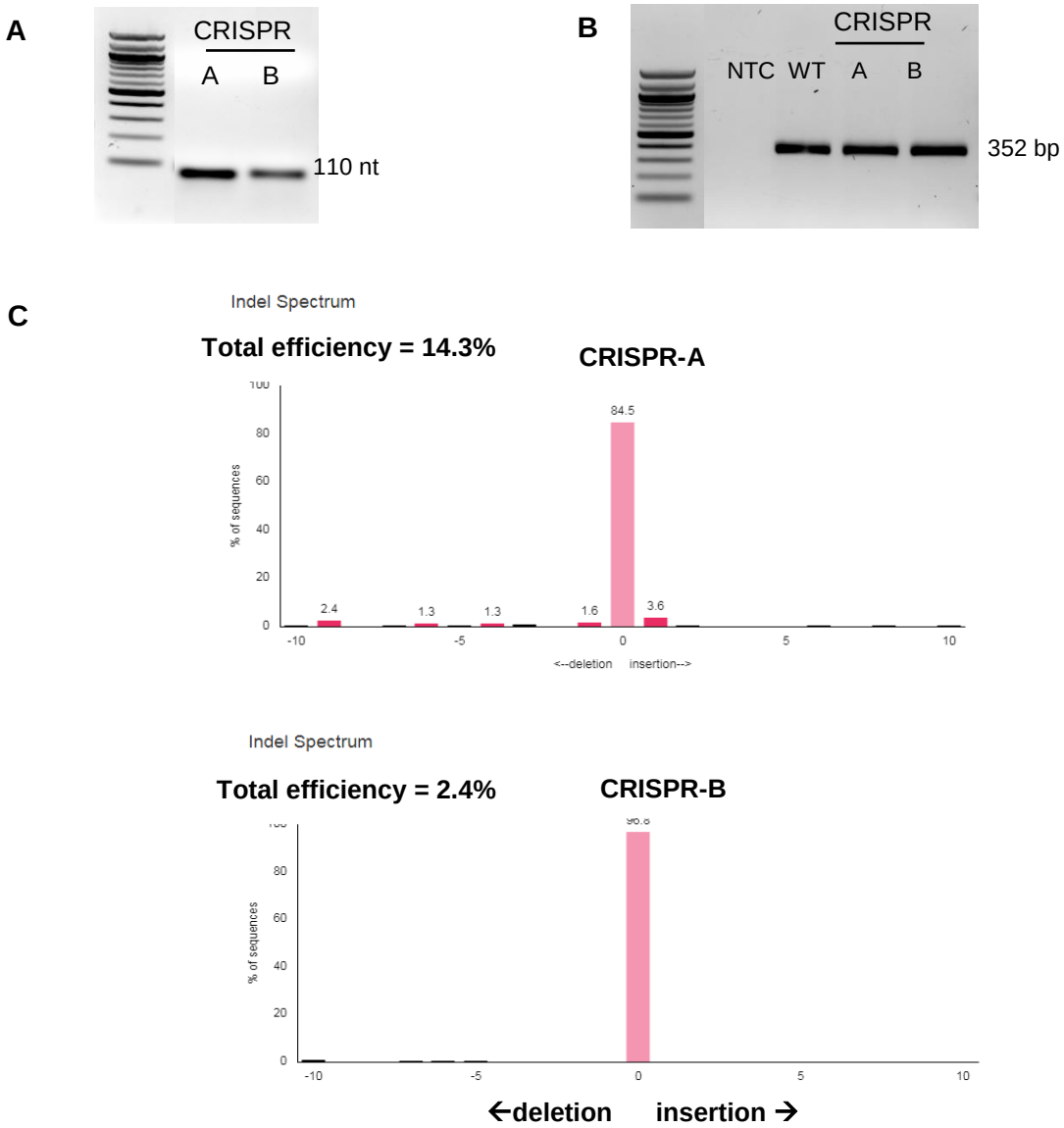


Figure 3-9 TIDE analysis for CRISPRs targeting the *PITX2* locus. A) *In vitro* transcribed gRNAs in a heat-denatured form containing the CRISPR A and B sequences. B) PCR products 48 hr after CRISPR/Cas9 transfection into KOLF2-C1 iPSCs using the primers PITX2-F/R which amplified the sequence flanking the stop codon. These PCR products were sequenced and analysed for CRISPR activity by TIDE as compared to the non-transfected WT control. C) TIDE score for each CRISPR design. NTC, no template control.

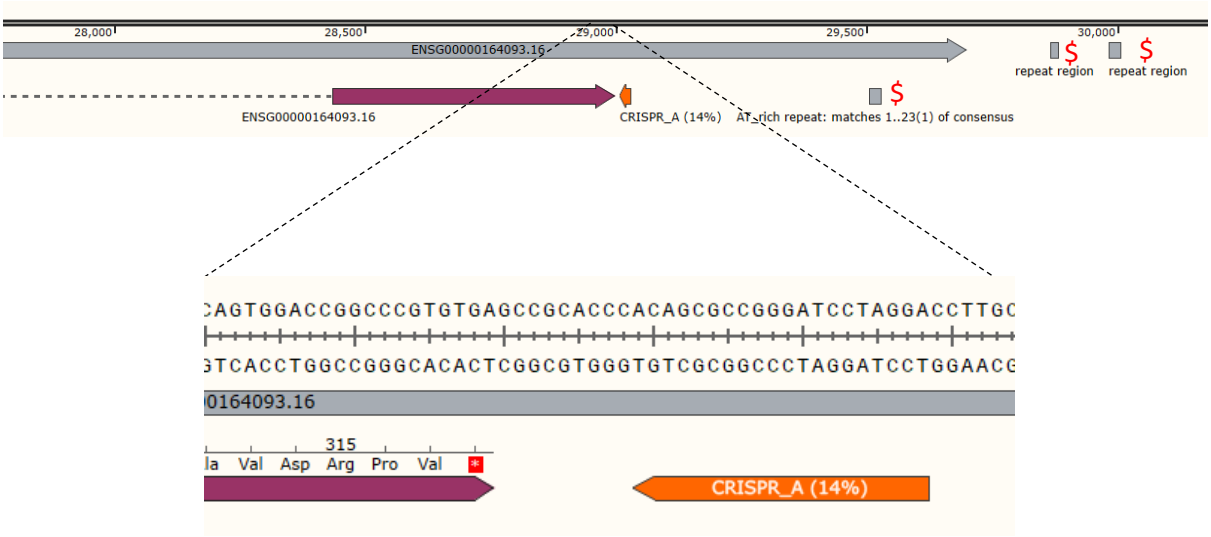


Figure 3-10 PITX2 CRISPR-A location relative to the *PITX2* locus. Approximately 1000 bp upstream and 1000 bp downstream of the stop codon contain only 3 small repeat motifs ($).

3.2.3.2 *PITX2-mClover* targeting vector cloning

The targeting vector was initially designed to be generated by Gibson Assembly of 4 fragments (Figure 3-11B). The first fragment (3111 bp) was generated by linearising the destination plasmid V924-neo-dTA via *Xho*I and *Hind*III digestion. The second fragment 5' HA (1475 bp) and the third fragment 3' HA (999 bp) were amplified by PCR from KOLF2-C1 gDNA using the primers PITX2-5HR-F2/R and PITX2-3HR-F/R2, respectively. The fourth fragment P2A-mClover (845 bp) was amplified by PCR from the *MYL2*-P2A-*mClover* gBlock template using the primers PITX2-mClover-F/R.

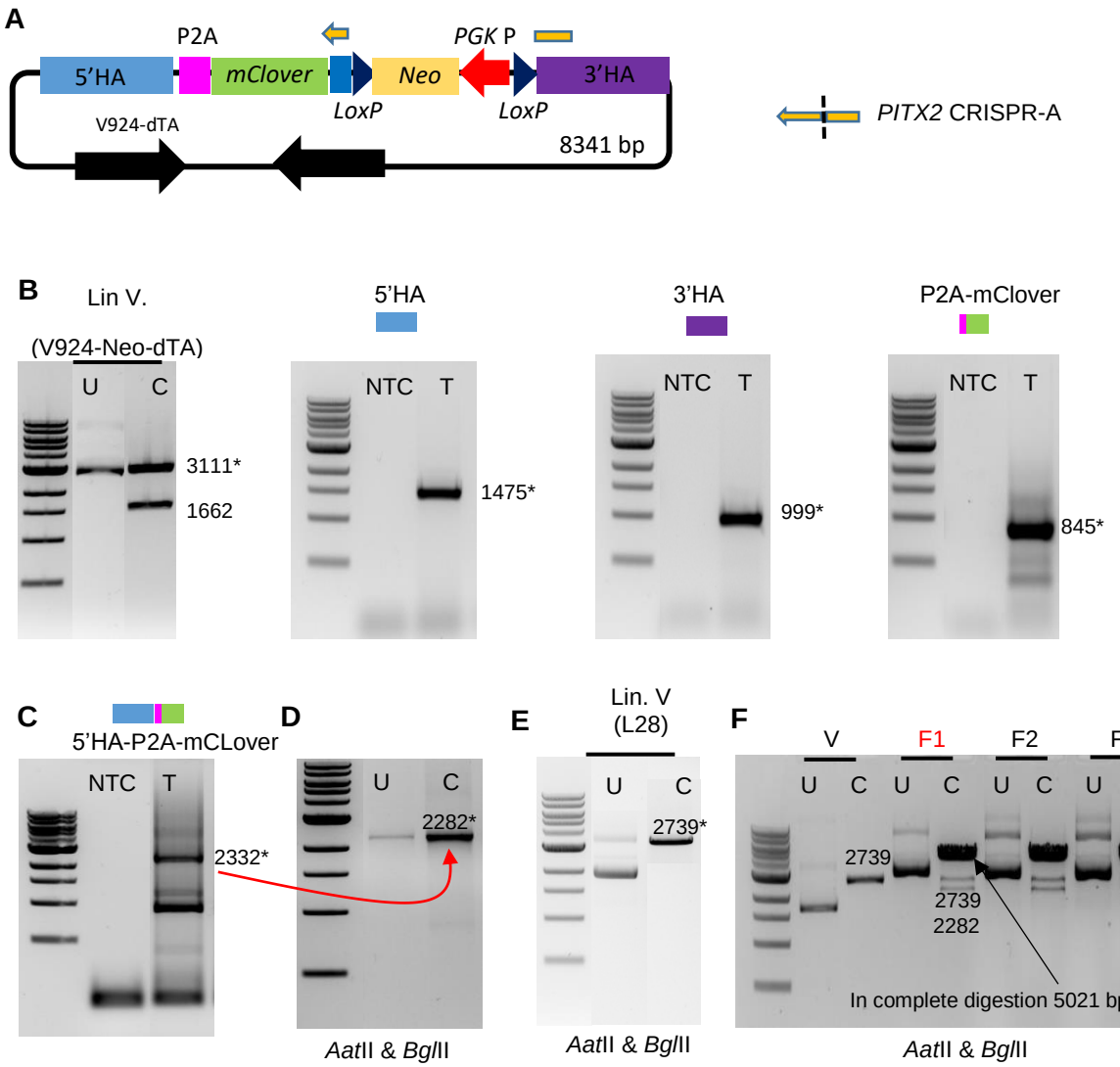
However, after several attempts, the assembly was unsuccessful. The plan for the targeting vector construction was then simplified by joining each pair of the Gibson Assembly fragments. The 5'HA fragment was combined with the P2A-mClover fragment by fusion PCR (2332 bp) (Figure 3-11C) using the primer PITX2-5HA-F2 and PITX2-mClover-R. The PITX2-5'HA-P2A-mClover fusion product was then trimmed by *Aat*II and *Bgl*II (2282 bp) (Figure 3-11D) before cloned into the cloning vector Litmus28 (L28) (Figure S3), previously digested with the same restriction enzymes (2739 bp) (Figure 3-11E). The succesful PITX2-5HA-P2A-mClover cloning was confimed by the *Aat*II and *Bgl*II digestion pattern (2739 and 2282 bp) and DNA

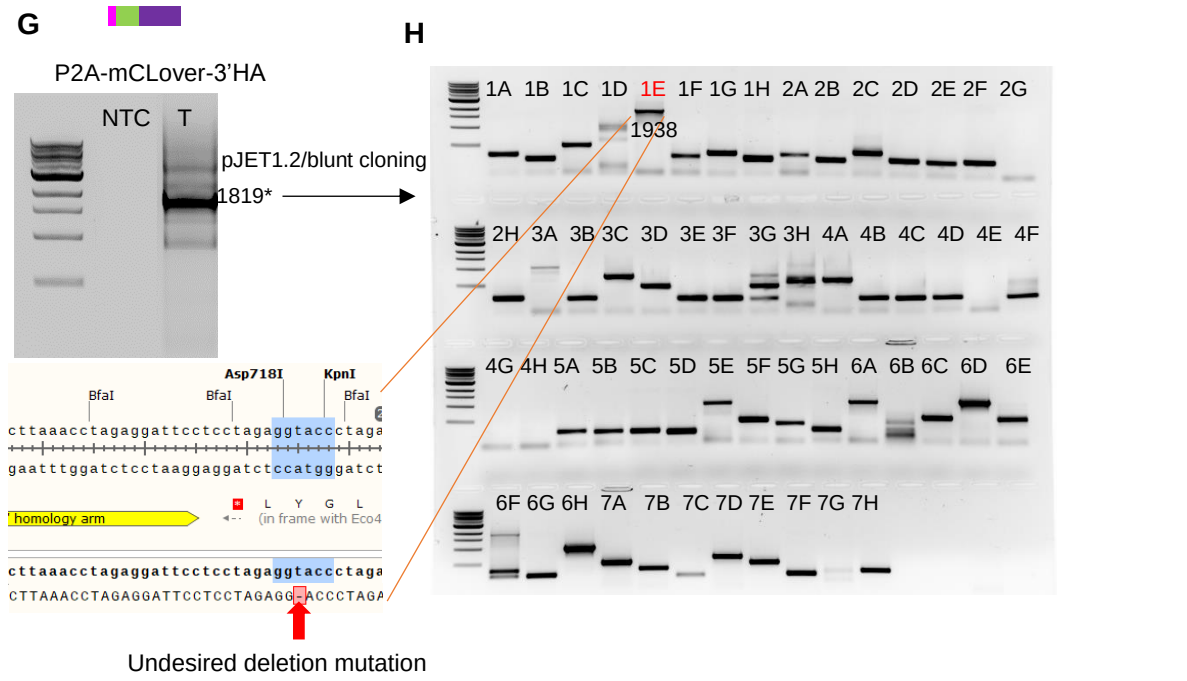
sequencing and the plasmid clone F1 was selected for the further targeting vector construction (Figure 3-11F). Similarly, the 3'HA fragment was joined to the P2A-mClover fragment by fusion PCR using the primers PITX2-mClover-F and PITX2-3HA-R2. The PITX2-P2A-mClover-3'HA fusion product (1819 bp) (Figure 3-11G) was cloned into the cloning vector pJET1.2/blunt (Figure S6). Interestingly, colony PCR screening using the primers pJET1.2-F/R found only 1 promising clone (1E, ~1938 bp) (Figure 3-11H) out of the total 56 clones, suggesting impurity of the fusion PCR products. DNA sequencing of the clone 1E revealed a single nucleotide deletion immediately downstream of the 3'HA region, which may partly explain the failed Gibson Assembly.

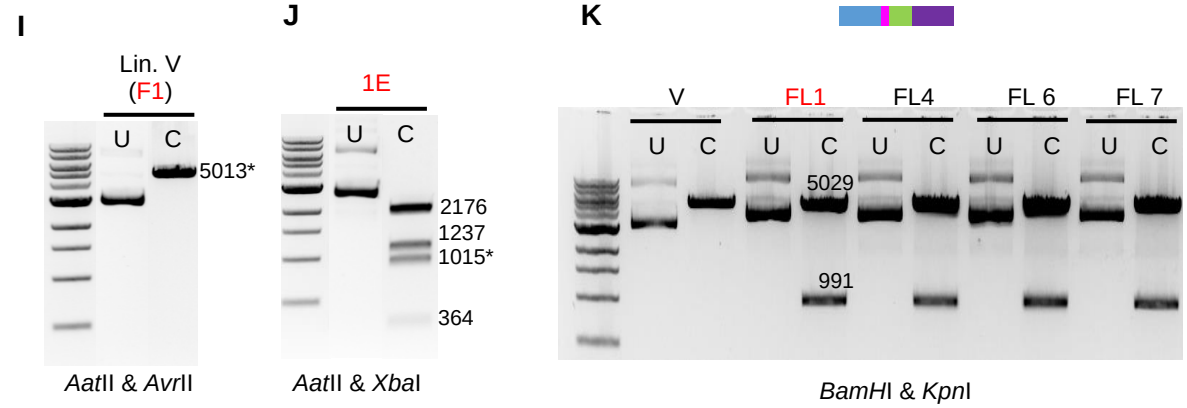
The next cloning step involved subcloning of a fragment from the plasmid clone 1E into the clone F1 to fuse the 3'HA to the 5'HA-mClover. The *Aat*II and *Avr*II linearised plasmid clone F1 (5013 bp) (Figure 3-11I) was used as a vector for cloning the 1015-bp fragment (Figure 3-11J) that was obtained by *Xba*I and *Aat*II digestion of the plasmid clone 1E. The successful subcloning from this step should have two products (5029 and 991 bp) when analysed by *Bam*HI and *Kpn*I digestion (Figure 3-11K). The clone FL1 with verified sequence was chosen for use in the next step.

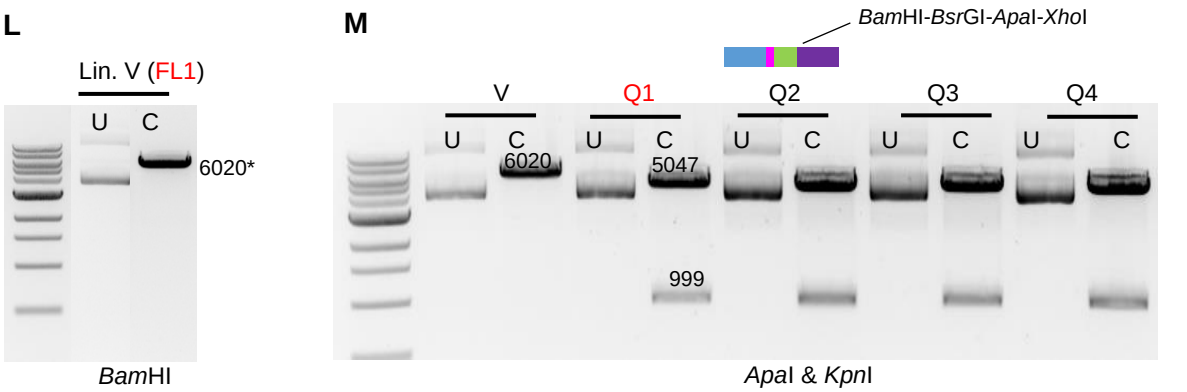
To prepare the plasmid for floxed *Neo* cassette insertion, the annealed oligo linker containing *Bam*HI-*Bsr*GI-*Apa*I-*Xho*I restriction sites (26 bp) (NS-TV-link1/2) was cloned into the plasmid clone FL1, previously digested with *Bam*HI (6020 bp) (Figure 3-11L) at the immediately downstream of the *mClover*. The resultant plasmid clone Q1 with the correct *Apa*I and *Kpn*I digestion pattern (5047 and 999 bp) (Figure 3-11M) and verified sequence was selected for the next step. Notably, the oligo linker insertion resulted in the CRISPR-A recognition sequence being interrupted (Figure 3-11A), which prevented the re-targeting of successful genomic insertion events. The *BamH*I and *Apa*I linearised plasmid clone Q1 (6028 bp) (Figure 3-11N) was used as a cloning vector for the *Neo* cassette fragment (1922 bp) (Figure 3-11O) obtained from digestion of the plasmid PL452 with the same restriction endonucleases. The

successful clone R1 with the correct *Xho*I digestion pattern (5071, 1901 and 962 bp) (Figure 3-11P) and confirmed sequence was selected for the final cloning step.




G
P2A-mCLover-3'HA
NTC
T
pJET1.2/blunt cloning
1819*
Asp718I
KpnI
BfaI
BfaI
BfaI
homology arm
L Y G L
(in frame with Eco4
Undesired deletion mutation
H
1A 1B 1C 1D 1E 1F 1G 1H 2A 2B 2C 2D 2E 2F 2G
1938
2H 3A 3B 3C 3D 3E 3F 3G 3H 4A 4B 4C 4D 4E 4F
4G 4H 5A 5B 5C 5D 5E 5F 5G 5H 6A 6B 6C 6D 6E
6F 6G 6H 7A 7B 7C 7D 7E 7F 7G 7H


I
Lin. V
(F1)
U
C
5013*
AatII & AvrII
J
1E
U
C
2176
1237
1015*
364
AatII & XbaI
K
V
FL1
FL4
FL 6
FL 7
U
C
5029
991
BamHI & KpnI


L
Lin. V (FL1)
U
C
6020*
BamHI
M
BamHI-BsrGI-ApaI-XhoI
V
Q1
Q2
Q3
Q4
U
C
6020
5047
999
ApaI & KpnI

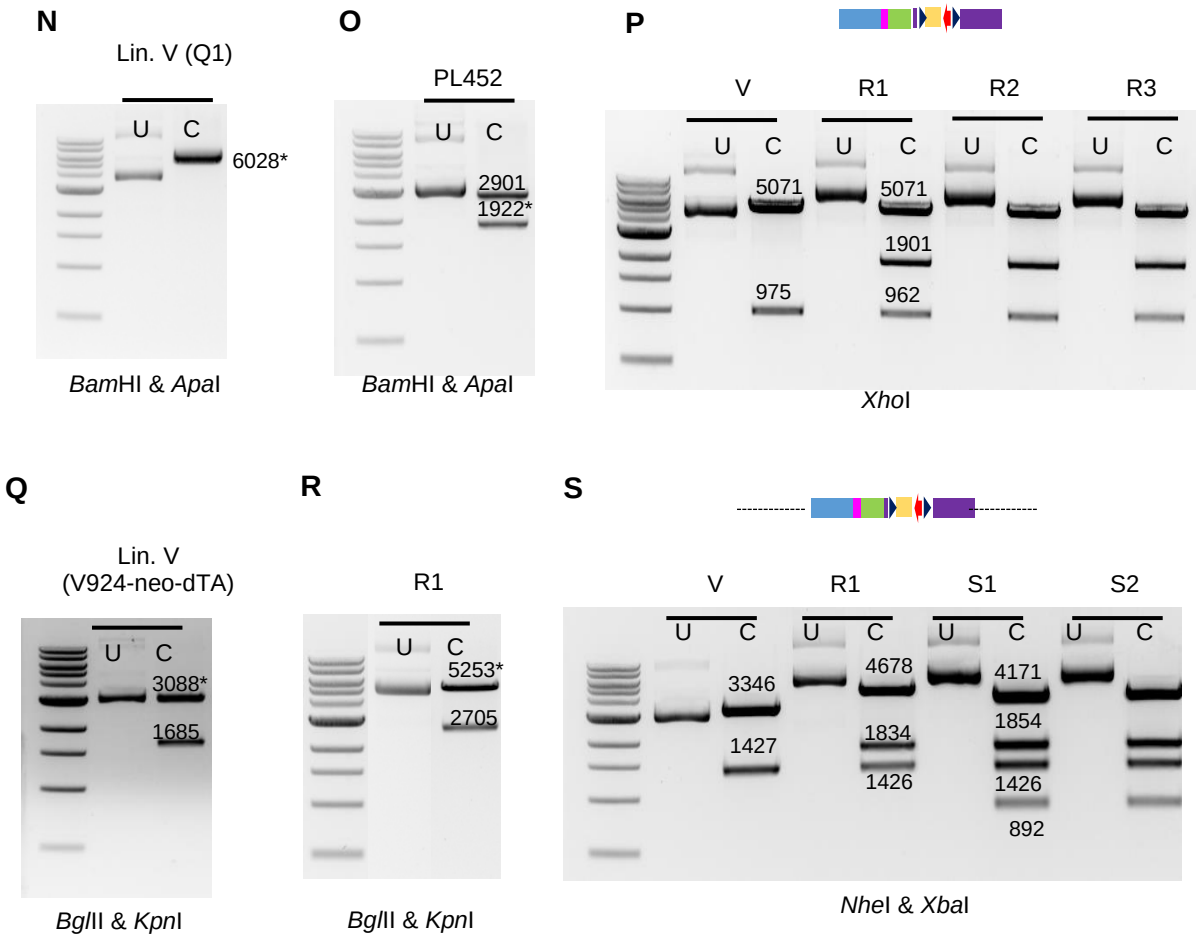


Figure 3-11 *PITX2-mClover* targeting vector cloning. A) A simplified illustration of the *PITX2-mClover* plasmid construct showing the CRISPR-A recognition site split between the major region of the 3'HA and the minor region of the 3'HA as a consequence of the cloning. The targeting plasmid was initially designed to be generated by Gibson Assembly which combined 4 DNA fragments. B) Preparation of Gibson Assembly fragments. Since the Gibson Assembly failed, these GA fragments were then used as substrates for PCR extension (C and G). C) Fusion PCR products of the 5'HA-P2A-mClover fragment. D-E) Preparation for cloning the 5'HA-P2A-mClover into the L28 vector F) Diagnostic screening for successful 5'HA-P2A-mClover cloning (F1). G) Fusion PCR products of the P2A-mClover-3'HA fragment. H) Colony PCR screening for successful cloning of P2A-mClover-3'HA into the cloning vector pJET1.2/blunt showing the promising clone 1E. However, DNA sequencing identified a single nucleotide deletion immediately downstream of the 3'HA. I-J) Preparation of fragments for cloning of the P2A-mClover-3'HA fragment into the plasmid clone F1. K) Diagnostic screening for successful cloning in I-J (FL1). L) Preparation of the plasmid clone FL1 for *Bam*HI-*Bsr*GI-*Apa*I-*Xho*I linker insertion at the downstream of the *mClover* region. M) Diagnostic screening for successful linker cloning (Q1). N-O) Preparation for cloning of floxed *Neo* into the plasmid clone Q1. P) Diagnostic screening of successful floxed *Neo* cloning (R1). Q-R) Preparation for cloning of the P2A-*mClover*-*Neo*-3'HA fragment S) Diagnostic screening for the final from of the targeting vector S1). *Amp*, ampicillin resistant gene; *PGK* P, phosphoglycerate kinase promoter; U, uncut; C, cut; NTC, no template control; T, with template; V, vector; Lin. V, linearised vector; *fragments used for relevant cloning(s).

The final cloning step was cloning of the whole targeting construct from the L28 backbone into the destination vector V924-Neo-dTA. The plasmid V924-Neo-dTA was firstly digested with *Bgl*II and *Kpn*I to remove the *Neo* cassette and produce a desired linearlised vector fragment

(3088 bp) (Figure 3-11Q). Secondly, the plasmid clone R1 was digested with the same enzymes to obtain a 5253-bp fragment (Figure 3-11R) containing the whole targeting construct, which was subsequently cloned into the linearised V924-dTA vector. The final targeting plasmid (8341 bp) when digested with *Nhe*I and *Xba*I should have 4 products (4171, 1854, 1426 and 892). The plasmid clone S1, which was called the *PITX2-mClover* targeting vector, with the above digestion confirmation and sequence verification was selected for use in generating the *PITX2-mClover* reporter line.

3.2.4 *MYL2-mClover* reporter line

3.3.4.1 MYL2 gRNA efficiency test

2 different gRNAs, which harboured 20-nt CRISPR target sequences complementary to the *MYL2* locus, were synthesised and tested for their efficiency. CRISPR-A yielded 17% TIDE (Figure 3-12C) and CRISPR-B yielded 19.4% TIDE. Due to the higher efficiency, the CRISPR-B was selected for generating the *MYL2-mClover* reporter line.

Because there is a considerable amount of repeat DNA motifs (6 relatively large and 4 relatively small motifs) reported approximately 1000 bp upstream and 1000 bp downstream of the stop codon (Figure 3-13), which may reduce the efficiency of targeted integration, the donor template was designed in a form of lssDNA, which required only ~100bp HAs.

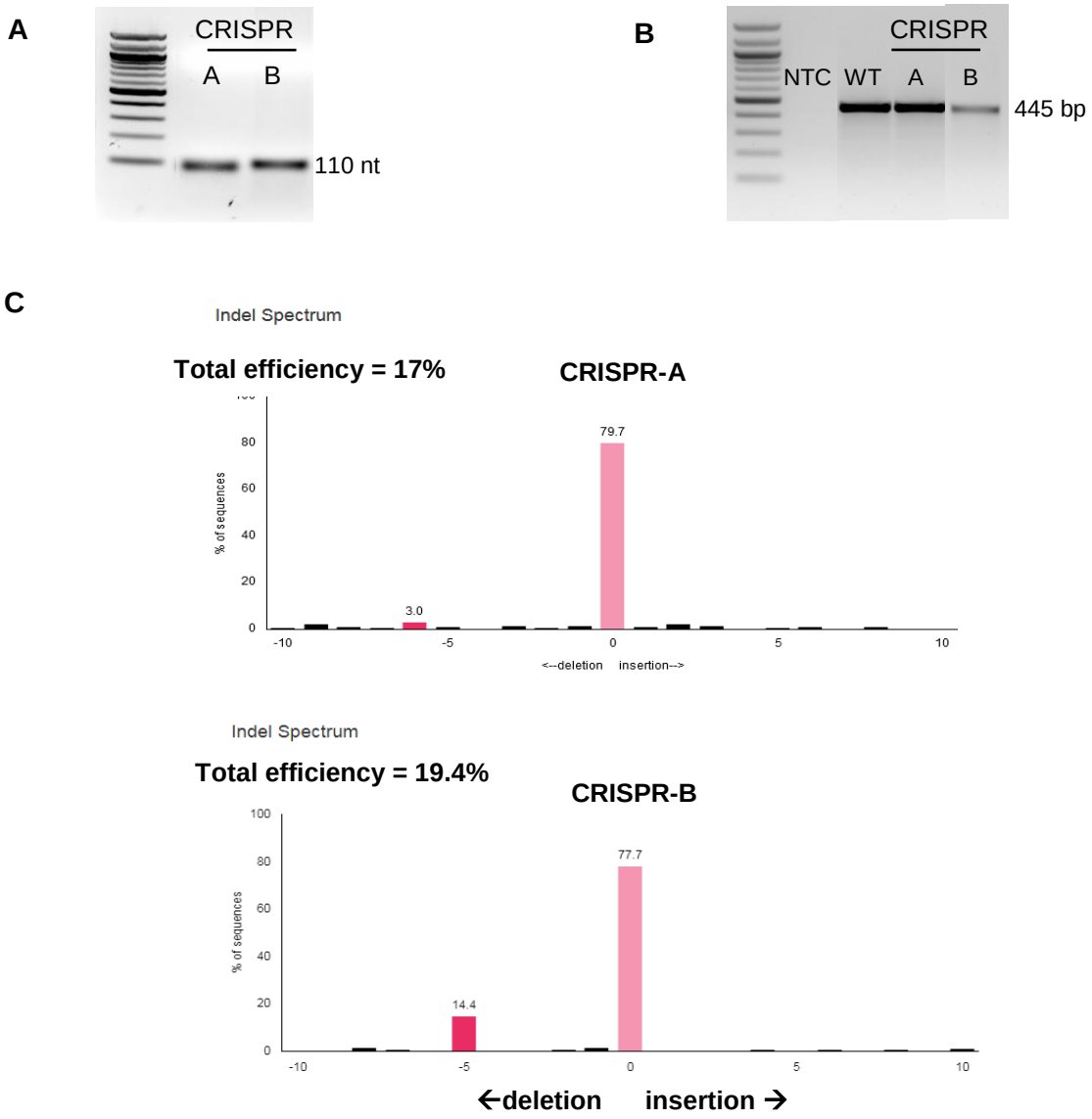


Figure 3-12 TIDE analysis for CRISPRs targeting the *MYL2* locus. A) *In vitro* transcribed gRNAs in a heat-denatured form containing the CRISPR A and B sequences. B) PCR products 48 hr after CRISPR/Cas9 transfection into KOLF2-C1 iPSCs using the primers MYL2-F/R which amplified the sequence flanking the stop codon. These PCR products were sequenced and analysed for CRISPR activity by TIDE as compared to the non-transfected WT control. C) TIDE score for each CRISPR design. NTC, no template control.

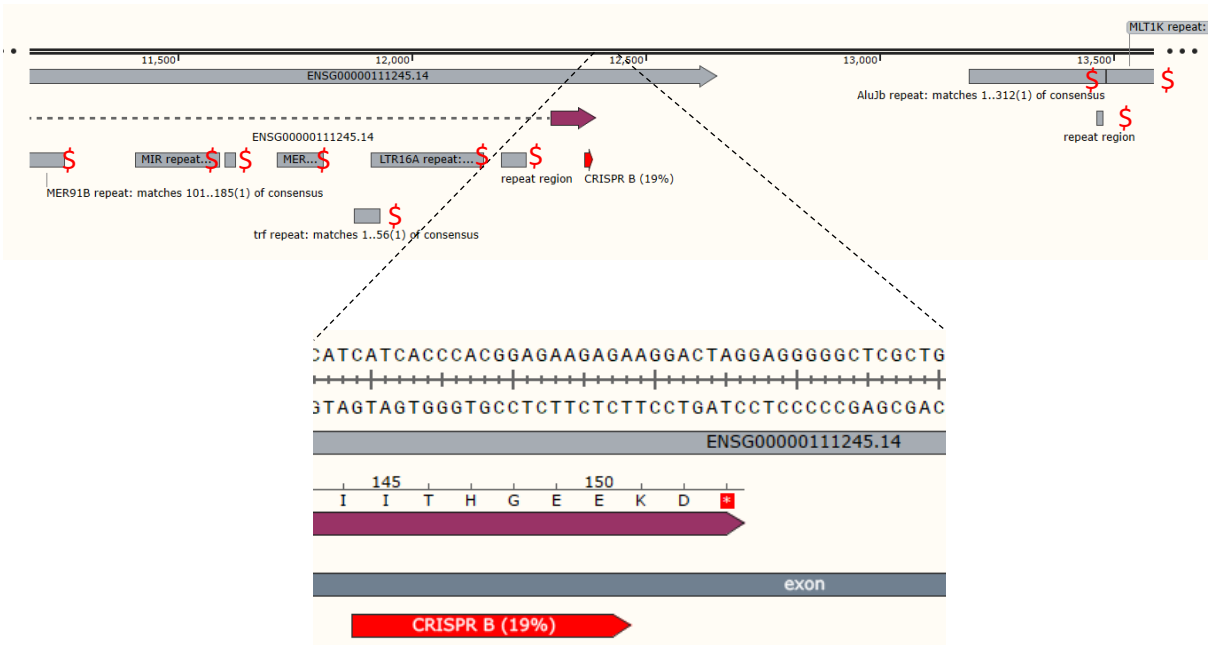


Figure 3-13 MYL2 CRISPR-B location relative to the *MYL2* locus. Approximately 1000 bp upstream and 1000 bp downstream of the stop codon contain 6 relatively large and 4 relatively small repeat motifs ($).

3.2.4.2 MYL2 gRNA vector cloning

From experiences in mouse ESCs in Professor Ben Davies' laboratory, a lssDNA was not compatible with RNP electroporation, the most efficient CRISPR/Cas9 delivery mode (Burger et al., 2016). In addition, CRISPR/Cas9 targeting efficiency was also relatively low when RNP was introduced by lipofection. To address these issues, the sequence encoding the MYL2 CRISPR-A was cloned into a CRISPR/Cas9 expression plasmid. This plasmid was later introduced together with the lssDNA into iPSCs by lipofection for generating the fluorescent *MYL2* reporter line.

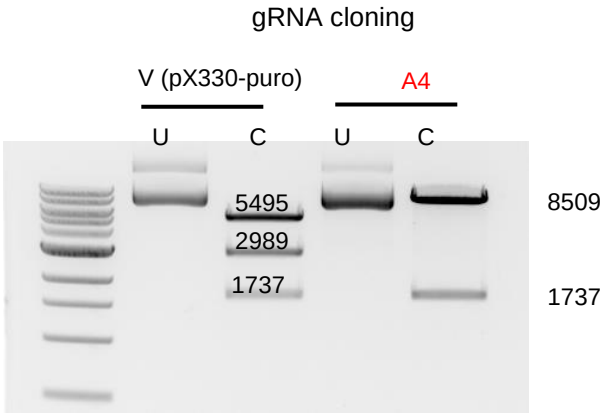
The annealed oligos containing the MYL2 CRISPR-A (MYL2-A1/A2) were cloned into a human codon-optimised CRISPR/Cas9 expression plasmid pX330-puro (Figure S4) by Golden Gate Assembly. The successful pX330-puro-MYL2 clone A4 was confirmed by *Bbs*I and *Not*I digestion (8509 and 1737 bp) (Figure 3-14A) and DNA sequencing and selected for generating the *MYL2-mClover* reporter line.

3.2.4.3 *MYL2-mClover* lssDNA donor synthesis

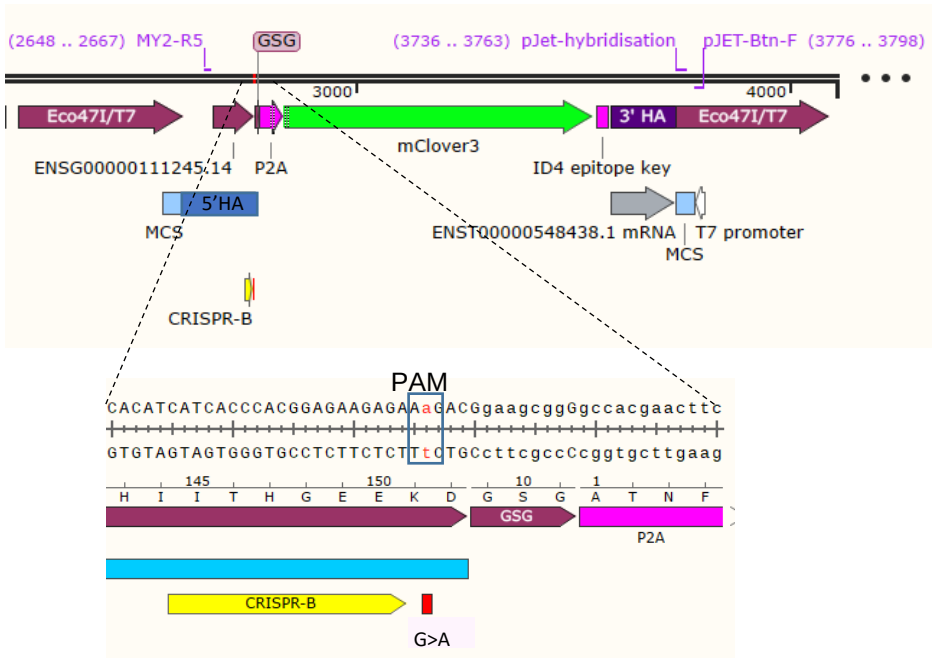
The *MYL2*-P2A-*mClover* gBlock template containing 120-bp 5'HA and 150-bp 3'HA (Figure S7) was cloned the pJET1.2/blunt vector. The PAM sequence downstream of the MYL2 CRISPR-A recognition sequence on this template was altered with silent mutations to prevent the CRISPR/Cas9 activity at the successfully targeted allele (Figure 3-14B). The resultant plasmid clone B3 with the correct *Not*I and *Aat*II digestion pattern (2176, 1309, 386 and 203 bp) (Figure 3-14C) and verified sequence was selected for the lssDNA synthesis.

The whole targeting construct was amplified by PCR from the plasmid clone B3 using the biotinylated primer pJET1.2-Btn-F and the primer MYL2-R5 (1151 bp, Figure 3-14B & C), which was purified before incubation with streptavidin beads. The streptavidin-captured dsDNA was denatured in an alkali condition, allowing bottom lssDNA elution (B-E1, B-E2 and B-E3; Figure 3-14D). The top lssDNA was eluted by reconstituting a region of dsDNA using the pJET-Hybridization primer. This dsDNA sequence contained an *Xho*I restriction site that could then be used to release and elute the top lssDNA (T-E1; Figure 3-14E). Each lssDNA molecule, including the top and the bottom, had a molecular weight of half that of the dsDNA (1151 bp) and was thus expected to migrate on an agarose gel at the equivalent distance of ~500 bp dsDNA fragment (Figure 3-14F). This approach yielded approximately 1000 and 3750 ng of the top lssDNA (T-E1) and the bottom lssDNA (B-E1), respectively, purified from 10 µg dsDNA template. From the BD's laboratory experience, top lssDNAs have higher purity than bottom lssDNAs after the described purification process. Therefore, the top lssDNA (1108 nt) was selected for use in generation of the *MYL2-mClover* reporter line rather than the bottom strand.

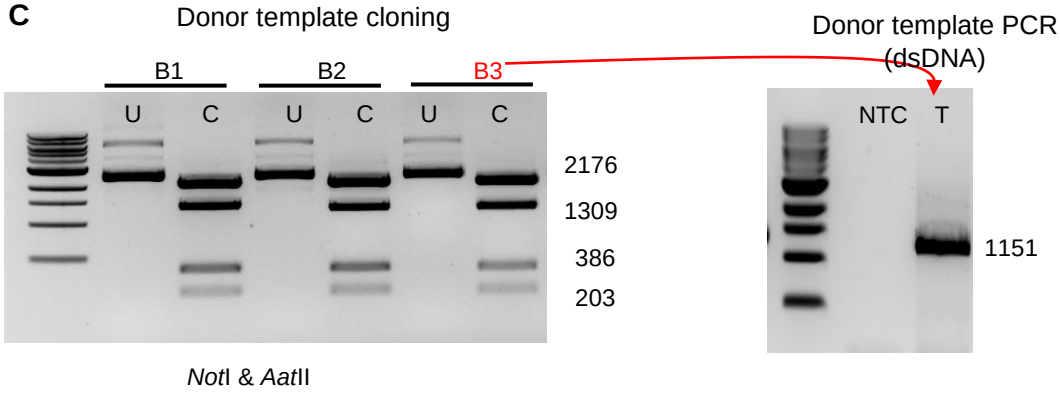
A


gRNA cloning
V (pX330-puro)
A4
U
C
U
C
5495
2989
1737
8509
1737
*Bbs*I & *Not*I

B


(2648 .. 2667) MY2-R5
GSG
(3736 .. 3763) pJet-hybridisation
pJET-Btn-F (3776 .. 3798)
3000
4000
Eco47I/T7
mClover3
3' HA
Eco47I/T7
ENSG00000111245.14
P2A
ID4 epitope key
5'HA
MCS
ENST00000548438.1 mRNA
T7 promoter
MCS
CRISPR-B
PAM
CACATCATCACCCACGGAGAAGAGAAaGACGgaagcggGgccacgaacttc
GTGTAGTAGTGGGTGCCTCTTCTCTTtCTGCcttcgcccCcggtgcttgaag
H I I T H G E E K D G S G A T N F
GSG
P2A
CRISPR-B
G>A

C


Donor template cloning
Donor template PCR
(dsDNA)
B1
B2
B3
U
C
U
C
U
C
2176
1309
386
203
NTC
T
1151
*Not*I & *Aat*II

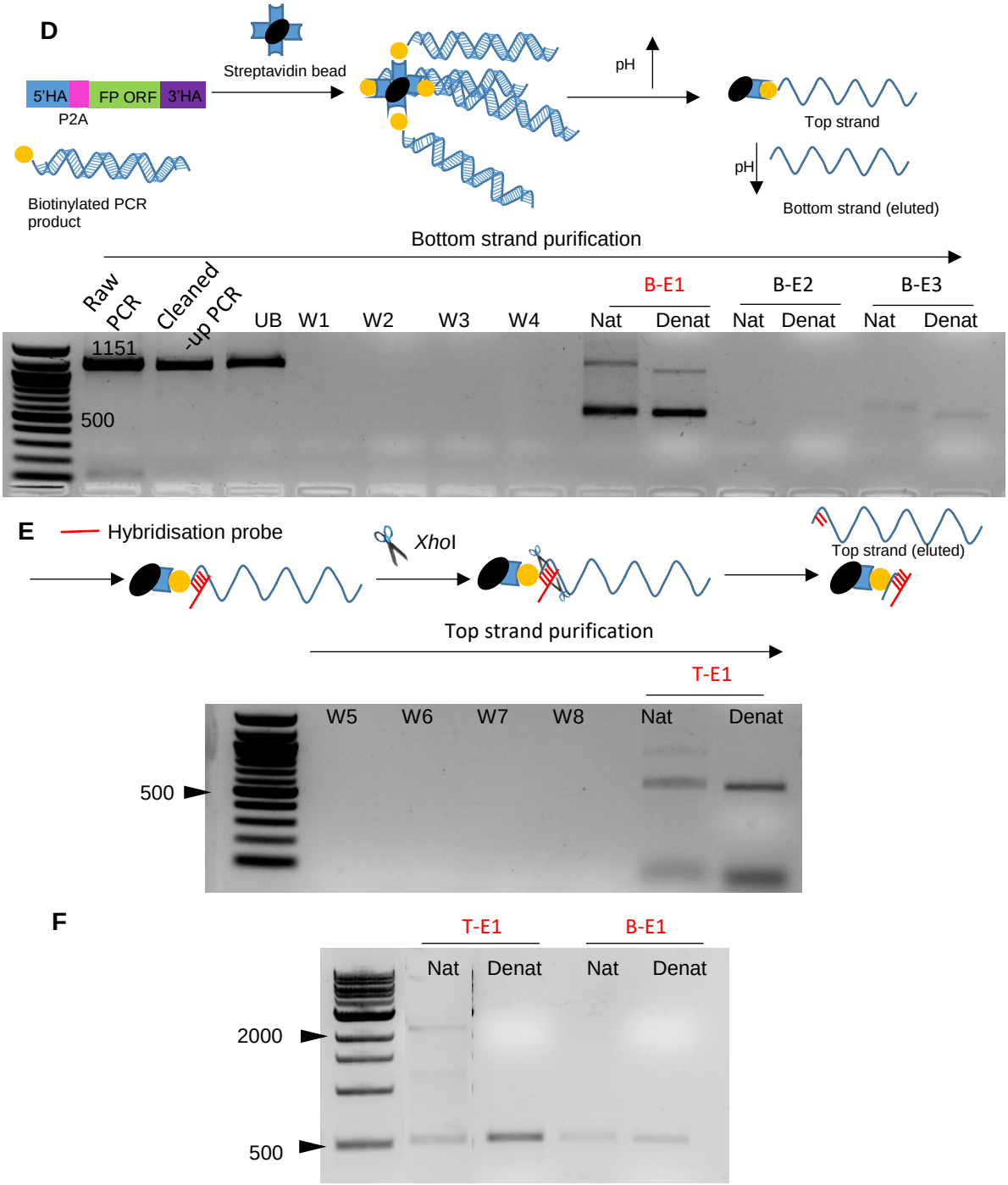


Figure 3-14 MYL2 CRISPR-B cloning and *MYL2-mClover* donor purification. A) Diagnostic screening for successful MYL2 CRISPR-B into the pX330-puro vector. B-F) Biotinylation-mediated *MYL2-mClover* lssDNA purification. B-C) Initially, the *MYL2-mClover* targeting construct was cloned into pJET1.2/blunt vector. B) A diagram highlighting the intentional silent mutation at the CRISPR-B PAM site (G>A) and the relevant primers/probes position relative to the resultant pJET1.2-*MYL2-mClover* plasmid. C) Diagnostic screening for successful cloning as in B (B3). The clone B3 was used as a template for PCR amplification of biotin-tagged dsDNA containing the *MYL2-mClover* construct. D-F) lssDNA purification D) Purification of the bottom lssDNA (B-E1, B-E2 and B-E3) from the dsDNA in C. E) Purification of the top lssDNA (T-E1). F) The final purified/concentrated products of the top (T-E1) and bottom (B-E1) lssDNAs. U, uncut; C, cut; NTC, no template control; T, with template; V, vector; UB, unbound fraction; W, wash; B-E1/2/3; bottom lssDNA-elutions 1/2/3; T-E1, top lssDNA-elution 1; Nat, native form; Denat, heat-denatured form.

3.3 CRISPR/Cas9-mediated generation of fluorescent reporter iPSC lines

In order to establish a fluorescent reporter, CRISPR/Cas9 is needed to be transfected into the parental KOLF2-C1 iPSCs together with a DNA donor template. The different subtype-specific fluorescent reporter relied on two different CRISPR/Cas9 delivery methods. For the plasmid donor templates harbouring a *Neo* cassette i.e. the *MYL7-mClover*, *PITX2-mClover* and *HCN4-mScarle*t lines, the relevant re-constituted gRNA/Cas9 RNP was introduced into the cells by electroporation. As described in 3.2.4.2, for the lssDNA template i.e. the *MYL2-mClover* line, the pX330-puro-MYL2 plasmid was reverse-transfected into the cells by lipofection.

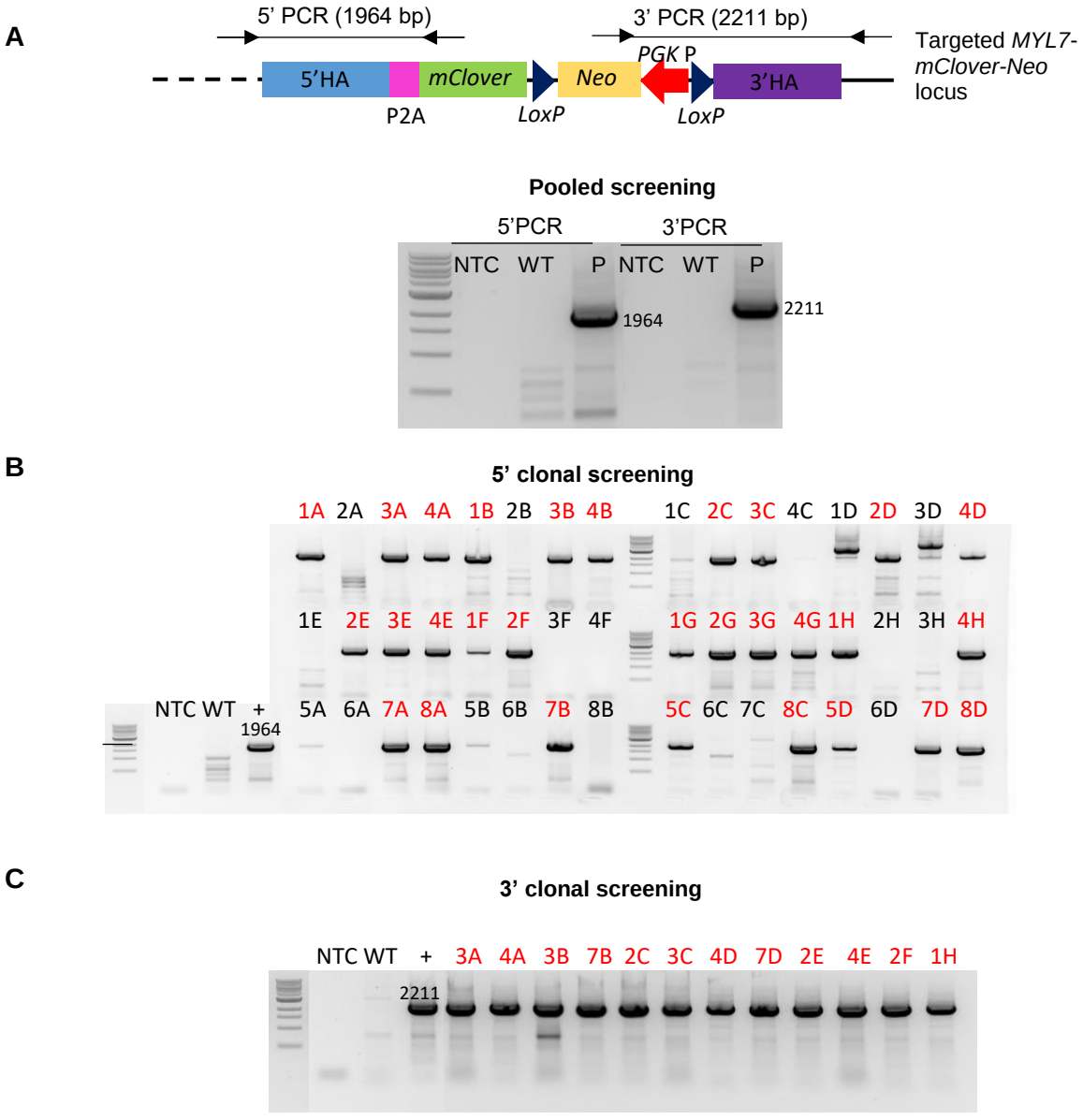
The transfected cells were allowed to recover for a few days prior to PCR screening for insertion events (pooled screening) of the fluorescent reporter, followed by low density seeding. For the *MYL7*, *HCN4* and *PITX2* reporter projects, approximately 5 days post-transfection, the cells were treated with 350 µg/ml G418 for 7 days followed by 200 ug/ml G418 for another 7 days. For the *MYL2* project, the cells were selected in E8 containing 350 ng/ml puromycin 48 hr post-transfection for 48 hr, allowing selection of cell that had up-taken the pX330-Puro-MYL2 plasmid. Subsequently, colonies were picked from the low-density cultures and analysed for individual clones (clonal screening) carrying the fluorescent reporter by PCR. Finally, the sequence of the targeted allele in each promising clone was confirmed by DNA sequencing.

3.3.1 Generation of *MYL7-mClover* line

5’ PCR screening using the primers MYL7-F3 and mClover5’R and 3’ PCR screening using the primers ExNeo2 and MYL7-R3 indicated successful integration of the P2A-*mClover-Neo* construct at the *MYL7* locus in the pool of KOLF2-C1 cells electroporated with the RNP/*MYL7-mClover* targeting vector (Figure 3-15A). Following low density plating, individual clones showing successful 5’ integration were also isolated, approximately 60% being positive

(Figure 3-15B). Subsequent 3' screening of 12 selected 5'-positive clones showed 100% positive integration at this end of the vector (Figure 3-15C).

Two promising clones (7B and 7D) were taken forward for the removal of the *Neo* cassette via Cre-Lox recombination. The Cre-Pac plasmid (Taniguchi et al., 1998) harbouring Cre recombinase was reverse transfected into the cells by lipofection. Some of the pool of the transfected cells were analysed for *Neo* cassette deletion by PCR using the primers mCloverF2 and MYL7-R2 while the remaining cells were subjected to a second round of low density plating and clonal establishment.



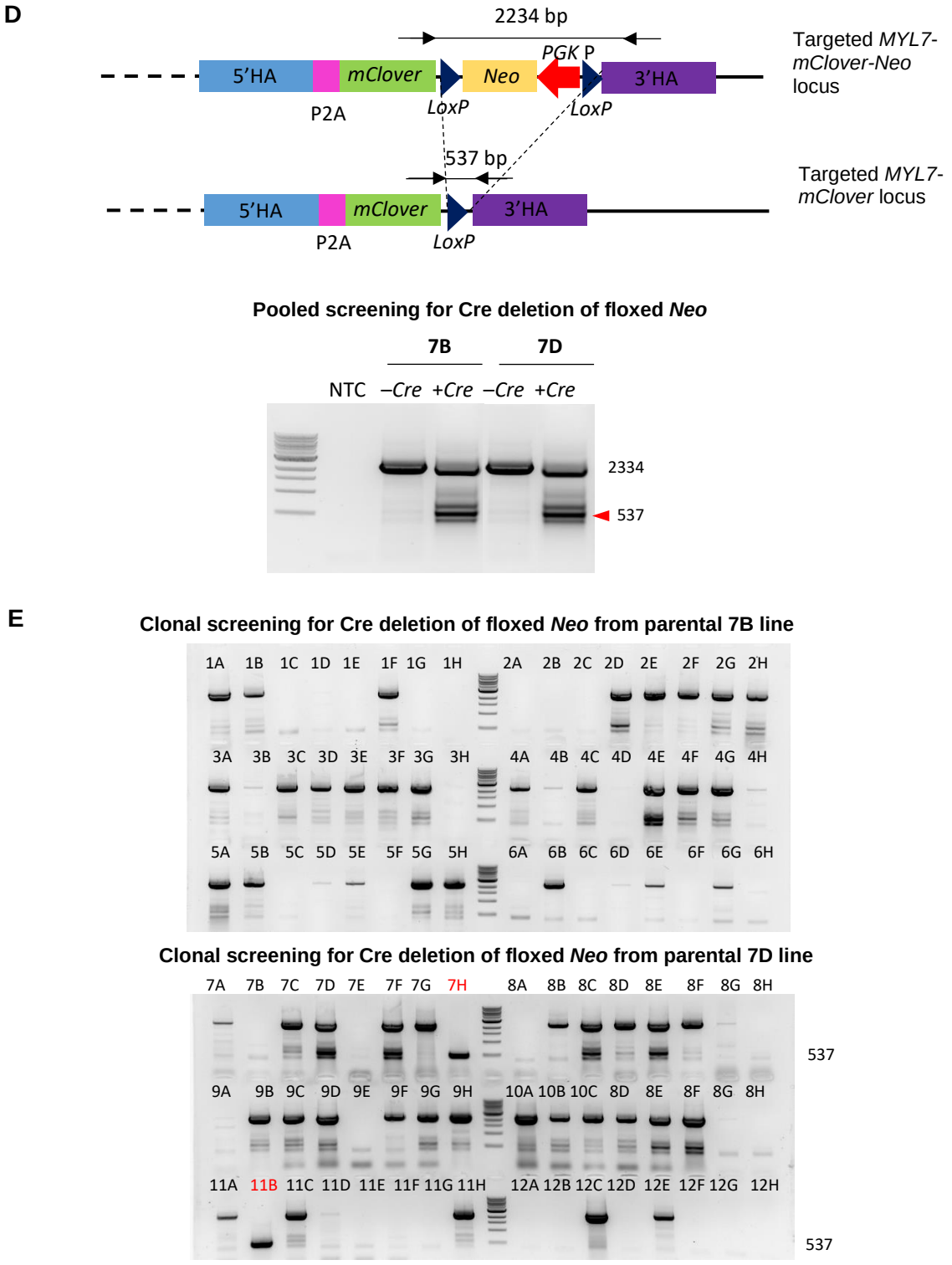


Figure 3-15 Genotypic screening for *MYL7-mClover* line. A) 5' and 3' PCR screenings in the pool of CRISPR/Cas9-transfected iPSCs. B) Clonal screening for 5' integration C) Clonal screening for 3' integration. The promising clone 7B and 7D were subsequently transfected with a *Cre* expression plasmid to excise the floxed *Neo* cassette. D) PCR screening for *Neo* deletion in the pool of *Cre*-transfected cells. E) Clonal screening for *Neo* deletion of clones (red) derived from the parental *Neo*-positive 7B and 7D clones. NTC, no template control; P, pool of transfected cells; *PGK* P, phosphoglycerate kinase promoter.

The successful *Neo* deletion was detectable in the pool of Cre transfected cells derived from the *Neo* positive clones 7B and 7D (Figure 3-15D). Further screening in isolated colonies confirmed 2 promising *Neo* negative clones 7H and 11B (Figure 3-15E) derived from the parental *Neo*-positive 7D, which were later simply called clones 7H and 11B, respectively, but none was found in *Neo* positive 7B-derived colonies. It is notable that the Cre-Pac plasmid harbours a puromycin selectable marker. However, the *Cre* transfection was performed without antibiotic selection, which may explain the small number of clones harbouring deleted *Neo* allele.

The clones 7H and 11B were expanded and reconfirmed for the integration of the P2A-*mClover* at the *MYL7* locus by 5' PCR using the primers MYL7-F3 and mClover5'R and 3' PCR using the primers mClover-F2 and MYL7-R3 (Figure 3-16). A further internal PCR assay was performed, amplifying from within the 5'HA to the 3'HA using the primers MYL7-F2/R2 to verify a complete insertion of the *P2A-mClover* sequence, which was later confirmed by DNA sequencing. In addition, both clones were identified as being heterozygous for the fluorophore, where the non-targeted alleles displayed no undesired CRISPR-related mutations. These two clones were selected for further quality checks.

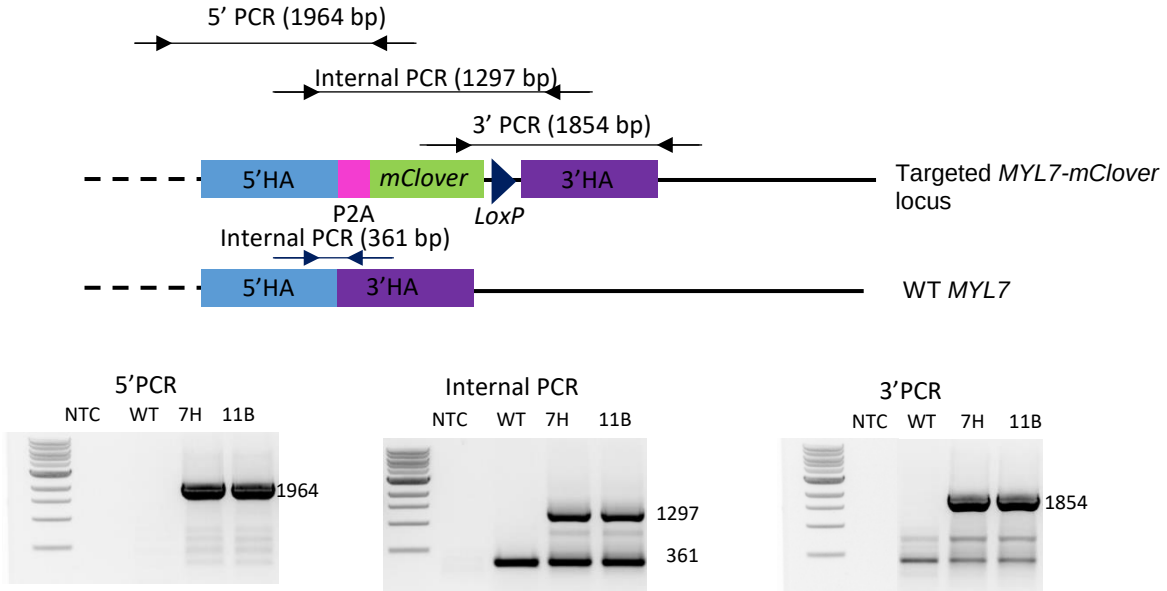
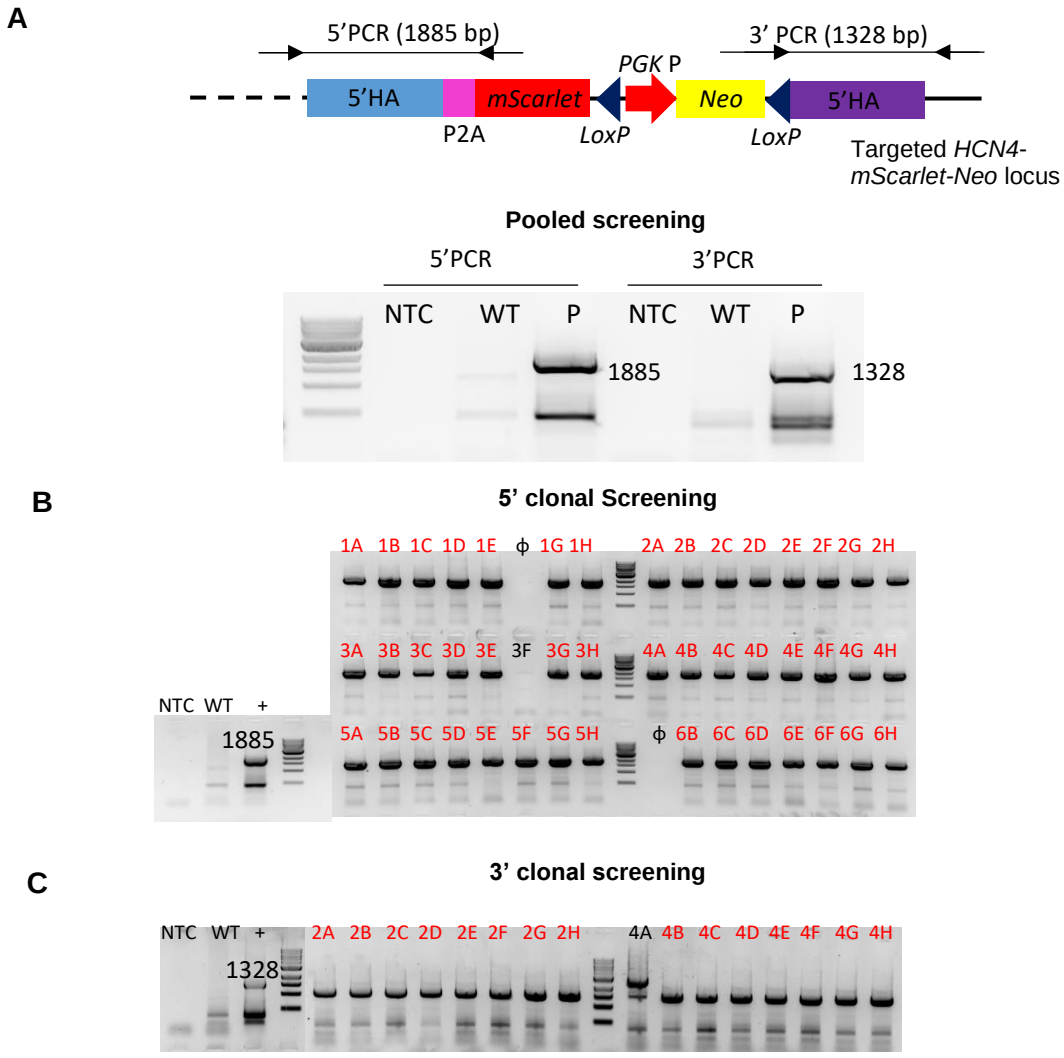


Figure 3-16 Conformational PCR of isolated *MYL7-mClover* iPSCs after clonal expansion. 5' and 3' PCR screenings indicated successful P2A-*mClover* insertion downstream of the *MYL7* locus devoid of the *Neo* cassette. Internal PCR confirmed the intact P2A-*mClover* insert with a heterozygous status. NTC, no template control.

3.3.2 Generation of *HCN4-mScarlet* line

5’ PCR screening using the primers HCN4-F2 and mScarlet-R and 3’ PCR screening using the primers bGH-F2 and HCN4-R4 indicated successful integration of the P2A-*mScarlet-Neo* at the *HCN4* locus in the pool of KOLF2-C1 cells electroporated with the RNP/*HCN4-mScarlet* targeting vector (Figure 3-17A). Following low density plating, individual clones showing successful 5’ integration were also isolated, approximately 98% being positive (Figure 3-17B). Subsequent 3’ screening of 16 selected 5’-positive clones showed 100% positive integration at this end of the vector (Figure 3-17C).



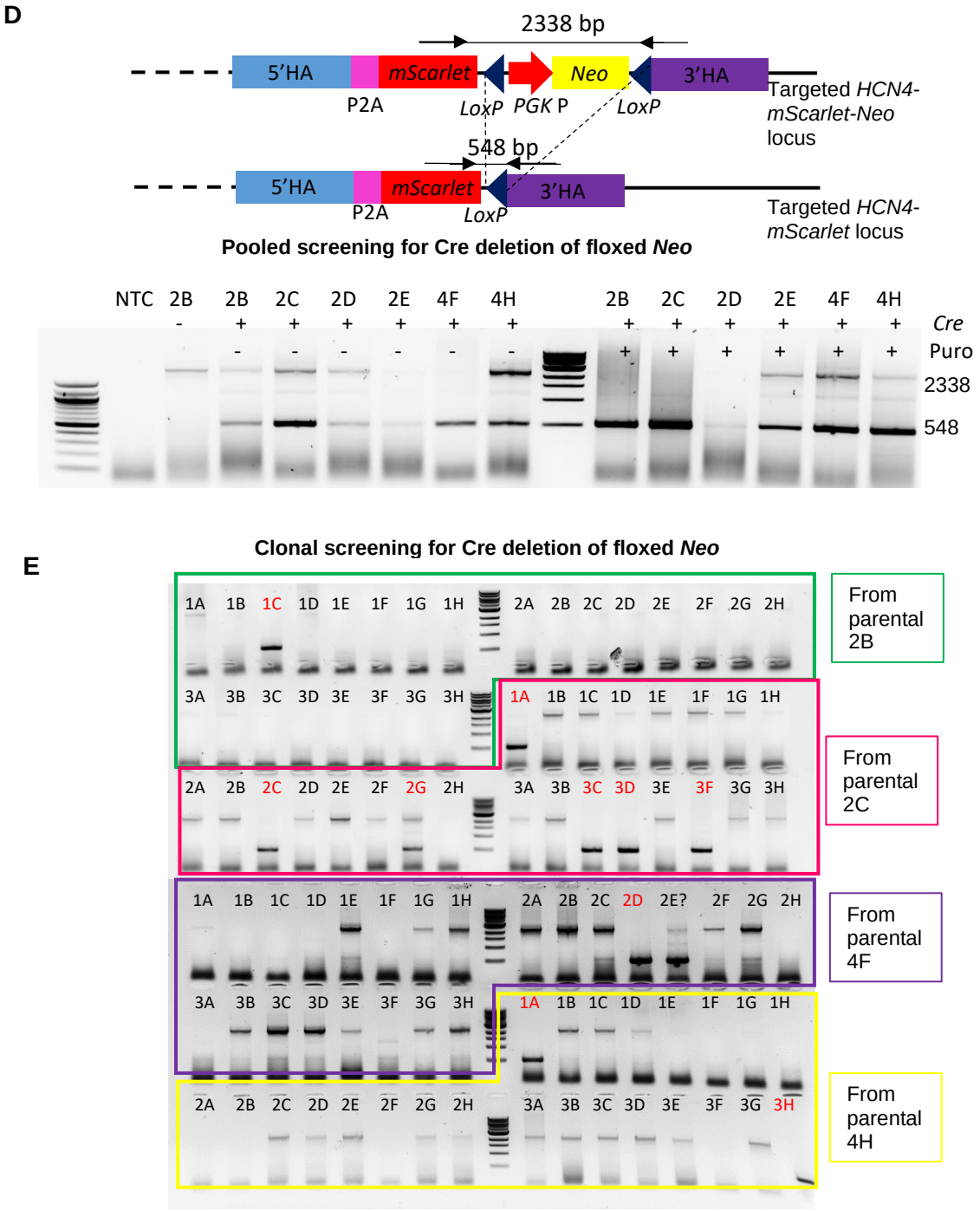


Figure 3-17 Genotypic screening for *HCN4-mScarlet* line. A) 5' and 3' PCR screenings in the pool of CRISPR/Cas9-transfected iPSCs. B) Clonal screening for 5' integration C) Clonal screening for 3' integration. The promising clone 2B, 2C, 2D, 2E, 4F and 4H were subsequently transfected with a *Cre* expression plasmid to excise the floxed *Neo* cassette. D) PCR screening for *Neo* deletion in the pool of *Cre*-transfected cells with and without puromycin selection (548 bp). E) Clonal screening for *Neo* deletion (red) of puromycin-treated clones derived from the parental *Neo*-positive 2B, 2C, 4F and 4H clones. NTC, no template control; P, pool of transfected cells; *PGK* P, phosphoglycerate kinase promoter.

DNA sequencing indicated heterozygosity of the *mScarlet* insertion event in clones 2B, 2D, 2E, 4F and 4H with no deleterious mutations on the non-targeted allele and likely homozygosity of the *mScarle*t in clone 2C. These clones were transfected with the Cre-Pac plasmid, 48 hr after which the cells were selected in 300 ng/ml puromycin for a further 48 hr. The pool of the transfected cells showed Cre-Lox-mediated *Neo* excision when assayed by PCR using the primers mScarlet-F and HCN4-R (Figure 3-17D). Further screening of isolated colonies having been selected against puromycin revealed i) 1/24 (1C) promising *Neo* negative clones derived from the parental *Neo* positive clone 2B, ii) 6/24 (1A, 2C, 2G, 3C, 3D, and 3F) promising *Neo* negative clones derived from the parental *Neo* positive clone 2C, iii) 1/24 (2D) promising *Neo* negative clones derived from the parental *Neo* positive clone 4F and iv) 2/24 (1A and 3H) promising *Neo* negative clones derived from the parental *Neo* positive clone 4H (Figure 3-17E).

The clones *Neo*-positive-2B-derived 2B-1C, *Neo*-positive-2C-derived 2C-3D and *Neo*-positive-4F-derived 4F-2D were expanded and reconfirmed for the integration of the P2A-*mScarlet* at the *HCN*4 locus by 5' PCR using the primers HCN4-F2 and mScarlet-R and 3' PCR using the primers mScarlet-F and HCN4-R4 (Figure 3-18). A further internal PCR assay using the primers HCN4-F/R confirmed the integrity of the full-length P2A-*mScarlet* integration event, which was later verified by DNA sequencing. In addition, both clones 2B-1C and 4F-2D were identified as being heterozygous for the fluorophore with the non-targeted alleles displaying no undesired CRISPR-related mutations. A long-range PCR (3763 bp) amplifying outside of the HAs i.e. further away from the CRISPR cut site using the primers HCN4-F3/R4 further suggested *mScarlet* heterozygosity in the clone 4F-2D. This PCR assay was designed to investigate whether the CRISPR/Cas9 introduced any large deletion events undetectable by the internal PCR. Since there were no products smaller than the expected amplicons containing the inserted fluorophore by both the internal and long-range PCR, it could be assumed that the clone 2C-3D was successfully edited for both alleles. Subsequently, the clones 2B-1C, 2C-3D and 4F-2D were selected for further quality checks.

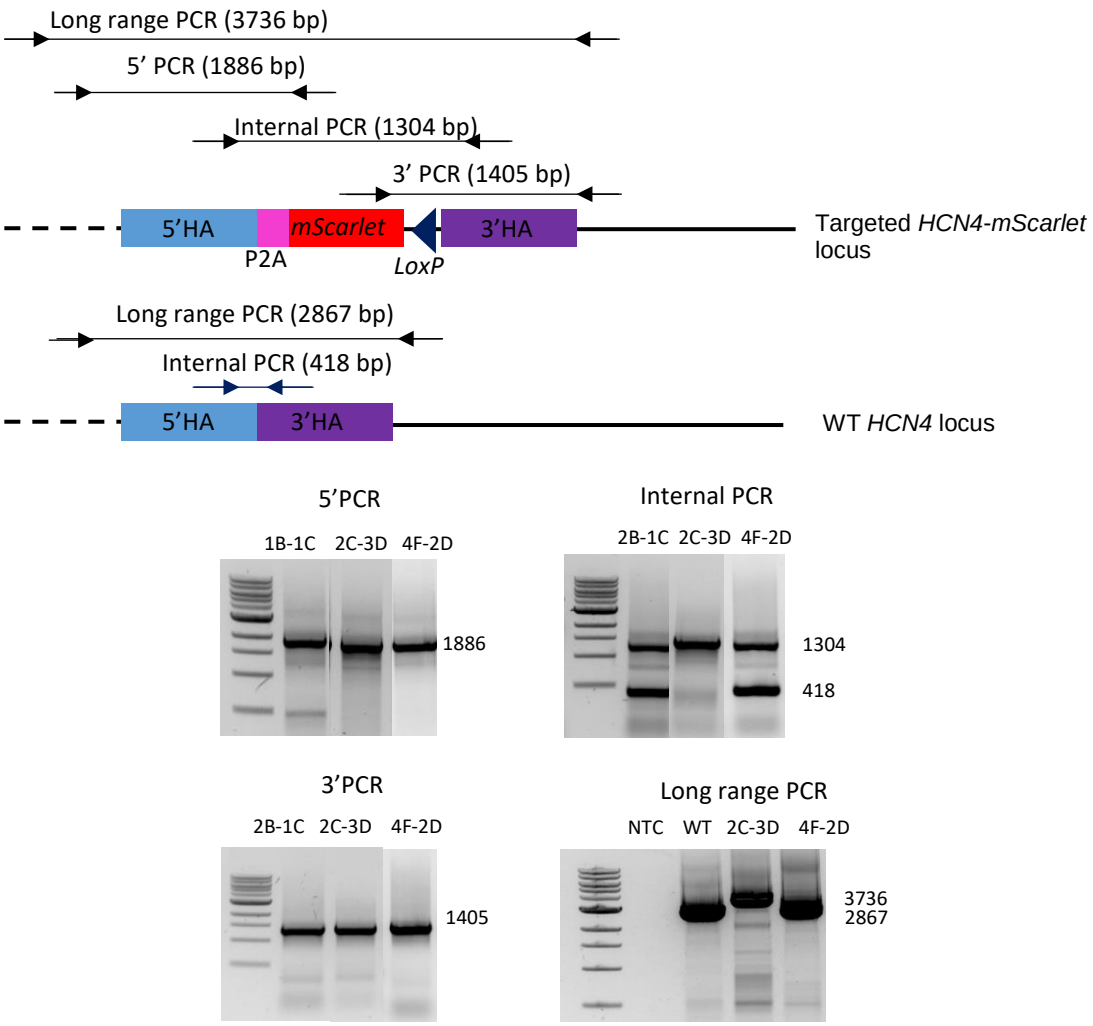
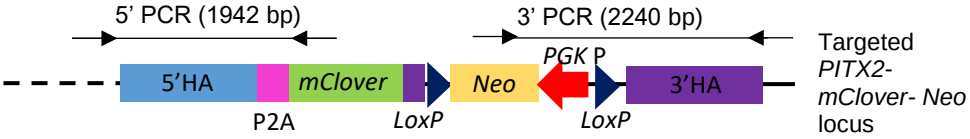


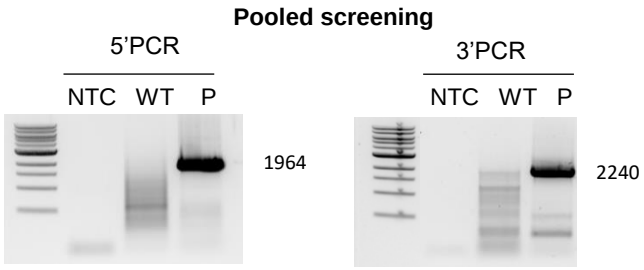
Figure 3-18 Conformational PCR of isolated *HCN4-mScarlet* iPSCs after clonal expansion. 5’ and 3’ PCR screenings indicated successful P2A-*mScarlet* insertion downstream of the *HCN4* locus devoid of the *Neo* cassette. Internal PCR confirmed the intact P2A-*mScarlet* insert with a heterozygous status in the clone 2B-1C and 4F-2D. Long range PCR suggested homozygous insertion of the *mScarlet* sequence at the *HCN4* locus in the clone 2C-3D. NTC, no template control.

3.3.3 Generation of *PITX2-mClover* line

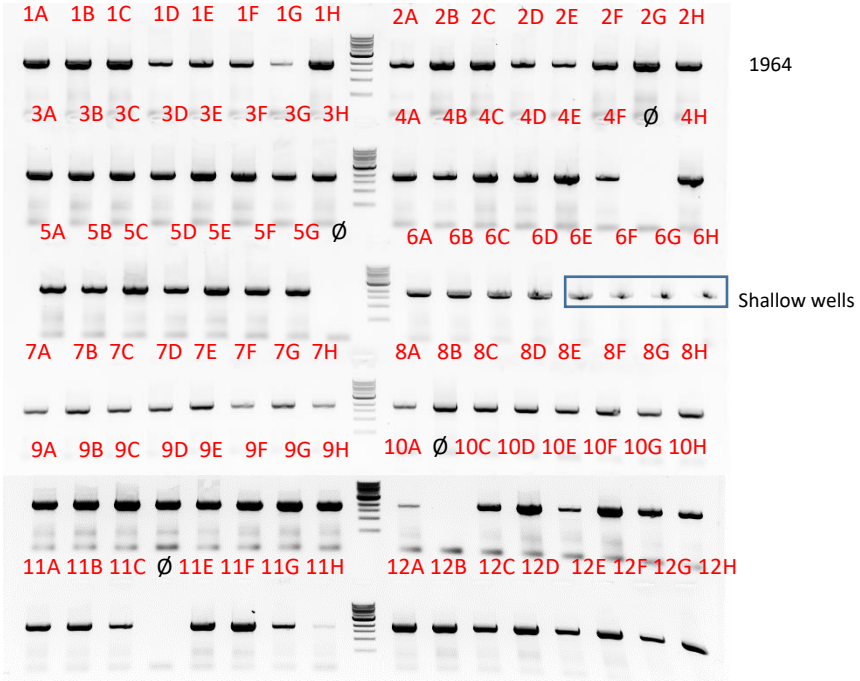
5’ PCR screening using the primers PITX2-F2 and mClover5’-R and 3’ PCR screening using the primers exNeo2 and PITX2-R2 indicated successful integration of P2A-*mClover-Neo* at the *PITX2* locus in the pool of KOLF2-C1 cells electroporated with the RNP/*PITX2*-mClover targeting vector (Figure 3-19A). Following low density plating, individual clones showing a successful 5’ integration were also isolated, 100% being positive (Figure 3-19B). Subsequent 3’ screening of 24 selected 5’-positive clones showed 100% positive integration at this end of the vector (Figure 3-19C).

A

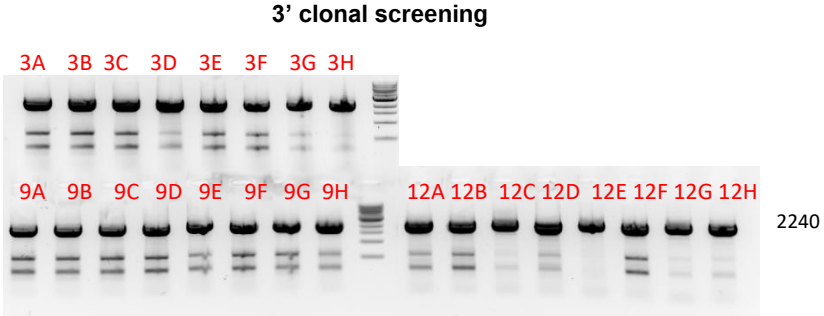

5' PCR (1942 bp)
3' PCR (2240 bp)
PGK P
5'HA
mClover
Neo
3'HA
P2A
LoxP
LoxP
Targeted PITX2-mClover- Neo locus


Pooled screening
5'PCR
NTC WT P
1964
3'PCR
NTC WT P
2240

B


5' clonal screening
1A 1B 1C 1D 1E 1F 1G 1H
2A 2B 2C 2D 2E 2F 2G 2H
1964
3A 3B 3C 3D 3E 3F 3G 3H
4A 4B 4C 4D 4E 4F Ø 4H
5A 5B 5C 5D 5E 5F 5G Ø
6A 6B 6C 6D 6E 6F 6G 6H
Shallow wells
7A 7B 7C 7D 7E 7F 7G 7H
8A 8B 8C 8D 8E 8F 8G 8H
9A 9B 9C 9D 9E 9F 9G 9H
10A Ø 10C 10D 10E 10F 10G 10H
11A 11B 11C Ø 11E 11F 11G 11H
12A 12B 12C 12D 12E 12F 12G 12H

C


3' clonal screening
3A 3B 3C 3D 3E 3F 3G 3H
9A 9B 9C 9D 9E 9F 9G 9H
12A 12B 12C 12D 12E 12F 12G 12H
2240

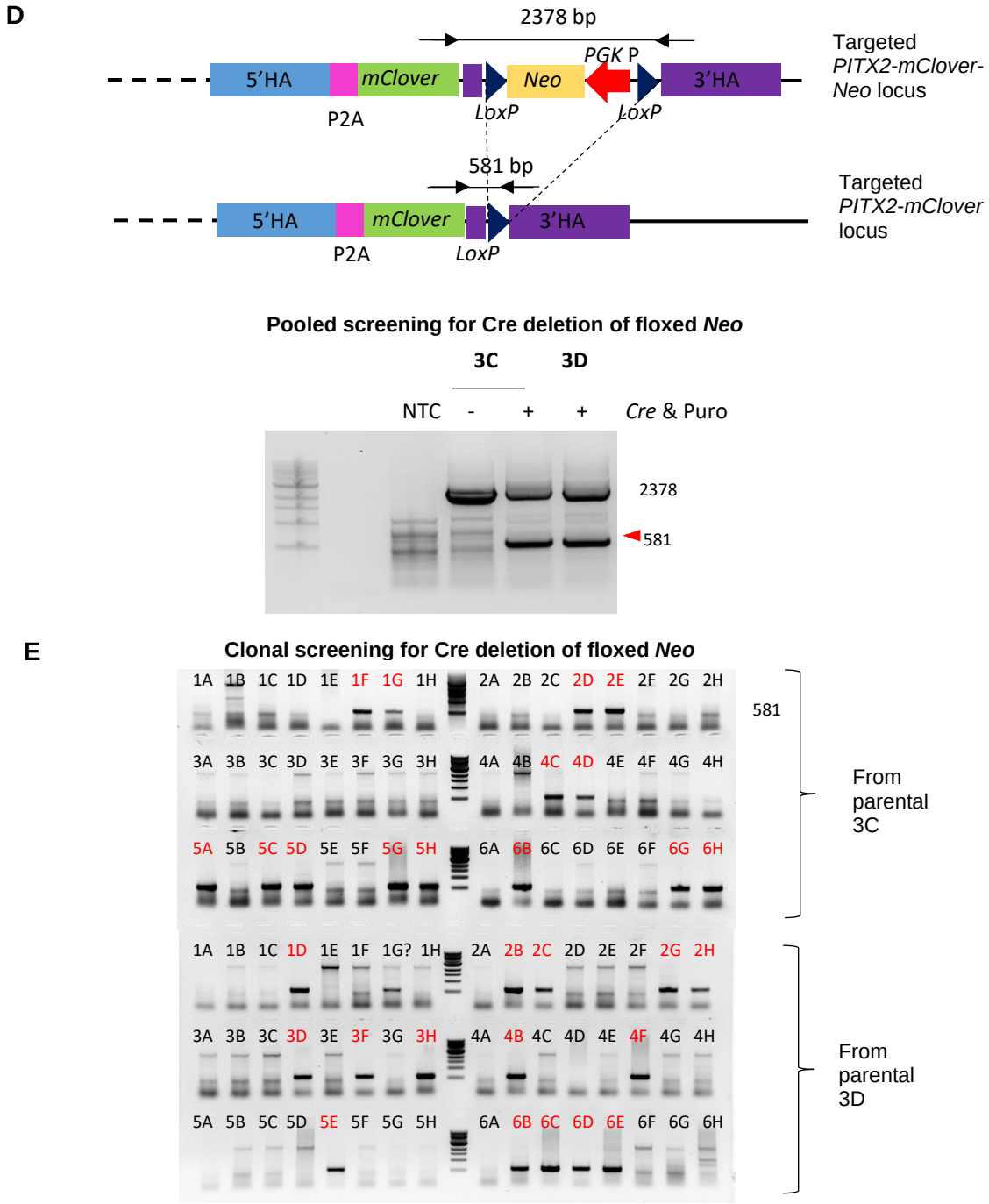


Figure 3-19 Genotypic screening for *PITX2-mClover* line. A) 5' and 3' PCR screenings in the pool of CRISPR/Cas9-transfected iPSCs. B) Clonal screening for 5' integration. C) Clonal screening for 3' integration. The promising clone 3C and 3D were subsequently transfected with a *Cre* expression plasmid to excise the floxed *Neo* cassette. D) PCR screening for *Neo* deletion in the pool of *Cre*-transfected cells followed by puromycin selection. E) Clonal screening for *Neo* deletion of puromycin-treated clones derived from the parental *Neo*-positive 3C and 3D clones. NTC, no template control; P, pool of transfected cells; *PGK* P, phosphoglycerate kinase promoter.

The promising *Neo*-carrying clones 3C and 3D were transfected with the Cre-Pac plasmid, followed by puromycin treatment. PCR analysis using the primers mClover-F2 and PITX2-R suggested successful Cre-Lox-mediated removal of the floxed *Neo* cassette in the pool of the transfected cells (Figure 3-19D) and in subsequent clonal progeny colonies (Figure 3-19E).

The promising clones *Neo*-positive-3C-derived 3C-4C, 3C-4D and 3C-6G and *Neo*-positive-3D-derived 3D-3D, 3D-3H and 3D-6C were expanded and reconfirmed for the integration of the P2A-*mClover* at the *PITX2* locus by 5' PCR using the primers PITX2-F2 and mClover5'-R and 3' PCR using the primers mClover-F2 and PITX2-R2 (Figure 3-20). A further internal PCR assay using the primers PITX2-F/R confirmed the integrity of the full-length P2A-*mClover* and suggested a homozygous status in the clones derived from the *Neo*-positive-3D, which was later confirmed by a long-range PCR using the primer PITX2-F2/R2 (Figure 3-20). However, the internal PCR indicated an unexpected Cre-Lox recombination in all the clones derived from the *Neo*-positive-3C which resembled a selective *Neo* excision on a single allele in a homozygous line, resulting in mixed amplicons with and without the *Neo* cassette. This homozygous targeting was also suggested by a long-range PCR using the primers PITX2F2/R2. Heterozygous clones may be obtainable from e.g. the *Neo* positive clone 3E (Figure 3-20, internal PCR), but the whole process of repeating the *Cre* transfection and the subsequent genotyping would take time. Therefore, the homozygous 3D-3D, 3D-3H and 3D-3C were selected for the iPSC quality control checks.

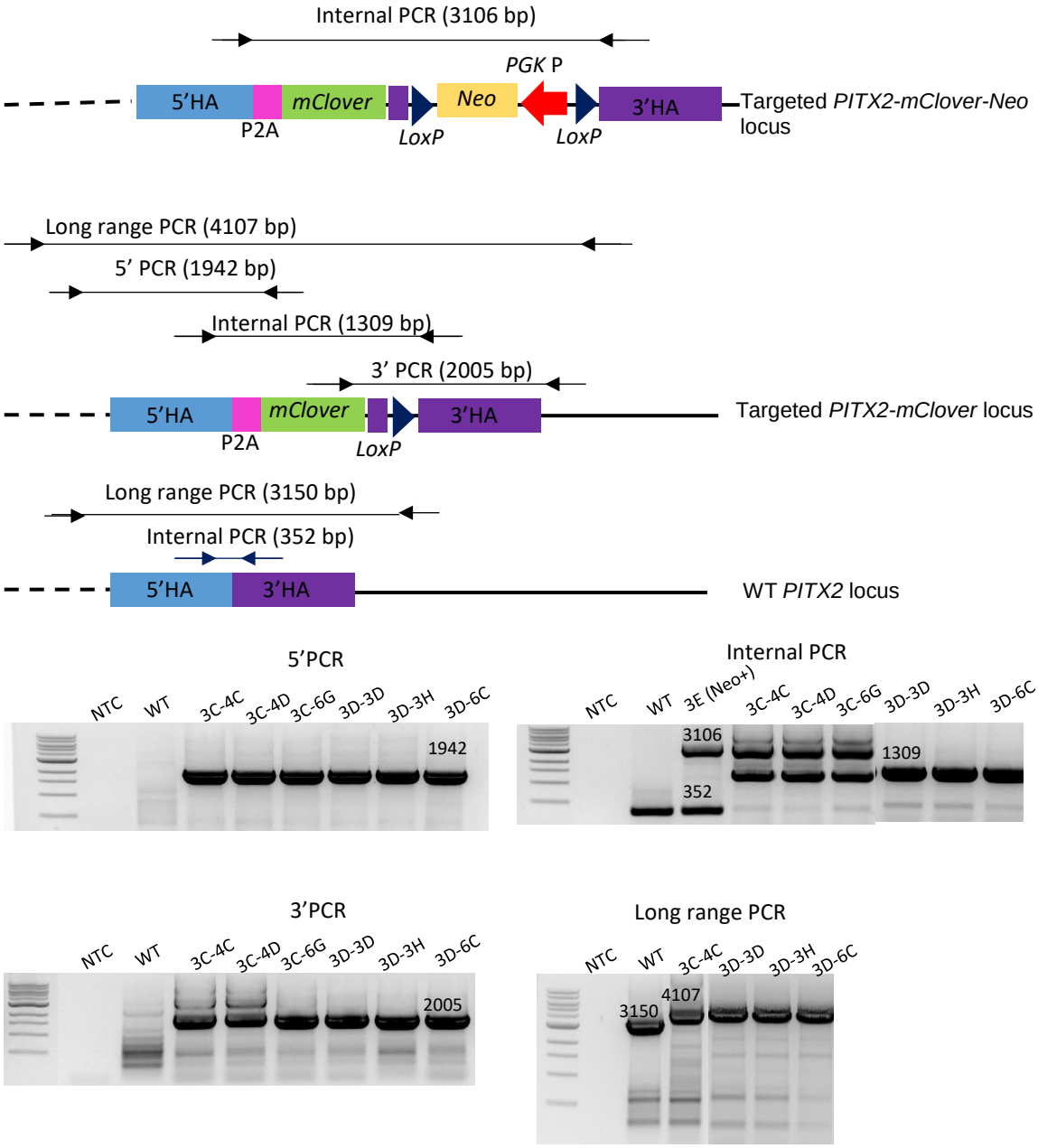
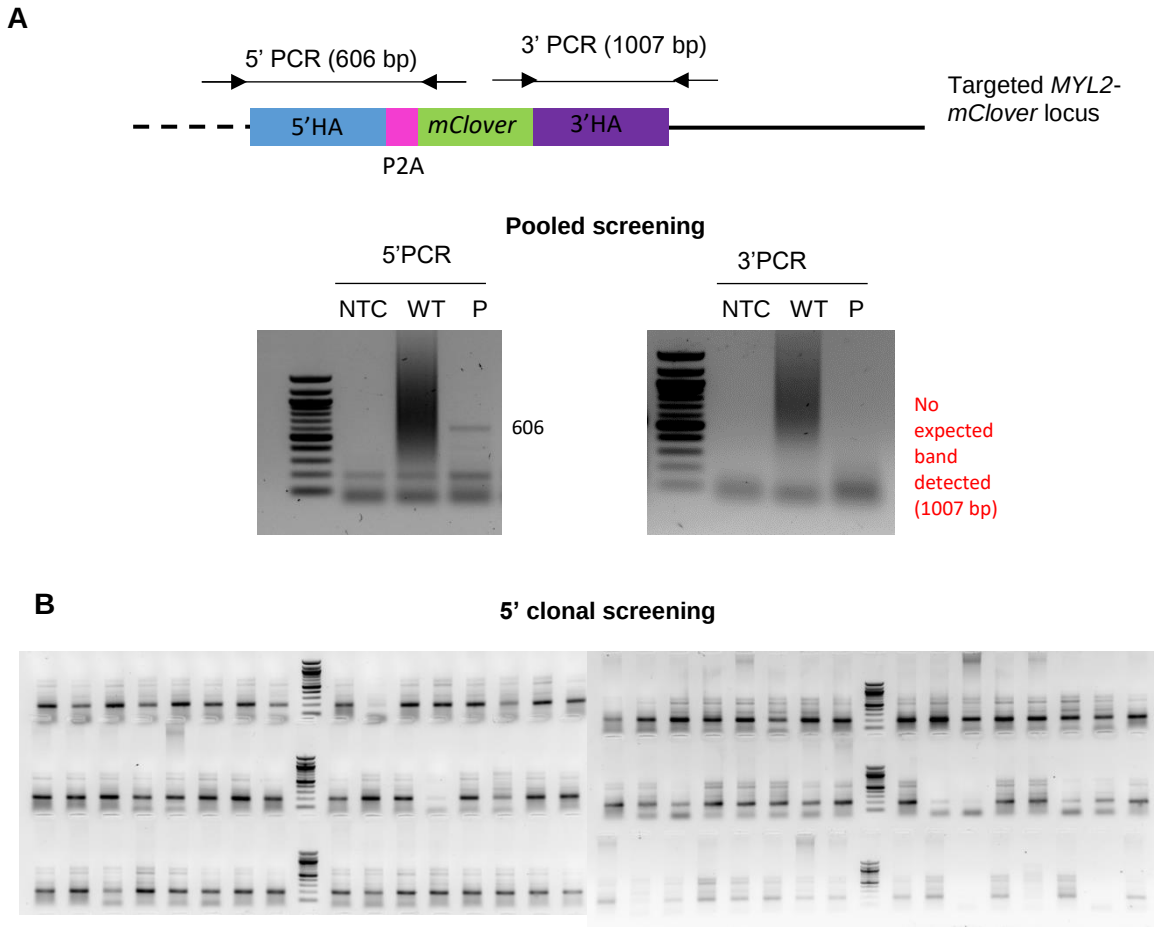


Figure 3-20 Conformational PCR of isolated *PITX2-mClover* reporter iPSCs after clonal expansion. 5' and 3' PCR indicated successful P2A-*mClover* insertion downstream of the *PITX2* locus in all the selected clones. Internal PCR clearly confirmed the intact P2A-*mClover* insert devoid of the *Neo* cassette with a homozygous status in the clones 3D-3D, 3D-3H and 3D-6C, consistent with the long range PCR results. Despite being suggested by the long range PCR as being homozygous for the fluorophore insert, the clones 3C-4C, 3C-4D and 3C-6G, showed an imperfect *Neo* excision with mixed *Neo*-positive and *Neo*-negative PCR amplicons. NTC, no template control; *PGK* P, phosphoglycerate kinase promoter.

3.3.4 Generation of *MYL2-mClover* line

The generation of this fluorescent reporter line was slightly different from the previous 3 lines as it relied on a transient transfection of 1.3 µg CRISPR/Cas9 expression plasmid pX330-puro-MYL2 and 250 ng lssDNA donor template carrying the P2A-*mClover* construct. A faint band of 5' PCR using the primers MYL2-F4 and mClover5'-R suggested a low efficiency of integration of the fluorescent reporter at this end of the donor in the pool of the transfected cells (Figure 3-21A). Unfortunately, 3' PCR using the primers mClover-F and MYL2-R3 showed no positive band at the expected size (1007 bp) in the pool of the transfected cells (Figure 3-21A). Further 5' and 3' genotyping in isolated colonies found only 2/512 promising clones (6F and 11C) (Figure 3-21B).



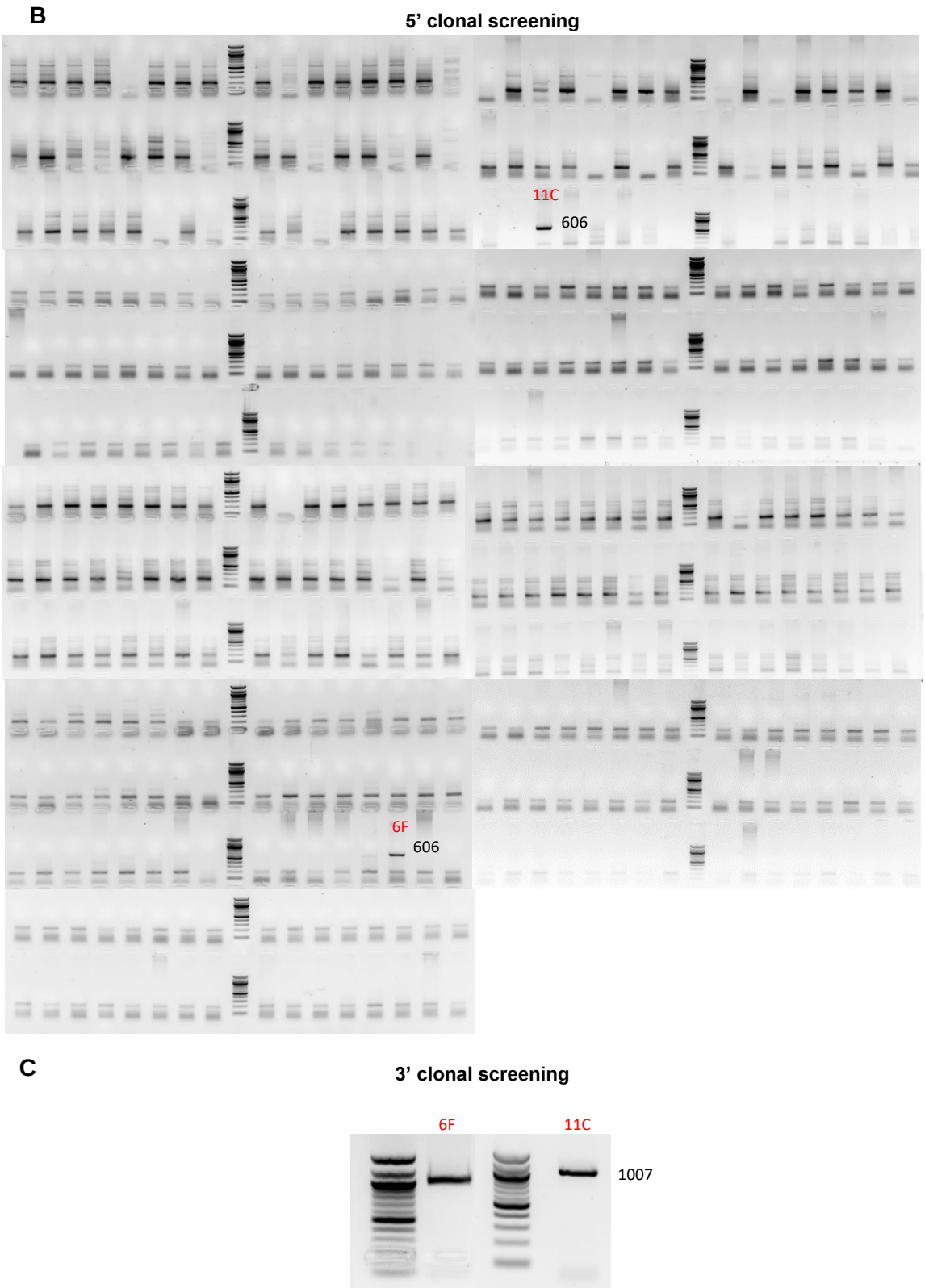


Figure 3-21 Genotypic screening for *MYL2-mClover* line. A) 5’ PCR indicated insertion of the 5’ end of the P2A-*mClover* at the *MYL2* locus in the pool of CRISPR/Cas9-transfected iPSCs, but 3’ PCR failed to confirm the insertion at the other end. B) Clonal screening for 5’ integration of the P2A-*mClover*. C) Clonal screening for 3’ integration of the P2A-*mClover* of the 5’ positive clones 6F and 11C. NTC, no template control; P, pool of transfected cells.

Both clones 6F and 11C were expanded and analysed for the integrity of the full-length P2A-*mClover* insertion by internal PCR using the primers MYL2F/R (Figure 3-22) and DNA sequencing. This strategy also suggested that clone 6F was heterozygous for the insertion while clone 11C may be homozygous. However, a long-range PCR using the primers MYL2F4/R3 indicated heterozygosity for both clones (Figure 3-22). Moreover, DNA sequencing identified mutations at the end of the *MYL2*'s coding sequence within the non-targeted allele. The clone 6F had 3 amino acid changes and 6 amino acid loss whilst the clone 11C showed a deletion of the whole last exon (exon 7).

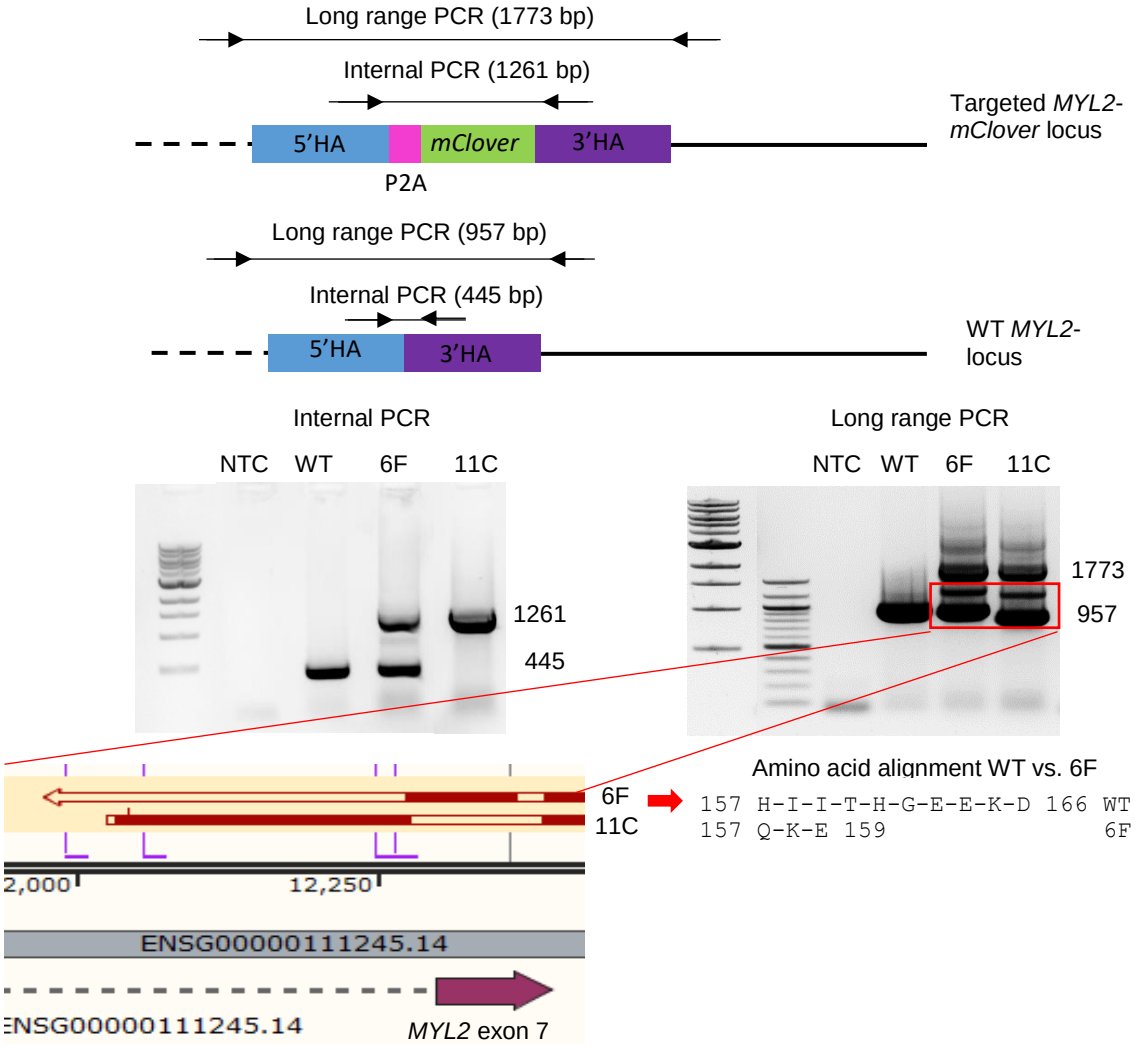


Figure 3-22 Conformational PCR of isolated *MYL2-mClover* reporter iPSCs after clonal expansion. Internal PCR confirmed a full-length integration of the P2A-*mClover* in both clones 6F and 11C. Long range PCR indicated heterozygosity of the insert for both clones where DNA sequencing suggested deletion mutations of the *MYL2* gene in the non-targeted alleles. The clone 6F had a small deletion mutation at the end of the *MYL2*'s coding sequence (6 amino acids) with amino acids 157-159 changed whereas the clone 11C showed a deletion of the whole coding sequence within exon 7.

Due to the extremely low targeting efficiency, it was considered not worth repeating the transfection and subsequent genotyping for this reporter line project. The undesired CRISPR-related deletions seen in the non-targeted alleles for both clones may be correctable by CRISPR/Cas9-mediated insertion of the WT sequence but the whole process would also be laborious. Therefore, both clones were subject directly to CM differentiation (Chapter 4) to investigate the consequence(s) of the truncations of this sarcomere protein.

3.4 iPSC quality checks

CRISPR/Cas9-mediated genome editing can lead to unintended chromosomal instability, not only by a megabase-scale deletion (Adikusuma et al., 2018, Alanis-Lobato et al., 2021), but also a variation in the chromosome number e.g. whole chromosome deletion (Papathanasiou et al., 2021), all of which may adversely influence gene expression and lead to unexpected pathogenic phenotypes. To ensure normal karyotype, the engineered fluorescent reporter lines were subject to metaphase chromosome counting

In addition, the genomic perturbation or simple prolonged clonal expansion may negatively alter expression of the stem cell markers, potentially supressing the iPSC-CMs differentiation potential. To investigate if the iPSC pluripotency was retained, the engineered fluorescent reporter lines were analysed for expression of the pluripotency marker including OCT4, SOX2, SSEA4 and NANOG by flow cytometry.

3.4.1 Quality checks for *MYL7-mClover* line

Both heterozygous clones 7H and 11B displayed a normal karyotype (46 chromosomes) (Figure 3-23A). It should be noted that the clone 11B has a sub-population (only 1/25 cells) with 47 chromosomes. Flow cytometry analysis of expression of the pluripotency markers including OCT4, SOX2, SSEA4 and NANOG confirmed that both clones retained pluripotent status (Figure 3-23B). Therefore, both were qualified iPSC clones.

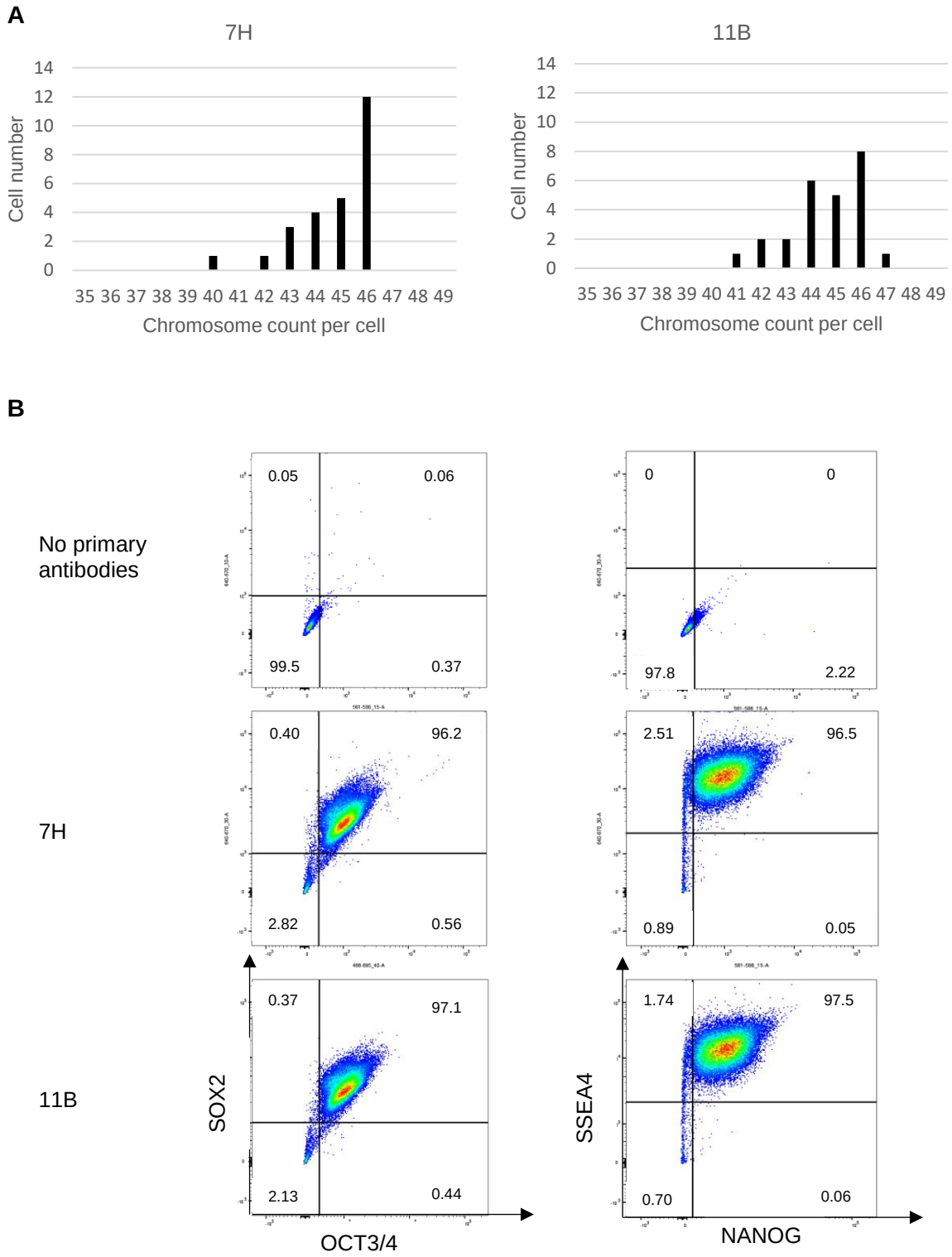
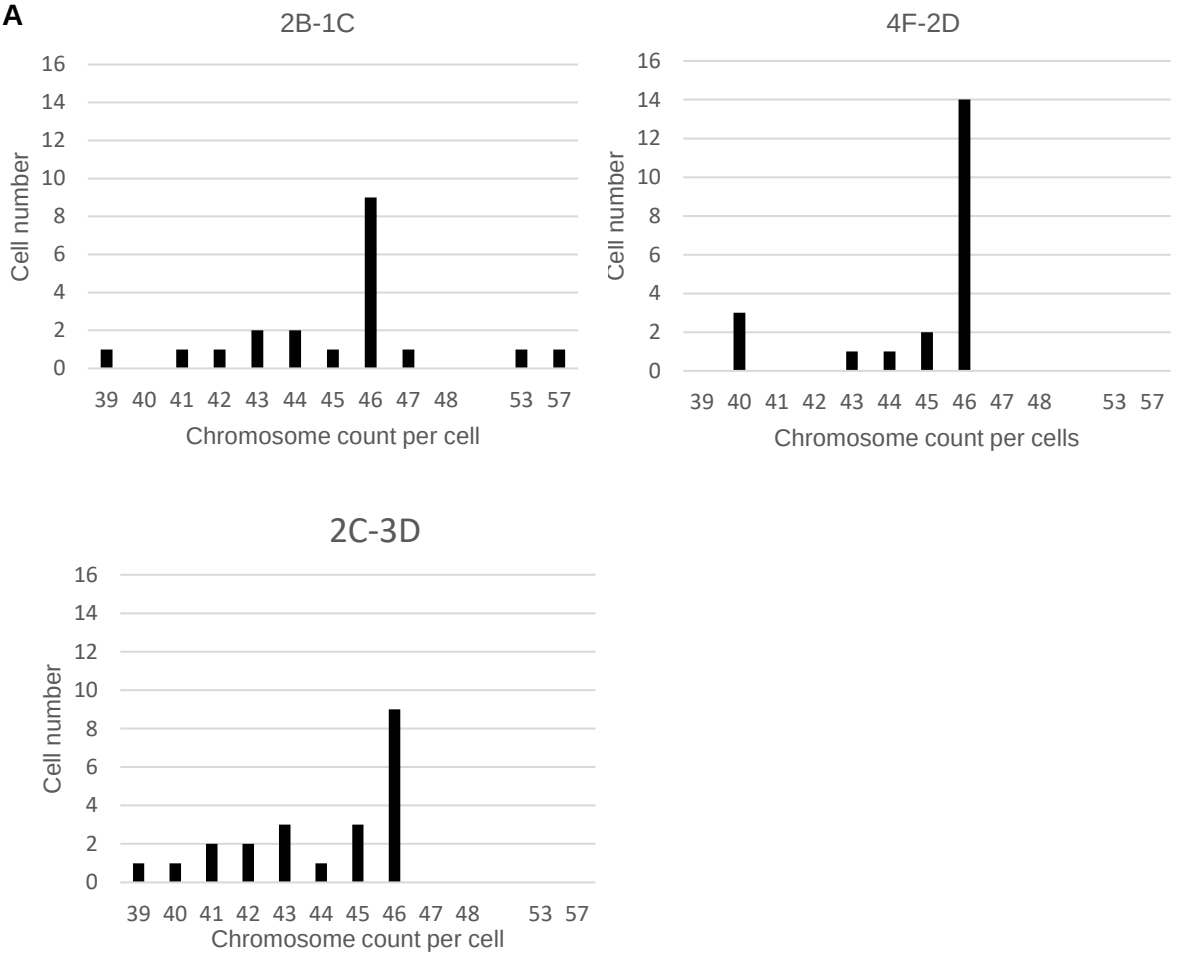


Figure 3-23 *MYL7-mClover* iPSC line's quality control. A) Distribution of chromosome count per cell. B) Flow cytometry analysis of pluripotency marker expression.

3.4.2 Quality checks for *HCN4-mScarlet* line

Both heterozygous clone 4F-2D and homozygous 2C-3D had a normal karyotype (Figure 3-24A). Meanwhile, the other heterozygous clone 2B-1C showed aneuploid sub-populations with 47 chromosomes (1/20 cells), 53 chromosomes (1/20 cells) and 57 chromosomes (1/20 cells). Flow cytometry analysis of pluripotency markers including OCT4, SOX2, SSEA4 and NANOG confirmed that all clones retained pluripotent status (Figure 3-24B). Due to the suboptimal karyotype, the clone 2B-1C was excluded from the further experiments.

A



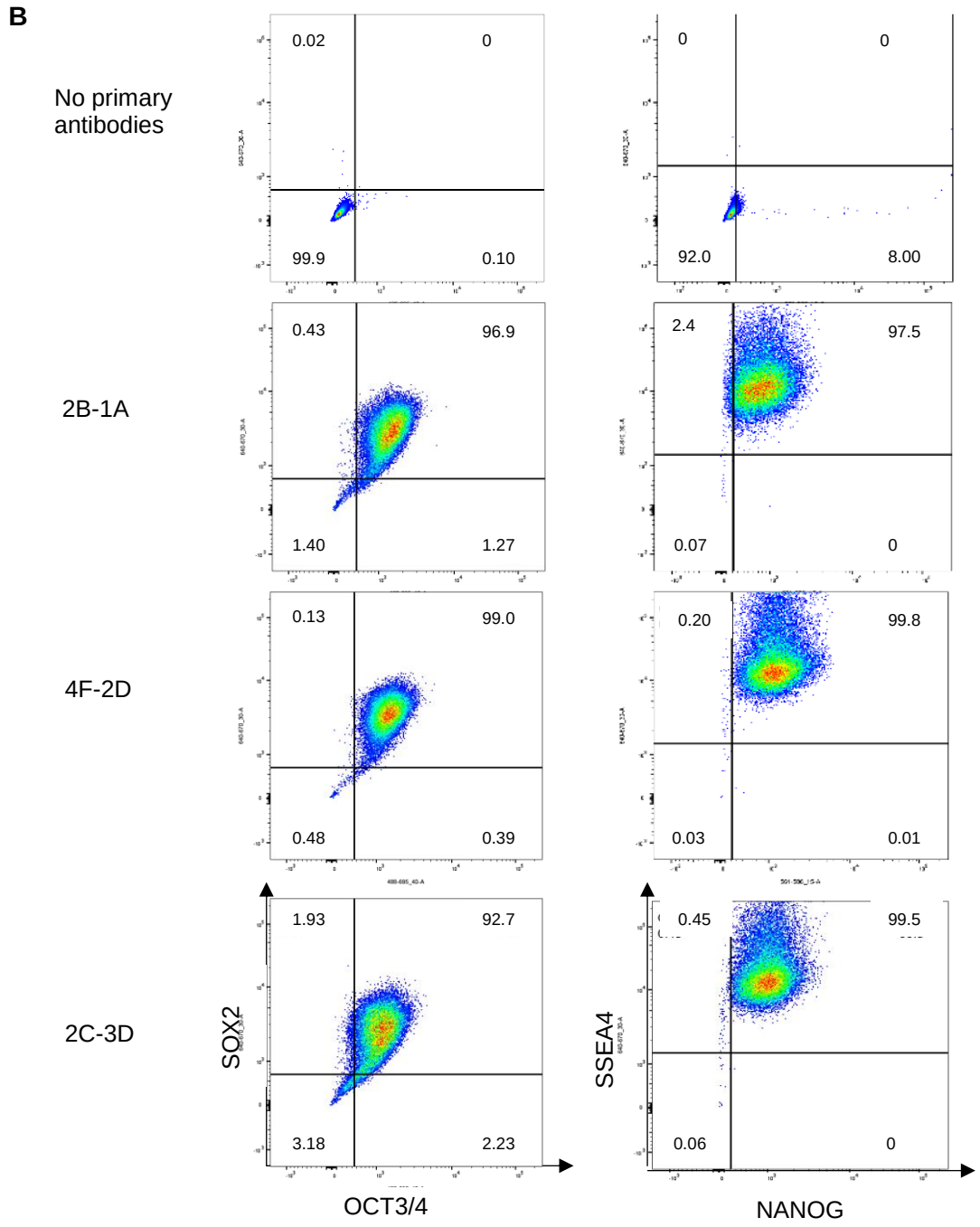
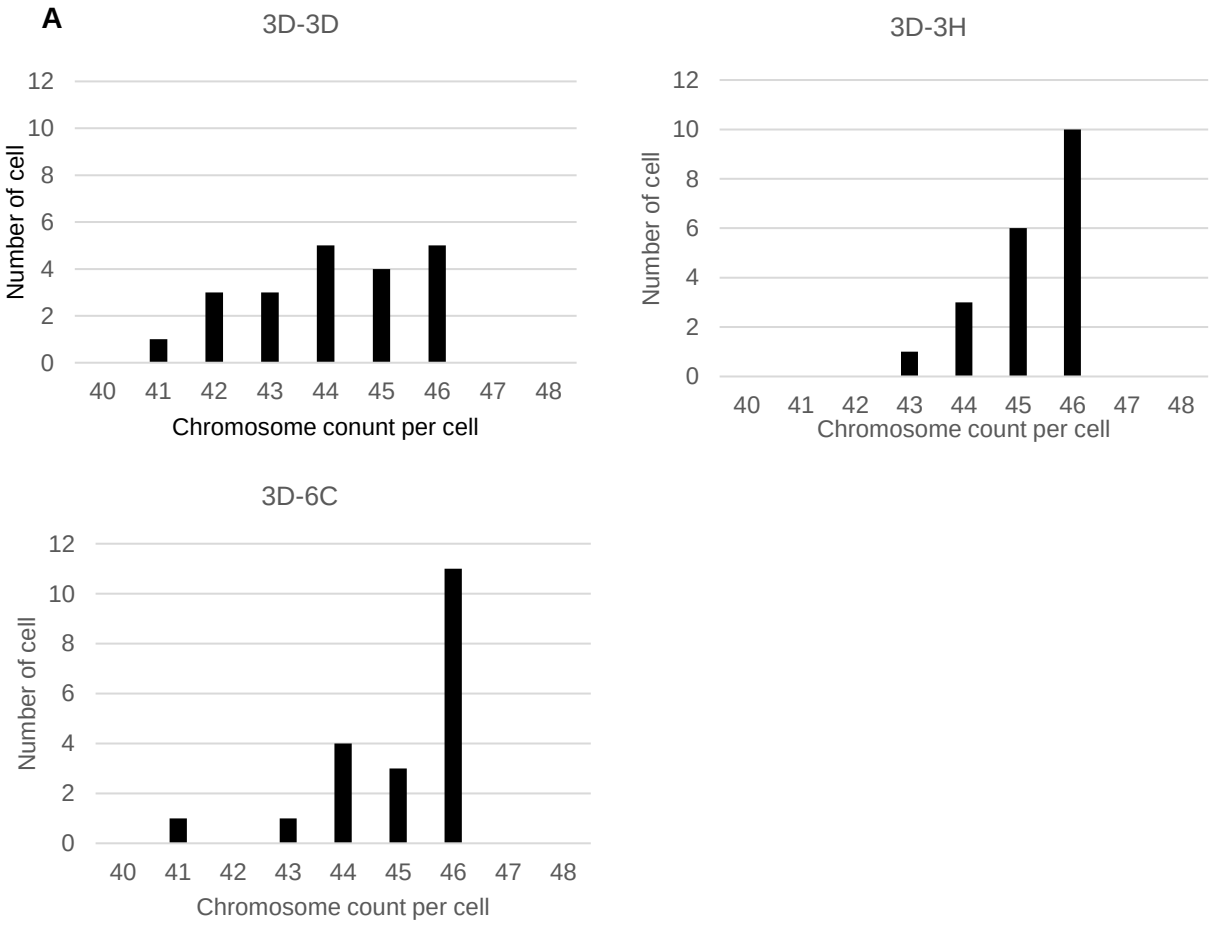


Figure 3-24 *HCN4-mScarlet* iPSC line's quality control. A) Distribution of chromosome count per cell. B) Flow cytometry analysis of pluripotency marker expression.

3.4.3 Quality checks for *PITX2-mClover* line

Both homozygous clones 3D-3H and 3D-6C had a normal karyotype (Figure 3-25A). Meanwhile, the clone 3D-3D showed two modal values (44 and 46 chromosomes), thus was considered karyotypically abnormal. Flow cytometry analysis of pluripotency markers including OCT4, SOX2, SSEA4 and NANOG confirmed that all clones retained pluripotent status (Figure 3-25B). Due to the suboptimal karyotype, the clone 3D-3D was excluded from the further experiments.



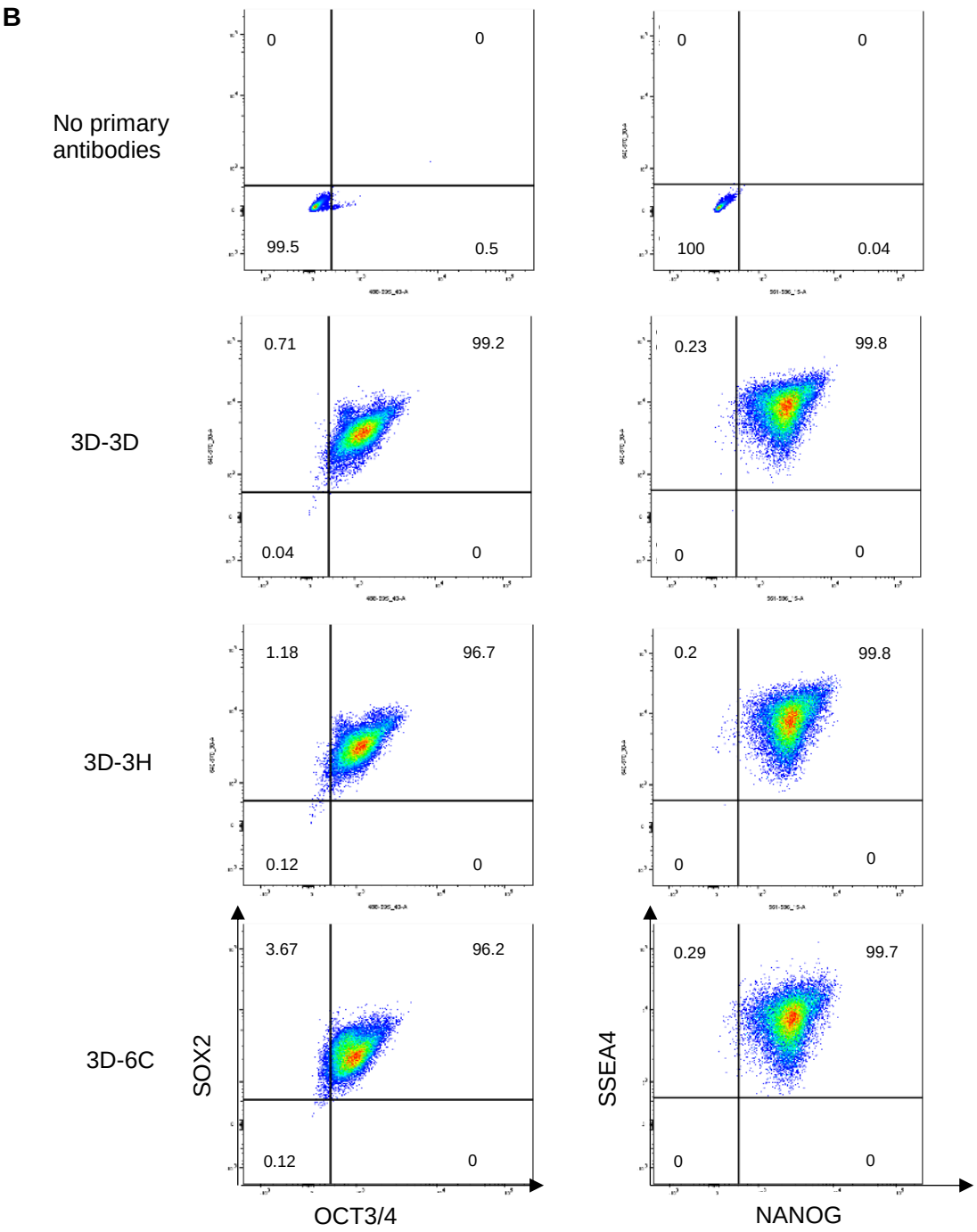


Figure 3-25 *PITX2-mClover* iPSC line's quality control. A) Distribution of chromosome count per cell. B) Flow cytometry analysis of pluripotency marker expression.

3.4.4 Quality check for *MYL2-mClover* line

The undesired *MYL2* mutations in both clones 6F and 11C raised a concern over their CM differentiation potential as the gene encodes the thick filament component. It will be discussed in Chapter 4 that defects in CM differentiation and/or expression of both mClover and MYL2 proteins were identified, disqualifying both clones from being fluorescent reporter iPSC-CMs. Hence, iPSC quality checks were no longer needed.

3.5 Discussion

In this chapter, 4 fluorescent reporter iPSC lines were generated using CRISPR/Cas9, each of which targeted a CM-subtype specific gene. Establishment of the atrial-specific *MYL7-mClover* and *PITX2-mClover* lines and the nodal-specific *HCN4-mScarlet* line relied on the same approach by RNP electroporation with a plasmid donor containing a selection marker *Neo*. The ventricular-specific *MYL2-mClover* line was produced by lipofection of a CRISPR/Cas9 expression plasmid and an lssDNA donor since the homology arm length was the limiting factor for use of a plasmid donor.

3.5.1 Modes of CRISPR/Cas9 delivery

One of the main advantages of CRISPR/Cas9 as a tool for genome editing is its simplicity, where the system's specificity depends on the gRNA component. The CRISPR/Cas9 design can therefore be tailored to the locus of interest by simple change of the gRNA sequence, making it a favourable research tool. Meanwhile, the other technologies e.g. zinc finger nucleases and transcription activator-like effector nucleases rely on a complicated protein engineering customised to each target sequence. Apart from the CRISPR sequence itself, modes of CRISPR/Cas9 delivery also determine the genome editing efficiency (Lino et al., 2018, Yip, 2020).

Firstly, a CRISPR/Cas9-encoding DNA plasmid can be introduced into the cells by either chemical transfection, electroporation or viral transduction. This delivery mode allows a long expression time for both gRNA and Cas9 and thus may increase the targeting efficiency.

However, the plasmid DNA format has a slow onset of gene editing as it takes time for transcription, translation and RNP complexing. The prolonged presence of the CRISPR/Cas9 RNP at high concentration in the cells also increases the risk of off-target events (Wu et al., 2014). Despite the intracellular plasmid stability, the efficiency of the *MYL2-mClover* line production was extremely low (2/512). It may partly be explained by the different lifetime of the CRISPR/Cas9 plasmid and the lssDNA donor. Compared to dsDNAs, the structure of single-stranded DNAs (ssDNAs) is highly labile (Murphy et al., 2004, McIntosh et al., 2014) and their linear form (required for HDR) is thermodynamically unfavourable. ssDNAs are also susceptible to enzymatic and chemical degradation.

Secondly, Cas9 can be delivered as mRNA together with relevant gRNA(s) to accelerate the onset of target editing and reduce the off-target effects (Wu et al., 2014). However, the RNA(s) short lifetime may compromise the editing efficiency. Furthermore, both mRNA and gRNA are prone to RNAse-mediated degradation, thus their delivery requires a careful optimisation.

Thirdly, a pre-complexed, readily active CRISPR/Cas9 RNP allows an immediate gene editing in the nucleus with minimised off-target effects (Burger et al., 2016, Wu et al., 2014). The gRNA stability is also ensured upon reconstitution with Cas9. RNP electroporation was proven to be highly efficient for HDR experiments in my hands. Delivery by electroporation might also be suitable for lssDNA donor templates. However, this method yielded no success in the *MYL2-mClover* reporter project (data not shown) and several other targeting projects, possibly due to lssDNA structural damage by electrical pulses or an inability for long DNA strands to transgress the membrane via the pore. The HDR efficiency can be further enhanced by covalently tethering the donor template with the RNP complex (Aird et al., 2018), thereby increasing the effective concentration of the donor template at the DSB site.

3.5.2 DNA donor template

The type of DNA donor templates in this chapter was guided by the available homology arm length at the targets. Repetitive sequences were avoided in homology arms as they can be

substrates for either random integration or chromosomal rearrangement (Pascarella et al., 2020, George and Alani, 2012). Both single stranded and double strand breaks, e.g. by CRISPR/Cas9 induction, may increase the risk of homologous recombination-mediated chromosomal structure alterations, resulting in loss of function of the target gene.

The sequences around the *MYL7*, *HCN4* and *PITX2* stop codons were minimally populated with repeat motifs and subsequently their fluorescent reporter donor were designed as a plasmid with ~1000-bp HAs. The high success of the targeted insertion was clearly due to the antibiotic selection. G418 (also known as Geneticin) treatment ensured a stable genomic insertion of *Neo*-containing sequences. G418 is an aminoglycoside antibiotic which inhibits the elongation step of protein translation in both prokaryotes and eukaryotes (Eustice and Wilhelm, 1984). Only cells acquiring and expressing *Neo* survive in the presence of G418 (Southern and Berg, 1982). While a transient *Neo* expression, e.g. from episomal vector, allows cells to grow for a certain period (as long as the *Neo* construct still persists), a stable *Neo* expression, e.g. from genomic incorporation, provides a permanent resistance against G418. In addition, dTA included in the plasmid backbone also eliminated a random integration of the vector independent of homologous recombination. Similarly to G418, dTA blocks protein synthesis by inhibiting the eukaryotic translation elongation factor 2 (Lange et al., 2019) and thus, any stable genomic integration of this cassette would kill the cells.

MYL2 was the only locus that showed a high abundance of repeat motifs surrounding the stop codon and thus, the *MYL2*-P2A-*mClover* DNA donor template was synthesised in a form of lssDNA with minimal HA length. The low yield of the targeted insertion may be attributable to the design of the lssDNA. Firstly, the large size of the insert (~900 bp) relative to the size of HAs (~100 bp) may compromise efficient recombination events. Secondly, there may be a lssDNA strand preference. The top strand of the *MYL2-mClover* lssDNA used in this study was the CRISPR/Cas9 target strand. A use of the non-target strand may provide a better substrate for HDR but this remains to be investigated. The strand dependency appears to vary among target loci (Ranawakage et al., 2021). Thirdly, the length of HAs plays a critical role for

an optimal HDR (Ranawakage et al., 2021). The length of either the 5' or 3' HA may need to be optimised to suit the locus as it may not necessary be the case that a long HA would be beneficial and vice versa. Nevertheless, the repeat motifs around the stop codon may not be as problematic for use of a plasmid donor template as predicted. (Chirikian et al., 2021), (Galdos et al., 2021) and (Luo et al., 2021) successfully used CRISPR/Cas9 to generate *MYL2* fluorescent reporter PSC-derived CM lines by relying on targeting plasmids with ~500-bp (or less) HAs.

3.5.3 Summary

In this chapter, 4 fluorescent reporter iPSC lines were established using CRISPR/Cas9, each of which had GSG-P2A-preceded fluorophore inserted downstream of each target gene in replacement of the original stop codon. The *MYL7-mClover*, *HCN4-mScarlet* and *PITX2-mCover* reporter lines displayed no CRISPR-related, undesired mutation. Meanwhile, the *MYL2-mClover* line acquired a deletion mutation at the *MYL2*'s C-terminus associated with the CRISPR activity on the non-targeted allele. All 4 lines were differentiated into CMs and characterised in Chapter 4.

Chapter 4

Characterisation of engineered fluorescent reporter iPSC-CM lines

4.1 Introduction

Advances in the ability to direct differentiation of iPSC into diverse cell lineages including CMs has expanded our knowledge of developmental biology and disease pathogenesis. However, the iPSC-CMs' population heterogeneity, including atrial-like, ventricular-like and nodal-like cells, presents limitations for their use in selective analysis. This thesis relied on WNT signalling modulation (Chapter 2) (Lian et al., 2012) to induce the cardiac differentiation, which favours ventricular-like CMs (Churko et al., 2018). Meanwhile, atrial-like CMs are the minor population and nodal-like cells are the rarest subtype (Churko et al., 2018). Subtype-specific fluorescent reporter lines could be useful systems to identify and purify these iPSC-CM subtypes for use in subtype-specific disease modelling.

This chapter mainly focuses on the characterisation of CMs derived from the fluorescent reporter iPSC lines described in Chapter 3, including atrial-specific *MYL7-mClover* and *PITX2-mClover*, ventricular-specific *MYL2-mClover* and nodal-specific *HCN4-mScarlet*.

I. Firstly, each iPSC line was differentiated using the conventional (ventricular) protocol (Lian et al., 2012) to test whether the differentiation potential was preserved after the CRISPR/Cas9-mediated genomic manipulation.

II. Secondly, an atrial differentiation protocol was established to facilitate characterisation of the atrial-specific reporter lines using the *MYL7-mClover* line as a model.

III. Thirdly, the CMs differentiated from each reporter line (by either the ventricular or atrial protocol) were analysed by qPCR, WB, fluorescent microscopy and FACS to investigate whether the expression of the fluorophore represents the expression of the target gene and faithfully indicates the respective CM subtype.

4.2 MYL7-mClover line

The heterozygous *MYL7-mClover* reporter iPSC line (clone 11B) showed a normal differentiation potential with spontaneous contraction visibly detected approximately from day 8-9 by the ventricular differentiation protocol.

4.2.1 Optimisation of retinoic acid (RA)-induced atrial differentiation

Establishment of an atrial differentiation protocol could be useful for characterisation of atrial reporter systems including the *MYL7-mClover*. The response of the reporter marker to the atrial differentiation could determine if the tested system represents the atrial lineage. Previous works suggested that retinoic acid (RA) signalling during the mesodermal stage (equivalent to day 3-5 in this study) is a prerequisite for atrial specification in iPSC-CM differentiation (Cyganek et al., 2018, Zhang et al., 2011, Lee et al., 2017).

4.2.1.1 Transcriptional expression of atrial and ventricular markers in response to RA

Four conditions of 1 μM RA supplementation were tested for optimisation for both the onset of the induction time point and the treatment length during the *MYL7-mClover* iPSC-CM differentiation, including days 3-7, days 3-9, days 4-7 and days 4-9 (Figure 4-1A). The differentiated cells were collected for analysis of gene expression at day 30 of differentiation by qPCR.

There was no change in pan-cardiac marker *ACTN2* expression in any treatment condition (Figure 4-1B), indicating that RA did not dramatically disturb the differentiation process. It was clearly shown that RA effectively suppressed the ventricular gene programme (*IRX4*, *HEY2* (hairy/enhancer-of-split related with YRPW motif protein 2) and *MYL2*) irrespective of the treatment conditions (Figure 4-1D). In contrast, there was a slight difference in atrial marker expression (*MYL7*, *NR2F2*, *KCNJ3* (inward rectifier K^+ channel subfamily J Member 3), *PITX2* and *MYH6*) (Figure 4-1C) between the RA conditions where the treatment on days 3-9 exerted the highest upregulation effect. Thus, days 3-9 condition was used as a standard condition for atrial differentiation in the subsequent experiments.

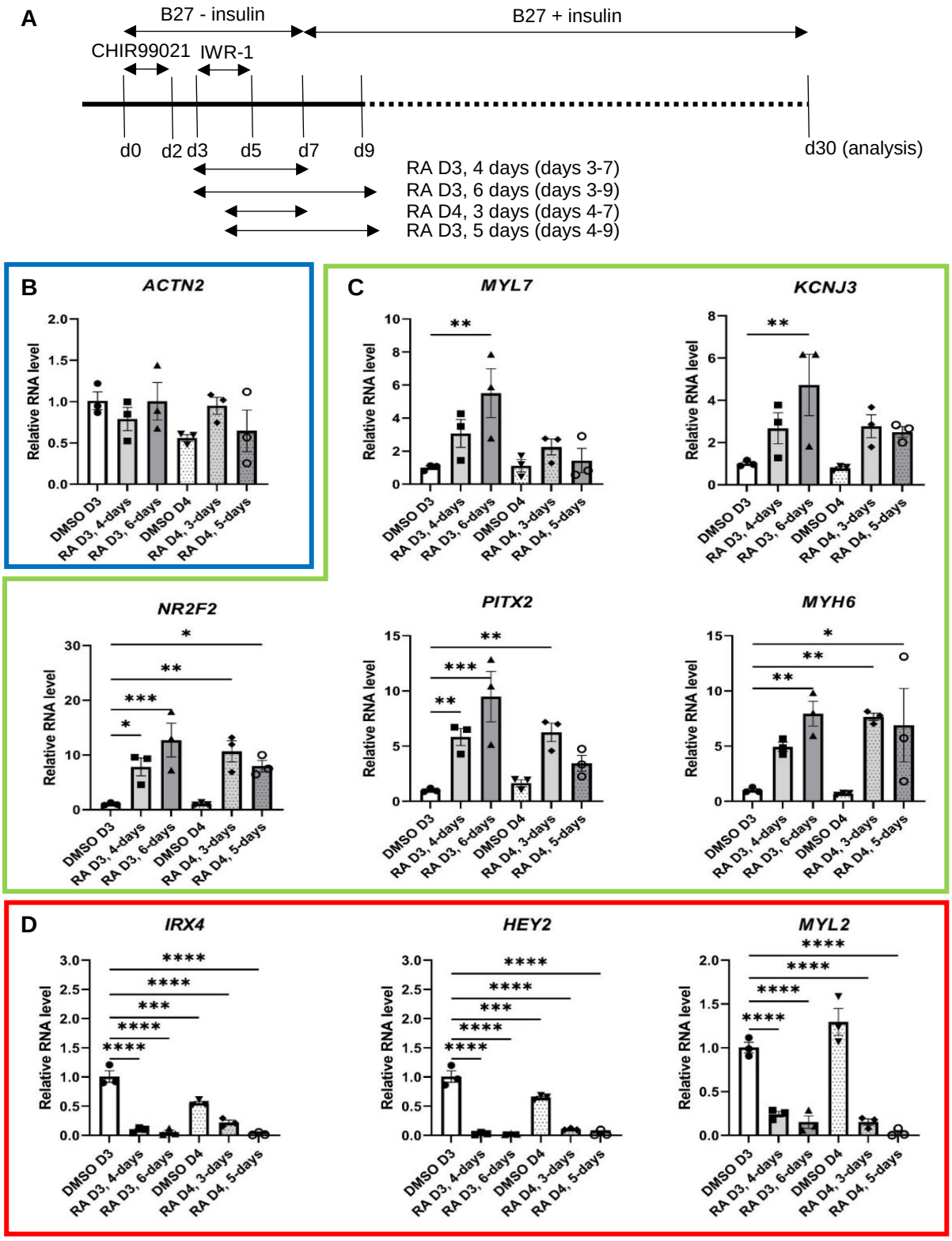


Figure 4-1 Gene expression profile for atrial differentiation in day 30 *MYL7-mClover* iPSC-CMs. A) Schematic of RA treatment conditions with respect to the days in the conventional (ventricular) iPSC-CM differentiation protocol. DMSO D3, DMSO treated on days 3-6; DMSO D4, DMSO treated on days 4-6; RA D3, 4-days, RA-treated on days 3-7; RA D3, 6-days, RA-treated on days 3-9; RA D4, 3-days, RA-treated on days 4-7; RA D4, 5-days, RA-treated on days 4-9. B) Pan-cardiac marker. C) Atrial markers D) Ventricular markers. Each dataset was obtained from different wells of differentiation using a single batch of iPSC. n=3, mean ± S.D. one-way ANOVA with Turkey's post-hoc. All expression levels were made relative to the DMSO D3 condition.

4.2.1.2 Protein expression of atrial and ventricular myosin light chains in response to RA

Protein from day 30 iPSC-CMs treated with RA (days 3-9 condition) was obtained for WB analysis (Figure 4-2). The ventricular light chain MYL2 expression level was almost completely undetectable, thereby confirming that ventricular gene programme was shut down. Meanwhile, the atrial light chain MYL7 was upregulated.

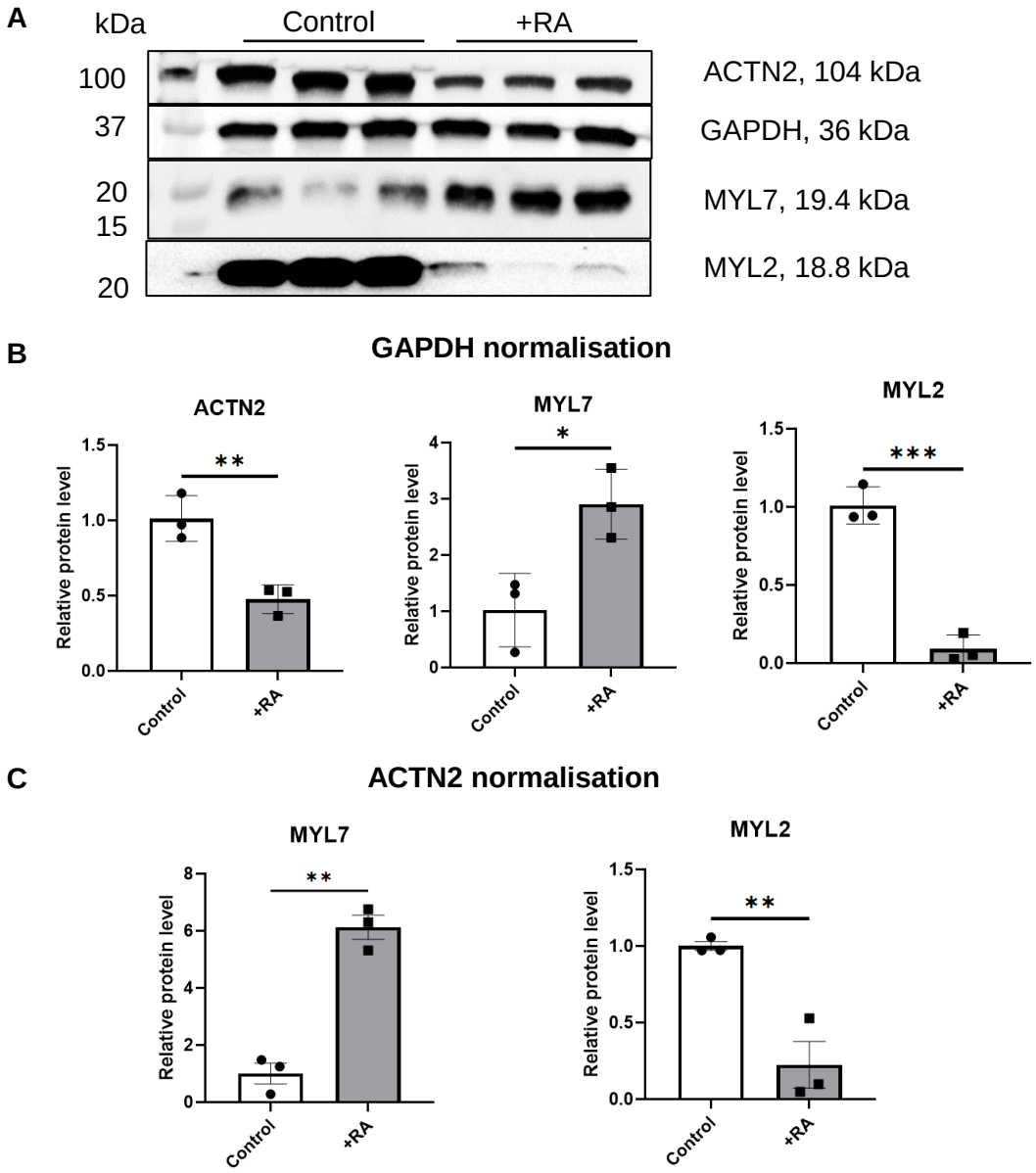


Figure 4-2 Protein expression analysis of atrial differentiation in day 30 *MYL7-mClover* iPSC-CMs. The cells were treated with RA (+RA) or without RA (control). A) WB. B) Quantification of protein from A as normalised by GAPDH. C) Quantification of protein from A as normalised by ACTN2. Since there was a reduction in ACTN2 by RA treatment observed in B, normalisation by ACTN2 may provide a more sensible comparison with respective to CM cell population. Each dataset was obtained from different wells of differentiation using a single iPSC batch. Mean ± S.D., n=3, Student t-test.

4.2.1.3 Effect of RA on cell size

Day 30 iPSC-CMs treated with RA were fixed and measured for cell area. RA induced a significant cell area reduction as compared to the non-treated control (Figure 4-3).

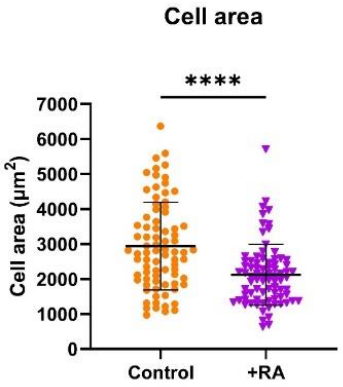


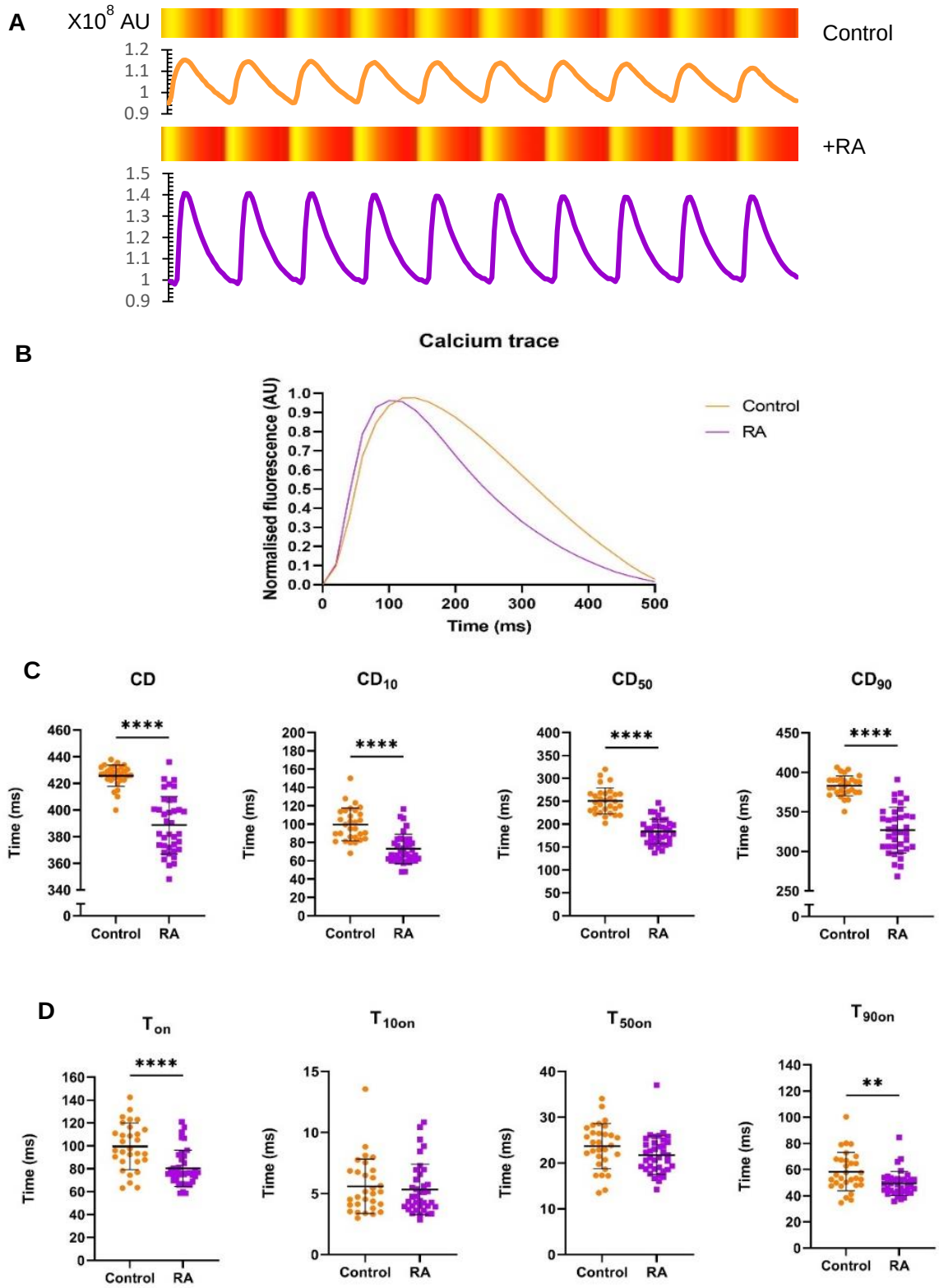
Figure 4-3 Quantification of cell area of day 30 *MYL7-mClover* iPSC-CMs upon atrial induction. +RA, RA treated; control, RA untreated. Each dataset was obtained from a single well of differentiation. Mean ± SD, n control = 80, n RA = 80, U-test.

4.2.1.4 Effect of RA on Ca^{2+} transient dynamics

To investigate the consequence of RA on the calcium dynamics, day 33 iPSC-CMs were transduced with adenoviruses carrying a fluorescent Ca^{2+} indicator *RGECO*. 48 hr after the viral transduction, videos of fluorescence oscillation behaviour were collected at 2 Hz pacing for CalTrack analysis (Psaras et al., 2021). It is worth noting that the RA treatment caused an apparent acceleration in spontaneous contraction rate beyond the capacity of the physiological pacing condition 1 Hz. The minimum working pacing frequency for the RA-treated cells was 2 Hz, which was used as a standard condition for the subsequent functional studies.

CalTrack suggested that RA significantly shortened the Ca^{2+} transient length (Figure 4-4A & B). This was supported by the reduction in the computed Ca^{2+} transient duration (CD), especially CD at 90% and 50% of Ca^{2+} decay (Figure 4-4C). The change in the CD parameters were further explained by a summation of both the faster Ca^{2+} release as illustrated by the shortened time to Ca^{2+} peak, especially the time to 90% of the peak (Figure 4-4D), and the

faster Ca^{2+} reuptake as suggested by the shortened time to Ca^{2+} decay, especially the times to 90% and 50% Ca^{2+} decay (Figure 4-4E).



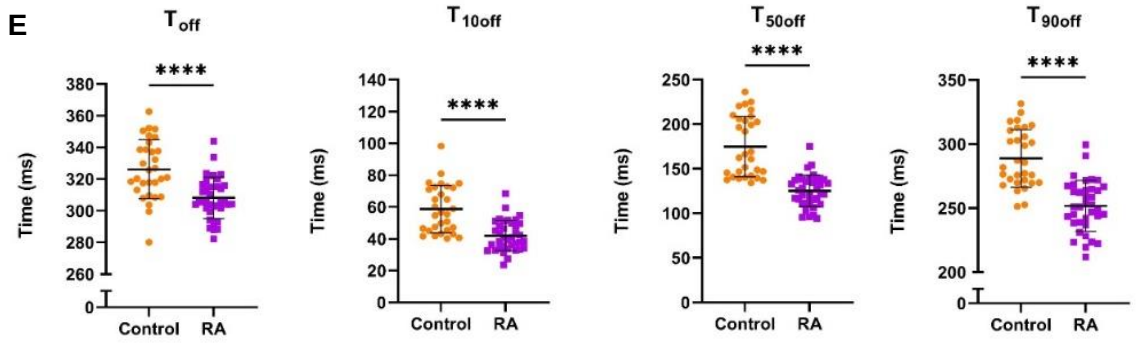


Figure 4-4 Ca^{2+} transient analysis of day 35 *MYL7-mClover* iPSC-CMs differentiated with or without RA treatment. A) Representative kymograph and intracellular, raw fluorescent Ca^{2+} traces. B) Normalised mean Ca^{2+} traces. C) Ca^{2+} transient duration (CD), CD at 10% decay (CD_{10}), CD at 50% decay (CD_{50}) and CD at 90% decay (CD_{90}). D) Time to Ca^{2+} peak (T_{on}), time to 10% Ca^{2+} peak (T_{10on}), time to 50% Ca^{2+} peak (T_{50on}) and time to 90% Ca^{2+} peak (T_{90on}). E) Time to Ca^{2+} decay (T_{off}), time to 10% Ca^{2+} decay (T_{10off}), time to 50% Ca^{2+} decay (T_{50off}) and time to 90% Ca^{2+} decay (T_{90off}). Each dataset was obtained from different cells from a single well of differentiation. Mean ± S.D., n control = 30, n RA = 37, Student t-test (CD_{10}, CD_{50}, T_{on}, T_{10on}, T_{50on}, T_{50off}, T_{90off}) or U-test (CD, CD_{90}, T_{90on}, T_{off}, T_{10off}).

4.2.1.5 Effect of RA on contractility

To investigate the consequence of RA on contractility, day 33 iPSC-CMs were transduced with adenoviruses carrying *mScarlet-ACTN2*. 48 hr after the viral transduction, videos of ACTN2-visualised sarcomeric contractile behaviour were collected at 2 Hz pacing (Figure 4-5A) for SarcTrack analysis (Toepfer et al., 2019).

SarcTrack identified no changes in the sarcomere length (Figure 4-5B & C) and the standard deviation of the sarcomere length (Figure 4-5D & E) between the cells with and without RA supplementation at both the contraction state and the relaxation state, indicating normal sarcomere maturation upon the RA treatment. The RA-treated cells displayed faster contraction (Figure 4-5F) and relaxation kinetics (Figure 4-5G) without a change in the fractional shortening (Figure 4-5H).

Overall, the observed increased atrial marker expression, suppressed ventricular marker expression, CM hypotrophy, shortened Ca^{2+} transient duration and accelerated sarcomere contraction and relaxation kinetics were suggestive of atrial-like characteristics of the RA-treated cells (Cyganek et al., 2018).

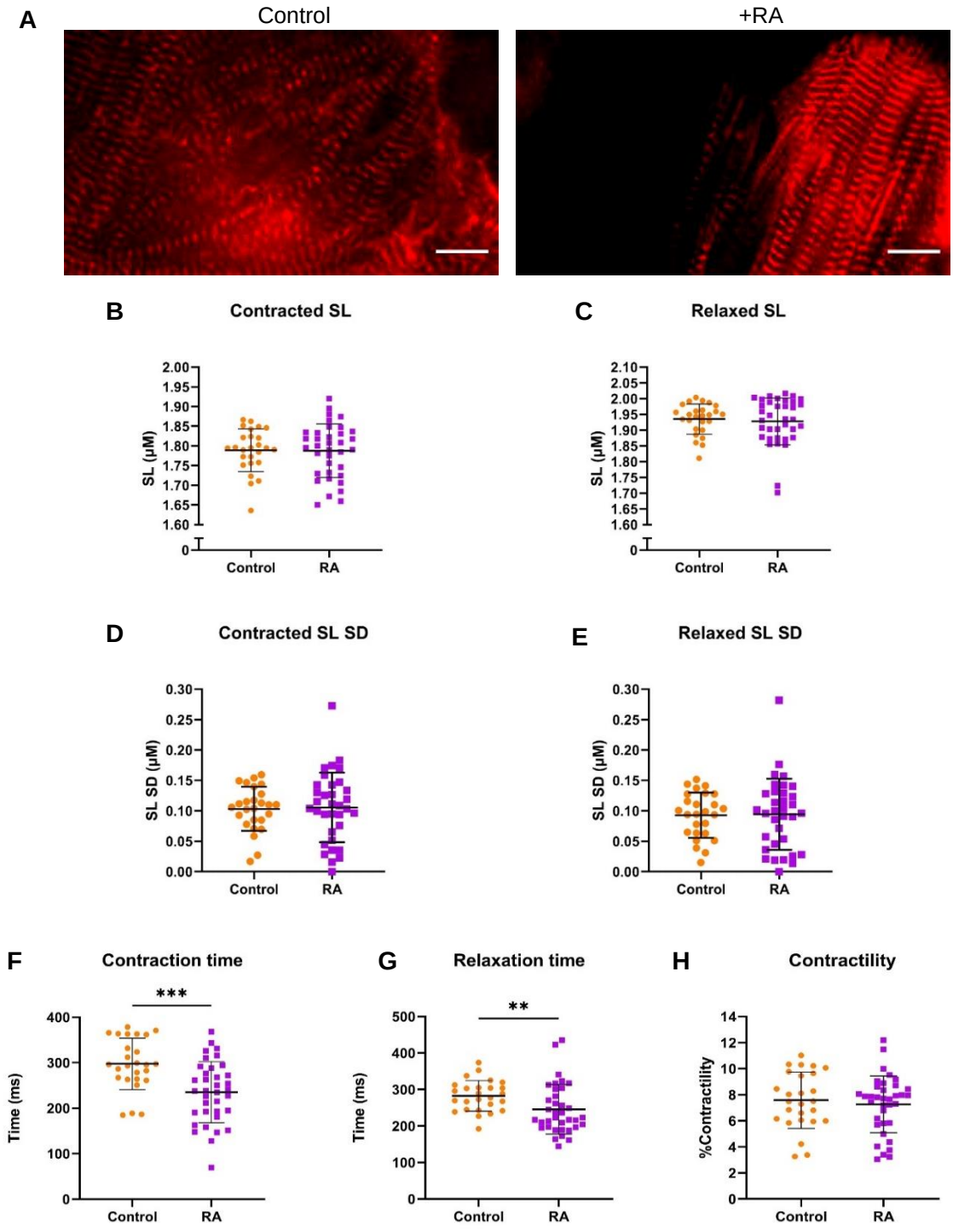


Figure 4-5 Contractility analysis of day 35 *MYL7-mClover* iPSC-CMs differentiated with or without RA treatment. A) Representative live cell images with fluorescently-labelled ACTN2. Scale bar = 10 µm. B) Minimum sarcomere length (SL) at contraction state. C) Maximum SL at relaxation state. D-E) S.D. of SL at D) contraction state and E) relaxation state. F) Contraction time. G) Relaxation time. H) Fractional shortening (contractility). Each dataset was obtained from different cells from a single well of differentiation. Mean ± S.D., n control = 26, n RA = 36, Student t-test (B, D, E, F, H) or U-test (C, G).

4.2.2 Correlation between *mClover* and *MYL7* expression

RNA samples from cells differentiated using the ventricular protocol for 0, 3, 5, 10, 15 and 30 days were harvested for qPCR. The expression pattern of *mClover* followed that of *MYL7* which became detectable from day 3 and continuously increased until reaching a plateau at day 30 (Figure 4-6A). Meanwhile, the onset of expression of pan-cardiac markers *ACTN2* and *TNNT2* was approximately at day 5 after which the expression steadily increased. The *mClover* and *MYL7* expression showed a positive correlation (Pearson correlation coefficient, r = 0.88) (Figure 4-6B), suggesting that *mClover* mirrored the *MYL7* expression at the RNA level.

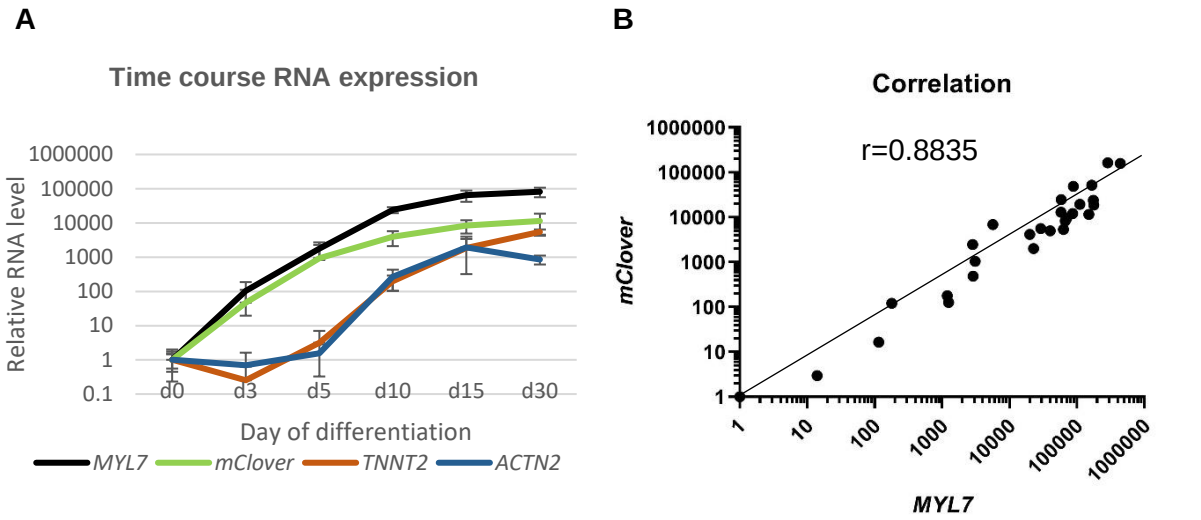


Figure 4-6 Relationship between *MYL7* and *MYL7*-driven *mClover* expression. A) Time course expression of *MYL7*, *mClover*, *TNNT2* and *ACTN2* between day 0 and day 30 of differentiation. n=3 for each time point acquired from different wells using a single iPSC batch, mean ± S.D. B) Correlation analysis of the *MYL7* and *mClover* expression. n=27, Pearson's correlation coefficient (r). Each datum point was collected from various time points from different wells of differentiation using a sing iPSC batch. All expression levels were made relative to the day 0 time point.

4.2.3 Response of *mClover* transcriptional and protein expression to atrial differentiation

RNA and protein samples from cells differentiated by the ventricular and atrial protocols for 30 days were harvested for qPCR and WB analysis, respectively.

qPCR suggested that the atrial induction increased the *mClover* expression (Figure 4-7). This effect was closely similar to the level observed in *MYL7*, highlighting *mClover* as being a marker for *MYL7* gene expression.

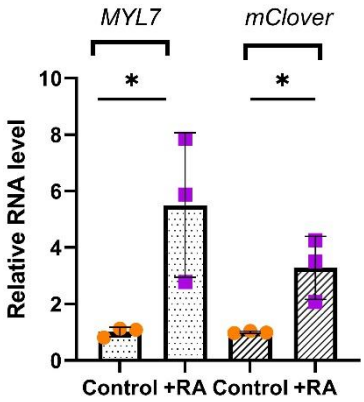


Figure 4-7 *mClover* transcriptional expression in day 30 *MYL7-mClover* iPSC-CMs upon RA-mediated atrial differentiation. Control, DMSO-treated; +RA, RA-treated. Mean ± S.D., n=3, Student t-test. Each dataset was obtained from different wells of differentiation using a single iPSC batch.

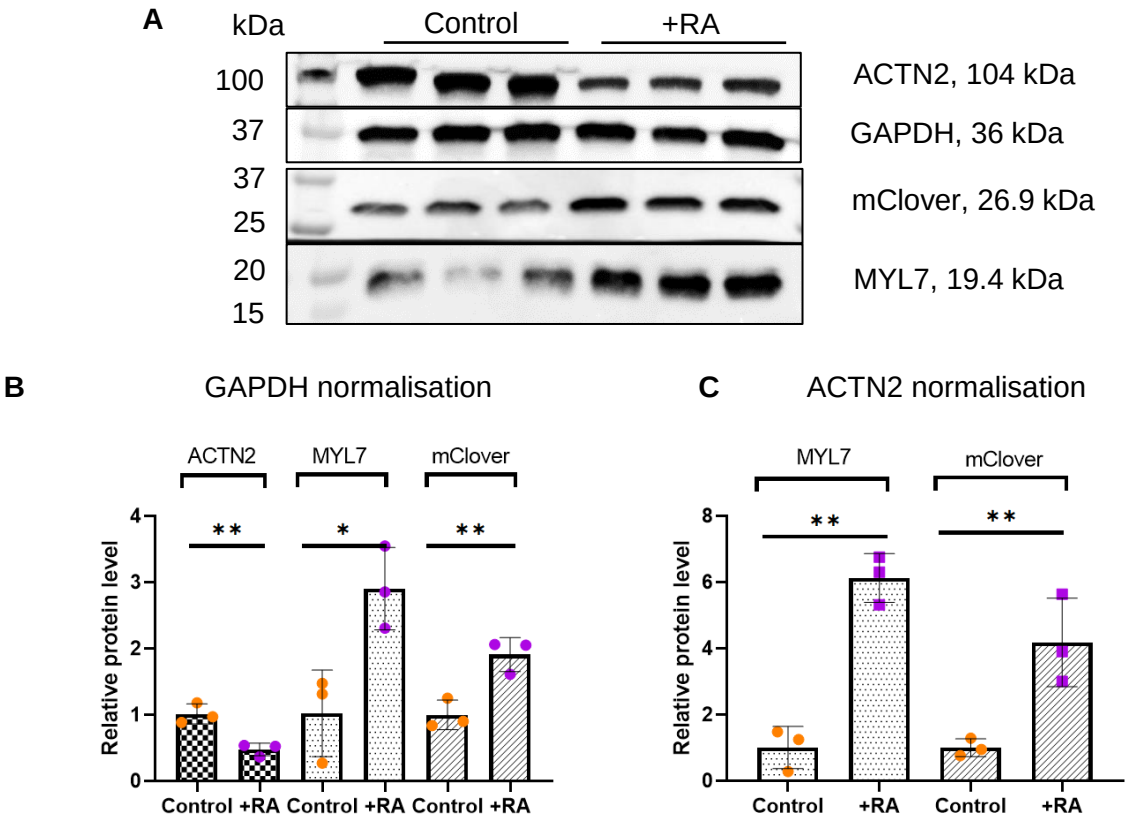


Figure 4-8 mClover protein expression in day 30 *MYL7-mClover* iPSC-CMs upon RA-mediated atrial differentiation. A) WB. B) Quantification of protein from A as normalised by GAPDH. C) Quantification of protein from A as normalised by ACTN2. Since there was a reduction in ACTN2 by RA treatment observed in B, normalisation by ACTN2 may provide a more sensible comparison with respective to CM cell population. Control, non-treated; +RA, RA-treated. Each dataset was obtained from different wells of differentiation using a single iPSC batch. Mean ± S.D., n=3, Student t-test.

At the protein level, RA also upregulated the mClover expression to the similar level observed in MYL7 (Figure 4-8). Notably, the data other than *mClover*/mClover were adapted from Figure 4-1 and Figure 4-2 as they were derived from the same cell source.

4.2.4 Investigation of mClover fluorescence expression by microscopy

The ultimate aim of generating the *MYL7-mClover* reporter line is purification of live atrial-like iPSC-CMs using FACS. Thus, it is important to first investigate if the mClover fluorescence can be visualised and whether its intensity is enhanced by the atrial differentiation protocol.

Live cell imaging (day 30) suggested a well-to-well variation of the mClover fluorescent intensity when the cells had been differentiated by the atrial protocol (Figure 4-9). Most of the cultures displayed average mClover expression, which was comparable to those differentiated by the ventricular protocol, whilst only 2 or 3/12 wells showed an obvious increase in the fluorescent signal. The mClover fluorescence diffused throughout the cytosol within the live cells.

As the *MYL7-mClover* reporter line was also used for the subsequent IF and FACS assays, it is worth investigating whether the mClover fluorescence is preserved upon fixation. Overall, the fluorescence was faintly detectable in fixed iPSC-CMs (day 30)(Figure 4-10), possibly explainable by a disruption of the mClover structure by paraformaldehyde (Park et al., 2019). There was no clear difference in the fluorescent intensity between the atrial and ventricular differentiation. The fluorescent signal was enriched within the sarcomere structure (marked by ACTN2 immunostaining) and nucleus (marked with DAPI), indicating that the monomeric mClover molecule was small enough to freely move into the nucleus in the fixed cells. mClover was specifically expressed in ACTN2-positive CMs but was unobservable in other cell types (ACTN2-negative), highlighting the (*MYL7*-driven) mClover as being a CM-specific marker.

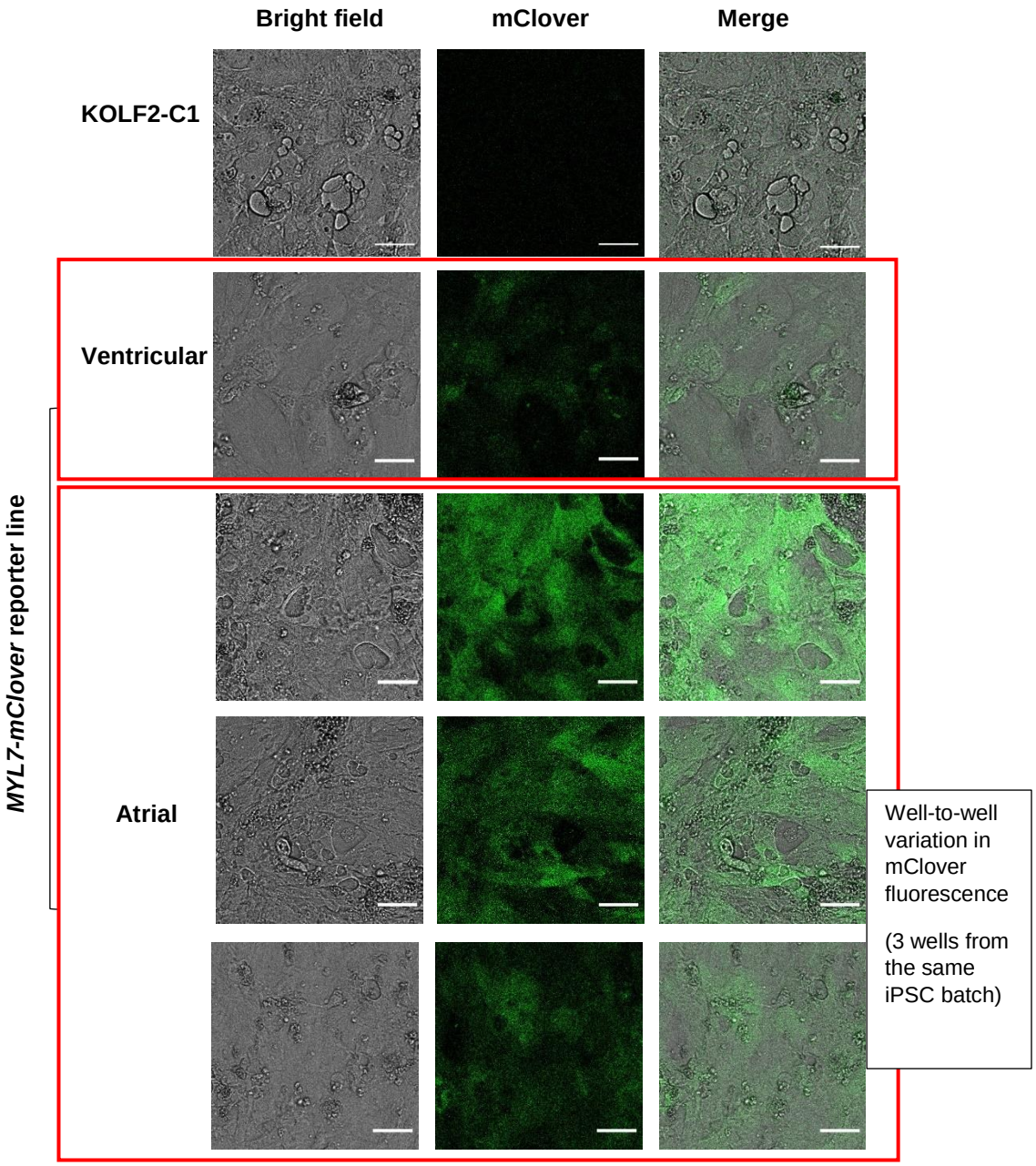


Figure 4-9 Live cell imaging of day 30 *MYL7-mClover* iPSC-CMs differentiated by the atrial and ventricular protocols. KOLF2-C1 cells were the non-fluorescent parental line where the *MYL7-mClover* reporter line was derived from. Scale bar 10 µm. Each Atrial set of images was acquired from different wells of differentiation using a single iPSC batch.

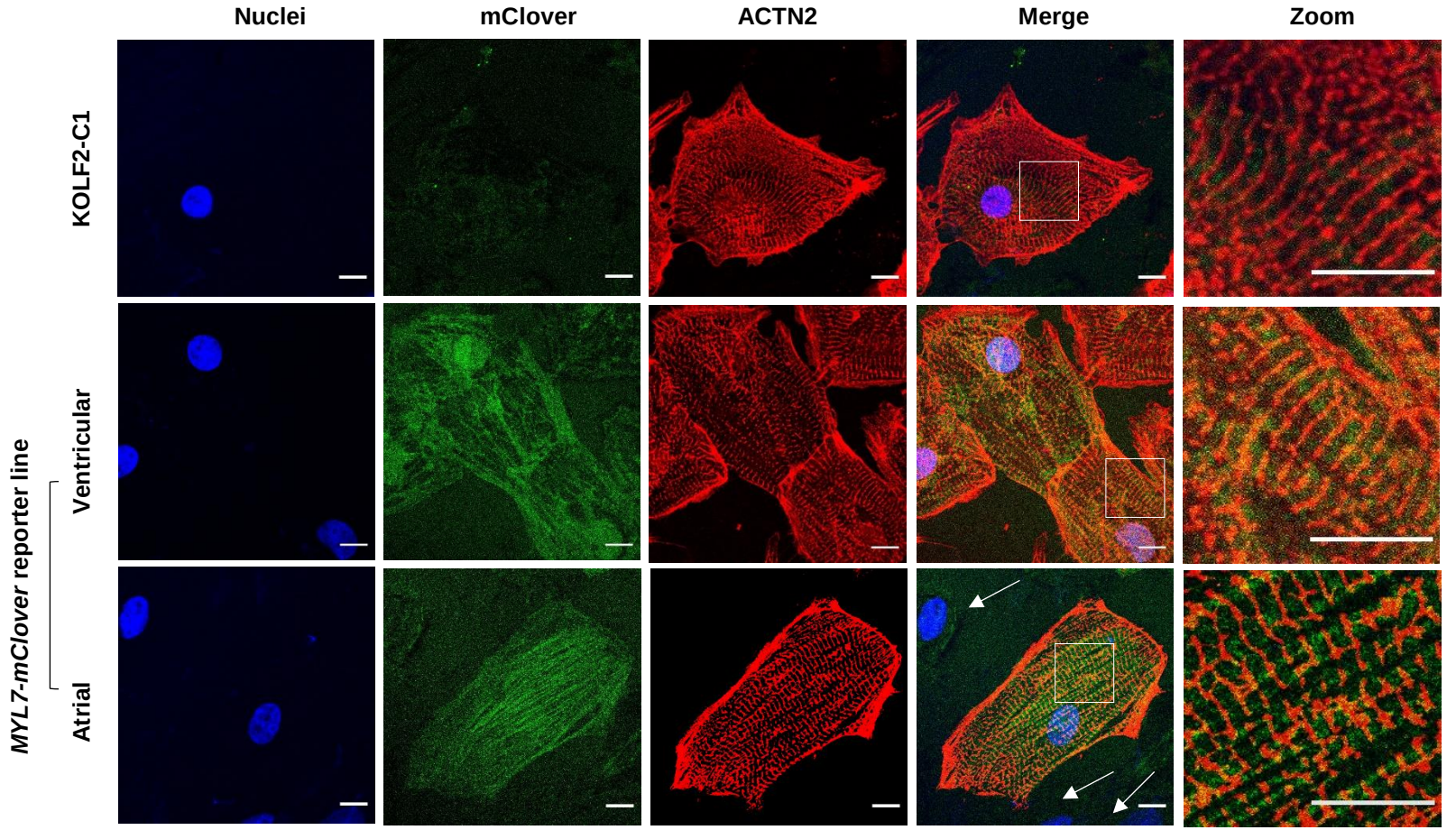


Figure 4-10 Immunofluorescent images of day 30 *MYL7-mClover* iPSC-CMs differentiated by the atrial and ventricular protocols. IF staining of sarcomeric ACTN2 and nuclei (DAPI). mClover fluorescence was directly visualised using a fluorescein-specific filter set. Merge is an overlay between mClover and ACTN2 images. Zoom images were digital zoom images from the square area destined in Merge. Scale bar 10 μm. Arrows indicate non-CMs (ACTN2-negative) with no expression of mClover.

4.2.5 Investigation of mClover fluorescence expression by FACS

Understanding how the mClover fluorescence correlates with the MYL7 expression especially upon the atrial induction ensures a success in live cell sorting for enrichment of the atrial-like subtype. However, the endogenous MYL7 proteins are distinguishable by the FACS machine only if fluorescently labelled. To avoid the potential masking effect of the fixative, mClover is also needed to be immunofluorescently labelled. In addition, the length of culture time seems to compromise with the iPSC-CM viability during cell preparation prior to live cell sorting. To address this concern, early stage iPSC-CMs were used in FACS analysis.

In this experiment, day 25 *MYL7-mClover* reporter iPSC-CMs from a single well of differentiation were fixed, stained for ACTN2, MYL7 and mClover and analysed by FACS. Almost all of the cells differentiated using the ventricular (97%) and atrial (94.5%) protocols were positive for MYL7 expression (Figure 4-11A & B). There was no clear difference in the MYL7 expression levels observed between the two conditions, except a slight shift of the population toward higher expression by the atrial induction (Figure 4-11C). Similarly, the mClover staining suggested no change in mClover expression level by the atrial induction, except a slight shift of the population toward higher expression (Figure 4-12). Overall, the expression of mClover was low and its pattern was not obviously separated into two distinct groups of high/positive expression and low/negative expression. The expression of MYL7 and mClover seemed to correlate with each other, highlighting mClover as being a marker for MYL7. The double staining of ACTN2 and mClover (Figure 4-13) agreed with the pattern seen in the mClover single staining. Almost all of the cells expressed ACTN2 either from the atrial differentiation (99.5%) or ventricular differentiation (99.1%), indicating no adverse effect by RA on differentiation potential. Within the ACTN2 positive population, there was no clear difference in the mClover expression observed between the differentiation conditions, except a slight shift of the population toward higher expression by the atrial induction. All in all, it is likely that the atrial induction by RA had a small or no effect on MYL7 and mClover expression in this fluorescent reporter line.

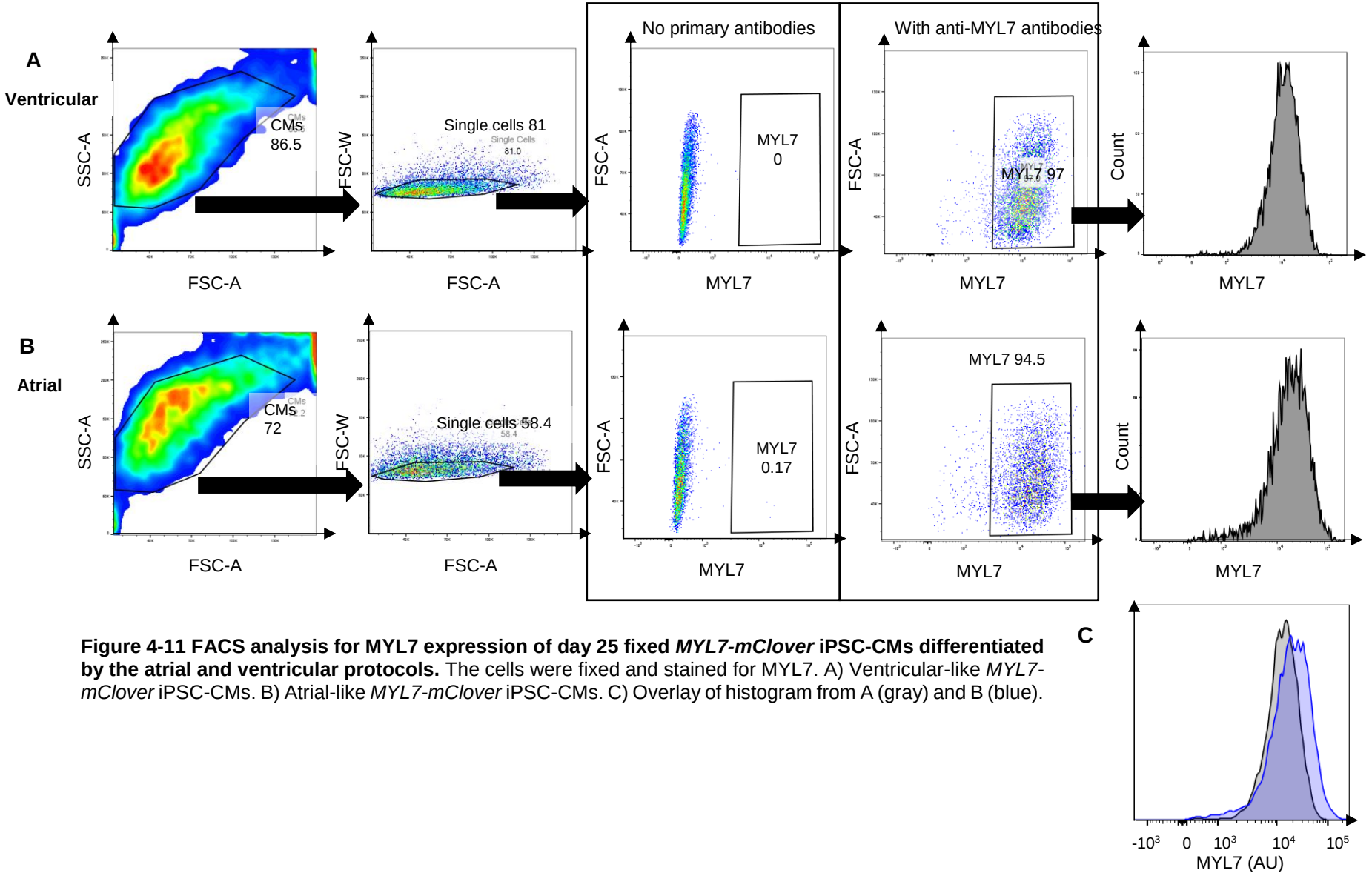


Figure 4-11 FACS analysis for MYL7 expression of day 25 fixed *MYL7-mClover* iPSC-CMs differentiated by the atrial and ventricular protocols. The cells were fixed and stained for MYL7. A) Ventricular-like *MYL7-mClover* iPSC-CMs. B) Atrial-like *MYL7-mClover* iPSC-CMs. C) Overlay of histogram from A (gray) and B (blue).

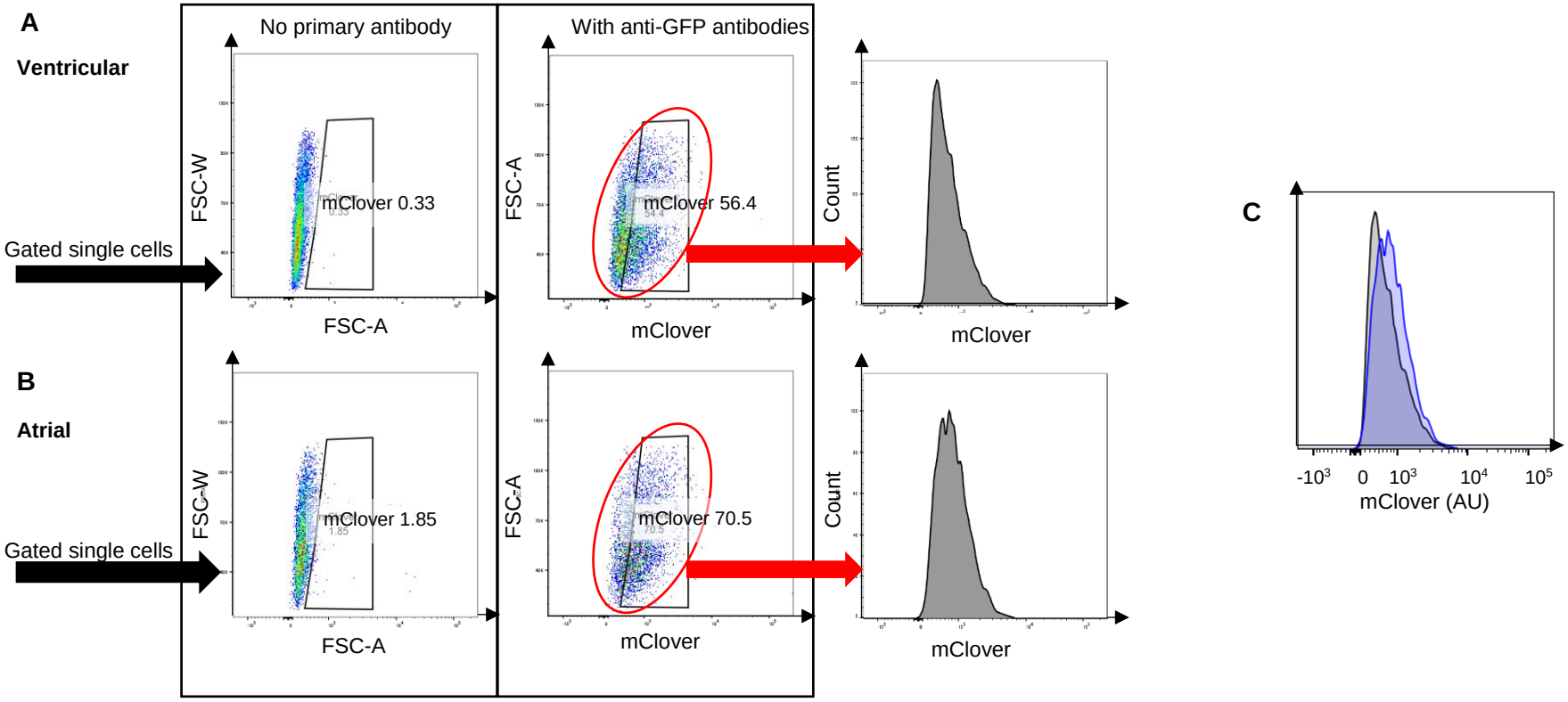


Figure 4-12 FACS analysis for mClover expression of day 25 fixed *MYL7-mClover* iPSC-CMs differentiated by the atrial and ventricular protocols. The cells were fixed and stained for mClover using anti-GFP antibodies. The first inputs were the gated single cells described in Figure 4-11. A) Ventricular-like *MYL7-mClover* iPSC-CMs. B) Atrial-like *MYL7-mClover* iPSC-CMs. C) Overlay of histogram from A (gray) and B (blue).

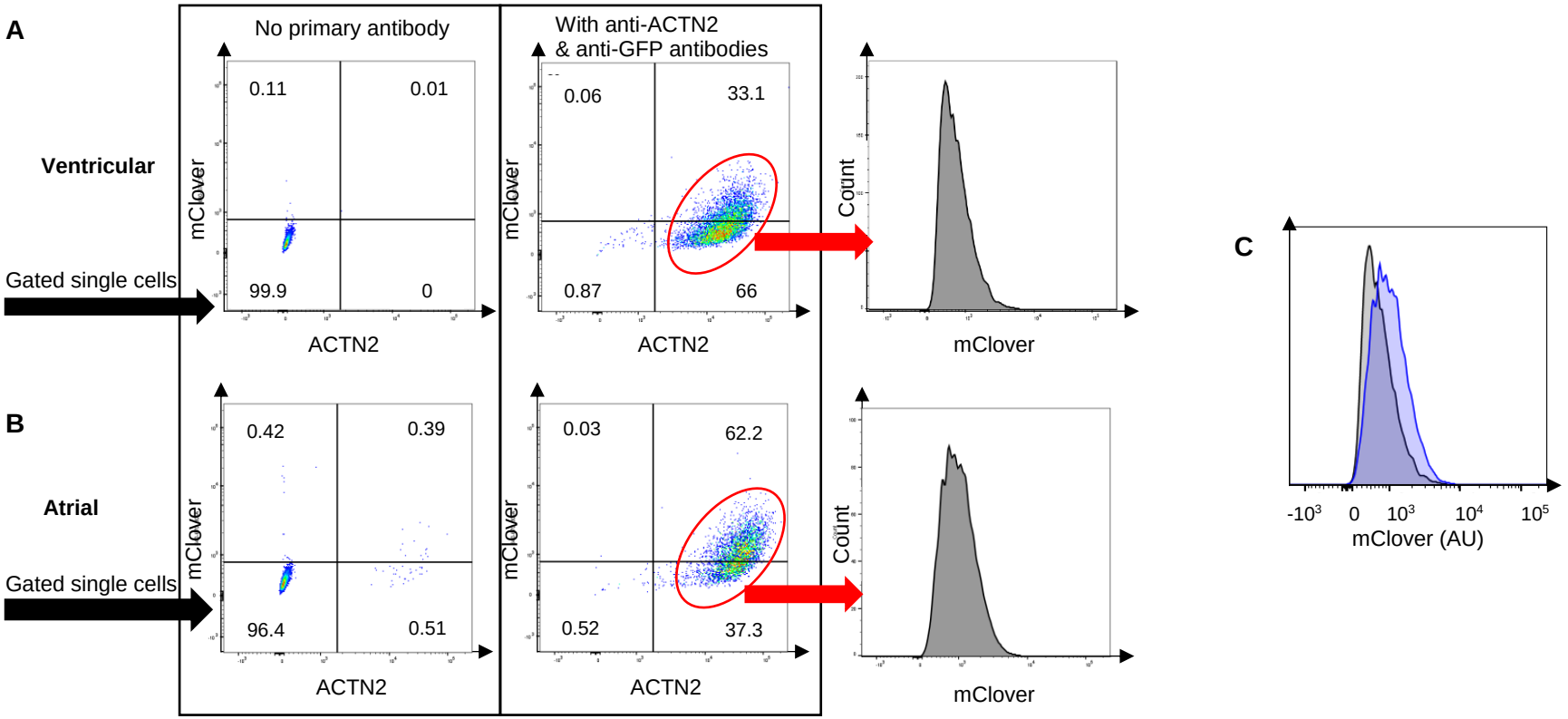


Figure 4-13 FACS analysis for ACTN2 and mClover expression of day 25 fixed *MYL7-mClover* iPSC-CMs differentiated by the atrial and ventricular protocols. The cells were fixed and stained for ACTN2 and mClover (using anti-GFP antibodies). The first inputs were the gated single cells described in Figure 4-11. A) Ventricular-like *MYL7-mClover* iPSC-CMs. B) Atrial-like *MYL7-mClover* iPSC-CMs. C) Overlay of histogram from A (gray) and B (blue).

FACS analysis in the live ventricular-like cells showed a similar pattern of mClover expression in day 20 *MYL7-mClover* iPSC-CMs (Figure 4-14) as compared to the fixed cells (Figure 4-14). Overall, the fluorescence was not strong and there was a gradient of the fluorescent intensity with no clear boundary between negative and positive fractions observed, challenging the mClover gating. The density plot may suggest two distinct populations but they cannot be simply gated using the parental KOLF2-C1 line as a negative control. However, using this gating strategy approximately 11% can be determined as mClover positive population.

It is worth nothing that the mClover fluorescent intensity of the live and fixed cells cannot be directly compared due to several factors. In the fixed cells, the mClover expression was measured indirectly using the GFP-specific antibodies and the signal was amplified by the secondary antibodies. There were also differences in the FACS machines used for the two experiments (Chapter 2) and voltage setting for each parameter measured. Further, only approximately 3.63% (Figure 4-14) can be gated as CMs in the *MYL7-mClover* reporter line based on size because there were many cell clumps (whose size was beyond the measurement window), debris and dead cells. Therefore, single cell preparation and maintenance prior to and during live cell sorting require a substantial optimisation by e.g. using a cell strainer and a contraction inhibitor “blebbistatin”, which has been shown to improve single iPSC-CM’s viability (Li et al., 2021), respectively.

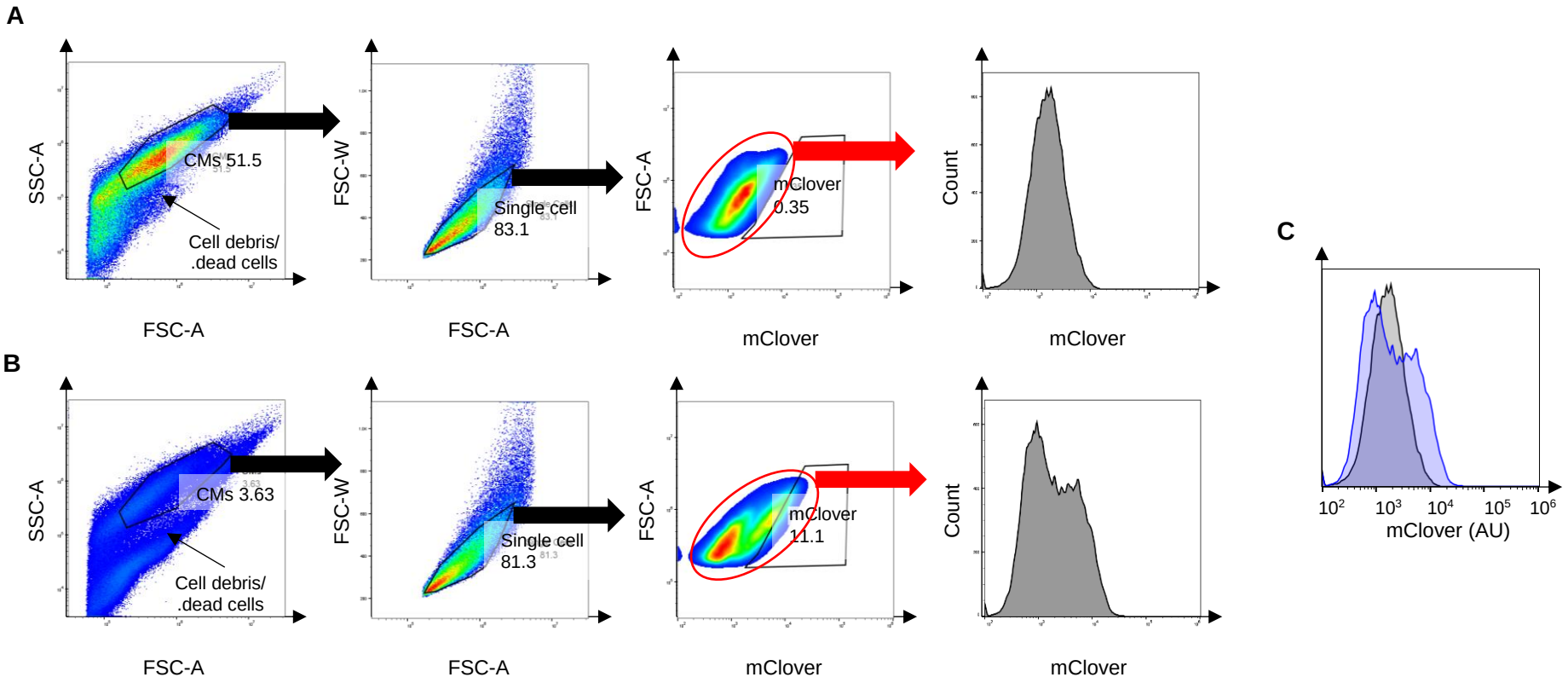


Figure 4-14 FACS analysis for mClover expression of day 20 live *MYL7-mClover* iPSC-CMs differentiated by the ventricular protocol. The mClover fluorescence was acquired by a FITC filter set. A) Parental KOLF2-C1 iPSC-CMs. B) *MYL7-mClover* iPSC-CMs. C) Overlay of histogram from A (gray) and B (blue).

4.3 HCN4-mScarlet line

The heterozygous *HCN4-mScarlet* reporter iPSC line (clone 4F-2D) showed limited cardiac differentiation potential. Approximately 50% of the cell culture area could not develop beating sheets of CMs after being differentiated using the ventricular protocol. The spontaneous contraction was normally visible approximately from day 8-9.

4.3.1 Correlation between *mScarlet* and *HCN4* expression

RNA samples from cells differentiated for 0, 3, 5, 10, 15 and 30 days were harvested for qPCR analysis. The expression pattern of *mScarlet* followed that of *HCN4*. It became detectable at day 5 and continuously increased until day 10, after which both expression levels declined (Figure 4-15A). The onset of the expression of pan-cardiac marker *ACTN2* was approximately day 5, after which a sigmoid-like growth was observed. The expression pattern of *ACTN2* in the *HCN4-mScarlet* reporter line was similar to those presented in the *MYL7-mClover* line. However, the maximum *ACTN2* expression level in the *HCN4-mScarlet* line was approximately 10 times lower than the level seen in the *MYL7-mClover* line (Figure 4-6), highlighting the low cardiac differentiation potential. The *mScarlet* and *HCN4* expression showed a positive correlation (r = 0.93) (Figure 4-15B), suggesting that *mScarlet* likely reflected the *HCN4* expression at the RNA level.

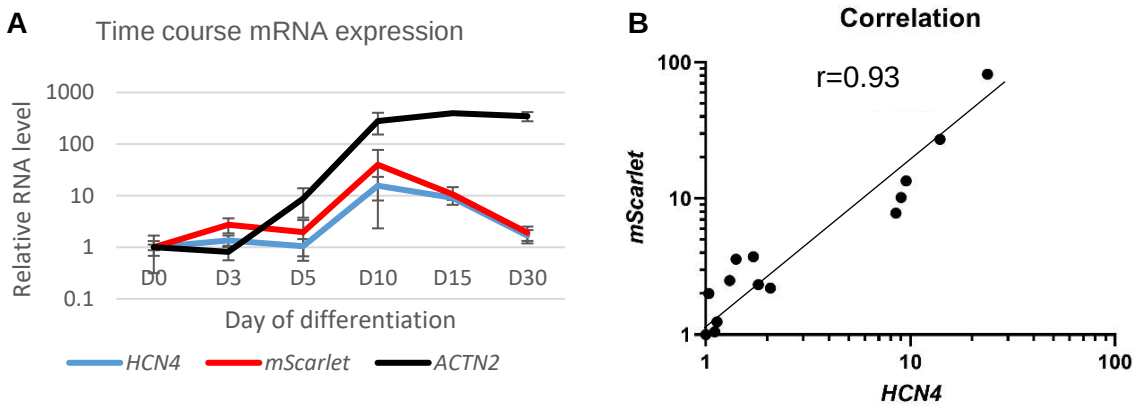


Figure 4-15 Relationship between *HCN4* and *HCN4*-driven *mScarlet* expression. A) Time course expression of *HCN4*, *mScarlet* and *ACTN2* between day 0 and day 30 of differentiation. n=3 for each time point acquired from different wells using a single iPSC batch, mean ± S.D. B) Correlation analysis of the *HCN4* and *mScarlet* expression. n=15, Pearson's correlation coefficient (r). Each datum point was collected from various time points from different wells of differentiation using a single iPSC batch. All expression levels were made relative to the day 0 time point.

4.3.2 Investigation of mScarlet fluorescence expression by microscopy

Live cell imaging of day 23 *HCN4-mScarle*t iPSC-CMs differentiated by the ventricular protocol showed that the mScarlet fluorescence can be detected in most of the contracting area of CMs with higher intensity in some specific area (Figure 4-16), likely suggesting that (*HCN4*-driven) mScarlet (or the endogenous HCN4) was commonly expressed in iPSC-CMs.

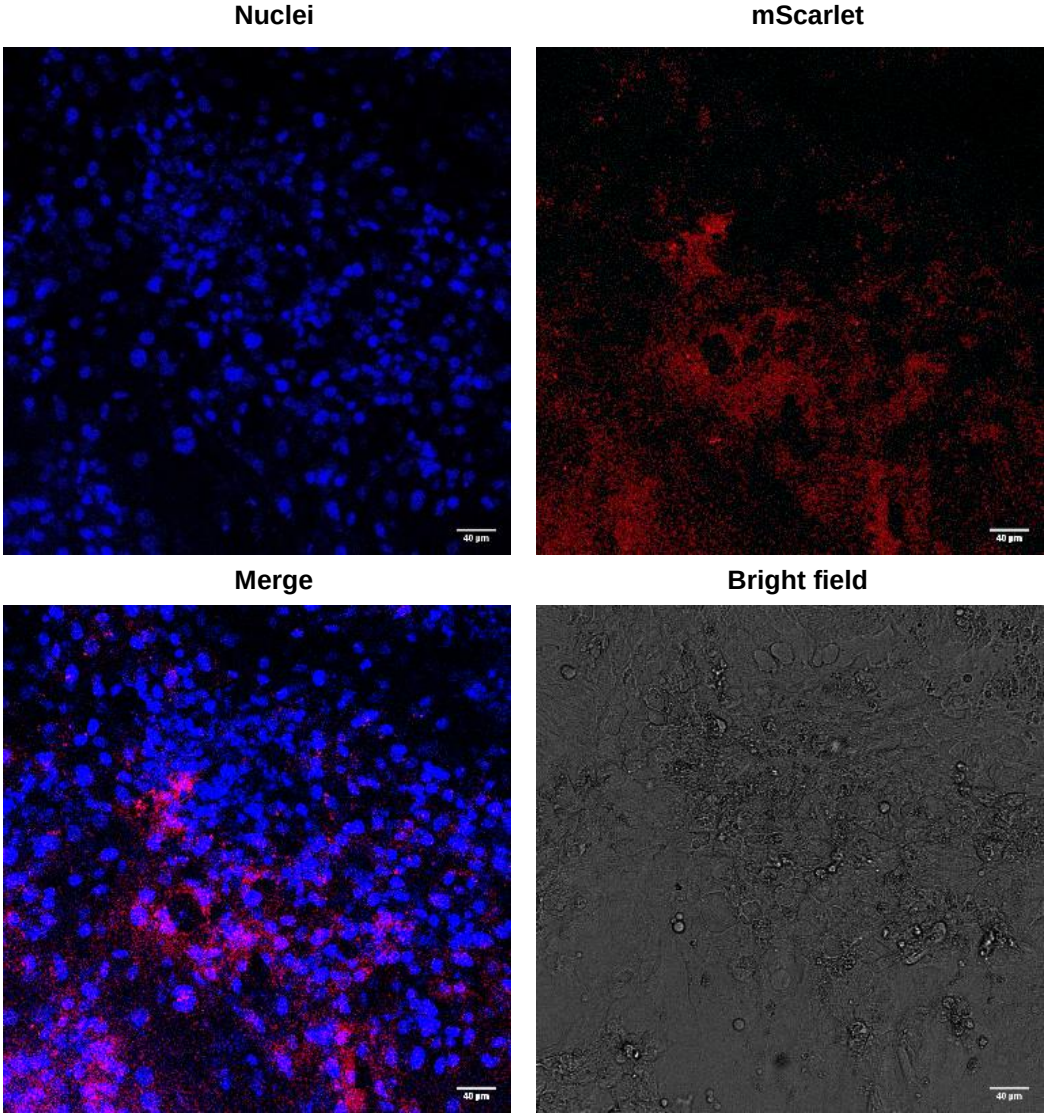


Figure 4-16 Live cell imaging of day 23 *HCN4-mScarlet* iPSC-CMs. Note: KOLF2-C1 iPSC-CMs were used a negative control for mScarlet fluorescence (not shown). Scale bar 40 µm.

4.3.3 Investigation of mScarlet fluorescence expression by FACS

Day 25 *HCN4-mScarlet* iPSC-CMs from a single well of ventricular differentiation were fixed, stained for ACTN2 and analysed by FACS. Only 30.3% of the total differentiated cells were gated as CMs based on size compared to 75.8% in the KOLF2-C1 iPSC-CMs. Within the gated CMs population, only 78.5% of the cells were stained positive for ACTN2 compared to 98.9% in the KOLFC2-C1 iPSC-CMs (Figure 4-17A&B), underlining the aforementioned differentiation defect. Within the ACTN2 positive population, there was a gradient of the mScarlet fluorescent signal detected but there was no clear boundary between mScarlet positive and mScarlet negative populations. Using the parental KOLF2-C1 as a negative control, approximately 29% of the population (gated based on the density plot, Figure 4-16A & B) or 40% (gated based on the histogram, Figure 4-17C) showed a positive expression for mScarlet, which may indicate those cells expressing higher level of (*HCN4*-driven) mScarlet. This mScarlet-high population may either represent nodal-like subtype population or generally reflect the variation of mScarlet or HCN4 expression within iPSC-CMs.

FACS analysis in the live cells showed a similar pattern of mScarlet expression in day 20 *HCN4-mScarlet* iPSC-CMs (Figure 4-18) as compared to the fixed cells (Figure 4-17). No clear boundary between mScarlet negative and positive fractions was observed, challenging the mScarlet gating. However, using the parental KOLF2-C1 as a negative control, approximately 25% (based on the density plot) or 18% (based on the histogram) can be gated as mScarlet positive population.

Notably, only approximately 12.5% can be gated as live CMs based on size in the *HCN4-mScarlet* reporter line. Single cell preparation and maintenance prior to and during live cell sorting need further optimisation to improve the yield of CMs, possibly by using a cell strainer and blebbistatin treatment, respectively.

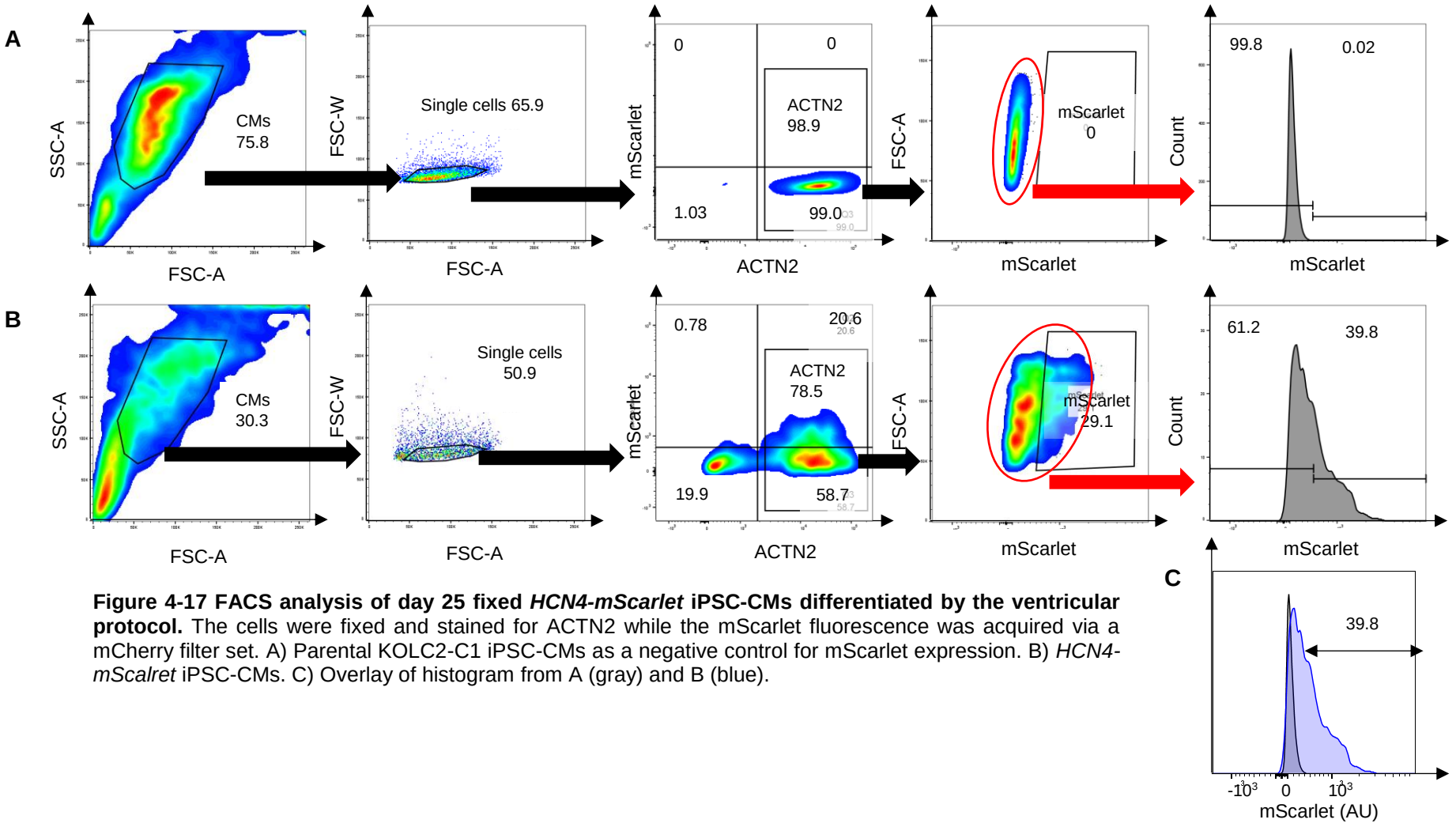


Figure 4-17 FACS analysis of day 25 fixed *HCN4-mScarlet* iPSC-CMs differentiated by the ventricular protocol. The cells were fixed and stained for ACTN2 while the mScarlet fluorescence was acquired via a mCherry filter set. A) Parental KOLC2-C1 iPSC-CMs as a negative control for mScarlet expression. B) *HCN4-mScalret* iPSC-CMs. C) Overlay of histogram from A (gray) and B (blue).

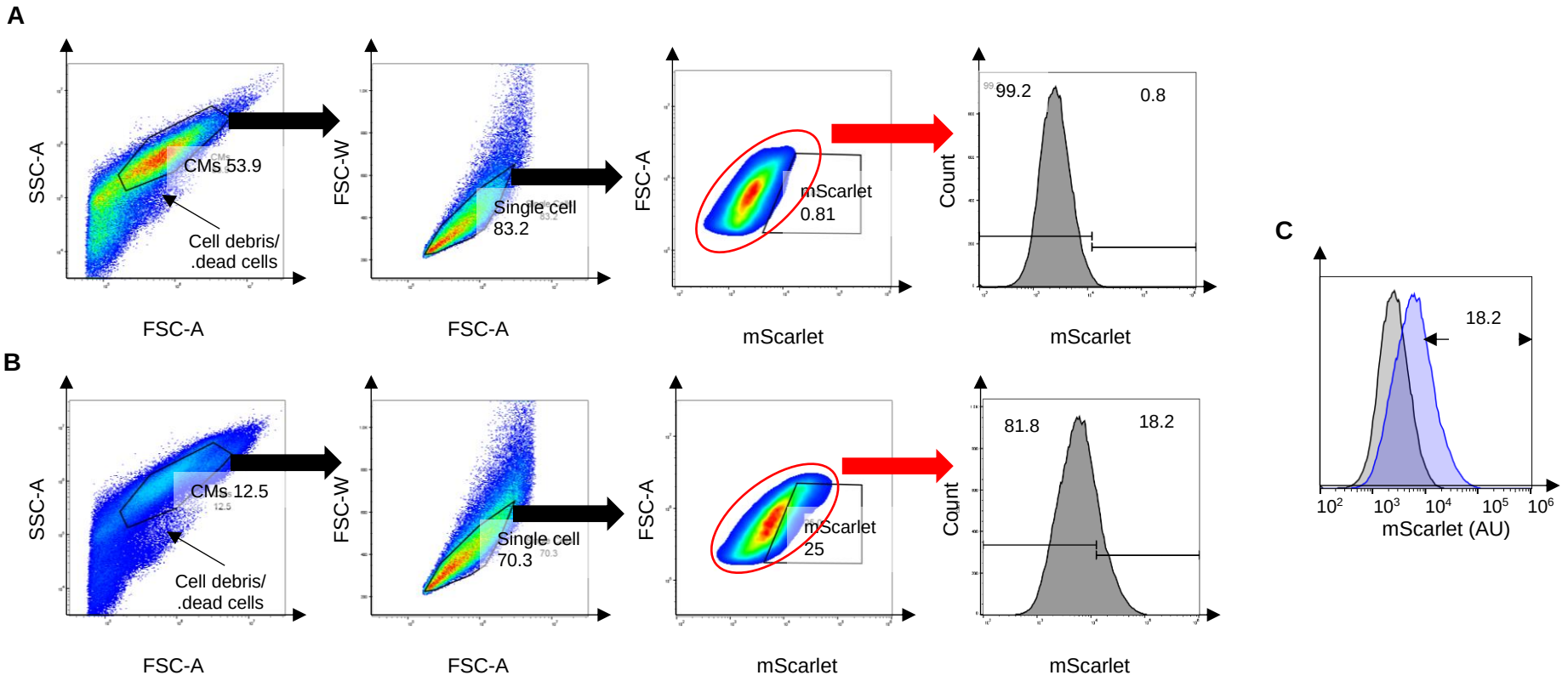


Figure 4-18 FACS analysis for mScarlet expression of day 20 live *HCN4-mScarlet* iPSC-CMs differentiated by the ventricular protocol. The mScarlet fluorescence was acquired by a mCherry filter set. A) Parental KOLF2-C1 iPSC-CMs. B) *HCN4-mScarlet* iPSC-CMs. C) Overlay of histogram from A (gray) and B (blue).

4.4 *PITX2-mClover* reporter line

Both homozygous *PITX2-mClover* line clones 3D-3H and 3D-6C were unable to differentiate into CMs using the ventricular protocol. Hence, this project was terminated at this stage.

4.5 *MYL2-mClover* reporter line

As presented in Chapter 3, both the heterozygous *MYL2-mClover* line clones 6F and 11C contained undesired deletions at the *MYL2* C-terminus on the non-targeted alleles, posing a concern over a loss of function.

Both clones were differentiated using the ventricular protocol. The clone 6F, with a minimal CRISPR-related deletion, was well differentiated with an onset of spontaneous contraction at ~ day 8-9. qPCR analysis of day 25 cells showed the expression levels of *MYL7*, *ACTN2* and *TNNT2* nearly comparable to those detected in the parental KOLF2-C1 line (Figure 4-19). Despite the presence of the transcripts, both MYL2 and mClover proteins were not detected by WB (Figure 4-20).

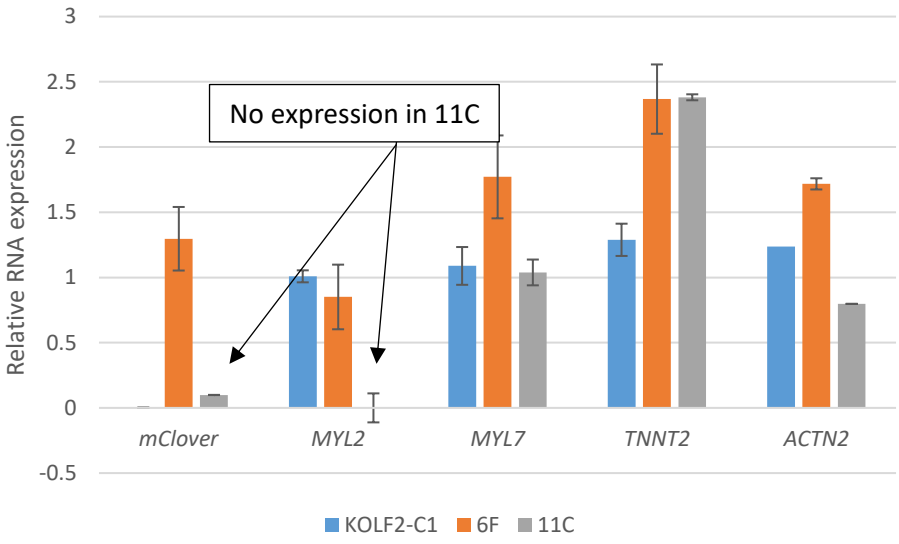


Figure 4-19 RNA expression of day 25 *MYL2-mClover* iPSC-CMs differentiated by the ventricular protocol. While the clone 6F had normal *MYL2* and *mClover* expression (orange bar), the clone 11C clearly lacked expression of such markers (grey bar with arrow). All expression levels were made relative to KOLF2-C1 (blue bar), except that analysis of *mClover* was made relative to the reporter clone 6F. Each dataset was obtained from different wells of differentiation using a single iPSC batch. n=3, mean ± S.E.

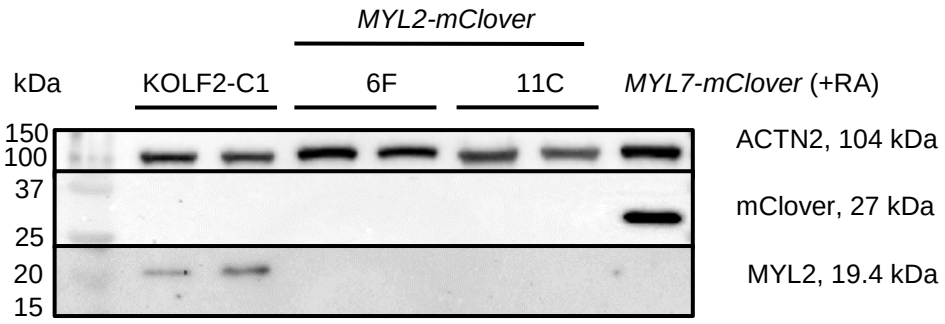


Figure 4-20 WB analysis of day 25 *MYL2-mClover* iPSC-CMs differentiated by the ventricular protocol. Both engineered clones 6F and 11C did not express mClover and MYL2. *MYL7-mClover* (+RA) was used as a positive control for mClover expression while KOLF2 C-1 was used as a positive control for MYL2 expression. Each dataset was obtained from different wells of differentiation using a single iPSC batch

The clone 11C, with a deletion of the whole exon 7, exhibited a severe defect in cardiac differentiation potential with only a few numbers of cells displaying a late onset spontaneous contraction (approximately day 11-13). qPCR analysis of day 25 cells, despite the normal *ACTN2*, *TNNT2* and *MYL7* expression profile, clearly indicated abolished *MYL2* and *mClover* expression (Figure 4-19), in agreement with the WB result (Figure 4-20).

Since both clones lacked the fluorophore expression at the protein level, they were no longer qualified as fluorescent reporter systems. Considering both the challenging targeting approach discussed in Chapter 3 and the incorrectly engineered 6F and 11C clones presented in this chapter, this fluorescent reporter iPSC-CM project was terminated at this stage.

4.6 Discussion

iPSC-CMs are increasingly exploited to model cardiac disease and hold promise in drug development. Although it is commonly accepted that iPSC-CMs behave differently from the adult CMs in certain aspects (including immaturity, high variability within the population and spontaneous contraction), the ability to work with human cells bearing a disease causing mutation on the correct genetic background, makes them a very attractive model. Still, their population heterogeneity poses a challenge for their subtype-specific phenotypic study. Fluorescent reporter lines are one of the tools that could be potentially used to address this problem. Genetic markers which are uniquely expressed in a certain CM-subtype can be

fluorescently labelled and hence, the fluorescent cells can be visibly distinguishable and isolated by FACS technique. However, it is imperative to validate if the selected genetic marker and the fluorescent reporter reliably allow identification of the desired subtype in the iPSC-CM setting.

In this chapter, the atrial-specific *MYL7-mClover* and *PITX2-mClover*, ventricular-specific *MYL7-mClover* and nodal-specific *HCN4-mScarlet* fluorescent reporter iPSC lines were differentiated into CMs by the ventricular differentiation protocol (Lian et al., 2012). These CMs were assayed by qPCR, WB, fluorescent microscopy and FACS to elucidate if each reporter system represents phenotype(s) resembling the respective CM subtype. For the *MYL7-mClover* line, an atrial differentiation protocol by RA induction was also established and used to help characterise the atrial-related phenotypes.

I) Establishment of atrial differentiation protocol

RA supplementation during the day 3 to 9 of differentiation by the ventricular protocol in the *MYL7-mClover* iPSC-CM was shown to suppress expression of the ventricular markers (*HEY2*, *IRX4* and *MYL2*) and promote expression of the atrial markers (*MYL7*, *KCNJ3*, *MYH6*, *NR2F2* and *PITX2*). At the functional level, it induced abbreviated Ca^{2+} transient kinetics, sarcomere contraction kinetics and sarcomere relaxation kinetics. Morphometrically, it caused a reduction in CM size. All of these changes represent atrial-like characteristics (Cyganek et al., 2018, Bird et al., 2003, Lemme et al., 2018).

II) Characterisation of MYL7-mClover line

The heterozygous *MYL7-mClover* line (clone 11B) displayed a normal CM differentiation potential when the cells were differentiated by the ventricular protocol. Correlation analysis suggested that the transcriptional expression of *mClover* mirrored the expression of *MYL7*. *MYL7* and *mClover* responded positively to the RA-induced atrial differentiation but the changes were not robust, perhaps due to the iPSC-CM immaturity (Gunawan et al., 2021). Despite the increase detected by WB, both the MYL7 and mClover protein expression profiles

could not be distinguished by FACS analysis between the atrial and ventricular differentiation conditions, questioning the reliability of MYL7 as an atrial-specific marker in iPSC-CMs. The comparable mClover fluorescence between the two differentiation conditions in the immunofluorescent images may be attributable to a loss of fluorescent signals because paraformaldehyde disrupts the structure of green fluorescent protein (Park et al., 2019). However, this does not explain the situation in the FACS analysis as the mClover expression was measured indirectly using the specific antibodies. As illustrated by live cell imaging, the atrial induction caused a well-to-well variation in the mClover fluorescence intensity even when the cells were derived from the same batch of iPSC. The majority of the cells differentiated by the atrial protocol displayed mClover fluorescence at the comparable level observed from the cells differentiated by the ventricular protocol. It is likely that the expression of *MYL7*/MYL7 in the cells acquired for the FACS analysis did not strongly respond to RA. It would be interesting to investigate if the atrial differentiation conditions e.g. RA concentration (Devalla et al., 2015) could be further tuned to enhance the *MYL7/mClover* expression. However, such an experiment would be challenging as it may require numerous batches of differentiation. In addition, the iPSC-CM variation in gene expression could be addressed by an automated culture system (Denning et al., 2016). This platform may ensure the consistency of the mClover fluorescence expression, facilitating population gating via the FACS assay. Lastly, the FACS analysis in the live ventricular-like cells showed a similar pattern of the mClover expression to the fixed cells and suggested a low fluorescent intensity which challenged the population gating and cell sorting. It would be interesting to investigate if the mClover signal could be enhanced by the future RA treatment optimisation. Comparison of the expression profiles of atrial/ventricular genes between the sorted mClover positive population and pre-sort sample using the KOLF2-C1 line as a negative control for fluorescence would clarify whether the *MYL7-mClover* reporter is useful for atrial-like cell purification.

III) Characterisation of HCN4-mScarlet line

The heterozygous *HCN4-mScarlet* line (clone 4F-2D) showed a compromised differentiation potential with half of the cell population failing to develop beating clusters when differentiated using the ventricular protocol. The differentiation defect was also confirmed by FACS, where the ACTN2 positive cells was less abundant as compared to the data from the parental KOLF2-C1 line. This abnormality may be explained by non-genetic factors. Albeit carefully avoided, high iPSC colony density and passage number can cause a variation in iPSC differentiation potential (Rivera et al., 2020, Volpato et al., 2018). Nevertheless, time course RNA expression analysis suggested that the expression of *mScarlet* positively correlated with *HCN4*, highlighting a faithful fluorescent reporter system. Live cell imaging and FACS analysis of the live and fixed cells indicated that the mScarlet fluorescence was expressed in the majority of the iPSC-CMs with a variation in expression level, emphasising that HCN4 is a common channel in immature iPSC-CMs (Yechikov et al., 2016). Mature iPSC-CMs tend to have lower rate of spontaneous contraction, which corresponds to the downregulation of *HCN4* in the majority of the cells (Bosman et al., 2013). However, its expression is still maintained in nodal-like cells (Bosman et al., 2013). Comparative FACS analysis of expression of the nodal-specific proteins (e.g. SHOX2 and ISLET1 (Ichimura et al., 2020)) of HCN4/mScarlet-positive population between iPSC-CMs differentiated conventionally and iPSC-CMs differentiated with a maturation-promoting condition e.g. 3-dimensional culture (Denning et al., 2016) would clarify whether HCN4/mScarlet expression from the *HCN4-mScarlet* reporter system enables identification and purification of mature, nodal-like cells. Alternatively, a further investigation of how the mScarlet expression responds to a nodal differentiation protocol (Yechikov et al., 2020) will indicate if the *HCN4-mScarlet* reporter is specific to nodal-like cells. Nevertheless, these studies will require new clonal *HCN4-mScarlet* reporter lines, which fully retain the cardiac differentiation potential, possibly obtainable by subcloning.

IV) Characterisation of PITX2-mClover line

The homozygous *PITX2-mClover* reporter line project (both clone 3D-3H and 3D-6C) was terminated due to failure in CM differentiation when using the ventricular protocol. This defect may be explained by poor handling or prolonged culture of iPSCs, loss of *PITX2* function or CRISPR off-target effects. Firstly, as discussed in Chapter 3, the additional P2A sequence in the fluorescent reporter construct introduces extra amino acid(s) either at the C-terminus of the target gene or at the N-terminus of the fluorophore. The C-terminus remnant of the P2A peptide may alter the configuration of PITX2 and lead to a non-functional allele. The influence of additional sequences, either N- or C-terminus, acquired from P2A cleavage, has been largely unexplored and may vary from gene to gene. (Reinhardt et al., 2020) found that the remaining N-terminal proline from P2A destabilised krupple like factor 4, leading to protein degradation and loss of iPSC reprogramming capacity. Constitutive *Pitx2* KO caused embryonic lethality and atrial isomerism in a murine model, signifying the important role of this gene in cardiac development and morphogenesis (Tessari et al., 2008). In addition, CM-specific *Pitx2* KO delayed CM maturation in the ventricles as illustrated by peri-membranous ACTN2 localisation in contrast to Z-disc localisation in the mature stage (Tessari et al., 2008). Nevertheless, *PITX2* KO iPSCs retained their cardiac differentiation potential (Marczenke et al., 2017). Therefore, loss-of-function less likely supports the defect in the *PITX2-mClover* line. Secondly, the *PITX2* CRISPR may exert off-target effects on the essential regulators of cardiac differentiation e.g. BMP4 and fibroblast growth factor (Clowes et al., 2014). However, all CRISPRs in this thesis were carefully selected based on criteria: high specificity and low off-target scores predicted by http://crispor.tefor.net/. Thirdly, as discussed earlier, poor iPSC handling or non-specific accumulation of karyotype aberrations in culture, can alter the differentiation potential. Those having a poor differentiation potential may predominate the iPSC population, leading to the complete loss of differentiation potential.

V) Characterisation of MYL2-mClover line

The heterozygous *MYL2-mClover* reporter line clones 6F and 11C showed loss of both MYL2 and mClover protein expression at day 25 when differentiated by the ventricular protocol. Therefore, they were no longer qualified as a fluorescent reporter. The loss of the protein expression in the clone 11C was supported by the loss of the transcripts. The absence of *MYL2* expression is associated with the delay in differentiation in the clone 11 as this marker tends to predominate in mature iPSC-CMs (Piccini et al., 2015). Both clones contained CRISPR-induced deletions at the C-terminus of *MYL2* on the non-targeted alleles. These undesired mutations truncate the third domain of Ca^{2+} binding EF hand (residues 130-165, Uniprot #P10916). A HCM-associated frameshift variant within this domain (c.431_432delCT) has been shown to destabilise the MYL2 structure, leading to protein degradation and aggregate accumulation (Manivannan et al., 2020). Thus, protein instability likely explains the loss of MYL2 protein expression from the non-targeted allele in the clone 6F where its transcript was present. However, the situation for MYL2 and mClover on the targeted allele is complicated. The loss of mClover expression is likely related to the "translation fall-off" mechanism. The translation may continue from the *MYL2* sequence until reaching the P2A cleavage site (i.e. the last proline residue). Here, the ribosome subunits disintegrate and the translation is halted (Liu et al., 2017), resulting in only the MYL2 peptide being produced. Given that the C-terminus is essential for the protein stability (Manivannan et al., 2020), the additional sequences acquired from P2A may also induce MYL2 protein degradation.

4.6.1 RA-induced atrial cell specification

RA signalling plays an essential role in cardiogenesis. RA receptors form complexes with retinoid X receptors. These complexes, in the presence of RA, function as transcription factors which activate the expression of genes featuring RA responsive elements (Figure 4-21). In the absence of RA, the complexes recruit transcription corepressors, which promote histone deacetylase and Polycomb repressive complex 2 interaction at the target genes to silence their expression (Zhao et al., 2020). RA biosynthesis disruption caused abnormal second heart

field development in a mouse model, leading to impaired heart formation lacking the atrial chamber(s) (Ryckebusch et al., 2008, Kruse et al., 2008).

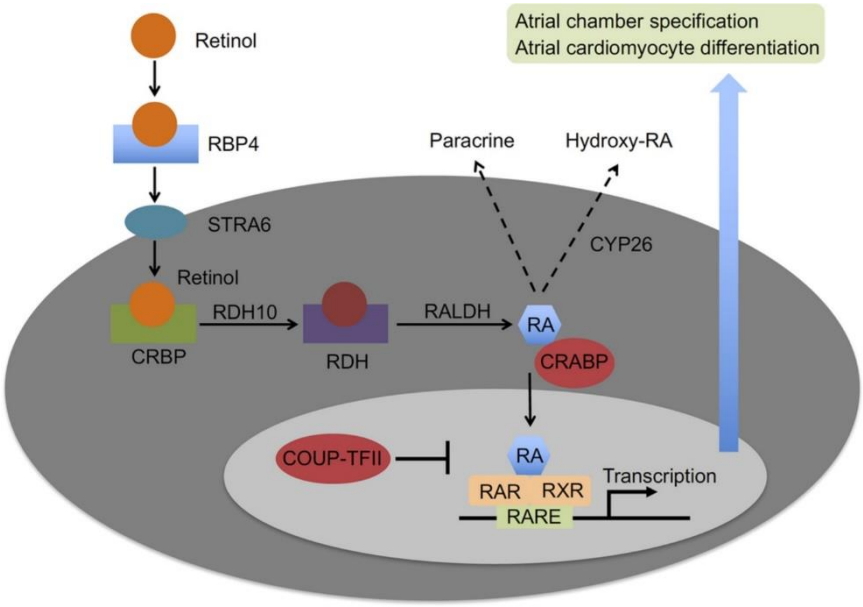


Figure 4-21 RA signalling and its regulatory role in atrial lineage specification. Circulating retinol binds to retinol binding protein 4 (RBP4) and is transported into the target cell by stimulated by retinoic acid gene 6 reporter (STRA4). Retinol dehydrogenase (RDH), with the cellular retinol bind protein (CRBP) aid, converts retinol into retinaldehyde, which can be further oxidised into RA by retinaldehyde dehydrogenase (RALDH). RA translocates into the nucleus, with the cellular retinoic acid binding protein (CRAPB) assistance, and binds to retinoic acid receptor (RAR). RAR forms an active complex with retinoid X receptor (RXR) and turns on expression of genes under the control of retinoic acid responsive element (RARE), including chicken ovalbumin upstream promoter transcription factor II (COUP-TFII) encoded by the *NR2F2* gene. Activated COUP-TFII is believed to mediate the atrial chamber morphogenesis and atrial CM differentiation and also potentially suppress the RXR-mediated RA signalling pathway. The excessive RA can be either secreted as a paracrine signal or degraded by cytochrome P450 family 26 (CYP26). Taken from (Zhao et al., 2020)

Studies have suggested that RA supplementation during the mesodermal stage enables atrial-like cell differentiation in iPSC-CMs and ESC-CMs (Cyganek et al., 2018, Zhang et al., 2011, Lee et al., 2017). The underlying RA-induced atrial cell commitment mechanism is not well understood but it may be associated with the activated NR2F2 pathway (Wu et al., 2013). CM-specific *Nr2f2* KO mice developed ventricularised atria, ventricular-like action potential and extensive T-tubules with downregulation in atrial markers (Wu et al., 2013). Similarly, *NR2F2* knockdown ESC-CMs showed downregulation of atrial-specific voltage-gated K^+ channel 1.5, inwardly rectifying K^+ channels 3.1 (encoded by *KCNJ3*) and inwardly rectifying K^+ channel 3.4 (Devalla et al., 2015)

In this chapter, RA was used to produce a specific phenotype in the *MYL7-mClover* fluorescent reporter iPSC-CM line. Treatment of RA between days 3-9 successfully steered the CM differentiation toward the atrial pathway as confirmed by the molecular, morphometric and functional characteristics.

Activation of the atrial gene programme, including *KCNJ3*, *PITX2*, *NR2F2* and *MYH6*, and inhibition of the ventricular gene programme, including *MYL2*, *HEY2* and *IRX4*, suggested that RA molecularly induced a fate change toward atrial-like iPSC-CMs, which was previously reported in other ESC-CMs and iPSC-CMs (Lemme et al., 2018, Lee et al., 2017, Cyganek et al., 2018, Gunawan et al., 2021, Devalla et al., 2015).

Both primary human and rodent ventricular CMs have a more rectangular shape whereas atrial CMs are smaller in size with a more spindle-like morphology (Bird et al., 2003, Blackwood et al., 2020, Plačkić and Kockskämper, 2018, Bögeholz et al., 2018).The significant cell area reduction by RA supplementation during the *MYL7-mClover* iPSC-CM differentiation therefore confirmed the morphometric change toward the atrial pathway.

Lastly, the RA-treated iPSC-CMs exhibited accelerated calcium cycling and sarcomere contraction and relaxation, resembling functional characteristics of atrial-like CMs (Cyganek et al., 2018, Lemme et al., 2018). These phenotypes can be explained by the role of RA in regulating the Ca^{2+} handling mechanism. KO of RA receptor α has been shown to modulate *SERCA* expression, PLN phosphorylation and Ca^{2+}/calmodulin-dependent protein kinase II δ phosphorylation in mice, leading to delayed Ca^{2+} reuptake and sarcomere relaxation (Zhu et al., 2016). Phosphorylation of Ca^{2+}/calmodulin-dependent protein kinase II δ is important for its functions in activating several proteins involving in the Ca^{2+} homeostasis such as RYR, SERCA and LTCC (Zhang, 2017). Thus, overall, RA increases both the Ca^{2+} release and reuptake activities, resulting in abbreviated Ca^{2+} transient kinetics. The Ca^{2+} transient kinetics positively correlates with sarcomere contraction and relaxation kinetics as troponin-calcium binding is a pre-requisite for actomyosin crossbridges and subsequent myosin power stroke (England and Loughna, 2013).

4.6.2 *MYL7* expression during *in vitro* atrial differentiation

The expression of *MYL7,* encoding the myosin regulatory light chain atrial isoform, is enriched in the adult atria (Guo and Pu, 2020). Many previous works conclusively agreed on the effect of RA supplement in suppressing the myosin regulatory light chain ventricular isoform *MYL2* in ESC-CMs and iPSC-CMs (Gunawan et al., 2021, Lemme et al., 2018, Devalla et al., 2015, Laksman et al., 2017, Cyganek et al., 2018, Schwach et al., 2022, Miao et al., 2020). However, the reverse effect is only reported to a limited number for *MYL7* (Lemme et al., 2018, Miao et al., 2020, Cyganek et al., 2018).

MYL7 is the predominant myosin regulatory light chain isoform in iPSC-CMs (Kamakura et al., 2013). In the *MYL7-mClover* line, there was a continuous increase of the expression till day 30 of differentiation by the ventricular protocol. The expression of *MYL7* in general ESC-CM and iPSC-CM populations, independent of atrial or ventricular subtype, is considered as an indicator for immaturity (Burridge et al., 2014, Fukushima et al., 2020, Zhang et al., 2009, Kim et al., 2017, Kamakura et al., 2013) and it is interesting to investigate further at which day its expression starts to decline in non-atrial subtypes. This situation recapitulates the embryonic cardiac development, where *Myl7* is expressed throughout the primitive heart tube (Kubalak et al., 1994, Franco et al., 1998) and uniformly expressed in all chambers of the early embryonic heart (mouse embryonic day 10.5) (Zammit et al., 2000). *Myl7* is downregulated in the ventricular chambers at approximately mouse embryonic day 12.5 (Zammit et al., 2000, Chen et al., 1998), after which it becomes confined to the atria into adulthood (Kubalak et al., 1994).

Based on the FACS analysis, there were no clear difference in MYL7 and mClover expression observed in the *MYL7-mClover* reporter line between the RA-treated and non-treated conditions. Similarly, (Gunawan et al., 2021) showed that RA supplementation drove the atrial differentiation in iPSC-CMs at both molecular and functional levels despite unaffected *MYL7* expression. It is believed that the differential *MYL7* expression may require additional maturation. In support, (Cyganek et al., 2018) reported that the expression of MYL7 protein in RA-treated iPSC-CMs increased with time in culture. Comparing *MYL7* expression levels between the control and RA-treated iPSC-CMs may need an observation at the optimal time point where its expression is downregulated in non-atrial subtypes while specifically elevated in the atrial subtype (Figure 4-22).

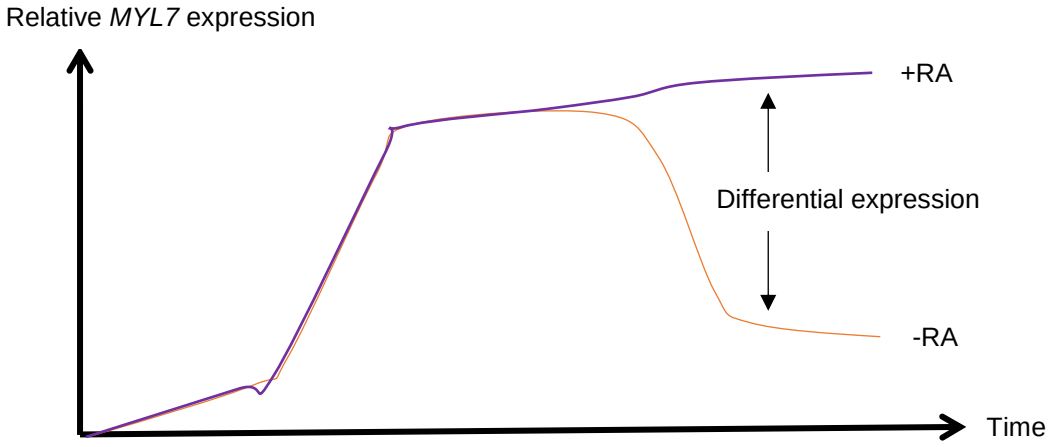


Figure 4-22 Proposed time course *MYL7* expression in iPSC-CMs differentiated with and without RA.

However, it would be challenging if live KOLF2-C1-derived iPSC-CMs need to be FACS-sorted at the late stage for further study of the utility of the *MYL7-mClover* reporter line for atrial-like cell isolation as they tend to be less viable during cell dissociation and post sorting. Maturation methods other than prolonged culture e.g. 3-dimentional culture (Lemme et al., 2018, Goldfracht et al., 2020) may be strategies worth exploring.

4.6.3 *HCN4* expression during iPSC-CM maturation

To my current knowledge, there is no single specific biomarker for nodal cells and fluorescent reporters for nodal-specific iPSC-CMs have not been extensively engineered and studied

compared to the contractile subtypes. Identification of nodal cells usually relies on a set of molecular markers such as HCN4, SHOX2 and ISLET1 and T-box 1 (Ichimura et al., 2020) (Goodyer et al., 2019).

HCN4 is responsible for the nodal cell's automaticity (Klabunde, 2011). Interestingly, the fluorescent mScarlet signal observed by microscopy in the day 25 live *HCN4-mScarlet* iPSC-CMs was detected in most of the spontaneously contractile cells, suggesting that HCN4 may be abundantly expressed in iPSC-CMs irrespective of subtype. This agrees with the FACS analysis which showed no distinct mScarlet positive population but rather a gradient fluorescent pattern, where those cells displaying higher mScarlet intensity may represent a nodal-like population. (Yechikov et al., 2016) illustrated that not only iPSC-CMs exhibiting the nodal-type electrophysiology were positive for HCN4 expression, but also those identified with the atrial-type and the ventricular-type electrophysiology.

HCN4 expression in ESC-CMs and iPSC-CMs tends to be downregulated with time in culture (Ichimura et al., 2020, Bosman et al., 2013, Yechikov et al., 2016, Sartiani et al., 2007), which is likely consistent with the time-course analysis in the *HCN4-mScarlet* reporter line. Reduction of *HCN4* expression at the late time point is likely associated with iPSC-CM maturation. It is suggested that *HCN4* is first expressed in the first heart field progenitor. Subsequently, its expression is downregulated in their atrial and ventricular CM progenies and becomes restricted to nodal CMs (Später et al., 2013). Caveolin 3 is a membrane protein whose interaction with ion channels influences the biophysical properties of the ion current. HCN4-calveolin3 compartmentalisation is only present in human adult CMs, not in foetal CMs (Bosman et al., 2013). In ESC-CMs, this molecular interaction progressively develops in aged cells. Loss of ubiquitous *HCN4* expression and increase of HCN4-caveolin 3 interaction correlate with slowed diastolic depolarisation and slowed spontaneous action potential firing (Sartiani et al., 2007, Bosman et al., 2013), resembling mature CM phenotypes.

4.6.4 Purification of iPSC-derived nodal-like cells

Purification of nodal-like iPSC-CMs may require further both maturation and nodal differentiation protocol. iPSC-CM maturation is likely to cause HCN4 expression to be more restricted to the nodal-like cells (Sartiani et al., 2007, Bosman et al., 2013), thus allowing the *HCN4*-derived mScarlet expression to be specific to nodal-like cells.

Nodal-like cells are the rarest subtype derived from the iPSC-CM differentiation by the ventricular protocol (Churko et al., 2018), making their purification yield relatively low. Development of selective differentiation approaches may be necessary to achieve a scalable iPSC-derived nodal-like cell production required for cardiovascular research but to date, such knowledge is still largely incomplete. For example, (Schweizer et al., 2017) illustrated that switching from serum-free medium to serum-rich medium coupled with visceral endoderm-like cell co-culture directed the iPSC-CM differentiation toward the nodal-cell lineage as confirmed by the nodal-type electrophysiology. A recent study has shown that inhibition of the NODAL signalling favoured the iPSC-CMs differentiation toward the nodal lineages by suppressing the *PITX2* pathway (Yechikov et al., 2020). PITX2 is known to mediate downregulation of an essential nodal transcription factor *SHOX2* and the development of sinoatrial node on the left atrium. (Liu et al., 2020) suggested an alternative way to promote nodal-like cell differentiation in iPSC-CMs by simultaneous modulating BMP4, fibroblast growth factor and RA signalling pathways. Inhibition of fibroblast growth factor and RA signalling are believed to effectively suppress ventricular and atrial specification, respectively (Liu et al., 2020). As presented in Chapter 1, BMP4 is a known factor enabling cardiac mesoderm induction. Maintained low concentration of BMP4 during the cardiac mesoderm stage was shown to be important for nodal-like cell population enrichment (Liu et al., 2020).

Even though the *mScarlet* expression mirrored the *HCN4* expression at the transcriptional level, the defective differentiation in the *HCN4-mScarlet* reporter iPSC-CM limited the downstream application(s). Since the defect was not found in the whole culture, it is likely that there was a heterogeneity in the differentiation potential. Poor iPSC handling may have

introduced such effect. The loss of differentiation potential decreased the yield of both iPSC-CMs in general and nodal-like cells, hindering the development of a viable nodal cell purification approach. iPSC subcloning may help to obtain a better performing clonal fluorescent reporter line.

To sum up, both CM maturation and nodal-selective differentiation methods may be useful for further development of an efficient iPSC-derived nodal-like cell purification method based on the fluorescent *HCN4-mScarlet* reporter line. While the former increases the yield qualitatively, the latter increases the yield quantitatively. The combination of using a nodal-specific reporter line and optimised nodal differentiation approach has been recently shown to enable a purification of nodal-like ESC-CMs in a *SHOX2-eGFP*/*MYH6-mCherry* dual reporter line (Ghazizadeh et al., 2022).

4.6.5 Summary

In this chapter, the engineered atrial-specific *MYL7-mClover*, atrial-specific *PITX2-mClover*, nodal-specific *HCN4-mScarlet* and ventricular-specific *MYL2-mClover* iPSCs were differentiated into CMs using the conventional (ventricular) protocol. The first question was whether these lines retained their cardiac differentiation potential after the genetic manipulation. If this was the case, they were further characterised by focusing on the question of whether the fluorescent reporter systems would be useful for representing the expected CM subtypes. To help characterise the atrial reporter lines, an atrial differentiation protocol was established by RA supplementation using the *MYL7-mClover* line as a testing model.

I. 1 µM RA supplementation during the days 3-9 was shown to exert highest upregulation effect on the atrial markers (*MYL7*, *MYH6*, *KCNJ3*, *NR2F2* and *PITX2*) and downregulation effect on the ventricular markers (*MYL2*, *HEY2* and *IRX4*). It also induced significant CM size reduction, accelerated Ca^{2+} transient dynamic, shortened sarcomere contraction time and shortened sarcomere relaxation time, recapitulating the atrial-like characteristics. Thus, this condition was used as a standard for the atrial differentiation.

II. The heterozygous *MYL7-mClover* reporter line was successfully differentiated into CMs by both the ventricular and RA-induced atrial protocols. Time course analysis confirmed that the *mClover* expression faithfully mirrored the expression of *MYL7* by the ventricular differentiation protocol. Though the RA induction was shown to mildly upregulate *MYL7* and *mClover* by qPCR and Western blot analyses, it failed to clearly demonstrate the differential expression when analysed by FACS. It remains to be elucidated whether both *MYL7* and *mClover* expression during atrial differentiation could be further enhanced by tuning RA conditions e.g. concentration or by iPSC-CM maturation. If such optimisations are possible, the *MYL7-mClover* reporter line would be a useful and powerful tool to identify atrial-like cells based on the mClover fluorescent Intensity. Isolating those cells displaying high fluorescent intensity would allow atrial-like cell enrichment.

III. Using the ventricular protocol, the heterozygous *HCN4-mScarlet* reporter line showed an intermediate degree of cardiac differentiation defect, which could be related to poor iPSC handling or non-specific karyotype aberrations caused by prolonged culture. Time course analysis confirmed that the *mScarlet* expression faithfully mirrored the expression of *HCN4.* Live cell imaging and FACS analysis suggested that the *HCN4*-driven mScarlet may be ubiquitously expressed in iPSC-CMs. It remains to study if either HCN4 or *HCN4*-driven mScarlet expression better represents the nodal phenotype when iPSC-CMs mature or when using nodal differentiation protocols. If such optimisations are possible, the *HCN4-mScarlet* reporter line would be a powerful system for FACS-aided nodal-like cell purification. Nevertheless, these future experiments need alternative clones of the fluorescent reporter line with a full cardiac differentiation potential, which may be obtainable by iPSC subcloning.

IV. The homozygous *PITX2-mClover* reporter line failed to differentiate into CMs. Several factors could possibly lead to such defect e.g. PITX2 loss-of-function, CRISPR off-target effects or poor iPSC handing. The targeting strategy may need a revision for generating a new line with retained cardiac differentiation potential before a correlation between the *PITX2* and *mClover* expression could be determined.

V. The heterozygous *MYL2-mClover* line was able to differentiate into contractile CMs using the ventricular protocol. However, the resultant iPSC-CMs did not express both MYL2 and mClover at the protein level, thus these engineered lines cannot serve as fluorescent reporter systems. In one clone, loss of the expressions can be explained by a delay in differentiation while the situation in the other clone is complex due to the presence of their transcripts. In addition, both clones contained CRISPR-related deletion mutations on the non-targeted alleles. I believe that the loss of MYL2 in the later clone is associated with protein instability induced by P2A (for the targeted allele) and the deletion mutation (for the non-targeted allele) while the loss mClover is mediated by the translation fall-off mechanism. Until the genome-editing is repeated to generate the initially intended reporter line, no conclusions can be drawn about the ability of *mClover* to track the *MYL2* expression.

In summary, out of the 4 engineered lines, only the *MYL7-mClover* and *HCN4-mScarlet* lines were able to differentiate into beating CMs where the fluorescence reporters seem to track expression of the endogenous genes. However, challenges around expression of the gene of interest in the target subtype of CMs require further optimisations before the reporter lines can be used for downstream applications.

Chapter 5

TTNtv modelling in KOLF2-C1 iPSC-CMs

5.1 Introduction

Titin is a gigantic protein which plays roles in CM contraction, mechanosensing and signal transduction (Herzog, 2018). Mutations that truncate this protein, or *TTN*tv, especially those affecting the A-band region, have been identified as a major genetic cause of DCM (Roberts et al., 2015, Herman et al., 2012). However, how *TTN*tv mediates the DCM pathogenesis is largely unknown. The titin's size, titin's multiple isoforms and incomplete knowledge of the protein in cardiac biology all have challenged the traditional ways for dissecting how *TTNtv* produce the clinical phenotype. This chapter aims to address this by using CRISPR/Cas9 technology to model *TTNtv* in the iPSC-CM system.

Two rare *TTN*tv identified in DCM patients were of interest in this study. Firstly, the *TTN* c.11952 C>A mutation introduces a premature stop codon into exon 49 (I-band) (Figure 5-1) (Ahlberg et al., 2018). All exon annotation throughout this study is according to the total exon numbering. Secondly, the *TTN* c.59926+1 G>A (rs553526525 C>T) mutation is located at the intron following *TTN* exon 303 (A-band) (Figure 5-1), where it is predicted to cause protein truncation through altered mRNA splicing (Hoorntje et al., 2018). Both mutations are associated with early onset atrial fibrillation and incorporated into both the *TTN* N2-B and N2-BA isoforms. gRNAs were designed to target both regions of interest. Only the one that targeted the A-band *TTNtv* showed a reasonable efficiency. Hence, iPSCs carrying heterozygous and homozygous *TTN* c.59926+1 G>A were generated and differentiated into CMs using the ventricular protocol (Chapter 2).

This chapter specifically investigates the following:

I. Whether *TTN* c.59926+1 G>A impairs sarcomere formation.

II. How *TTN* c.59926+1 G>A truncates titin protein.

iii.) Whether *TTN* c.59926+1 G>A is associated with foetal gene programme remodelling.

Figure 5-1 Schematic illustration of *TTN* c.11952 C>A and c.59926+1 G>A locations with respect to the *TTN* locus. *TTN* c.11952 C>A is a stop-gain variant in exon 49, causing I-band truncation. *TTN* c.59926+1 G>A is a splice variant identified within the intron immediately downstream of exon 303, causing A-band truncation.

5.2 CRISPR/Cas9-assisted generation of TTNtv iPSC line

5.2.1 CRISPR design

Each CRISPR was co-designed with Max Cumberland (University of Birmingham) to target *TTN* c.11952 and c.59926+1 (Figure 5-2) and their respective gRNAs were synthesised by IDT, with chemical modification to enhance their stability.

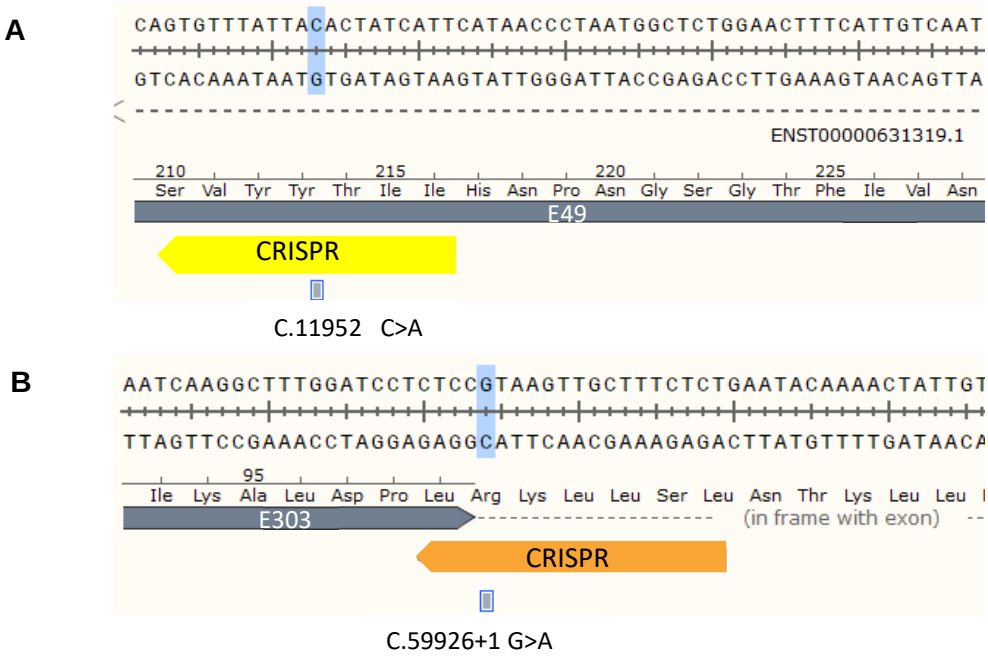


Figure 5-2 TTN CRISPR locations. A) TTN c.11952 CRISPR recognised the sequences within exon 49 and encompassed the target nucleotide. B) TTN c.59926+1 CRISPR recognised the sequences within exon 303 and the intron lying immediately downstream of exon 303, covering the target nucleotide.

Each gRNA was reconstituted with either EnGen® Spy Cas9 NLS (NEB) or Alt-R® S.p. HiFi Cas9 Nuclease (IDT) and electroporated into KOLF2-C1 cells. 48 hr post transfection, PCR covering the CRISPR target regions using the primers hTTNF5/R5 (*TTN* c.11952) or hTTNF6/R6 (*TTN* c.59926+1) were analysed for TIDE (Figure 5-3). The TTN c.59926+1 CRISPR had a higher efficiency than the TTN c.11952 CRISPR when using NEB Cas9 (15.2 vs. 1.7%) with enhanced performance when using IDT Cas9 (30.7%). Based on this, the *TTN* c.59926+1 site was easier to target by CRISPR/Cas9 and thus was selected for modelling *TTNtv* in iPSC-CMs using the tested gRNA and IDT Cas9.

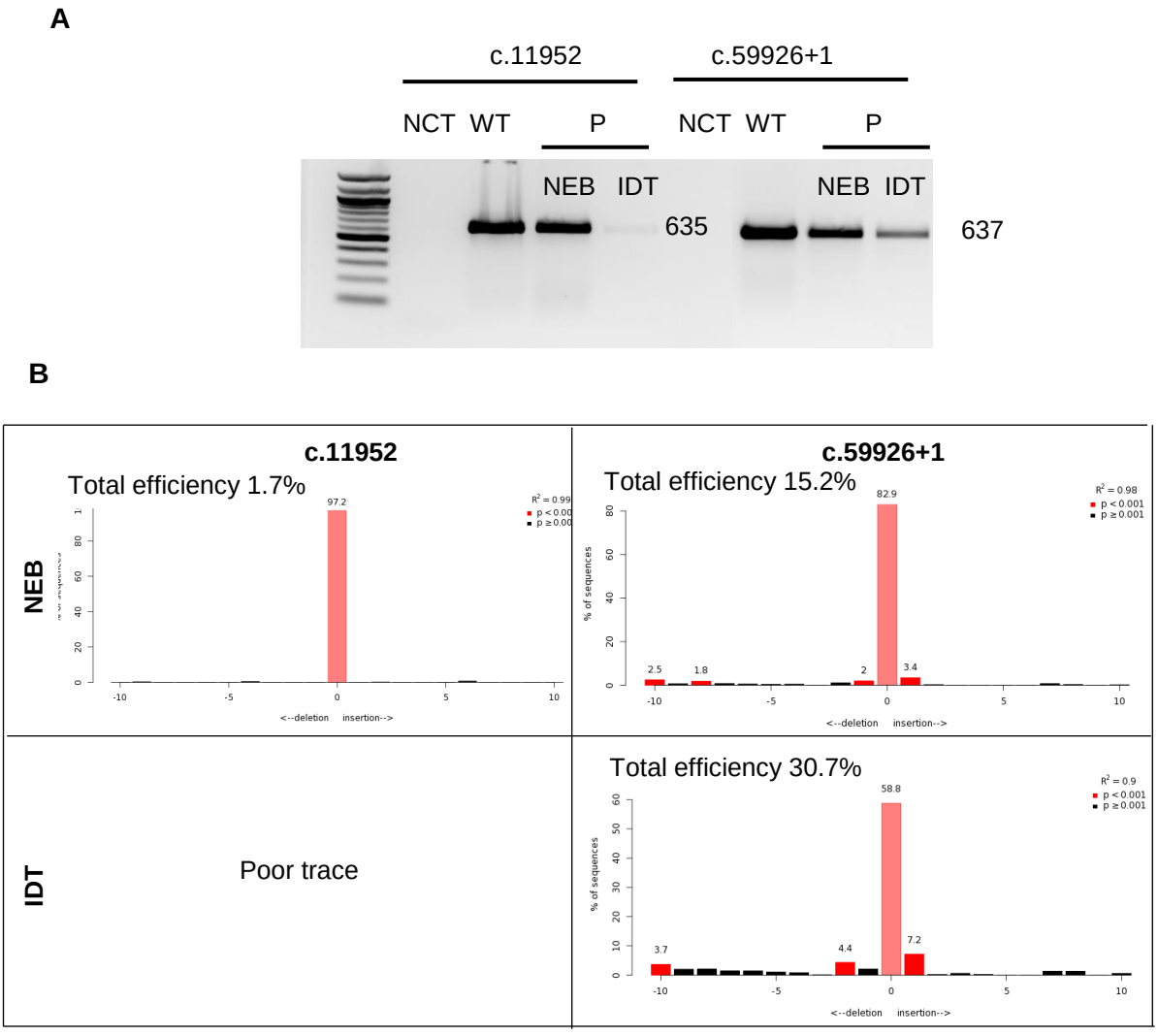


Figure 5-3 TIDE analysis for gRNAs containing TTN c.11952 and c.59926+1 CRISPRs. A) PCR products (635 bp for c.11952 CRISPR and 637 bp for c.59926+1 CRISPR) flanking the CRISPR target regions 48 hr post RNP electroporation using EnGen® Spy Cas9 NLS (NEB) or Alt-R® S.p. HiFi Cas9 Nuclease (IDT). B) TIDE efficiency.

5.2.2 *TTN* c.59926+1 G>A donor design

200 nt of sequences flanking the CRISPR recognition site (sense strand) were chosen as a ssODN donor for the targeted *TTN* c.59926+1 G>A change. A silent mutation (c.59910 G>A) was additionally introduced into exon 303 to allow the recombinant mutant allele to be tagged by a *de novo* *Hind*III restriction site (Figure 5-4). The ssODN template was synthesised by ThermoFisher.

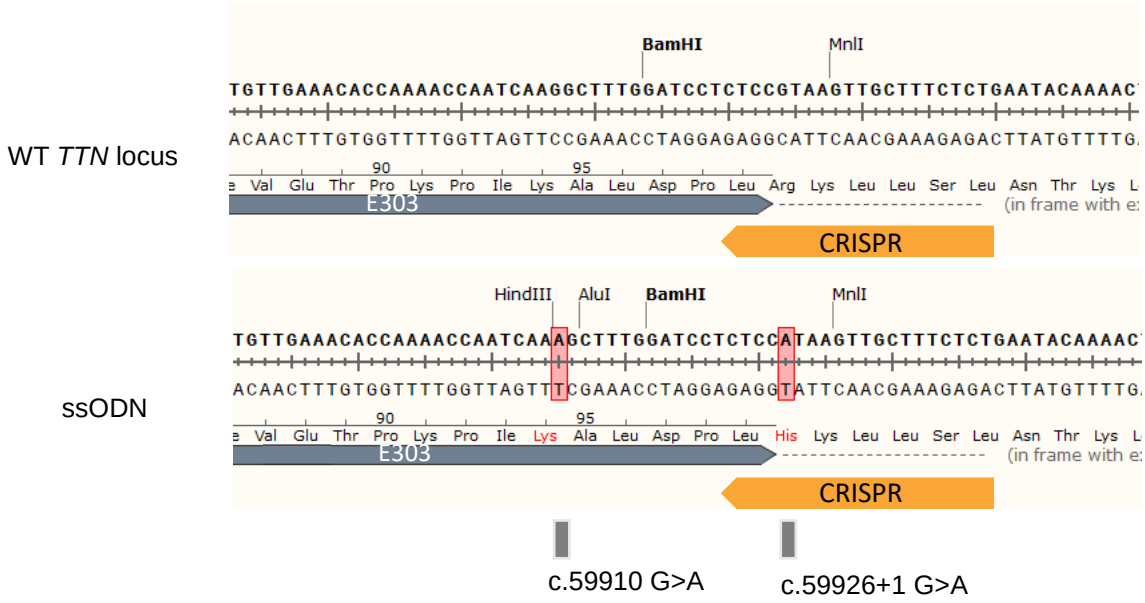


Figure 5-4 *TTN* c.59926+1 G>A ssODN donor template design. The ssODN contained the *TTN* c.59926+1 G>A mutation localised within the selected CRISPR recognition region and the additional c.59910 G>A silent mutation within exon 303 for *Hind*III digestion.

5.2.3 CRISPR/Cas9-assisted generation of *TTN*tv iPSC line

48 hr after RNP + ssODN electroporation, the pool of gDNA was amplified by PCR using the primers hTTNF6/R6 and digested by *Hind*III (Figure 5-5A). The two digested products (387 and 280 bp) directly indicated a successful ssODN incorporation and were indirectly suggestive of a *TTN* c.59926+1 G>A change. Further analysis in isolated colonies found 15/89 (~17%) being positive for the integration, 14 (93.33%) of which showed incomplete digestion (Figure 5-5B), indicative of heterozygous editing. A complete digestion was only observed in the clone 1F, presumably indicative of homozygous editing (or heterozygous editing with a

large deletion mutation on the non-targeted allele, which could not be detected by the screening primers (Kosicki et al., 2018)).

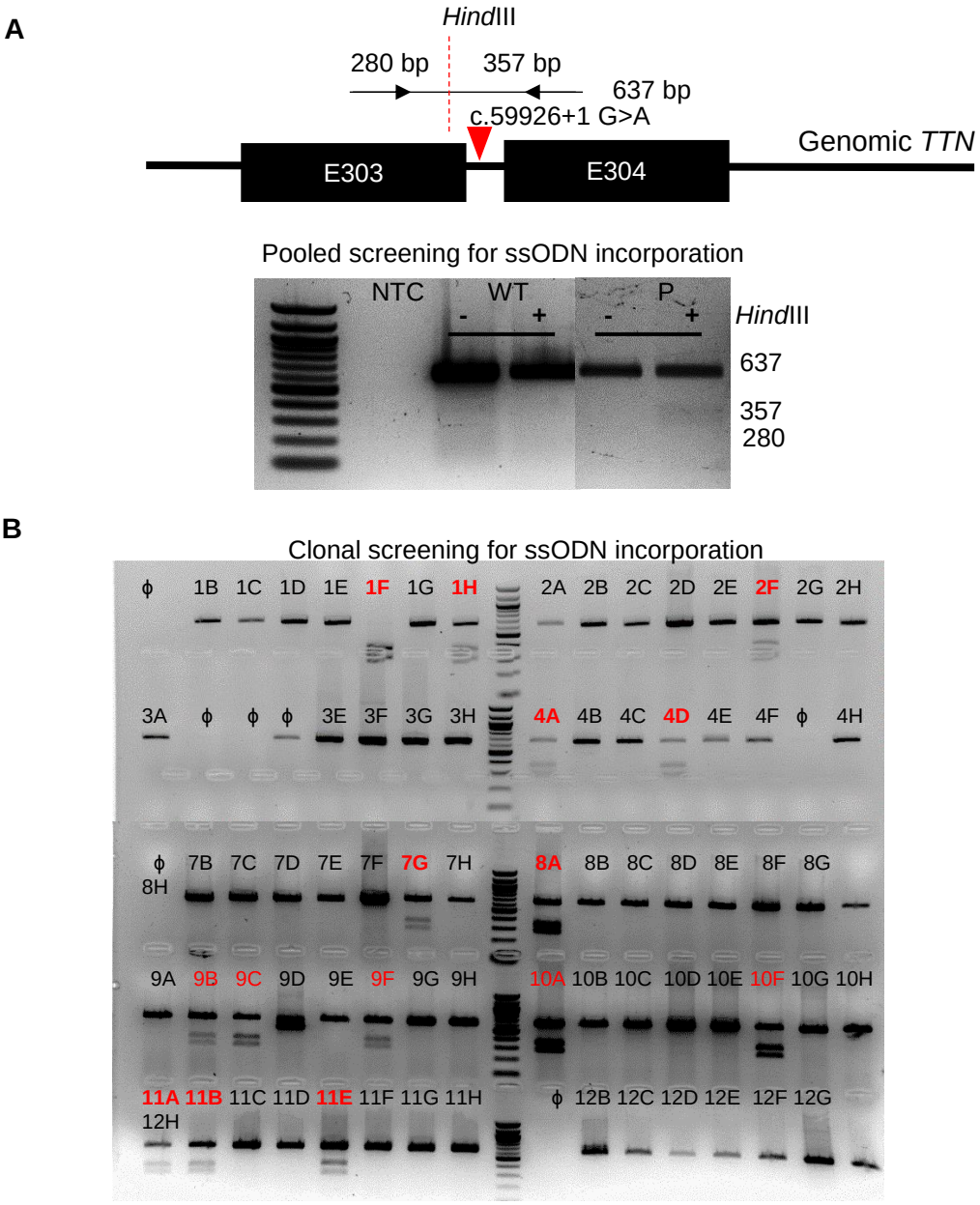


Figure 5-5 Genotypic screening for *TTN* c.59926+1 G>A change. A) *Hind*III digestion screening on PCR product encompassing the genomic *TTN*tv site in the pool of transfected iPSCs. Only PCR products acquiring ssODN donor were cleavable by *Hind*III. B) *Hind*III digestion screening in isolated iPSC colonies. NTC, no template control; P, pool of electroporated cells.

13 positive clones were expanded and genotypically confirmed by *Hind*III digestion (Figure 5-6A). Of these, 4 clones were sequenced to clarify the *TTNtv* zygosity (Figure 5-6B). Both clones 10F and 11B displayed mixed nucleotides (C/T, anti-sense strand) at both *TTN* c.59926+1 and c.59910, suggesting heterozygous clones. The clone 1F had a complete

change at both sites, thus considered as a homozygous clone. The clone 11A had an unusual HDR outcome – the *TTNtv* site was biallelically incorporated yet the *Hind*III site was likely only present in one allele.

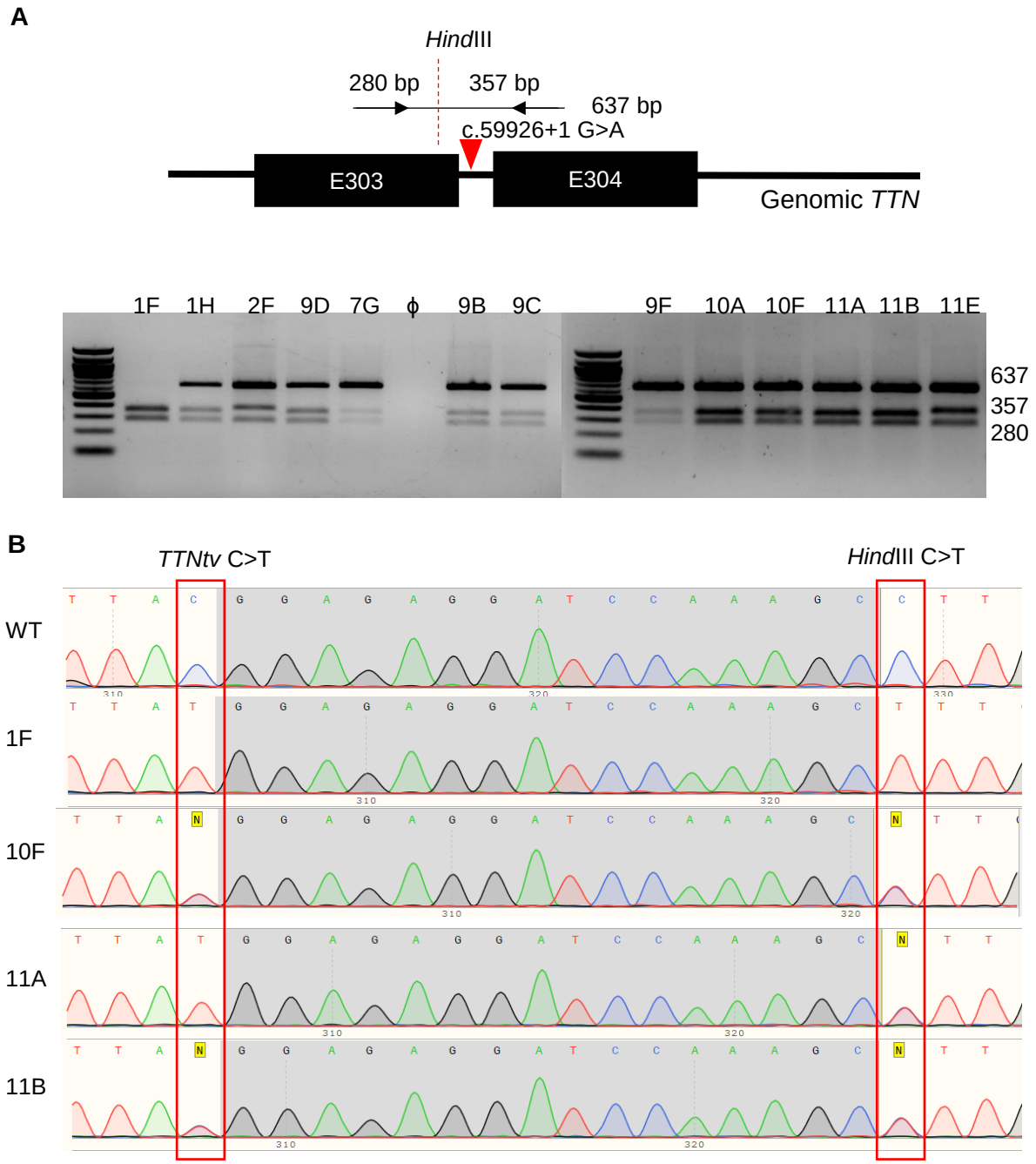
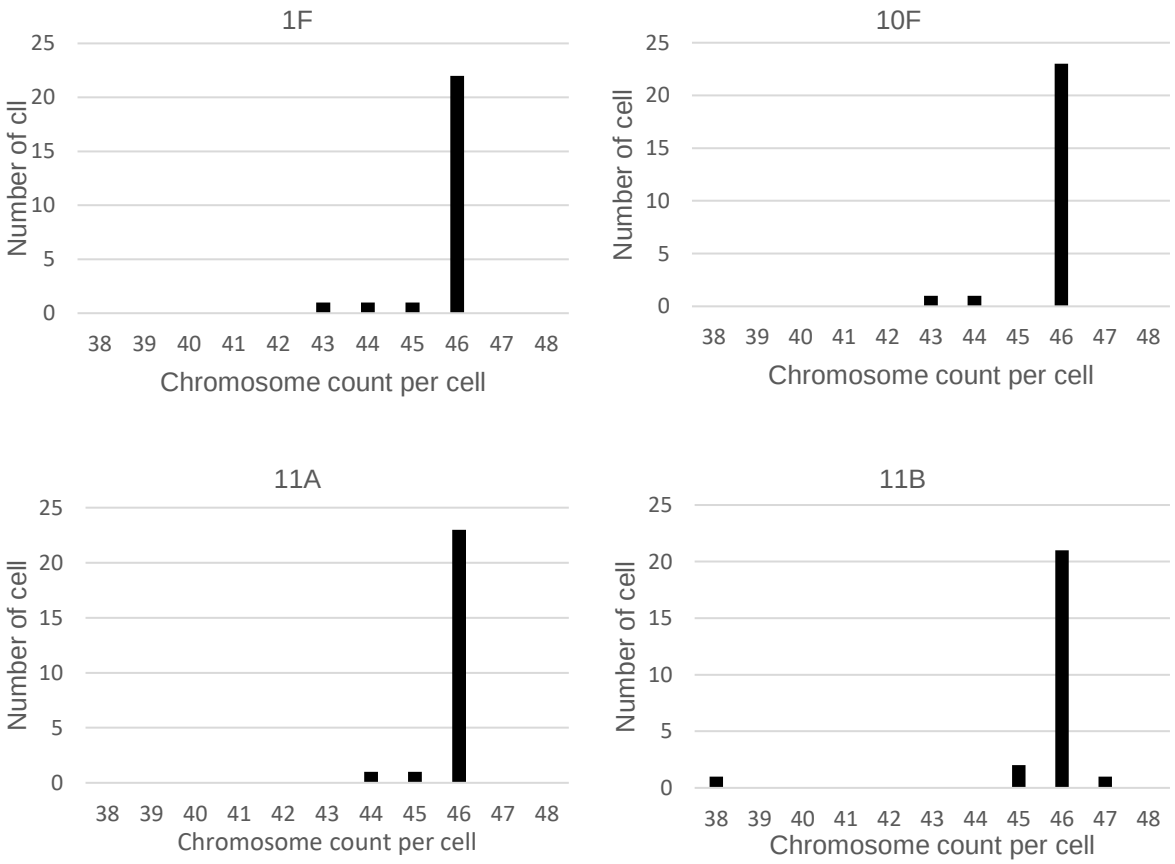


Figure 5-6 Genotypic confirmation for *TTN* c.59926+1 G>A editing. A) *Hind*III digestion on PCR products encompassing the genomic *TTN*tv site i.e. spanning *TTN* exon 303-304 in isolated iPSC colonies. Only PCR products acquiring ssODN donor were cleavable by *Hind*III (357 and 280 bp). B) Sequencing results of selected clones from A. 1F was homozygous for both the *TTNtv* and *Hind*III sites while 10F and 11B were heterozygous for both the *TTNtv* and *Hind*III sites. 11A was homozygous for the *TTNtv* site but heterozygous for the *Hind*III site.

5.2.4 iPSC quality control

All clones (1F, 10F, 11A and 11B) passed iPSC quality checks. They had a normal karyotype (Figure 5-7A) and retained expression of the pluripotency markers (OCT3/4, SOX2, SSEA4 and NANOG) (Figure 5-7B). It should be noted that the clone 11B had 1/25 cells with 47 chromosomes, but was considered karyotypically normal. The clone 1F, which was later called the *TTNtv* homozygous line, and 11B, which was later called the *TTNtv* heterozygous line, were selected for the subsequent experiments. The clone 11B was chosen rather than 10F because 10F showed a slightly late onset of differentiation (11B ~ day 9-10 vs. 10F ~day 11-12).

A



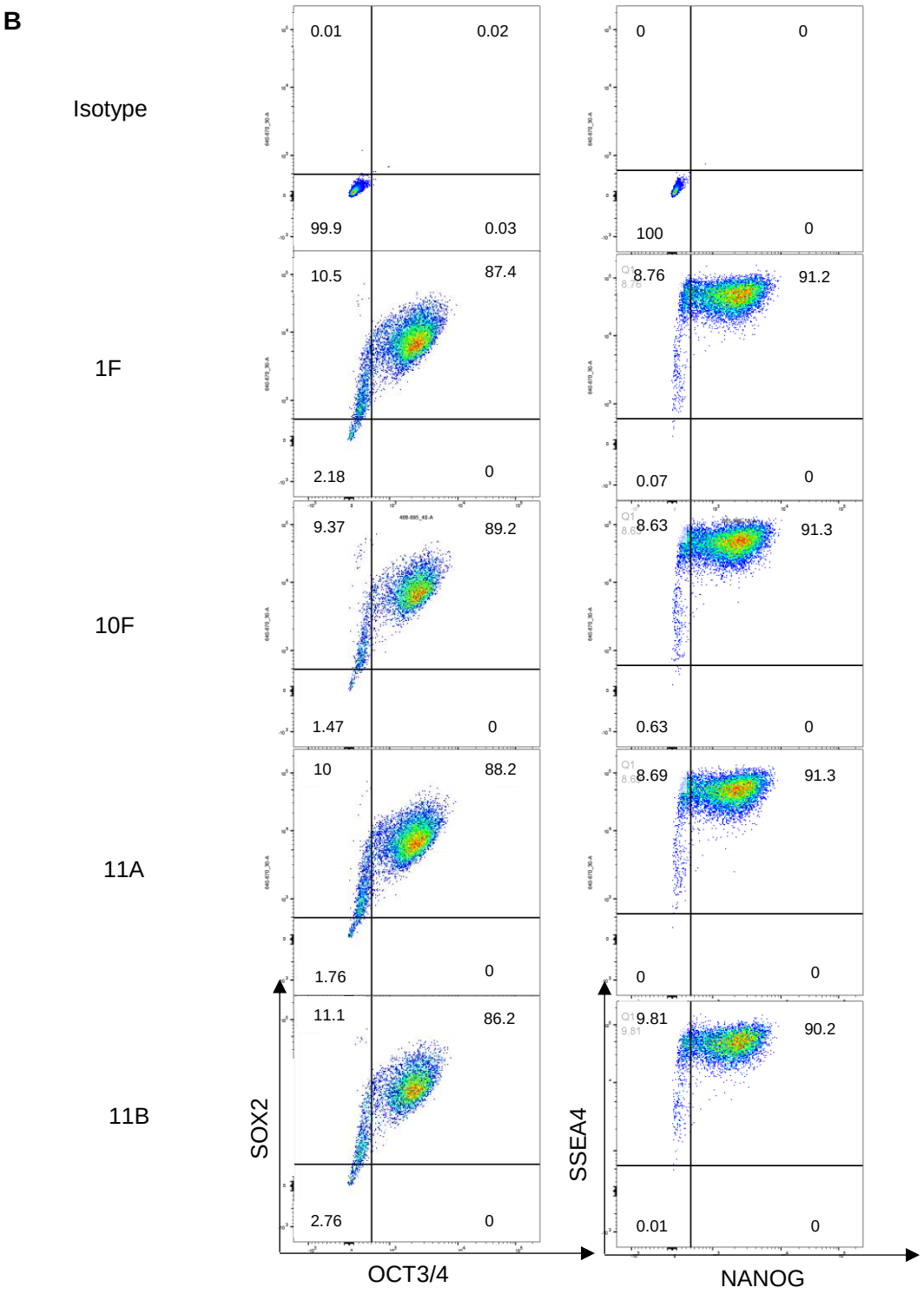


Figure 5-7 *TTN* c.59926+1 G>A iPSC quality control. A) Distribution of chromosome count per cell. B) Flow cytometry analysis of pluripotency marker expression. The smeary flow analysis plots were considered as an antibody staining problem rather than a loss of expression as every clone shared the same pattern.

5.3 Characterisation of *TTNtv* iPSC line

5.3.1 Molecular characterisation of *TTNtv*

5.3.1.1 Effect of *TTNtv* on cardiac gene expression

Both heterozygous and homozygous *TTNtv* iPSCs were differentiated into CMs using the ventricular protocol. RNA expression was quantified as compared to the WT background for a number of cardiac genes at day 21 (Figure 5-8 & Figure 5-10) and day 31 (Figure 5-9 & Figure 5-11) of differentiation.

Both time points showed no significant changes in the expression of pancardiac markers (*ACTN2*, *TNNT2* and *NKX2.5*) for both heterozygous and homozygous lines (Figure 5-8A & Figure 5-9A), suggesting that *TTNtv* did not inherently silence genes involved in the CM differentiation.

However, the homozygous line displayed changes in the expression of myofibrillar protein isoforms. The early myosin heavy chain isoform *MYH6* was found to be significantly upregulated whilst the late isoform *MYH7* was downregulated especially in day 21 (i.e. reduced *MYL7/MYH6* ratio) (Figure 5-8B & Figure 5-9B), indicating delayed CM development. The slight *MYH7* upregulation in the day 21 samples and the slight *MYH6* downregulation in the day 31 samples of the heterozygous line were considered as a consequence of iPSC-CM phenotypic variation.

Interestingly, the heart failure-associated foetal genes atrial natriuretic peptide (*NPPA*) and brain natriuretic peptide (*NPPB*) (Chien et al., 1991, Rademaker and Richards, 2005) were downregulated in the homozygous line at both time points (Figure 5-8C & Figure 5-9C). This observed reduced expression may suggest that the *TTNtv* caused a delay in iPSC-CM development. The heterozygous line showed a mild-but-significant *NPPB* upregulation (2-4 fold) in the day 31 samples which may be explained by variation within iPSC-CMs. With regards to the lack of a support from *NPPA*, such change was not considered to reflect a molecular remodelling associated with failing CMs.

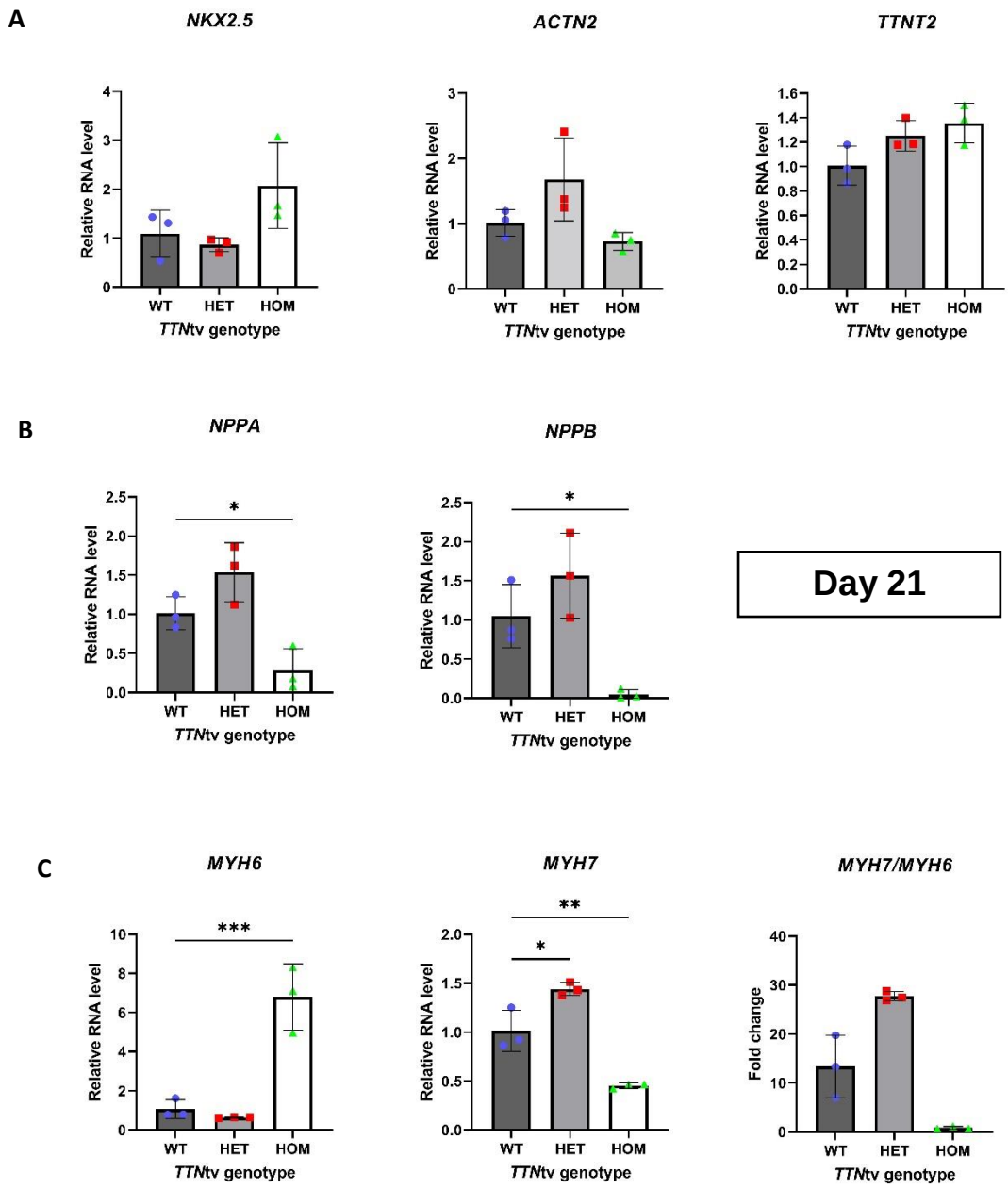


Figure 5-8 Transcriptional expression profile of day 21 *TTN* c.59926+1 G>A KOLF2-C1 iPSC-CMs. A) Pan-cardiac markers *NKX2.5, ACTN2*, and *TNNT2*. B) Heart failure markers *NPPA* and *NPPB*. C) Myosin heavy chain isoforms: *MYH6* and *MYH7*. All quantifications were made relative to the WT control, except for the *MYH7/MYH6* ratio. Each dataset was obtained from different wells of differentiation using a single iPSC batch. n=3, mean ± S.D., one-way ANOVA with Turkey's post-hoc. HET, heterozygous *TTNtv*; HOM, homozygous *TTNtv*.

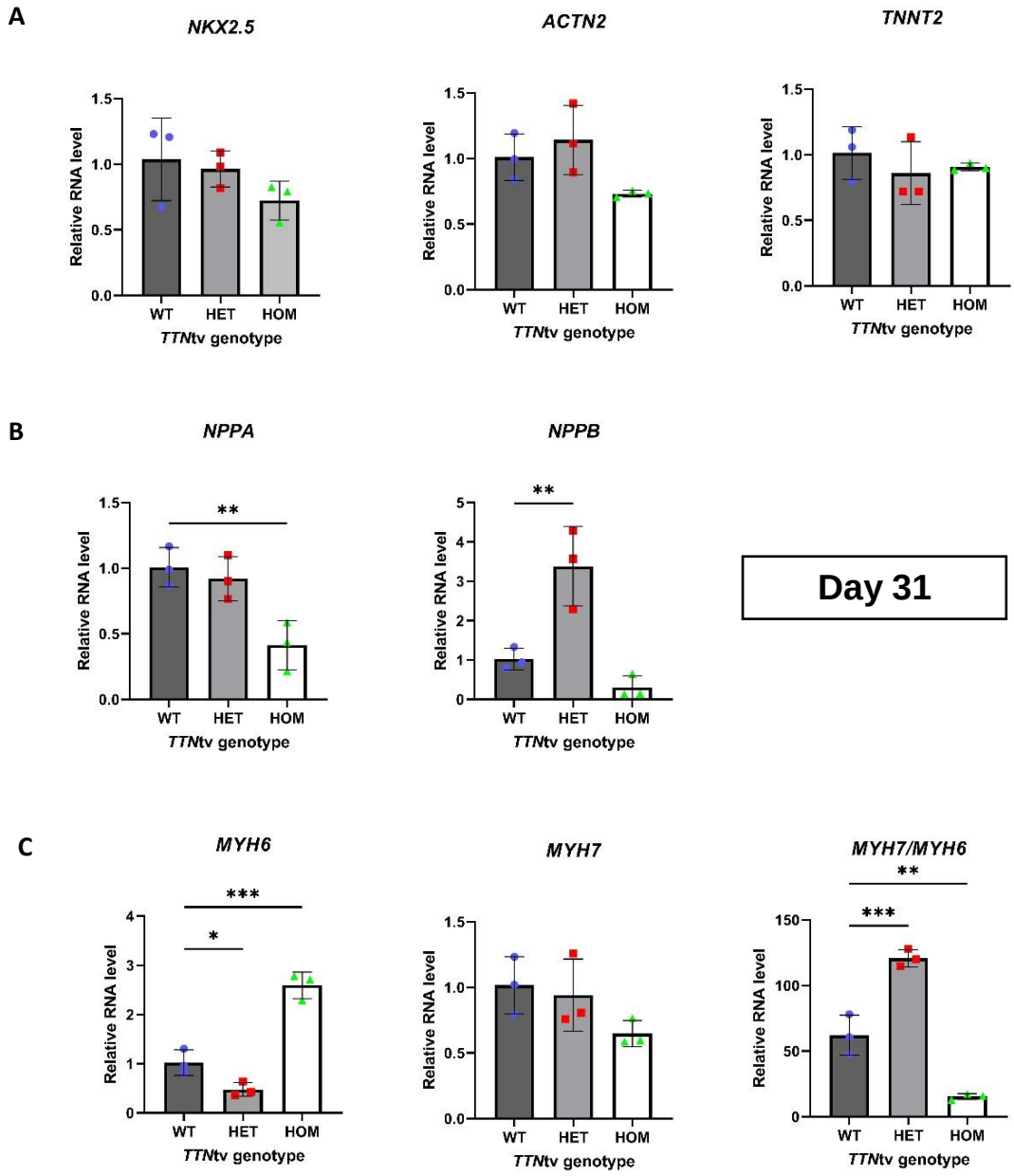


Figure 5-9 Transcriptional expression profile of day 31 *TTN* c.59926+1 G>A KOLF2-C1 iPSC-CMs. A) Pan-cardiac markers *NKX2.5, ACTN2*, and *TNNT2*. B) Heart failure markers *NPPA* and *NPPB*. C) Myosin heavy chain isoforms: *MYH6* and *MYH7*. All quantifications were made relative to the WT control, except for the *MYH7/MYH6* ratio. Each dataset was obtained from different wells of differentiation using a single iPSC batch. n=3, mean ± S.D., one-way ANOVA with Turkey's post-hoc. HET, heterozygous *TTNtv*; HOM, homozygous *TTNtv*.

TTN can be expressed in many isoforms and it is interesting to study whether *TTN* c.59926+1 G>A has an effect on *TTN* isoform switching. To quantify *TTN* transcriptional levels, several PCR probes were designed to target different constitutive exons (Figure 5-10 & 5-11). The Z-disc probes (hTTN-qF1/R1) span exons 6-7, which were used to quantify all *TTN* isoforms except the Cronos. The hTTN-qF3/R3 probes which span exons 296-297 were used for quantification of the A-band region upstream of the *TTNtv* site. The hTTN-qF4/R4 probes which span exons 317-318 were used for quantification of the A-band region downstream of the *TTNtv* site. The M-line probes (hTTN-qF6/R6) span exons 360-361. As there is no specific exon for the Cronos isoform, the A-band probes and M-line probes were together used for identification of alteration in the Cronos isoform expression. The hTTN-qF7/R7 probes spanning exons 11-12 were used for quantification of the N2-BA isoform. The hTTN-qF9/R9 probes spanning exons 49-50 were used for quantification of overall cardiac-specific isoforms (N2-B and N2-BA). Changes in N2-BA in combination with N2-BA/N2-B + N2-BA ratio were used a determinant of changes in titin stiffness/compliance.

There was no dramatic change in *TTN* levels when observed using specific probes for the Z-disc, A-band upstream of the *TTNtv* site, A-band downstream of the *TTNtv* site and M-line regions in both the heterozygous and homozygous lines for both time points (Figure 5-10 & 5-11), suggesting no alteration in the Cronos isoform expression or overall *TTN* expression. The slightly increased Z-disc *TTN* expression in day 21 heterozygous samples and the slightly decreased M-line *TTN* expression in day 31 homozygous samples were considered a consequence of iPSC-CM phenotypic variation. There was an increase in *TTN* N2-BA level or N2-BA/N2-B + N2-BA ratio in the homozygous line for both time points, which may reflect a delay in CM development.

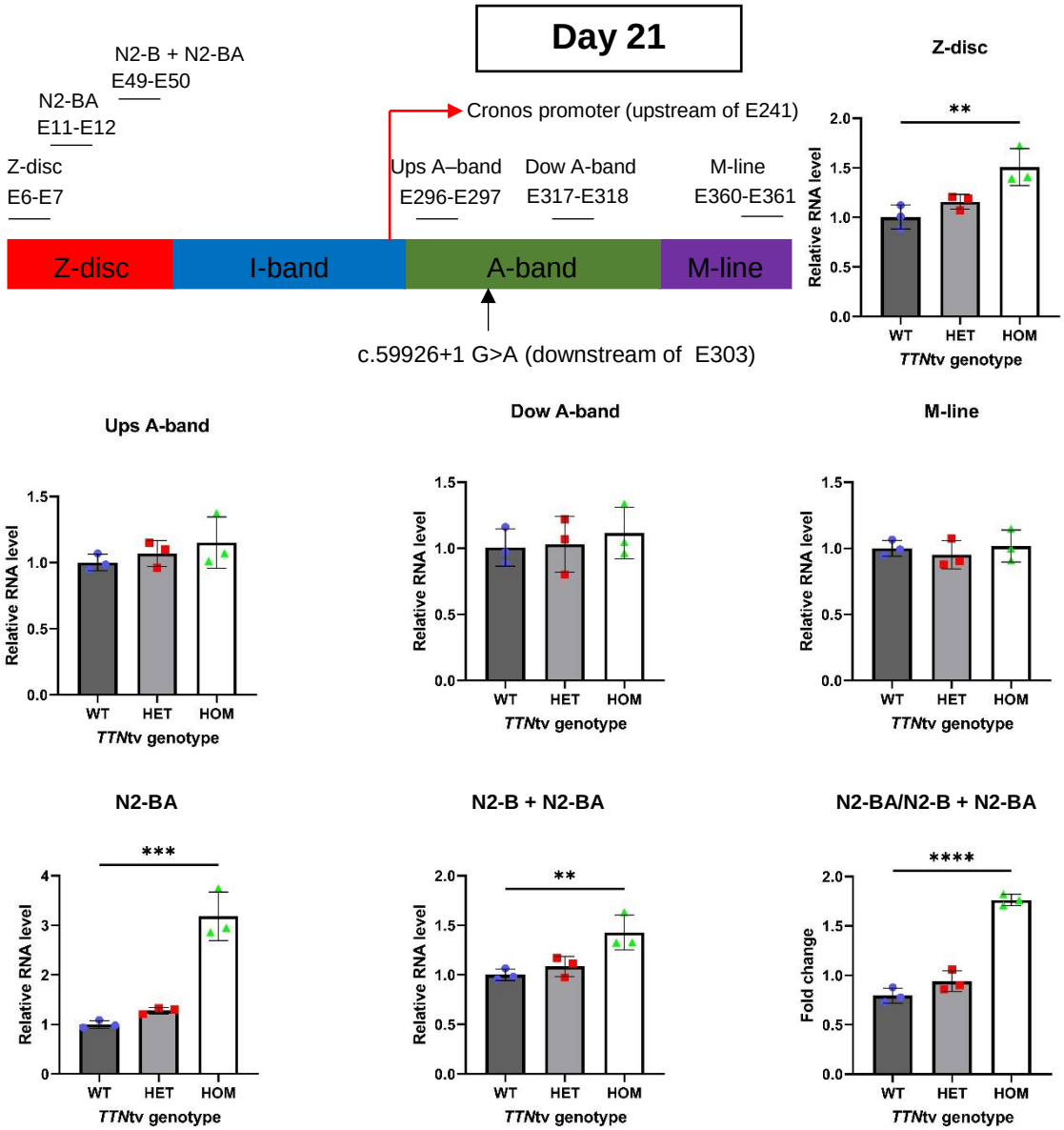


Figure 5-10 *TTN* transcriptional levels of day 21 *TTN* c.59926+1 G>A KOLF2-C1 iPSC-CMs. PCR probes were designed to be specifically spanning different pairs of exons responsible to the *TTN* regions; Z-disc (exons 6-7), A-band upstream of the *TTNtv* site (exons 296-297), A-band downstream of the *TTN*tv site (exons 317-318), and M-line (exons 360-361). The N2-BA probes (exons 11-12) were localised within the Z-disc region. The probes spanning exons 49-50 were designed to target both the cardiac specific N2-B and N2-BA isoforms. Each dataset was obtained from different wells of differentiation using a single iPSC batch. All quantifications were made relative to the WT control, except for the N2-BA/N2-B + N2-BA ratio. n=3, mean ± S.D., one-way ANOVA with Turkey's post-hoc. HET, heterozygous *TTNtv*; HOM, homozygous *TTNtv*.

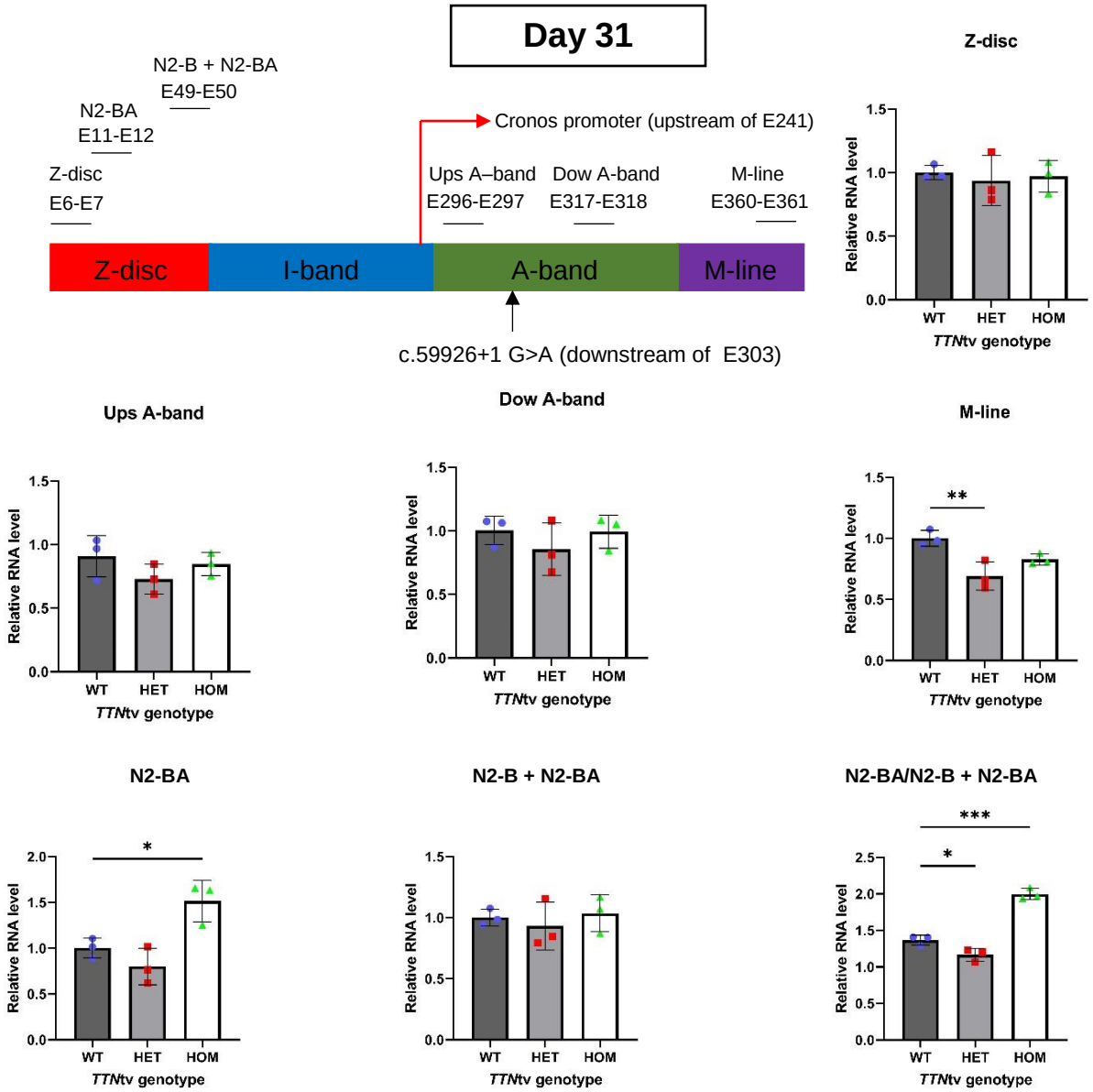


Figure 5-11 *TTN* transcriptional levels of day 31 *TTN* c.59926+1 G>A KOLF2-C1 iPSC-CMs. PCR probes were designed to be specifically spanning different pairs of exons responsible to the *TTN* regions; Z-disc (exons 6-7), A-band upstream of the *TTN*tv site (exons 296-297), A-band downstream of the *TTN*tv site (exons 317-318), and M-line (exons 360-361). The N2-BA probes (exons 11-12) were localised within the Z-disc region. The probes spanning exons 49-50 were designed to target both cardiac specific N2-B and N2-BA isoforms. Each dataset was obtained from different wells of differentiation using a single iPSC batch. All quantifications were made relative to the WT control, except for the N2-BA/N2-B + N2-BA ratio. n=3, mean ± S.D., one-way ANOVA with Turkey's post-hoc. HET, heterozygous *TTNtv*; HOM, homozygous *TTNtv*.

5.3.1.2 Effect of *TTNtv* on *TTN* exon splicing

Day 31 cDNAs were also used to investigate the consequence of *TTN* c.59926+1 G>A on *TTN* exon 303-304 splicing. Reverse transcription PCR using the primers hTTNF6/F6 spanning exon 303-304 identified a different species of a *TTN* transcript from the cells that harboured the *TTNtv*. This mutant *TTN* species was slightly higher in molecular weight as compared to the WT *TTN* species (552 bp) (Figure 5-12A&B). The homozygous line exclusively expressed the mutant *TTN* species while the heterozygous line expressed approximately half WT and half mutant *TTN* species without affecting the overall *TTN* level (Figure 5-12C). The equivalent expression of the two *TTN* species in the heterozygous line indicated that there was no allelic imbalance observed and thus NMD may not play a role in this titinopathy.

The PCR band of the mutant *TTN* transcript was also sequenced. The presence of *TTN* c.59926+1 G>A retained the whole intron lying immediately downstream of exon 303 in the transcript (Figure 5-13). The inclusion of the intron was predicted to introduce a premature stop codon, thereby truncating the titin protein from the early A-band region.

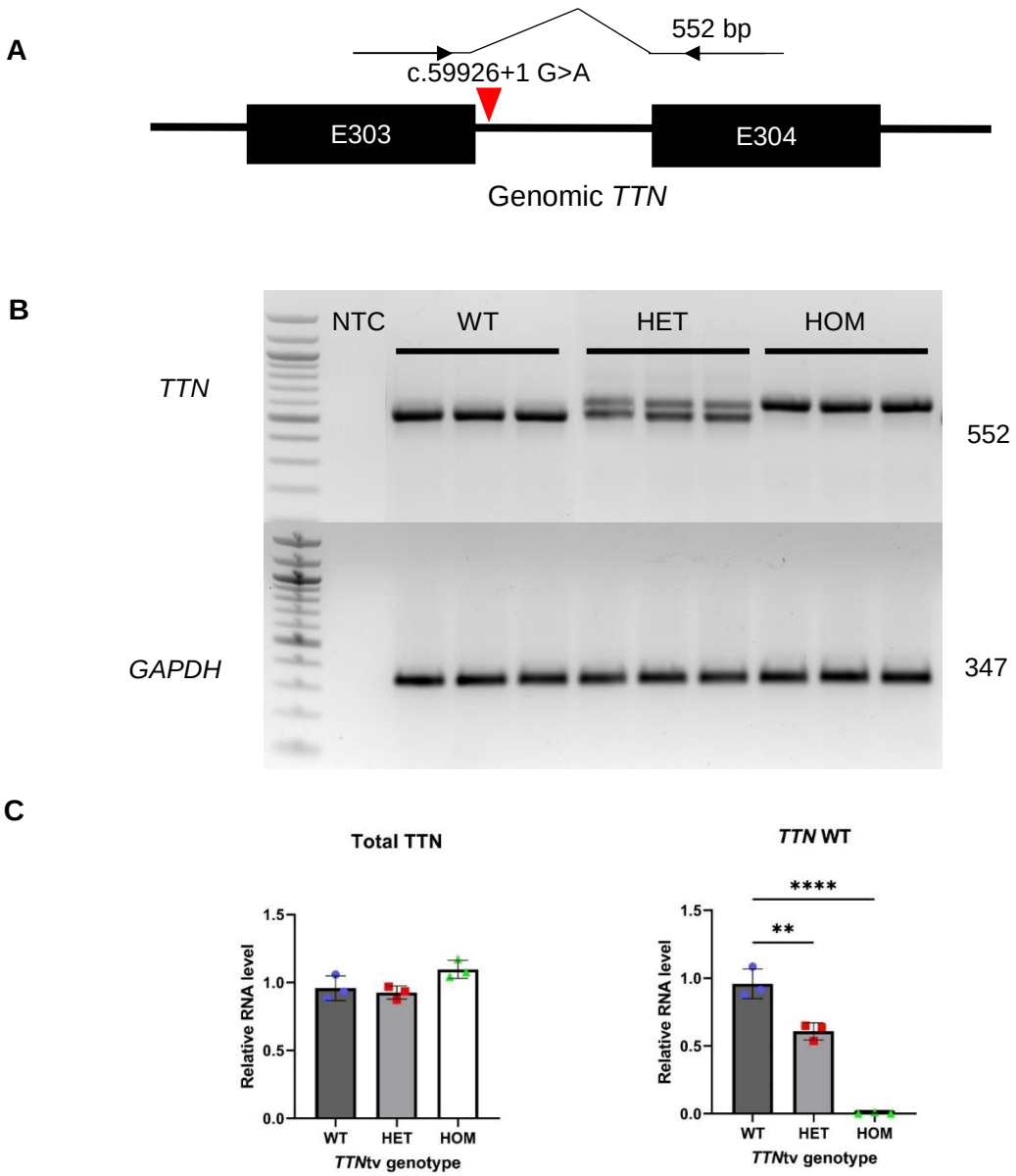


Figure 5-12 *TTN* c.59926+1 G>A caused changed *TTN* transcript species in KOLF2-C1 iPSC-CMs. A) Simplified genomic map over the *TTN*tv site illustrating PCR probes spanning exon 303-304. B) Agarose gels showing amplified cDNA products of day 21 WT *TTN*, heterozygous *TTNtv* (HET) and homozygous *TTNtv* (HOM) iPSC-CMs. The HOM line produced a higher molecular weight *TTN* transcript (mutant, top band) than those of the WT *TTN* transcript (552 bp, lower band) whilst the HET line expressed both the WT and the mutant *TTN* transcripts. C) Quantification of *TTN* transcriptional levels from B as normalised to *GAPDH*. Total *TTN* was a summation of both the WT and mutant *TTN* species. All quantifications were made relative to the WT control. Each dataset was obtained from different wells of differentiation using a single iPSC batch. n=3, mean ± S.D, one-way ANOVA with Tukey's post-hoc.

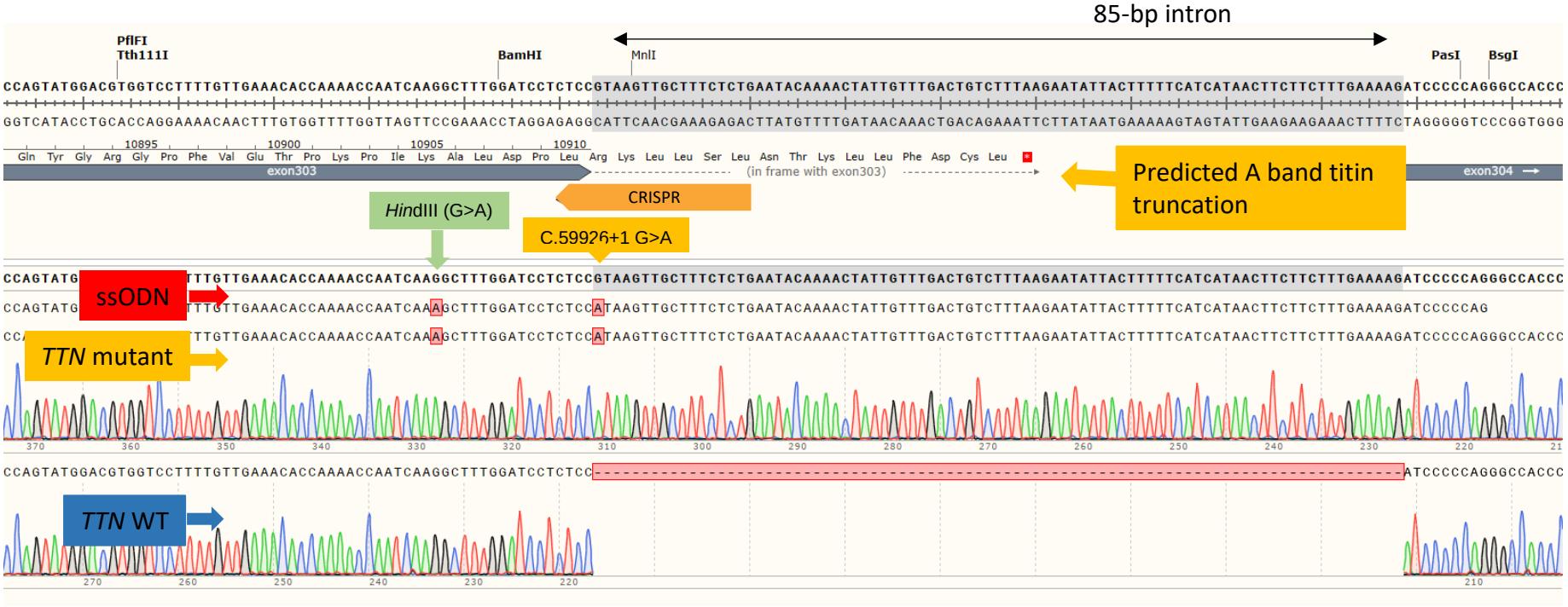


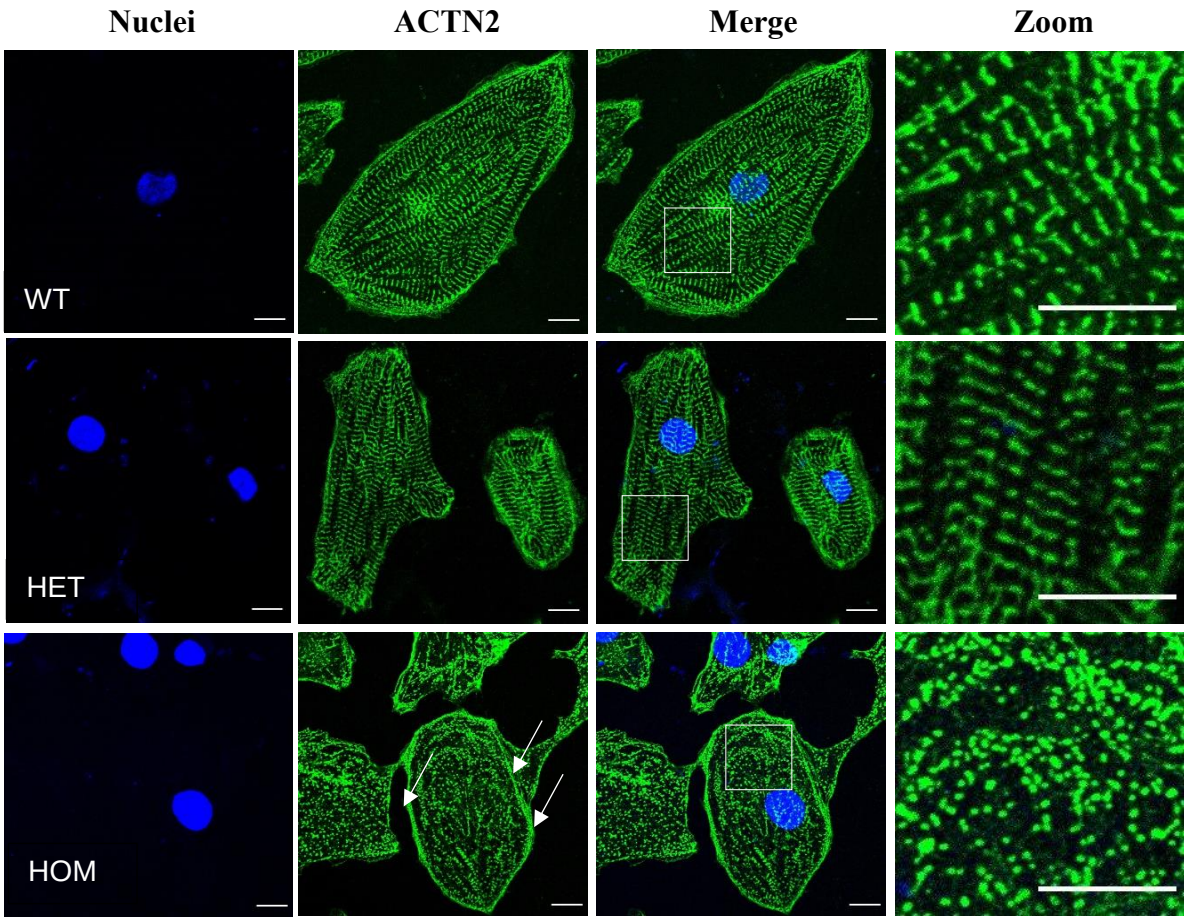
Figure 5-13 *TTN* c.59926+1 G>A caused intronic retention in KOLF2-C1 iPSC-CMs. cDNA sequence from the mutant *TTN* transcript contained additional 85 nt downstream of exon 303, which was predicted to mediate A-band truncation by a premature stop codon following additional 15 amino acids (R-K-L-L-S-L-N-T-K-L-L-F-D-C-L). The sequence was aligned against the WT *TTN* transcript and ssODN highlighting the additional G>A silent mutation for the *Hind*III site.

5.3.2 Functional and structural consequences of *TTNtv*

Both the heterozygous and homozygous lines were capable of being differentiated toward CMs. However, no spontaneous contraction was observed in the homozygous line. Meanwhile, the heterozygous line showed spontaneous contractile activity indistinguishable from the WT control with the onset of contraction approximately at day 9-11.

Further Z-disc staining on day 31 samples using ACTN2 antibodies suggested no intact sarcomere structure formed in the homozygous line (Figure 5-14A). Occasionally, there was enriched staining detected at the sarcolemma. Meanwhile, there was no distinction in the Z-disc striation between the heterozygous and the WT lines. Quantification of the cell size based on this ACTN2 staining found a significant cell area reduction in the homozygous line but such change was not detected in the heterozygous line (Figure 5-14B). The observed hypotrophy may be associated with the lack of a proper myofibrillar alignment or immature phenotype of iPSC-CMs (Jiang et al., 2018).

A



B

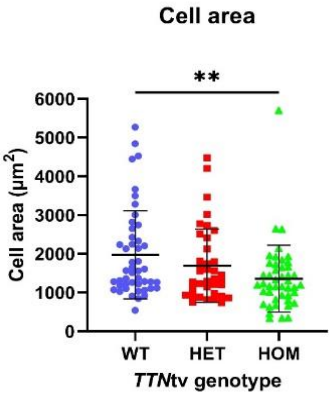


Figure 5-14 *TTN* c.59926+1 G>A impaired sarcomere formation in KOLF2-C1 iPSC-CMs. A) Representative sarcomeric ACTN2 IF images (green) of day 30 WT *TTN*, heterozygous *TTN*tv (HET), and homozygous *TTNtv* (HOM) iPSC-CMs counter-stained with nuclear-specific DAPI (blue). Zoom are digital zoom images from the square area destined in Merge. The HET line formed a relatively normal sarcomere. The HOM line had no sarcomere but peri-membranous ACTN2 localisation (arrow). Scale bar 10 µm. B) Quantification of cell area of ACTN2 immuno-positive iPSC-CMs. The HET line showed a relatively normal cell size compared to WT control while the HOM line had a significantly reduced cell area. Each dataset was obtained from different cells from a single batch of differentiation. n WT=80, n HET=80, n HOM=80, mean ± S.D., one-way ANOVA with Tukey's post-hoc.

5.4 Discussion

The study of titinopathy is challenging due to the gigantic size and isoform diversity of titin. The conventional approaches to study *TTNtv*, the most common cause of DCM, have relied on the availability of tissue biopsy, which is limited, or using animal models, which requires ethical approval and may not faithfully recapitulate the human biology. Thanks to the advance in iPSC technology, patient somatic cells can be reprogrammed into iPSCs from which CMs bearing the disease causing mutation can be differentiated and studied. Alternatively, pathogenic *TTNtv* mutations can be introduced into control iPSCs derived from healthy individual(s) using genome engineering tools including CRISPR/Cas9.

In this chapter, heterozygous and homozygous *TTN* c.59926+1 G>A KOLFC2-C1 iPSC lines were engineered using CRISPR/Cas9. Notably, the homozygous status described here was based on genotyping using probes amplifying ~500 bp over the CRISPR target site, which may not able to detect a CRISPR-related large deletion event (Kosicki et al., 2018). Thus, the homozygous line could have either two targeted alleles or one targeted allele with a large deletion on the non-targeted allele. Genotyping by a long range PCR, as described in Chapter 3, would further clarify the zygosity (Kosicki et al., 2018). Nevertheless, the *TTNtv* mutation prevented RNA splicing between exons 303 and 304 and caused intronic retention, generating a premature stop in the *TTN* transcript. This mutant transcript was predicted to bring about an A-band titin truncation. The mutation did not absolutely abolish CM differentiation even in the homozygous line but rather disrupted sarcomerogenesis, contributing to the lack of spontaneous contraction. In addition, the homozygous *TTNtv* mutation appeared to be associated with retained CM immaturity as suggested by the upregulation of foetal isoforms *MYH6* and *TTN* N2-BA and the downregulation of *NPPA* and *NPPB*.

There were no prominent phenotypes observed in the heterozygous line. Both the WT and the mutant *TTN* transcripts were equally expressed without changing the total *TTN* transcription levels. Haploinsufficiency by NMD mechanism thus may not be the main pathomechanism of this specific titinopathy. There were no obvious alterations in sarcomere structure,

spontaneous contraction capability and molecular remodelling. Since iPSC-CMs behave more like immature CMs, it is possibly that the observations in this chapter may not completely represent the phenotype in the patient's adult cells. Modelling at a more mature stage may enable better understanding of this disease-causing mutation. Moreover, the WT allele may be capable of fully rescuing the functional performance in the optimal culture condition. There may be (unexplored) hidden phenotypes which become explicit only upon pharmacological or electromechanical challenges.

5.4.1 *TTNtv* pathogenesis

Although *TTNtv*, especially A-band truncation, have been identified as the major genetic causes of DCM (Roberts et al., 2015, Herman et al., 2012), the mechanisms by which they modulate DCM phenotype are debatable. The previous failures to detect truncated titin proteins in affected DCM hearts had suggested titin haploinsufficiency as the disease mechanism. However, quantification of titin protein has been proven challenging due to its gigantic size and susceptibility to degradation. In addition, titin electrophoresis needs to be performed in a specialised low percent acrylamide gel or agarose gel at low voltage for several hours (Jiang et al., 2021, Vikhorev et al., 2017, Schafer et al., 2017). An investigation of titin protein in isolated human heart tissues using gel electrophoresis by (Vikhorev et al., 2017) did not show evidence of haploinsufficiency. In this work, there was no significant titin reduction (normalised to MYBPC3) in *TTNtv*-positive DCM samples as compared to *TTNtv*-negative DCM samples or healthy donor samples in both isolated myofibrillar fraction and whole tissue. This was in agreement with titin protein gel quantification normalised to myosin heavy chain by (Schafer et al., 2017) in a rat model. However, (Schafer et al., 2017) believed that titin haploinsufficiency happened in the modelled animal but the protein deficit was beyond the detection limit. This notion was based on the findings of the absence of truncated proteins coupled with a significant reduction in ribosomal occupancy on the WT allele. Nevertheless, (Fomin et al., 2021) illustrated a significantly decreased titin content quantified from a protein gel in *TTNtv*-positive DCM samples as compared to *TTNtv*-negative DCM samples or healthy

donor samples but this was made relative to a summation of proteins other than titin, suggesting overall reduction in sarcomere proteins. The ability to detect titin haploinsufficiency in this work suggests that improved titin gel electrophoresis and protein quantification methods would be advantageous. Due to time constraints, I was unable to apply the technique for titin electrophoresis so, intriguingly, it remains to be to elucidated if *TTN* c.59926+1 G>A also leads to loss of titin protein in iPSC-CMs.

One of the explanations for titin haploinsufficiency is NMD. It is postulated that the truncated titin provides a substrate for the NMD machinery which targets the mutant titin transcripts for degradation. In this chapter, the heterozygous *TTNtv* iPSC-CMs expressed a comparable level of the WT and mutant *TTN* species, confounding NMD. This finding was based on a conventional semi-quantitative PCR and hence, a more sensitive method e.g. RNA-sequencing could be employed to confirm this. There is a bias on quantification of PCR band from an agarose gel as the smaller product (the WT *TTN* allele) tends to be more amplified by PCR compared to the bigger product (the mutant *TTN* allele). Nevertheless, similar to the finding from this chapter, most of the previous RNA-sequencing results agreed on allelic balance (or a mild imbalance) in both human tissue (Fomin et al., 2021, Roberts et al., 2015, McAfee et al., 2021) and iPSC-CMs carrying heterozygous *TTNtv* (Hinson et al., 2015) except the study in a rat model by (Schafer et al., 2017). This may emphasise the differences in disease phenotypes between human and animal models.

The second proposed mechanism states that a truncated titin protein is stably expressed and exerts negative effects on the CM function. A recent success in identifying truncated titin proteins in patient's cardiac tissues (McAfee et al., 2021), with the content increasing with disease severity (Fomin et al., 2021), supports this argument. The actual functional consequences of this mutant protein have not been determined, but it was shown to be associated with deregulated PQC, especially UPS impairment and aggregate formation (Fomin et al., 2021). This aspect of *TTN* c.59926+1 G>A phenotyping will be discussed further in Chapter 6.

5.4.2 Titin isoforms and diseases

Upregulation of the titin N2-BA isoform at the expense of the N2-B isoform has been found in patients with cardiac diseases (Nagueh et al., 2004, LeWinter and Granzier, 2013), including heart failure with preserved ejection fraction and DCM. N2-BA confers less passive tension thanks to additional extensible I-band exon incorporation. The switching towards the compliant titin isoform represents a compensatory adaptation to cope with the increased stiffness of the extracellular matrix (Nagueh et al., 2004).

Nevertheless, there is no concrete supporting evidence for the association of *TTNtv* with titin isoform switching. Investigations using titin gels or WB in both patient-derived cardiac tissue (Vikhorev et al., 2017, Roberts et al., 2015) and iPSC-CMs (Hinson et al., 2015) showed no significant increase in titin N2-BA. This is consistent with the observed phenotype in the heterozygous model in this chapter although it was analysed by a different technique (qPCR using primers spanning the N2-BA isoform-specific exons 11-12). However, work by (Schafer et al., 2017) has contradicted this notion by illustrating an increased I-band exon usage in a rat model and human cardiac tissue by RNA-sequencing and ribosome profiling. The difference may be explained by the different sensitivity of each analysis tool.

5.4.3 *TTNtv* and CM development

N2-BA is the major titin isoform in foetal CMs and immature iPSC-CMs/ESC-CMs (Hinson et al., 2015, Lahmers et al., 2004).The slight upregulation of *TTN* N2-BA in the homozygous *TTN* c.59926+1 G>A iPSC-CMs may be indicative of restricted early CM development, supported by the upregulation of *MYH6* and downregulation of *MYH7* together with the downregulation of *NPPA* and *NPPB*. However, there were no such changes in the heterozygous model.

MYH6 is the predominant myosin heavy chain isoform expressed during early development of the human heart (Gacita et al., 2021). In the late developmental stage, there is an isoform switching mechanism which leads to an increased expression of *MYH7* in the adult heart. Thus, the low *MYH7/MYH6* ratio in the homozygous *TTN* c.59926+1 G>A iPSC-CMs may

reflect a delayed myosin heavy chain switching, which resembles an embryonic-like, immature state (Guo and Pu, 2020). In iPSC-CMs, an increased *MYH7/MYH6* ratio is linked with HCM but such an association was not found in DCM models (Eschenhagen and Carrier, 2019)

NPPA and NPPB have roles on both cardiac development and diseases. They are expressed highly during early development in both atrial and ventricular CMs. *NPPA*, however, is strongly downregulated in the ventricles after birth and the expression becomes confined to the atria (Christoffels et al., 2000, Bloch et al., 1986). Reactivation of both genes in ventricular CMs is associated with hypertrophic and failing hearts and is thus considered as pathological and prognostic biomarkers (Chien et al., 1991, Rademaker and Richards, 2005)

Similar to the *TTNtv* model in this chapter, a downregulation of *NPPA* and *NPPB* expression and an increase in the *MYH7/MYH6* ratio were observed in heterozygous and homozygous *TTN* P22582fs iPSC-CMs in a dose dependent manner (Hinson et al., 2015). This report highlighted that the *TTNtv* was associated with diminished activation of growth factors (e.g. transforming growth factor β1 and vascular endothelial growth factor) and mitogen-activated protein kinases, as well as force production deficits, impaired mechanical load and β-adrenergic stimulation. Notably, the *TTNtv* P22582fs models were studied as an engineered heart tissue (3-dimentional format) which promotes iPSC-CM maturation. In contrast, the *TTN* c.59926+1 G>A iPSC-CMs in this study were cultured as a monolayer (2-dimentional format). It would be interesting to further investigate whether the heterozygous *TTN* c.59926+1 G>A phenotype(s) were hidden by the maturity status, which could be unmasked by CM maturation methods including 3-dimensional culture.

5.4.4 *TTNtv* disrupts sarcomerogenesis

The mechanism of sarcomerogenesis is very complex and incompletely understood. It is proposed that CMs begin myofibrillar assembly at the cell periphery (Figure 5-15). Here, premyofibrils containing non-muscle myosin II and ACTN2 provide templates for actin filament growth and recruitment of other sarcomeric proteins. Subsequently, the premyofibrils develop

into nascent myofibrils, which replace non-muscle myosin II with muscle-specific myosin II. At this stage, titin molecules are also inserted to form the Z-discs and stabilise the sarcomere structure. Finally, the individual myofibrils coalesce into organised sarcomere with the prominent Z-disc, I-band, A-band and M-line regions.

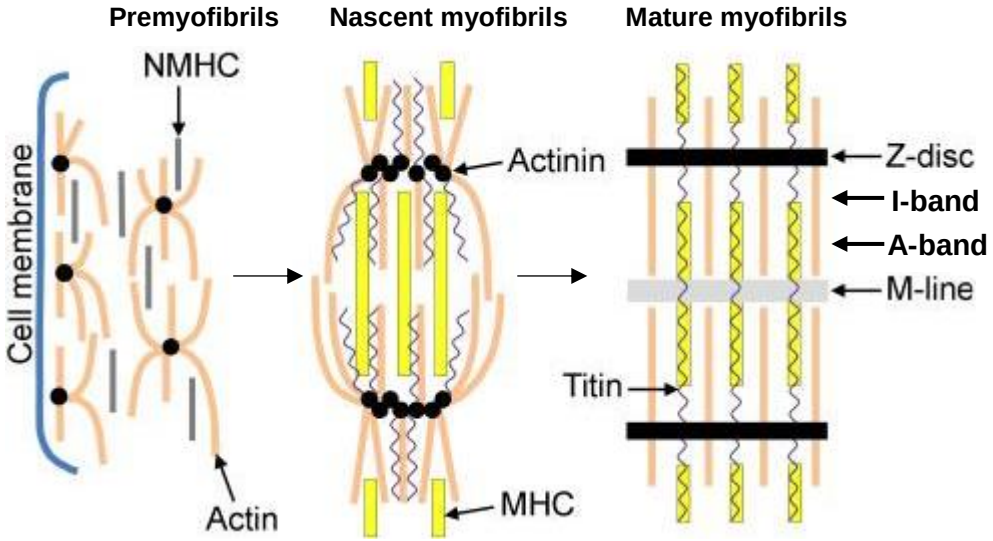


Figure 5-15 Simplified illustration of sarcomerogenesis in CMs. Premyofibrils, which consist of ACTN2 and actin filament intercalated with non-muscle myosin heavy chain (NMHC), assemble along the cell periphery. Subsequently, other sarcomere proteins including titin are recruited to the premyofibrils to form nascent myofribrils. At this stage, muscle-specific myosin heavy chain (MHC) substitutes NMHC. Finally, the individual nascent myobrifrils fuse and mature. Here, the 4 regions of the sarcomere structure (Z-disc, I-band, A-band and M-line) become distinguishable. Modified from (England and Loughna, 2013).

The complete disruption of sarcomere formation and loss of contractile activity in the homozygous *TTN* c.59926+1 G>A iPSC-CMs indicated an essential role of titin in sarcomerogenesis, though the presence of the truncated titin protein remains to be proven. (Chopra et al., 2018) suggested that titin coupled tension generated by MYH7 to protocostameres. Protocostameres are sites for cell-extracellular matrix adhesion which are required for initiation of the fusion of Z-disc precursors. The titin/MYH7 interaction induced sarcomerogenesis, which progressed from the cell periphery towards the cell interior (Chopra et al., 2018). The lack of the A-band and M-line regions of titin associated with the homozygous *TTN* c.59926+1 G>A mutation thus can remove the mechanical link required for initiating the sarcomere formation. The peri-membranous ACTN2 localisation in some cells likely indicated the remnant centripetal assembly of ACTN2-containing fibres which could not undergo

myofibrillogenesis because of the absence of the full-length titin. However, this does not explain the normal sarcomere structure observed in the heterozygous line. It is interesting, albeit challenging, to further investigate whether truncated titin proteins in the heterozygous model are stably expressed and interfere with the sarcomere assembly. The disorganised sarcomere structure with reduced sarcomere content seen in the heterozygous *TTN* P22582fs iPSC-CMs (Hinson et al., 2015, Chopra et al., 2018) was derived from patient-reprogrammed cells, which may have a different genetic and/or epigenetic background exacerbating sarcomere impairment.

5.4.5 Concluding remarks

TTN c.59926+1 G>A mutation was successfully integrated into KOLF2-C1 iPSCs. This mutation inhibited the normal exon 303-304 splicing, leading to intronic retention and A-band titin truncation. The heterozygous *TTN* c.59926+1 G>A iPSC-CMs showed no changes in sarcomere organisation and foetal gene expression (*NPPA*, *NPPB*, *TTN* N2-BA and *MYH6*) compared to the WT control. Meanwhile, investigations in the homozygous *TTN* c.59926+1 G>A iPSC-CMs suggested an embryonic/immature restriction (downregulation of *MYH7/MYH6*, *NPPA*, *NPPB*, upregulation of *TTN* N2-BA and CM hypotrophy) with completely disrupted sarcomerogenesis.

Chapter 6

Subtype-specific characterisation of *TTNtv* in *MYL7-mClover* iPSC-CMs

6.1 Introduction

TTNtv is associated with arrhythmic risks including familial atrial fibrillation (Ahlberg et al., 2018, KM et al., 2022, Haggerty et al., 2019, Andreasen et al., 2020, Olesen et al., 2018) and ventricular arrhythmias in DCM (Corden et al., 2019, Tayal et al., 2017). In Chapter 5, *TTN* c.59926+1 G>A was introduced into the parental KOLF2-C1 iPSCs and partly characterised in differentiated ventricular-like iPSC-CMs. In Chapter 4, an RA-induced atrial differentiation protocol was established in the *MYL7-mClover* reporter iPSC-CMs, which was engineered in Chapter 3. Since DCM patients carrying the heterozygous *TTN* c.59926+1 G>A mutation developed either atrial fibrillation or ventricular tachycardia (Hoorntje et al., 2018), it is intriguing to further investigate the *TTNtv* phenotypes in atrial like cells, which could provide new insight(s) into the arrhythmic aspect of titinopathy with respect to the CM subtype. Subtype-specific CM characterisation of *TTNtv* would exemplify how subtype-specific fluorescent reporter system(s) could be applied in the study of cardiac disease.

This chapter focuses primarily on exploring the utility of the atrial-specific *MYL7-mClover* reporter iPSC line in disease study. *TTN* c.59926+1 G>A was introduced into this reporter line and the resultant cell lines were differentiated via both ventricular and atrial pathways. Because the characterisation of the *MYL7-mClover* reporter line was not complete, where the question of whether the fluorescent reporter allows atrial-like cell isolation remains to be addressed (Chapter 4), the ventricular and atrial differentiation protocols were used as a means for sub-subtype study of *TTN* c.59926+1 G>A instead of relying the fluorescent reporter.

Specifically, this chapter aims to achieve the following:

I. Using CRISPR/Cas9 technology to incorporate *TTN* c.59926+1 G>A into the *MYL7-mClover* iPSC line.

II. To characterise the *TTN* c.59926+1 G>A phenotype in ventricular-like *MYL7-mClover* iPSC-CMs. This includes a study of molecular remodelling (changes in expression of sarcomeric genes and heart failure markers), a CM structural investigation (sarcomere formation and sarcomeric protein localisation) and functional characterisations (measurement of Ca^{2+} transient and sarcomere contraction).

III. To characterise *TTN* c.59926+1 G>A phenotype in atrial-like *MYL7-mClover* iPSC-CMs. It includes the same investigations as for ventricular-like cells (see II)

6.2 CRISPR/Cas9-assisted generation of TTNtv in MYL7-mClover iPSC line

6.2.1 CRISPR and *TTN* c.59926+1 G>A donor design

TTN c.59926+1 CRISPR and *TTN* c.59926+1 G>A ssODN designs were as described in Chapter 5.

6.2.2 CRISPR/Cas9-assisted generation of *TTNtv* iPSC line

The CRISPR/Cas9 RNP targeting *TTN* c.59926+1 and the c.59926+1 G>A ssODN were transfected into the *MYL7-mClover* iPSCs by electroporation. 48 hr after transfection, the pool of gDNA was analysed for genomic incorporation of the ssODN by PCR using the primers hTTNF6/R3 followed by *Hind*III digestion (Figure 6-1A). Subsequent genotyping of isolated colonies found 33/91 (~34%) to be positive for the integration event, 31 (94%) of which showed incomplete digestion (Figure 6-1B), indicative of heterozygous editing. Only two clones (4B and 12E) were completely digested by the restriction enzyme, indicative of homozygous targeting (or heterozygous editing with a large deletion event on the non-targeted allele (see Discussion, Chapter 5)).

16 promising clones were clonally expanded and genotypically confirmed by *Hind*III digestion (Figure 6-2A). The clone 1C, 4B and 12E were further analysed by DNA sequencing (Figure 6-2B). It suggested that the clone 1C was heterozygous for both the *TTN* c.59926+1 G>A

(*TTNtv*) and c.59910 G>A (*Hind*III) sites while the other two clones were homozygous for the both sites.

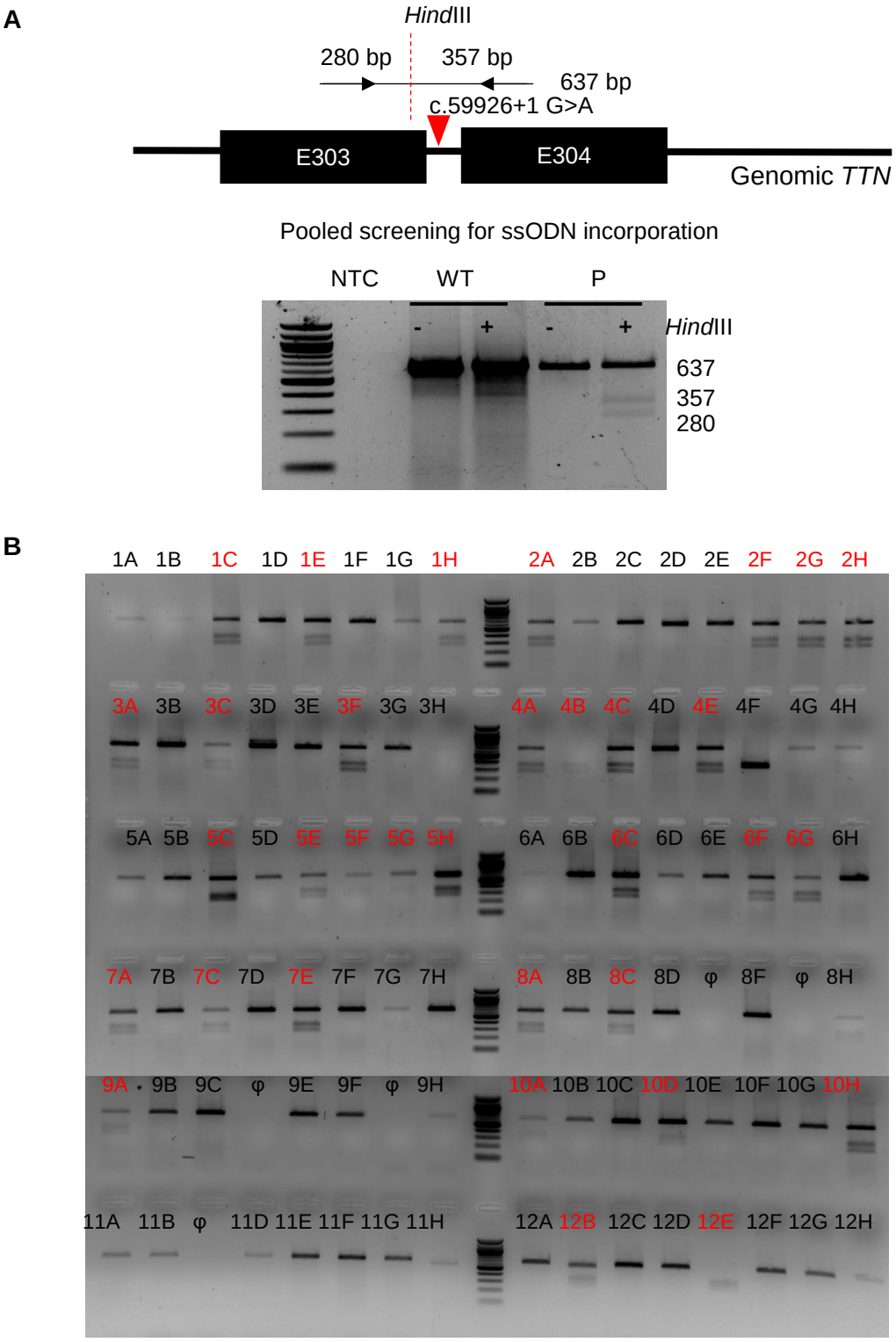


Figure 6-1 Genotypic screening for *TTN* c.59926+1 G>A in *MYL7-mClover* iPSCs. A) *Hind*III digestion screening of PCR products encompassing the genomic *TTN*tv site i.e. spanning *TTN* exons 303-304 in the pool of transfected iPSCs. Only PCR products acquiring the inserted ssODN donor were cleavable by *Hind*III (357 and 280 bp). B) *Hind*III digestion screening in isolated iPSC colonies. NTC, no template control; P, pool of electroporated cells.

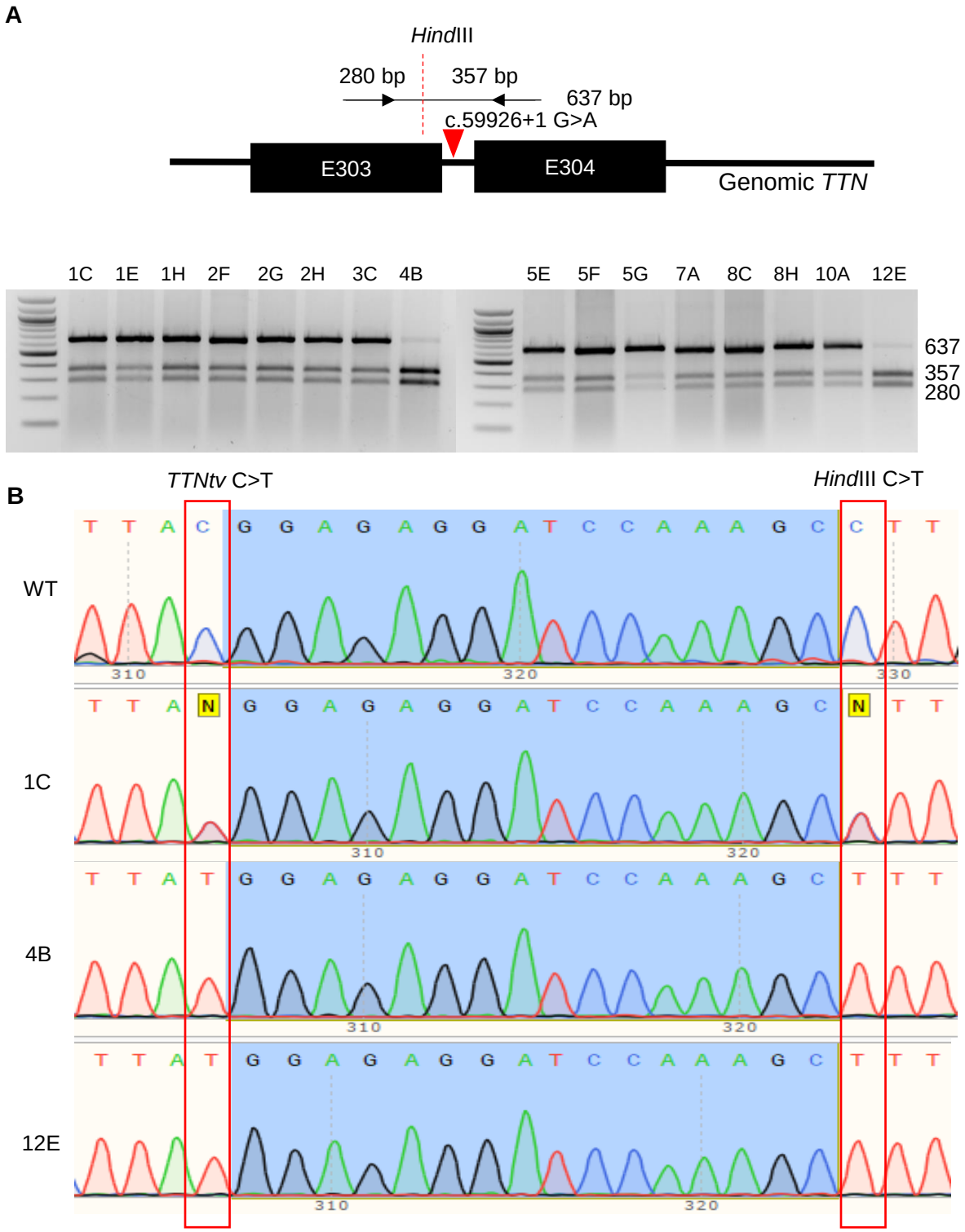
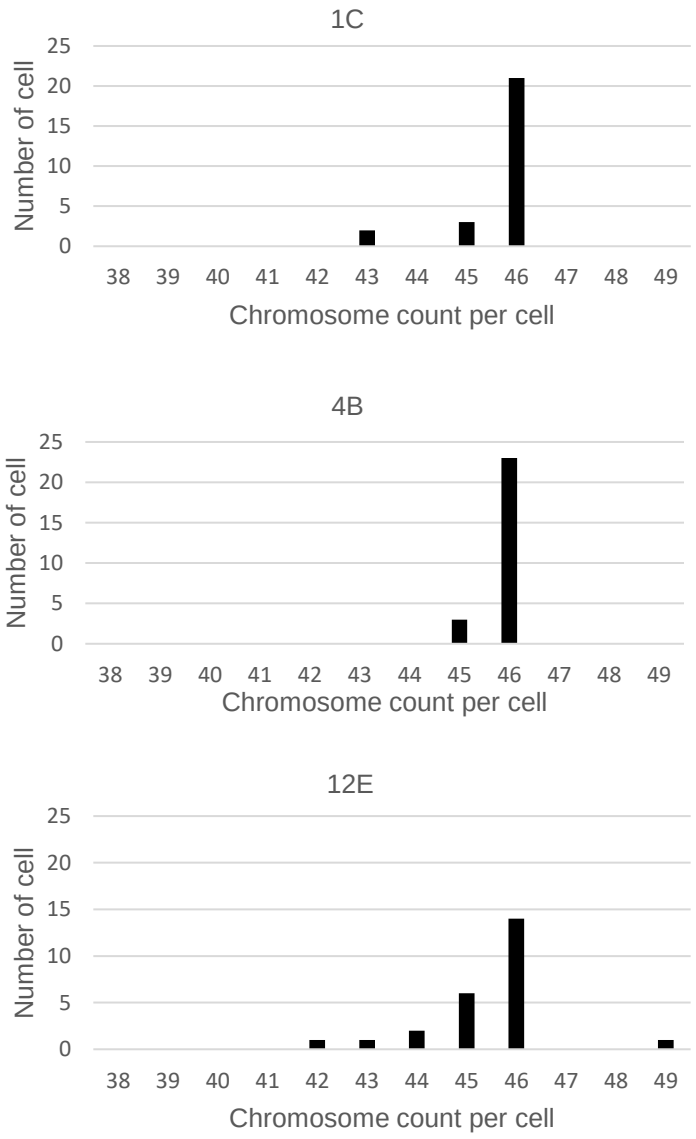


Figure 6-2 Genotypic confirmation for *TTN* c.59926+1 G>A in *MYL7-mClover* iPSCs. A) *Hind*III digestion on PCR products encompassing the genomic *TTN*tv site i.e. spanning *TTN* exon 303-304 in isolated iPSC colonies. Only PCR products acquiring the ssODN donor were cleavable by *Hind*III (357 and 280 bp). B) Sequencing results of selected clones from A. The clones 4B and 12E were homozygous for both the targeted *Hind*III and *TTN*tv sites whilst the clone 1C was heterozygous for both the *Hind*III and *TTN*tv sites.

6.2.3 iPSC quality checks

The clones 1C and 4B had a normal karyotype (Figure 6-3A) and showed retained expression of the pluripotency markers (NANOG, OCT3/4, SOX2 and SSEA4) (Figure 6-3B). Meanwhile the clone 12E had an aneuploid subpopulation with 49 chromosome (1/25) despite showing good expression of the pluripotency markers; hence, this clone was disqualified from being a normal iPSC line. The clones 1C, which was later called the heterozygous *TTNtv* line, and 4B, which was later called the homozygous *TTNtv* line, were selected and characterised in the subsequent experiments.

A



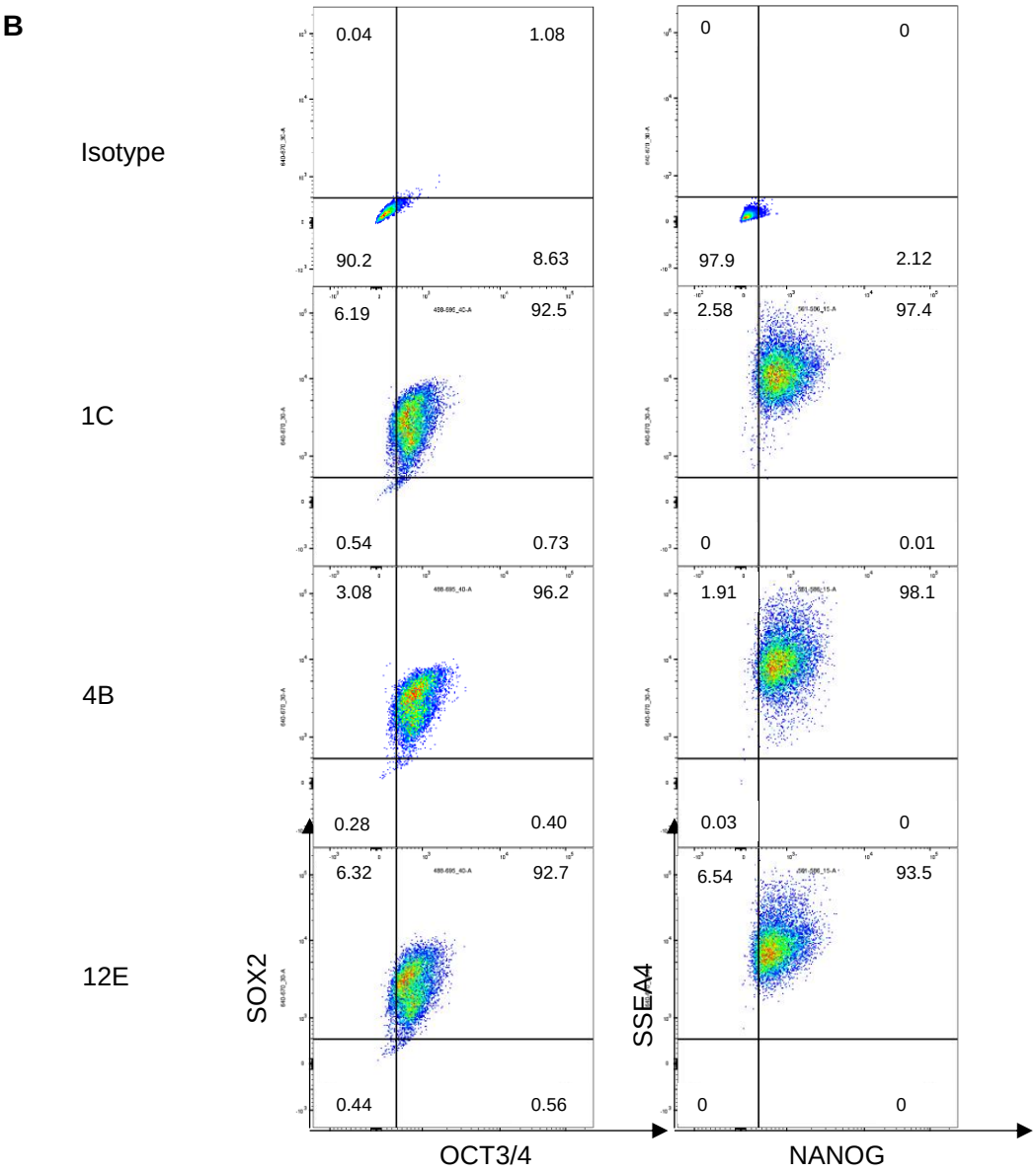


Figure 6-3 *MYL7-mClover, TTN* c.59926+1 G>A iPSC line's quality control. A) Distribution of chromosome count per cell. B) Flow cytometry analysis of pluripotency marker expression.

6.3 Characterisation of TTNtv in ventricular-like MYL7-mClover iPSC-CMs

Both the heterozygous and homozygous *TTNtv* lines were differentiated into CMs using the conventional (ventricular) protocol as described in Chapter 2.

6.3.1 Molecular analysis

6.3.1.1 Effect of *TTN*tv on transcriptional expression of cardiac genes

RNA expression level was calculated relative to the WT line for a number of genes including *TTN*, other sarcomeric markers and heart failure markers at day 31 of differentiation. (Figure 6-4 and Figure 6-5).

No obvious *TTN* transcriptional changes were observed in both the *TTNtv* heterozygous and homozygous lines (Figure 6-4). The unaltered *TTN* Z-disc expression indicated that the expression of all *TTN* isoforms except the Cronos, were not influenced by *TTN* c.59926+1 G>A. The Cronos *TTN* expression was also likely unaffected as highlighted by the unaltered *TTN* A-band upstream and downstream of the *TTNtv* site and the unaltered *TTN* M-band expression. The slightly decreased expression of the *TTN* A-band upstream of the *TTNtv* site and the *TTN* M-line in the heterozygous line and the slightly decreased expression of the *TTN* M-band in the homozygous line likely reflected the primer binding efficiency (not quantified) rather than an indication of downregulation of the Cronos isoform since there was no support from changes in expression of the *TTN* A-band downstream of the *TTNtv* site. The *TTN* long isoform N2-BA expression (or N2-BA/N2-B + N2-BA ratio) was also relatively normal in both the heterozygous and homozygous lines

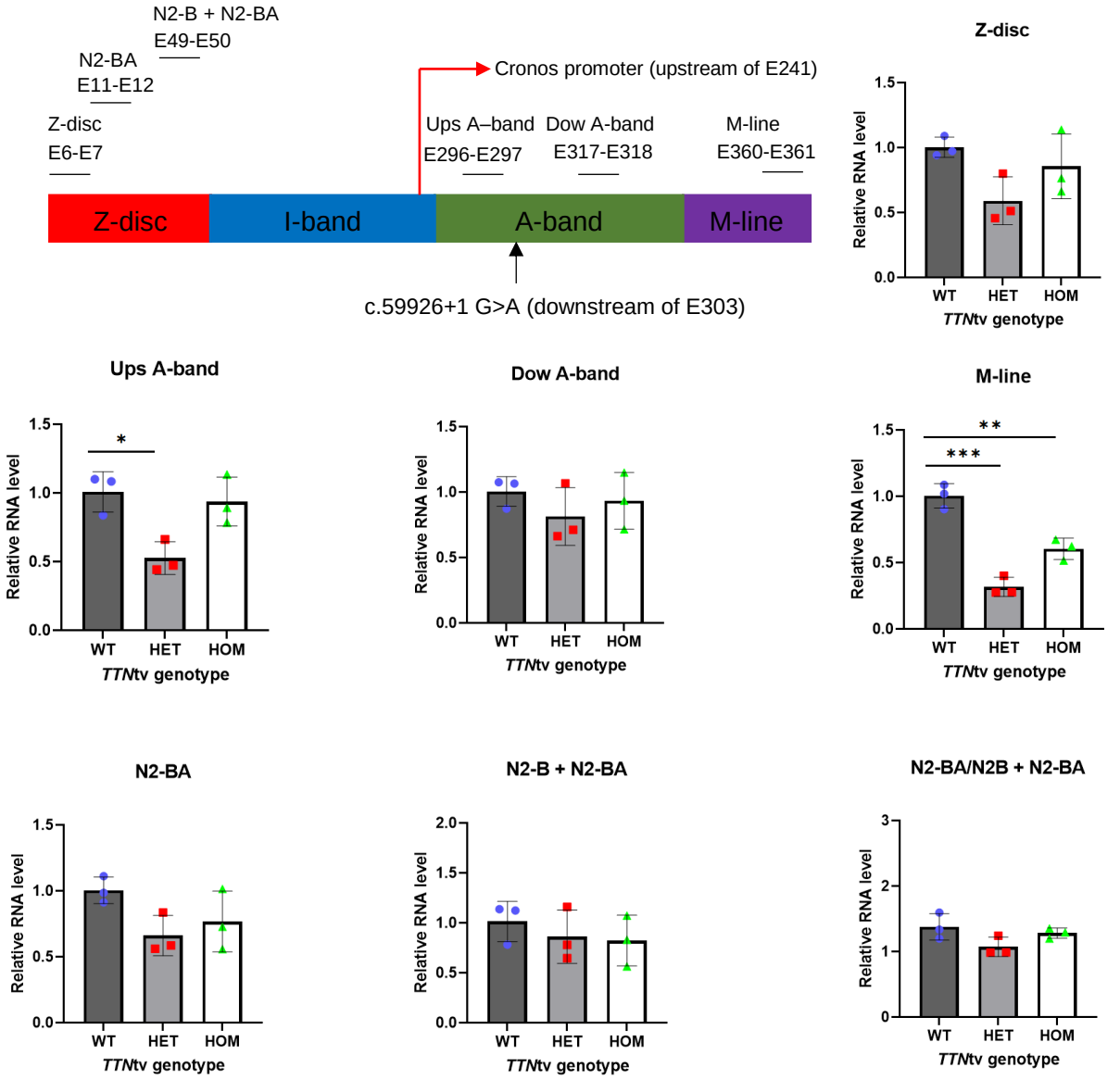
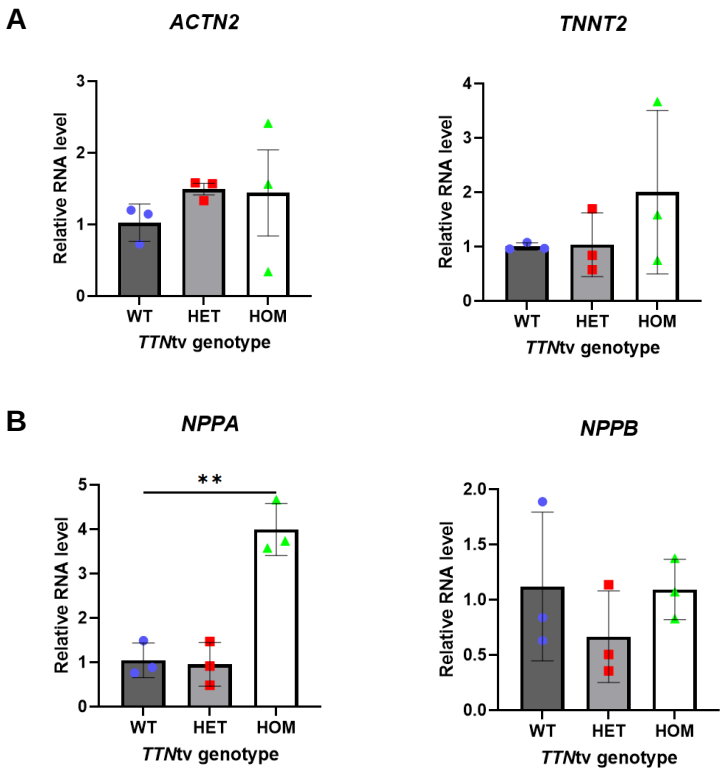


Figure 6-4 *TTN* transcriptional levels of day 31 ventricular-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A. PCR probes were designed to be specifically spanning different pairs of exons responsible to the *TTN* regions; Z-disc (exons 6-7), A-band upstream of the *TTN*tv site (exons 296-297), A-band downstream of the *TTN*tv site (exons 317-318), and M-line (exons 360-361). The N2-BA probes (exons 11-12) were localised within the Z-disc region. The probes spanning exons 49-50 were designed to target both cardiac specific N2-B and N2-BA isoforms. Each dataset was obtained from different wells of differentiation using a single iPSC batch. All quantifications were made relative to the WT control, except for the N2-BA/N-2B + N2-BA ratio. n=3, mean ± S.D., one-way ANOVA with Turkey's post-hoc. HET, heterozygous *TTNtv*; HOM, homozygous *TTNtv*.

No significant changes in expression of pancardiac markers (*ACTN2* and *TNNT2*) were detected in both the heterozygous and homozygous lines, suggesting retained cardiac differentiation potential (Figure 6-5A). Also, there were no obvious signs of molecular remodelling with respect to heart failure-associated foetal genes (*NPPA* and *NPPB*) revealed in both the heterozygous and homozygous lines (Figure 6-5B). The observed upregulation in *MYH7* (or *MYH7/MYH6* ratio) in the heterozygous line and the observed upregulation in *NPPA* in the homozygous line were interpreted as a consequence of the variation of gene expression in iPSC-CMs. In the homozygous line, there were a significant upregulation of *MYH6* (Figure 6-5C) and a significant downregulation of *MYL2* (or *MYL2/MYL7* ratio), indicative of impaired CM development. The expression of *mClover* was slightly upregulated in the heterozygous line (Figure 6-5D).



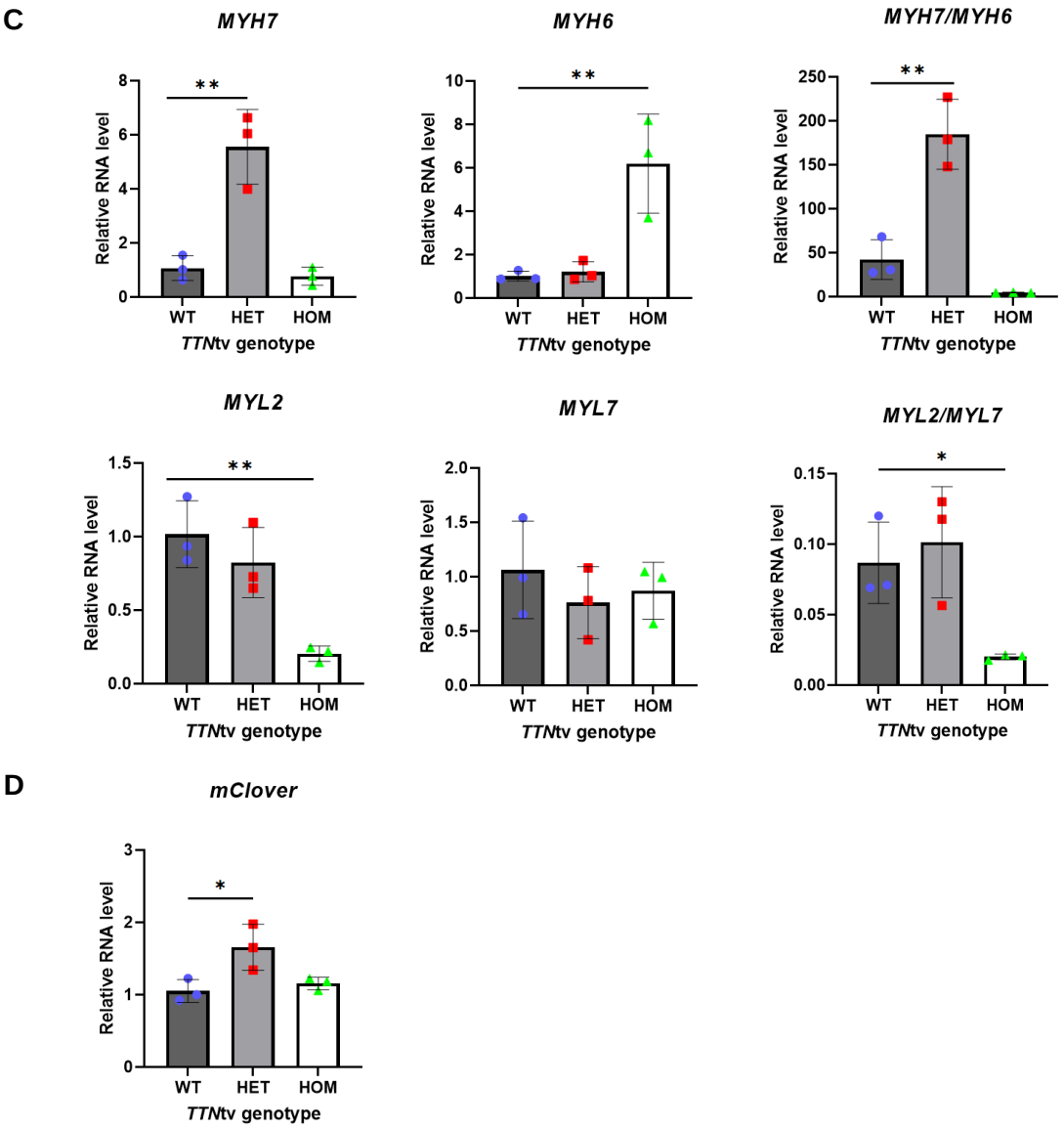


Figure 6-5 Transcriptional expression profile of day 31 ventricular-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A. A) Pan-cardiac markers: *ACTN2* and *TNNT2*. B) Heart failure markers *NPPA* and *NPPB*. C) Myosin heavy chain (*MYH6* and *MYH7)* and light chain (*MYL2* and *MYL7*) isoforms. D) *mClover*. Each dataset was obtained from different wells of differentiation using a single iPSC batch. All quantifications were made relative to the WT control, except for the *MYH6/MYH7* and *MYL2/MYL7* ratios. n=3, mean ± S.D., one-way ANOVA with Turkey's post-hoc. HET, heterozygous *TTNtv*; HOM, homozygous *TTNtv*.

6.3.1.2 Effect of *TTNtv* on *TTN* splicing

To establish whether the effect of *TTN* c.59926+1 G>A on *TTN* splicing observed in the parental KOLF2-C1 line (Chapter 5) was preserved in the *MYL7-mClover* line, day 31 cDNAs were analysed by PCR using the primers hTTNF6/R6 (Figure 6-6A). As expected, the heterozygous line produced two PCR products, one with a correct size corresponding to the WT species (552 bp) and one with a slightly higher molecular weight (Figure 6-6B). Meanwhile, the mutant *TTN* species was exclusively present in the homozygous line. Quantification of the PCR products suggested that there was no change in overall *TTN* expression at the RNA level between the *TTNtv* genotypes. Also, the WT *TTN* species was expressed at approximately 50% of total *TTN* level in the heterozygous line (Figure 6-6C), suggesting that NMD is not a molecular mechanism underlying this titinopathy.

The PCR product of the mutant *TTN* transcript from the homozygous line was also clarified by DNA sequencing (Figure 6-7). *TTN* c.59926+1 G>A retained the addition 85 bp intronic region immediately downstream of exon 303 and as previously discussed Chapter 5, was predicted to be causative of A-band truncation.

Overall, the *TTN* c.59926+1 G>A effect on *TTN* splicing was consistent between the *MYL7-mClover* reporter and parental iPSC-CM lines. Thus, as expected, the integration of the fluorophore at the *MYL7* locus may not inherently hinder *TTNtv* c.59926+1 G>A phenotyping with respect to *TTN* expression/splicing.

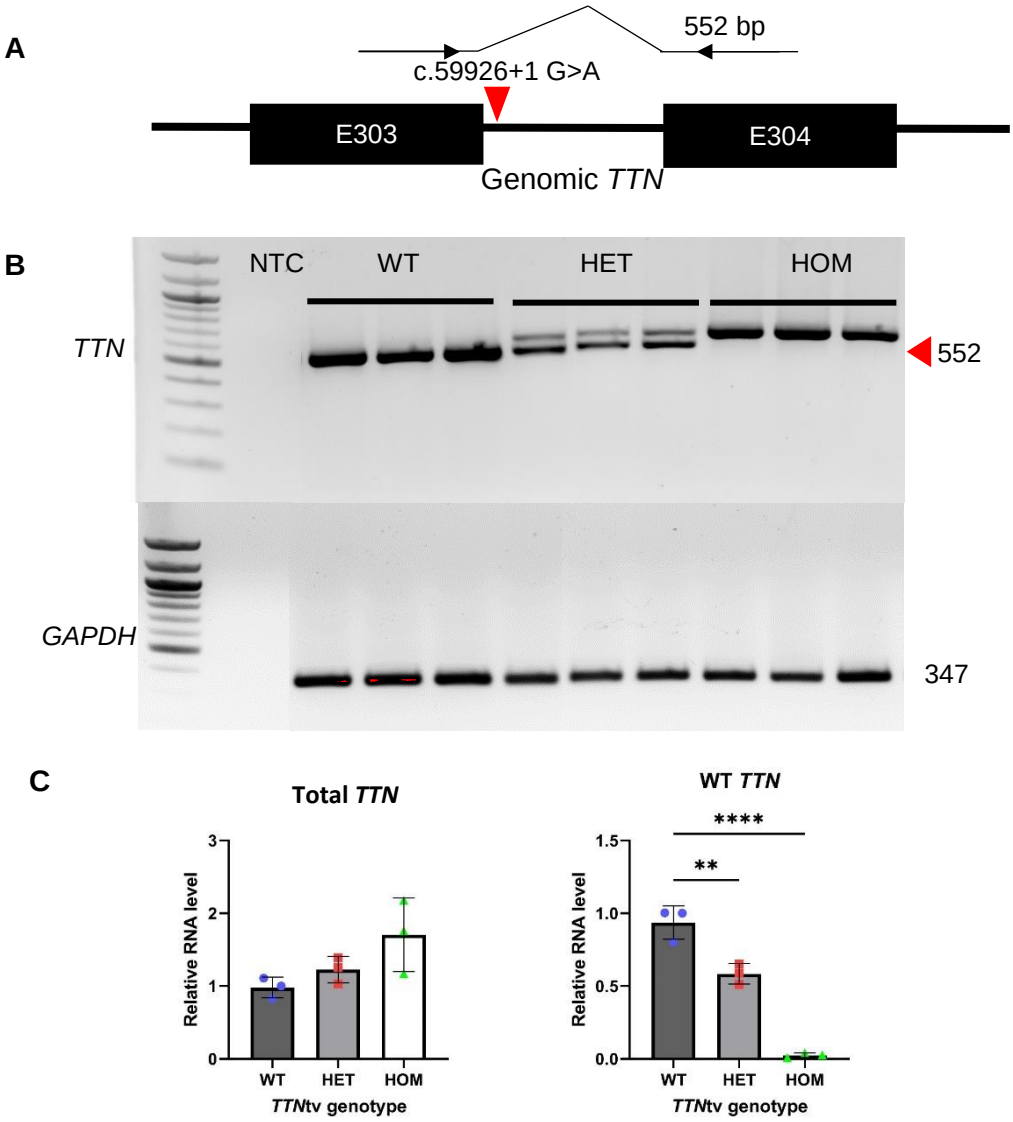


Figure 6-6 *TTN* c.59926+1 G>A caused changed *TTN* transcript species in ventricular-like *MYL7-mClover* iPSC-CMs. A) Simplified genomic map over the *TTN*tv site and PCR probes spanning exon 303-304. B) Agarose gels showing PCR products of day 31 iPSC-CMs. The homozygous *TTNtv* (HOM) line produced a higher molecular weight *TTN* transcript (mutant, top band) than those of the WT species (552 bp, lower band) whilst the heterozygous *TTNtv* (HET) line expressed both the WT and the mutant *TTN* transcripts. C) Quantification of *TTN* transcriptional levels based on B as normalised to *GAPDH* and as relative to the WT control. Total *TTN* was a summation of both the WT and the mutant species. Each dataset was obtained from different wells of differentiation using a single iPSC batch. n=3, mean ± S.D., one-way ANOVA with Tukey's post-hoc.

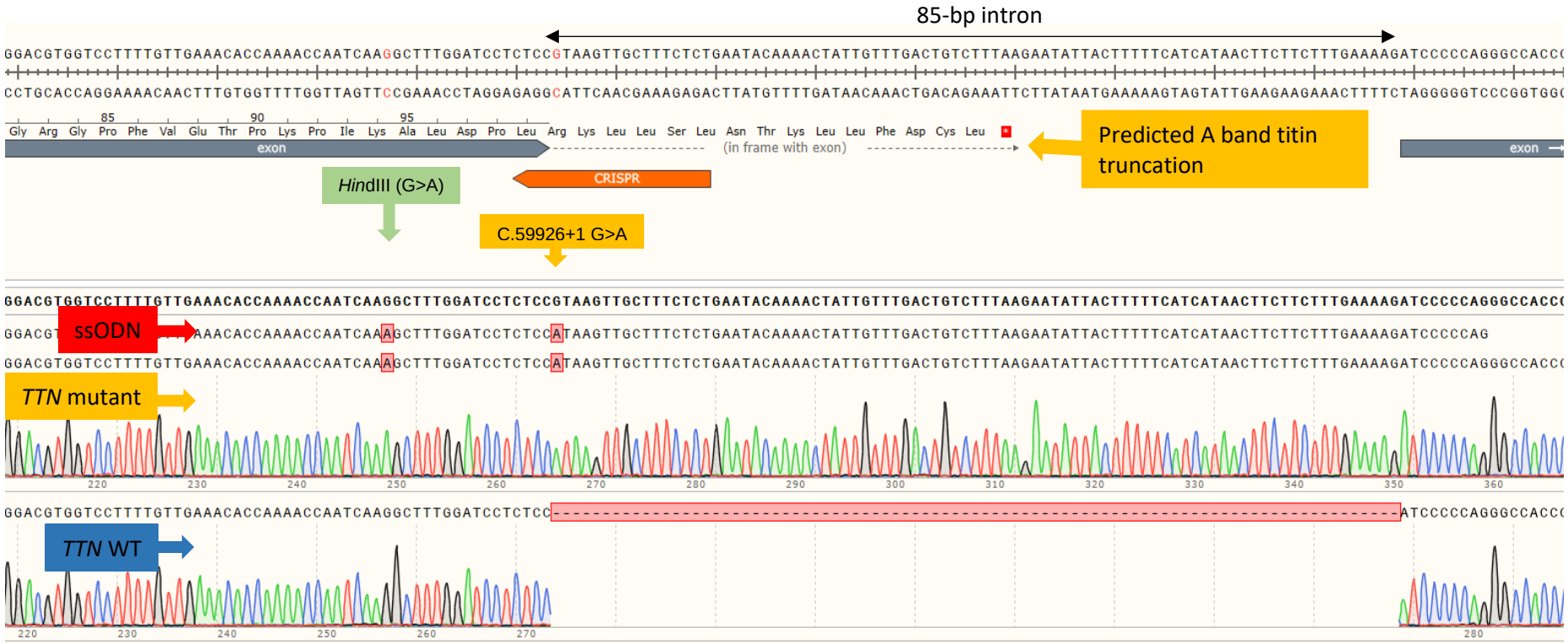


Figure 6-7 *TTN* c.59926+1 G>A caused intronic retention in ventricular-like *MYL7-mClover* iPSC-CMs. cDNA sequence from the mutant titin transcript contained additional 85 nt immediately downstream of exon 303, which was predicted to mediate A-band titin truncation by a premature stop codon following additional 15 amino acids (R-K-L-L-S-L-N-T-K-L-L-F-D-C-L). The sequence was aligned against the WT *TTN* transcript and ssODN highlighting the additional G>A silent mutation creating the *Hind*III site.

6.3.2 Functional analysis

6.3.2.1 Effect of *TTNtv* on spontaneous contraction

The *TTNtv* heterozygous line retained normal contractile activity comparable to the WT control line, showing an onset of spontaneous contraction at approximately day 9-11 (Figure 6-8). However, the homozygous line displayed no visible spontaneous contraction. These observations repeated the findings in the parental KOLF2-C1 line (Chapter 5), thus indicating no inhibitory effect of the *MYL7-mClover* fluorescent reporter on *TTNtv* phenotyping with respect to the spontaneous contraction.

On day 30 of differentiation, no difference in live cells' mClover fluorescence was apparent between the *TTNtv* genotypes (Figure 6-8), suggesting that the (*MYL7*-driven) mClover expression was not influenced by the presence of the *TTNtv*.

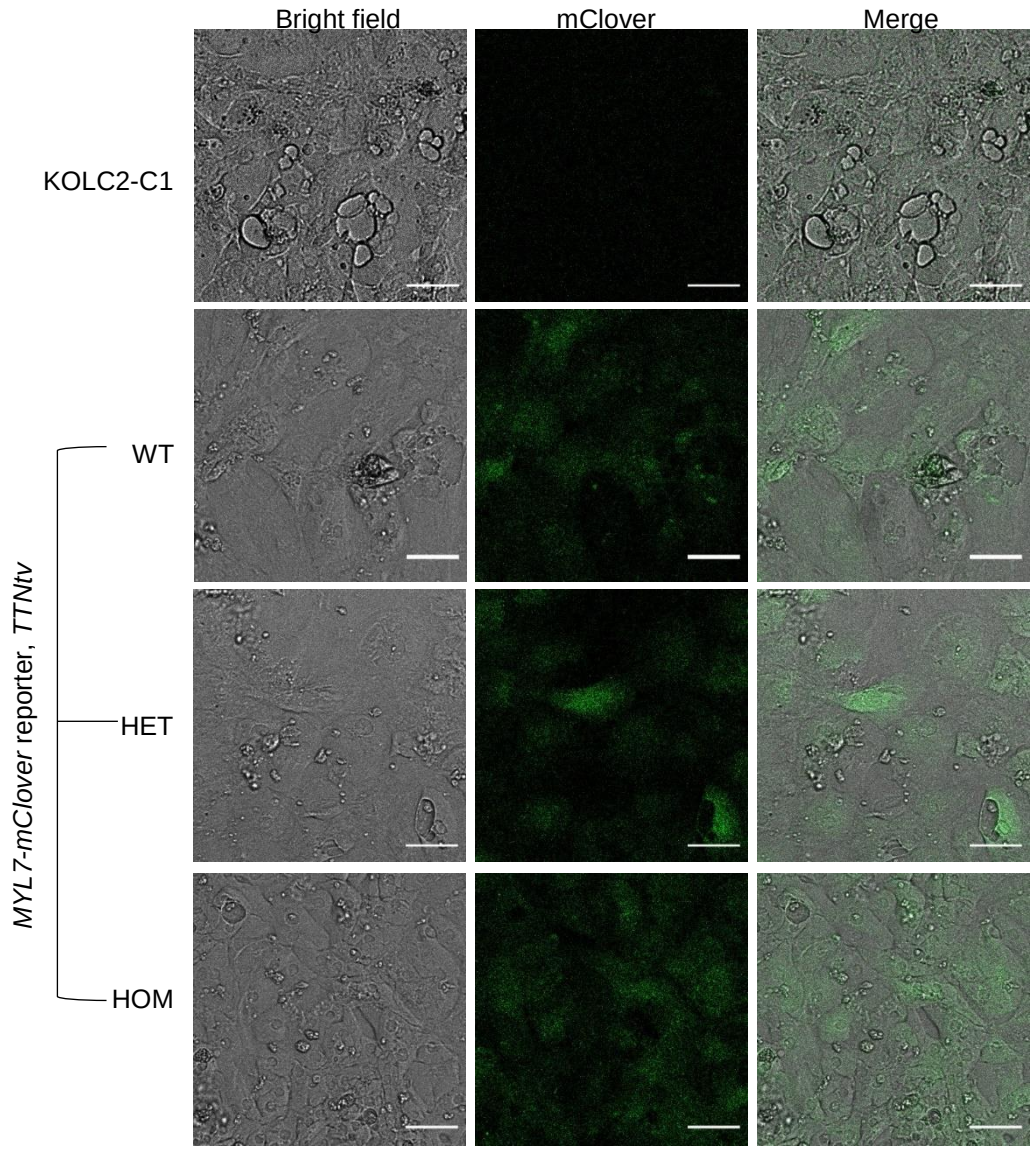
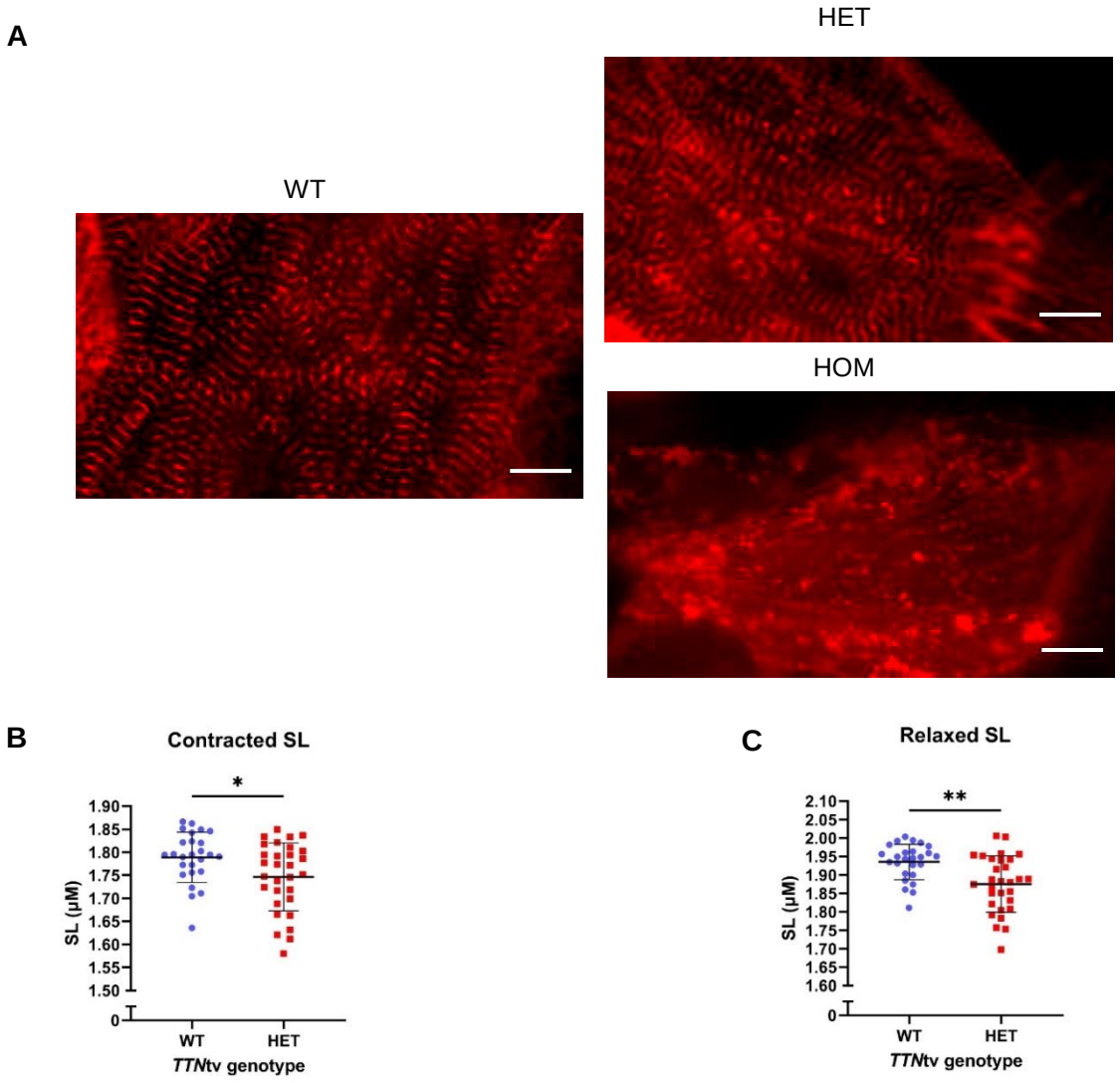


Figure 6-8 Live cell imaging of day 30 ventricular-like *MYL7-mClover* iPSC-CMs bearing *TTN* c.59926+1 G>A. KOLF2-C1 iPSC-CMs were used as a negative control for the mClover fluorescence. No difference in the mClover fluorescence detected between the *TTNtv* genotypes. WT, WT *TTN*; HET, heterozygous *TTNtv*; HOM, homozygous *TTNtv*. Scale bar 10 µm.

6.3.2.2 Effect of *TTNtv* on contractility

In order to elucidate whether sarcomere contractility is affected by *TTN* c.59926+1 G>A, day 33 ventricular-like iPSC-CMs were transduced with adenoviral viruses carrying *mScarlet-ACTN2* and subsequently analysed for sarcomere contractility by SarcTrack at day 35. Since the *TTNtv* homozygous line displayed no sarcomere structure, it was excluded from SarcTrak analysis (Figure 6-9A).



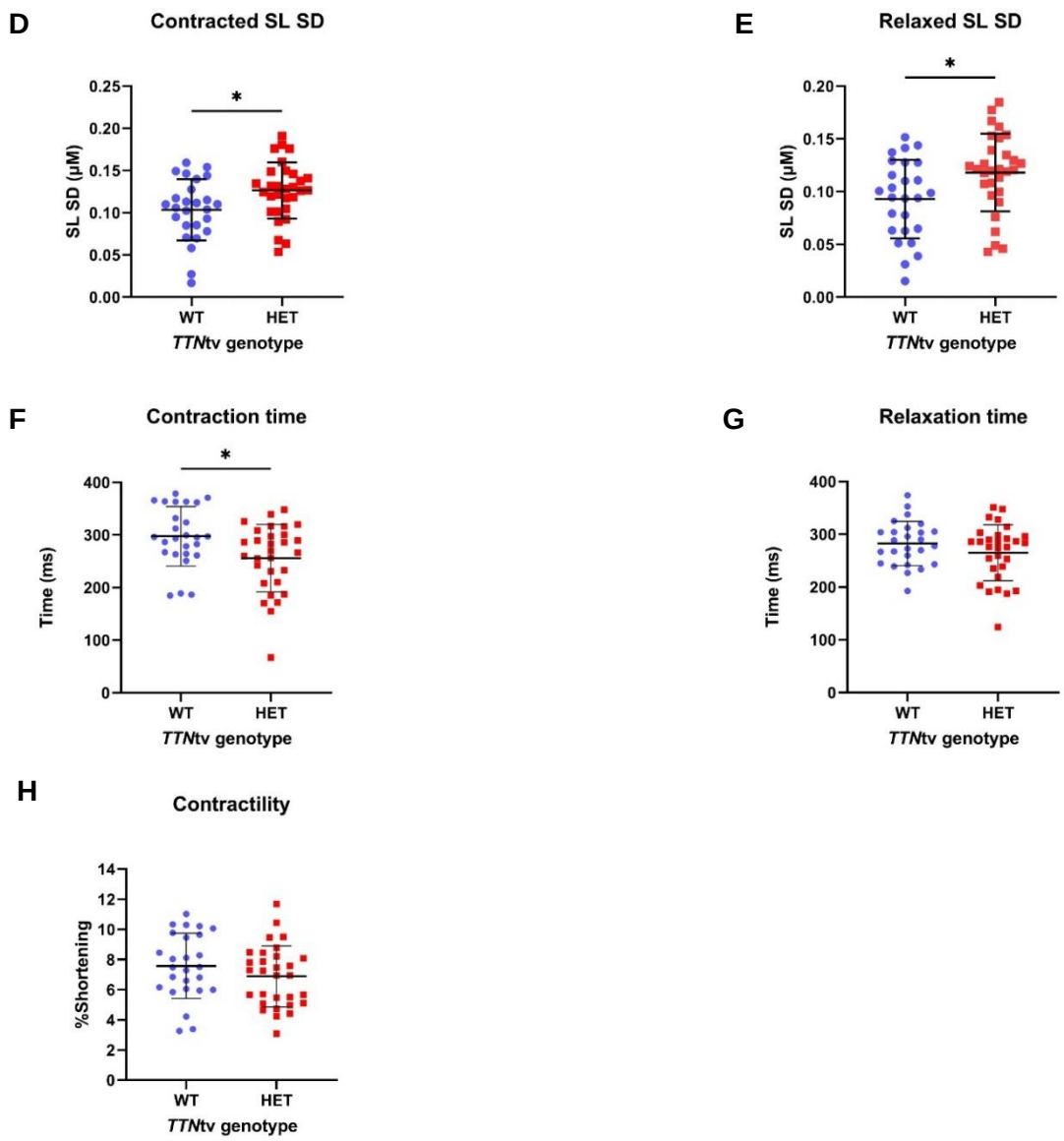


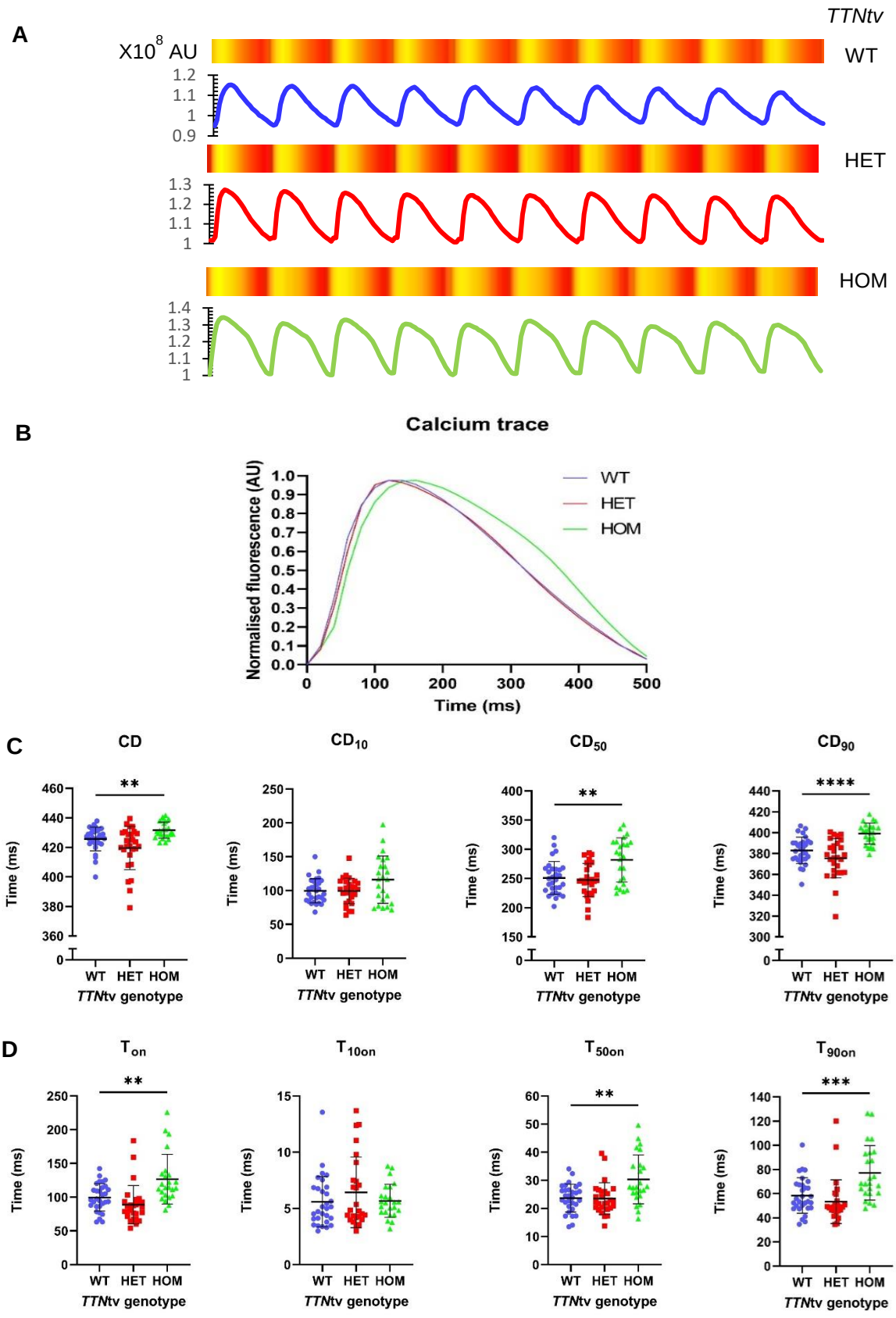
Figure 6-9 Contractility analysis of day 35 ventricular-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A. A) Representative live cell images with fluorescently-labelled ACTN2. Scale bar = 10 µm. B) Minimum sarcomere length (SL) at contraction state. C) Maximum SL at relaxation state. D-E) S.D. of SL at D) contraction state and E) relaxation state. F) Contraction time. G) Relaxation time. H) Fractional shortening (contractility). Each dataset was obtained from different cells from a single batch of differentiation. Comparative analysis was made relative to the WT *TTN* control line. HET, homozygous *TTNtv*; HOM, homozygous *TTNtv*. Mean ± S.D.; n WT = 26, n HET = 30; Student t-test (B, C, F and H) or U-test (D, E and G).

Compared to the WT line, the *TTNtv* heterozygous line presented a slightly shorter sarcomere length at both the contracted state (Figure 6-9B) and the relaxed state (Figure 6-9C) coupled with an increased variation in the sarcomere length at both the contracted state (Figure 6-9D) and the relaxed state (Figure 6-9E), indicating a delay in sarcomere maturation. The heterozygous line also showed faster contraction kinetics (Figure 6-9F) without a change in the fraction shortening (Figure 6-9H). There may be a trend of faster sarcomere relaxation kinetics, but it was not statistically significant (Figure 6-9G).

6.3.2.3 Effect of *TTNtv* on Ca^{2+} transient dynamics

Analysis of Ca^{2+} transients can provide further insights into functional consequences of *TTN* c.59926+1 G>A, which may correspond to the observed shortened sarcomere length (section 6.3.2.2).

Day 33 ventricular-like iPSC-CMs were transduced with adenoviruses harbouring *RGECO* and analysed by CalTrack algorithm for Ca^{2+} transients 2 days post transfection. The heterozygous line showed no deviation of any Ca^{2+} transient parameters from the WT control (Figure 6-10A & B). It is intriguing to highlight that the homozygous line maintained rhythmic, pace-able Ca^{2+} transients (Figure 6-10A) even though the cells lacked sarcomere organisation. The homozygous line developed a lengthened Ca^{2+} transient phenotype (Figure 6-10A & B) with a longer CD, especially CDs at 90% and 50% of the Ca^{2+} decay (Figure 6-10C), and lengthened time to Ca^{2+} peak, especially times to 90% and 50% of the Ca^{2+} peak (Figure 6-D). Even though the calculated time to Ca^{2+} decay was slightly shorter in the homozygous line (Figure 6-10E), it was not supported by the other parameters e.g. times to 90% and 50% of the Ca^{2+} decay. Overall, the observed Ca^{2+} transient dynamics in the homozygous line resembled the characteristics of immature, embryonic-like CMs (Funakoshi et al., 2021, Koivumäki et al., 2018, Pioner et al., 2019).


A
TTNtv
X10^8 AU
WT
HET
HOM
B
Calcium trace
Normalised fluorescence (AU)
Time (ms)
C
CD
CD10
CD50
CD90
Time (ms)
TTNtv genotype
D
Ton
T10on
T50on
T90on

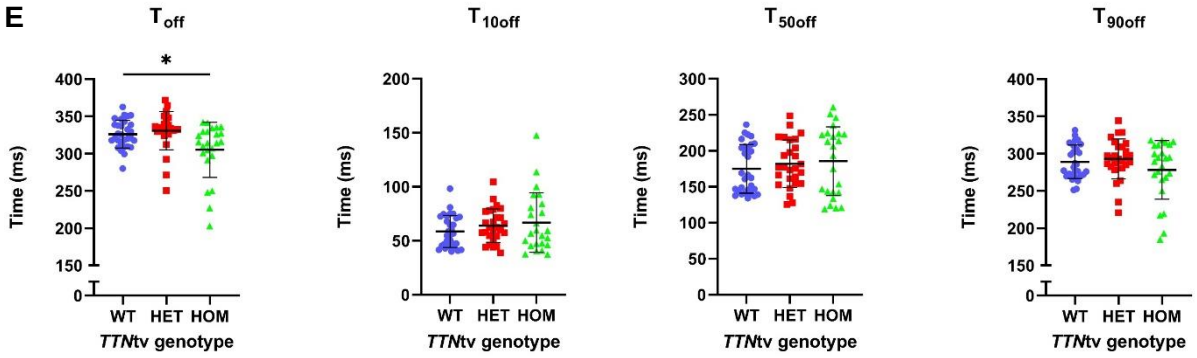


Figure 6-10 Ca^{2+} transient analysis of day 35 ventricular-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A. A) Representative kymograph and intracellular, raw fluorescent Ca^{2+} traces. B) Normalised Ca^{2+} traces. C) Ca^{2+} transient duration (CD), CD at 10% decay (CD_{10}), CD at 50% decay (CD_{50}) and CD at 90% decay (CD_{90}). D) Time to Ca^{2+} peak (T_{on}), time to 10% Ca^{2+} peak (T_{10on}), time to 50% Ca^{2+} peak (T_{50on}) and time to 90% Ca^{2+} peak (T_{90on}). E) Time to Ca^{2+} decay (T_{off}), time to 10% Ca^{2+} decay (T_{10off}), time to 50% Ca^{2+} decay (T_{50off}) and time to 90% Ca^{2+} decay (T_{90off}). Each dataset was obtained from different cells from a single batch of differentiation. Comparative analysis was made relative to the WT *TTN* control line. HET, homozygous *TTNtv*; HOM, homozygous *TTNtv*. Mean ± S.D.; n WT = 30, n HET = 27, n HOM = 24; one-way ANOVA with Tukey's post hoc (CD_{50}, T_{90on} and T_{50off}) or Kruskal-Wallis with Dunn correction post-hoc (CD, CD_{10}, CD_{90}, T_{on}, T_{10on}, T_{50on}, T_{off}, T_{10off} and T_{90off}).

6.3.3 Effect of *TTNtv* on sarcomeric protein localisation

The shortened sarcomere length observed in section 6.3.2.2 suggested that there may be a modification of sarcomere component(s) related to sarcomere maturation. To clarify this notion, day 31 ventricular-like iPSC-CMs were fixed and immuno-stained for several sarcomeric proteins.

In order to observe sarcomere formation, the cells were stained for Z-disc-localising ACTN2 (Figure 6-11A). Similar to the observations in the live cell imaging experiment (Figure 6-9A), it appeared that the heterozygous line formed an organised sarcomere structure comparable to the WT line whereas the homozygous line apparently lacked a sarcomere structure. The homozygous line also displayed hypotrophy when compared to the WT control while the heterozygous line maintained a normal cell size (Figure 6-11B). These findings confirmed the functional consequence of *TTN* c.59926+1 G>A in disrupting sarcomerogenesis previously presented in the parental iPSC-CM line (Chapter 5). There was no difference in the mClover fluorescence between the *TTNtv* genotypes (Figure 6-11A). As discussed in Chapter 4, the mClover fluorescence in fixed cells could not be used for a comparative analysis as the fixative

potentially disrupts the fluorophore's structure. Also presented in Chapter 4, the mClover protein localised mainly on the sarcomere structure and within the nucleus. In the absence of the sarcomere, the mClover fluorescence was primarily detected within the nuclei of the homozygous *TTNtv* iPSC-CMs.

To clarify whether *TTN* c.59926+1 G>A generated truncated titin proteins lacking the majority of the C-terminus, the cells were stained for titin epitopes at two different ends. Using T12 antibodies, which targets the early region of the titin's I-band, it was shown that the titin N-terminus was stably expressed at the sarcomere in the heterozygous line (Figure 6-12); in the homozygous line, there was distinct punctate staining (Figure 6-12), suggesting the expression of the truncated protein, possibly present as aggregates with ACTN2 (Figure 6-15). However, the titin C-terminus' epitope was exclusively absent in the homozygous line as shown by immunofluorescence using M8 antibodies recognising the M-line region of titin (Figure 6-13). Together, it could be concluded that truncated titin was produced and stably maintained in the homozygous setting.

To investigate changes within cardiac thin filaments and in myofribrillogenesis progress, the cells were probed for TNNT2 and F-actin. The heterozygous line showed a normal pattern of TNNT2/F-actin striation similar to the WT line (Figure 6-14). Meanwhile, fragments of TNNT2/F-actin positive components were abundantly observed in the homozygous line.

Delay in sarcomere maturation may reflect particular proteins missing from either the Z-disc or the M-line structures. Nevertheless, the M-line protein MYOM1 and the Z-disc protein T-CAP were normally detected in the heterozygous line (Figure 6-16). It is noteworthy that interpretation of T-CAP expression and its Z-disc deposition in the heterozygous setting is challenging because the T-CAP striation was occasionally undetectable even in the WT line. Since the homozygous line failed to establish sarcomeres, it is not surprising that both proteins were not detectable.

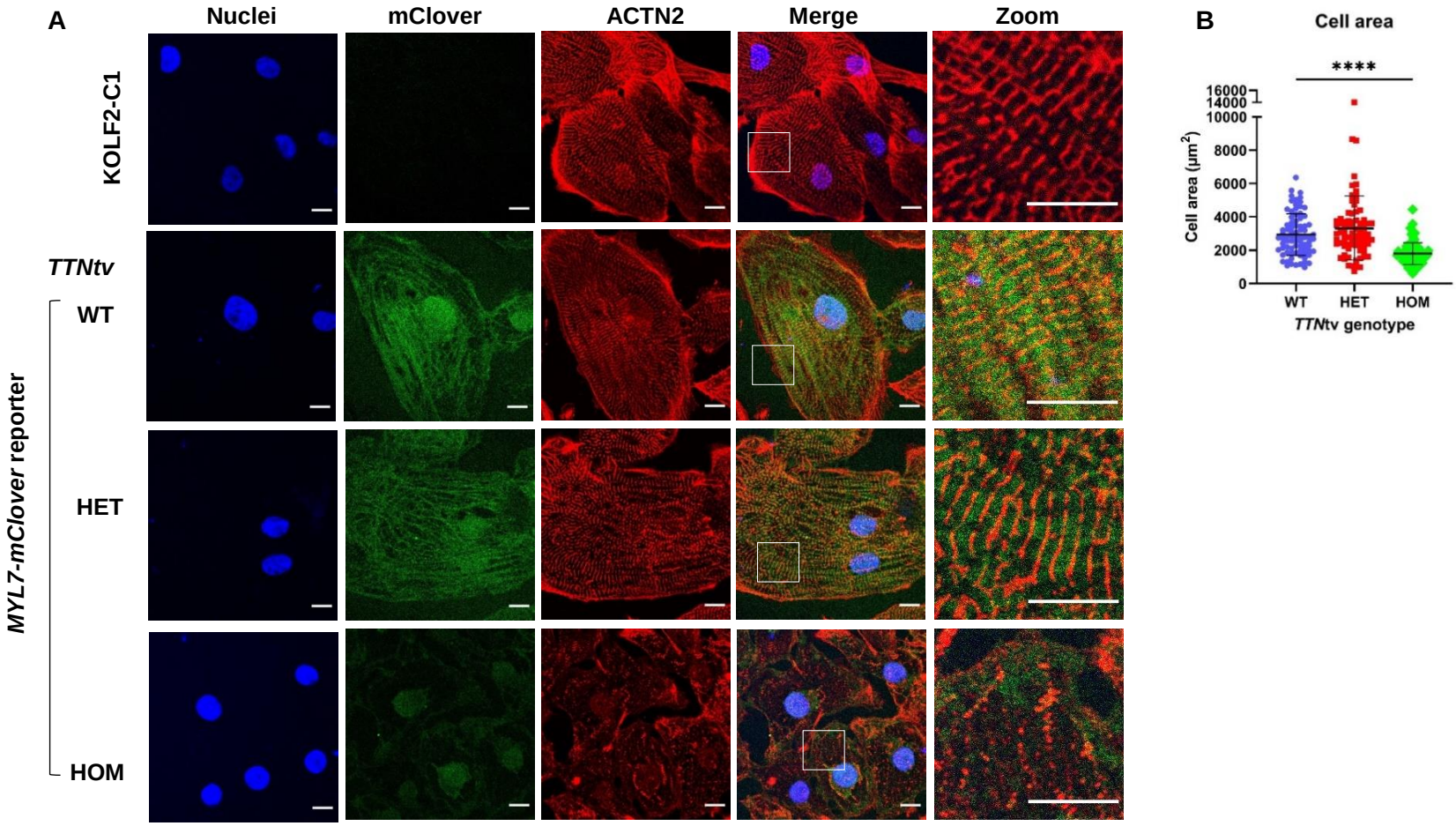


Figure 6-11 *TTN* c.59926+1 G>A impaired sarcomere formation in ventricular-like *MYL7-mClover* iPSC-CMs. A) Representative ACTN2 IF images of day 31 KOLF2-C1 and *MYL7-mClover* iPSC-CMs harbouring WT *TTN*, heterozygous *TTNtv* (HET) and homozygous *TTNtv* (HOM) counter-stained with nuclear-specific DAPI dye. mClover fluorescence was acquired via a fluorescein-specific filter set. Merge is an overlay of ACTN2, mClover and nuclei images. Zoom images are digital zoom images from the square area destined in Merge. Scale bar 10 µm. B) Quantification of cell area of ACTN2-positive iPSC-CMs. Each dataset was obtained from different cells from a single batch of differentiation. Mean ± S.D., n WT=80, n HET=80, n HOM=80, Kruskal-Wallis with Dunn correction post-hoc.

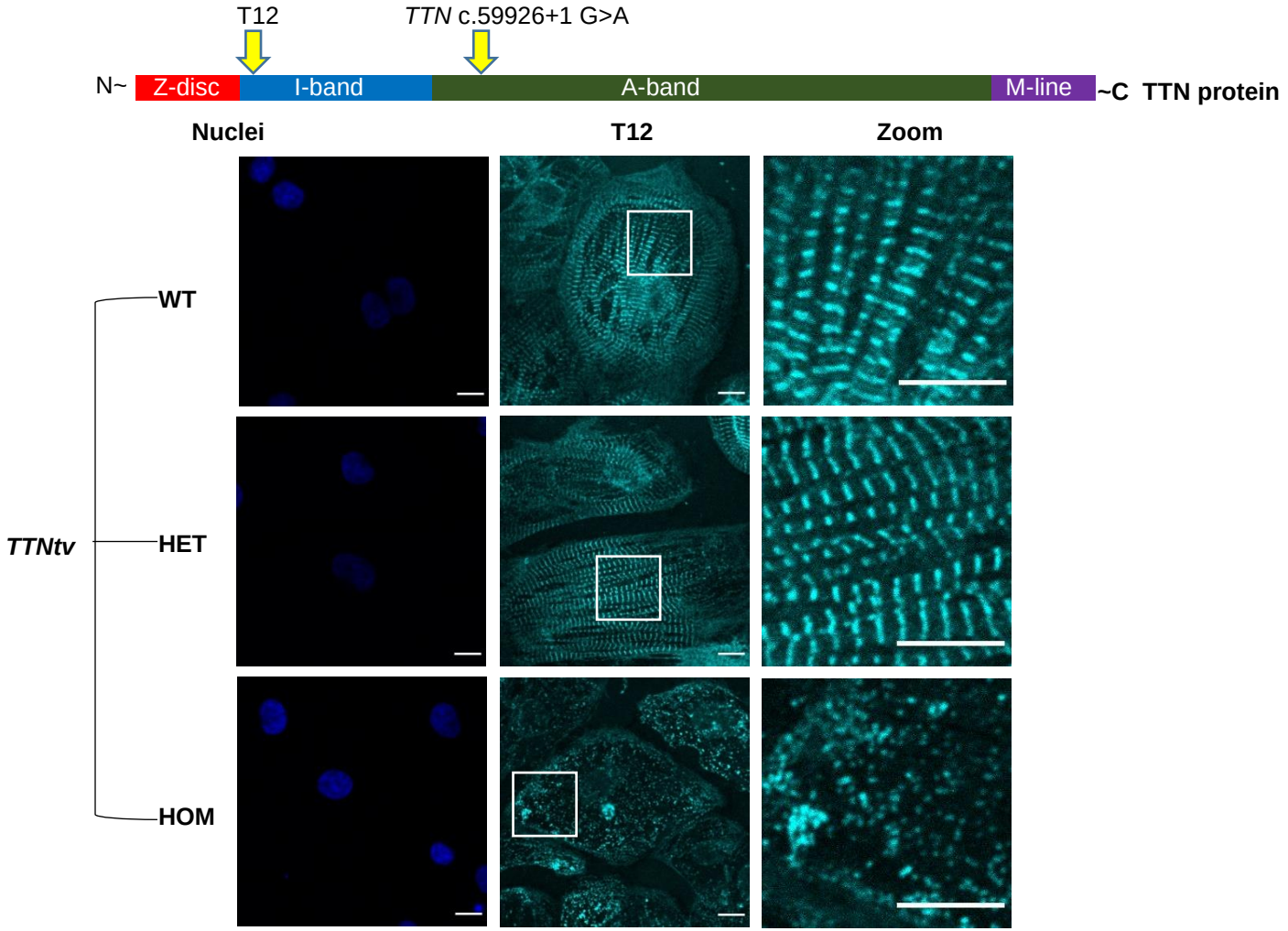


Figure 6-12 N-terminus titin was expressed in ventricular-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A. IF staining for TTN T12 (recognising the early I-band region) and nuclei (DAPI) of day 31 iPSC-CMs. Zoom images are digital zoom images from the square area destined in Merge. Scale bar 10 μm.

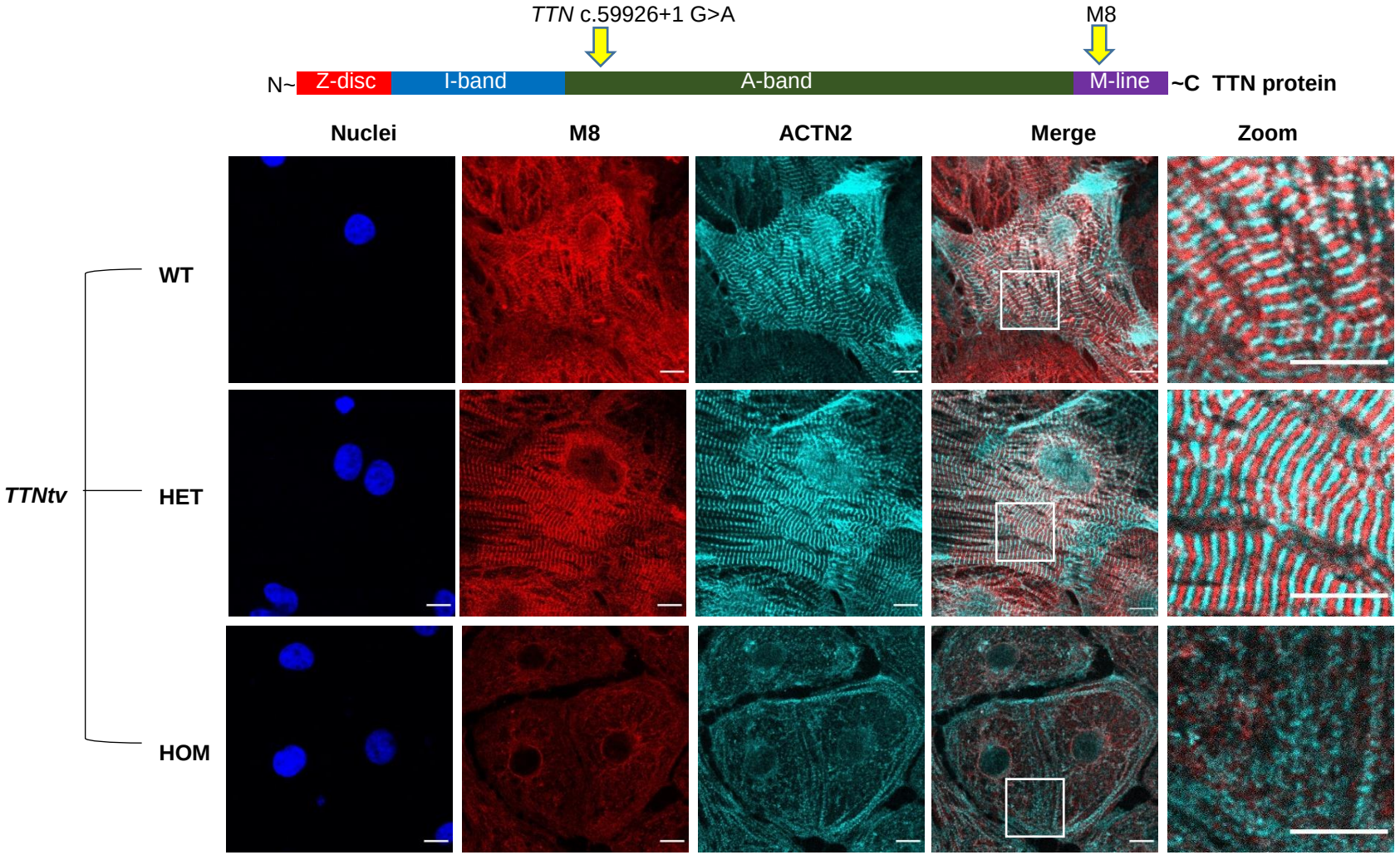


Figure 6-13 *TTN* c.59926+1 G>A generated truncated titins lacking the M-line region in ventricular-like *MYL7-mClover* iPSC-CMs. IF staining for TTN M8 (recognising the M-band titin), ACTN2 and nuclei (DAPI) of day 31 iPSC-CMs. Merge is an overlay of M8 and ACTN2 images. Zoom images are digital zoom images from the square area destined in Merge. Scale bar 10 μm.

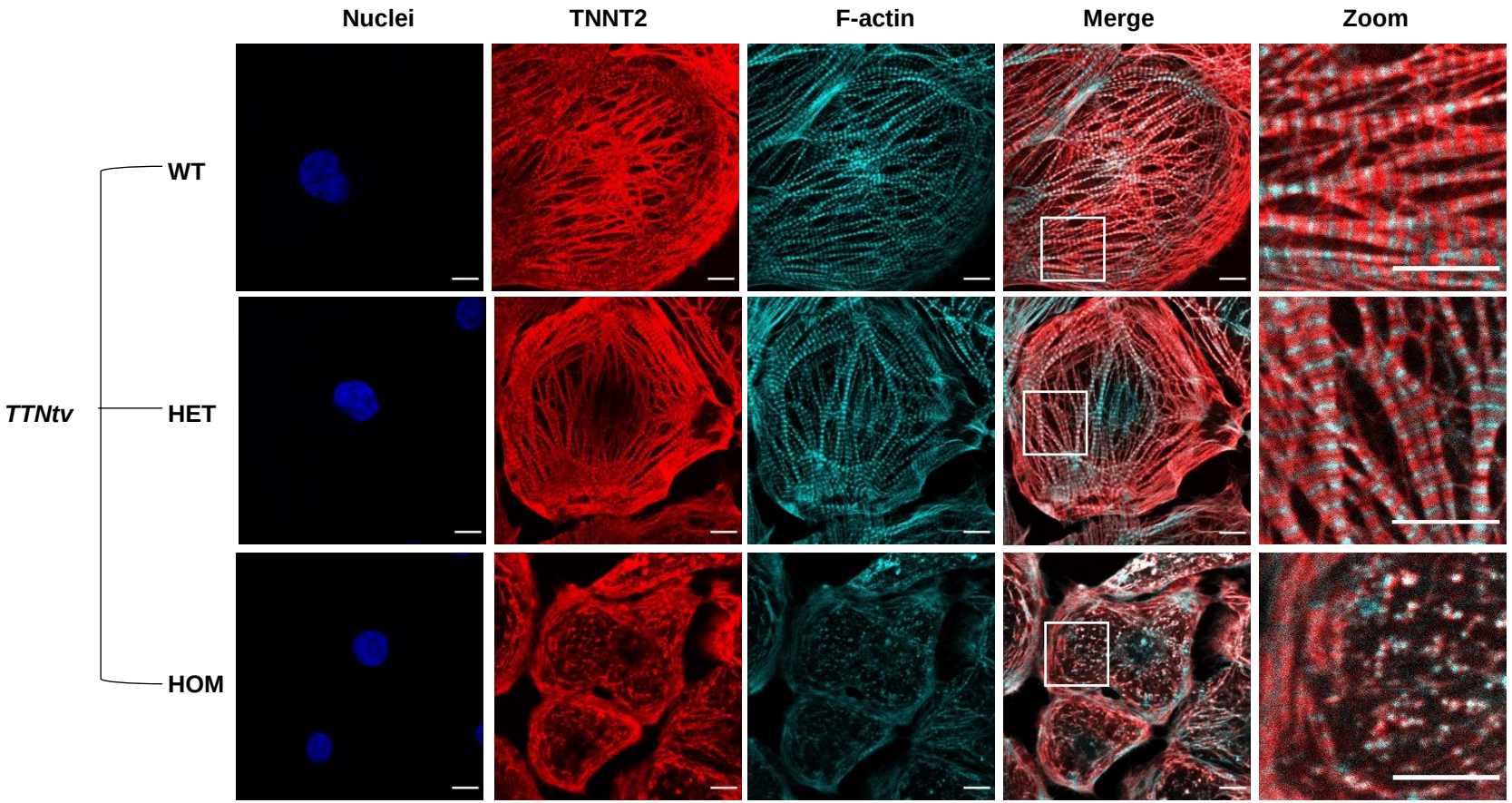


Figure 6-14 *TTN* c.59926+1 G>A impeded myofibrillogenesis in ventricular-like *MYL7-mClover* iPSC-CMs. IF staining for TNNT2, F-actin and nuclei (DAPI) of day 31 iPSC-CMs. Merge is an overlay of TNNT2 and F-actin images. Zoom images are digital zoom images from the square area destined in Merge. Scale bar 10 µm.

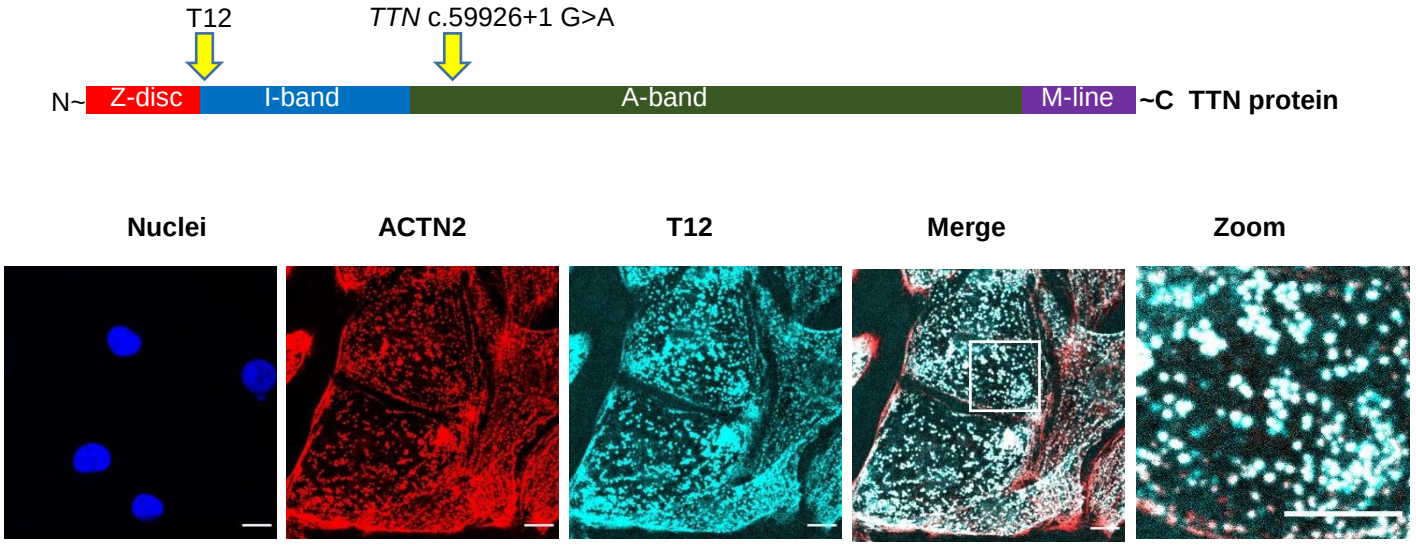


Figure 6-15 Truncated titin may interact with ACTN2 in ventricular-like *MYL7-mClover* iPSC-CMs carrying homozygous *TTN* c.59926+1 G>A. IF staining for ACTN2, N-terminus titin T12 epitope and nuclei (DAPI) of day 31 iPSC-CMs. Merge is an overlay between ACTN2 and T12 images. Zoom images are digital zoom images from the square area destined in Merge. Scale bar 10 µm

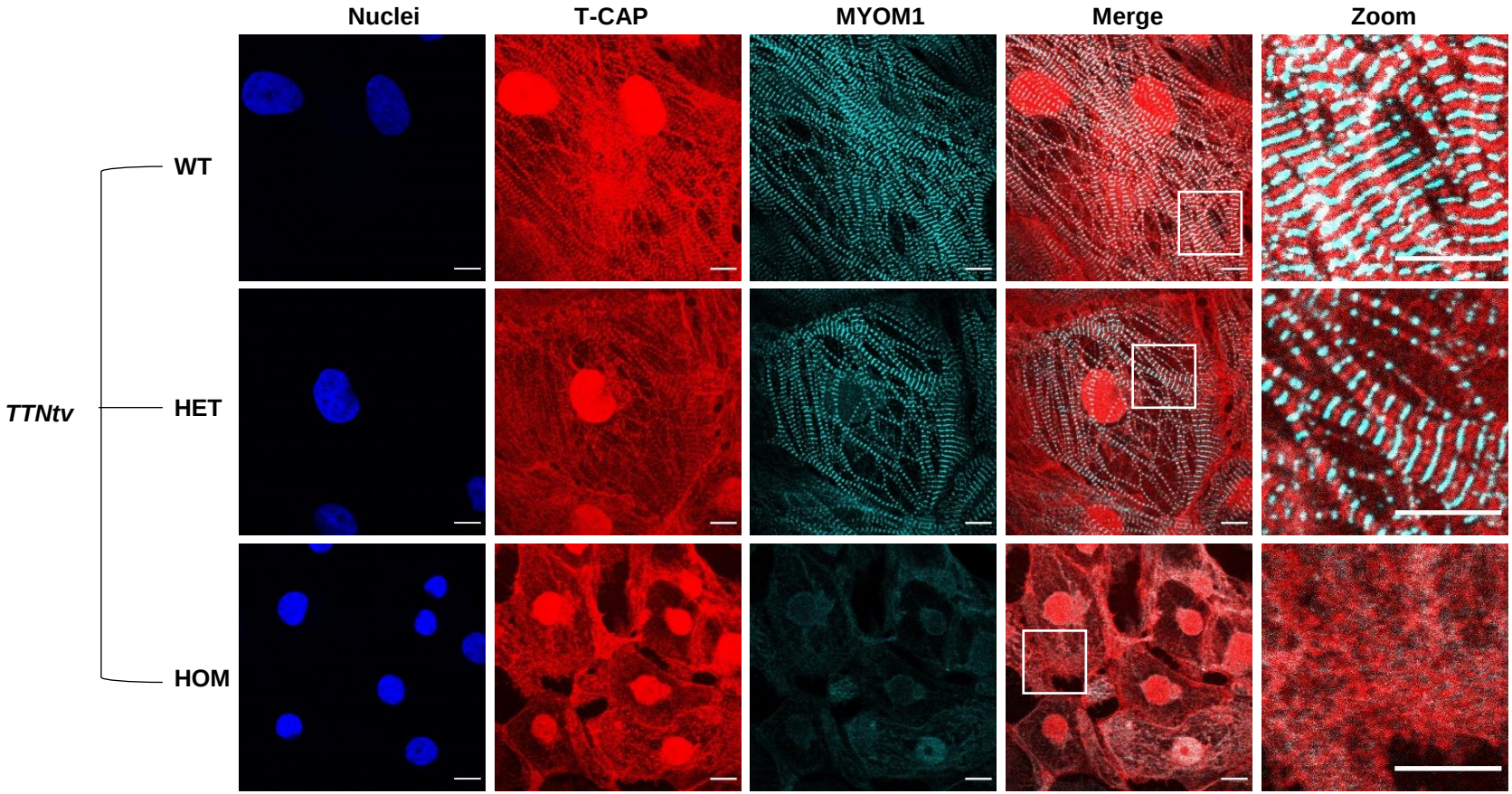


Figure 6-16 Normal Z-disc and M-line protein localisation in ventricular-like *MYL7-mClover* iPSC-CMs harbouring heterozygous *TTN* c.59926+1 G>A. IF staining for T-CAP, MYOM1 and nuclei (DAPI) of day 31 iPSC-CMs. Merge is an overlay between T-CAP and MYOM1 images. Zoom images are digital zoom images from the square area destined in Merge. Scale bar 10 µm.

6.4 Characterisation of TTNtv in atrial-like MYL7-mClover iPSC- CMs

6.4.1 Molecular analysis

Both the heterozygous and homozygous *TTNtv* lines were differentiated using the RA-induced atrial differentiation protocol described in Chapter 3.

6.4.1.1 Effect of *TTNtv* on transcriptional expression of cardiac genes

RNA expression was calculated relative to the WT line for a number of genes including *TTN*, other cardiac markers and heart failure markers at day 31 of differentiation.

Overall, there were no apparent changes in *TTN* transcriptional levels for any isoform between the *TTNtv* genotypes when probing different constitutive exons (Figure 6-17). The slight decrease in expression of the A-band *TTN* upstream or downstream of the *TTNtv* site in the heterozygous or homozygous line was not considered as an indicator of an alteration of the Cronos titin expression because there was no supporting evidence from either the A-band *TTN* upstream of the *TTNtv* or the M-line *TTN*. *TTN* N2-BA along with N2-BA/ N2-B + N2-BA ratio were not different between the *TTNtv* genotypes. Meanwhile, the negligible upregulation of N2-BA/ N2-B + N2-BA ratio in the heterozygous line was a result of a summation of the slight increase in N2-BA and the slight decrease in N2-B + N2-BA, which may not significantly influence the titin compliance.

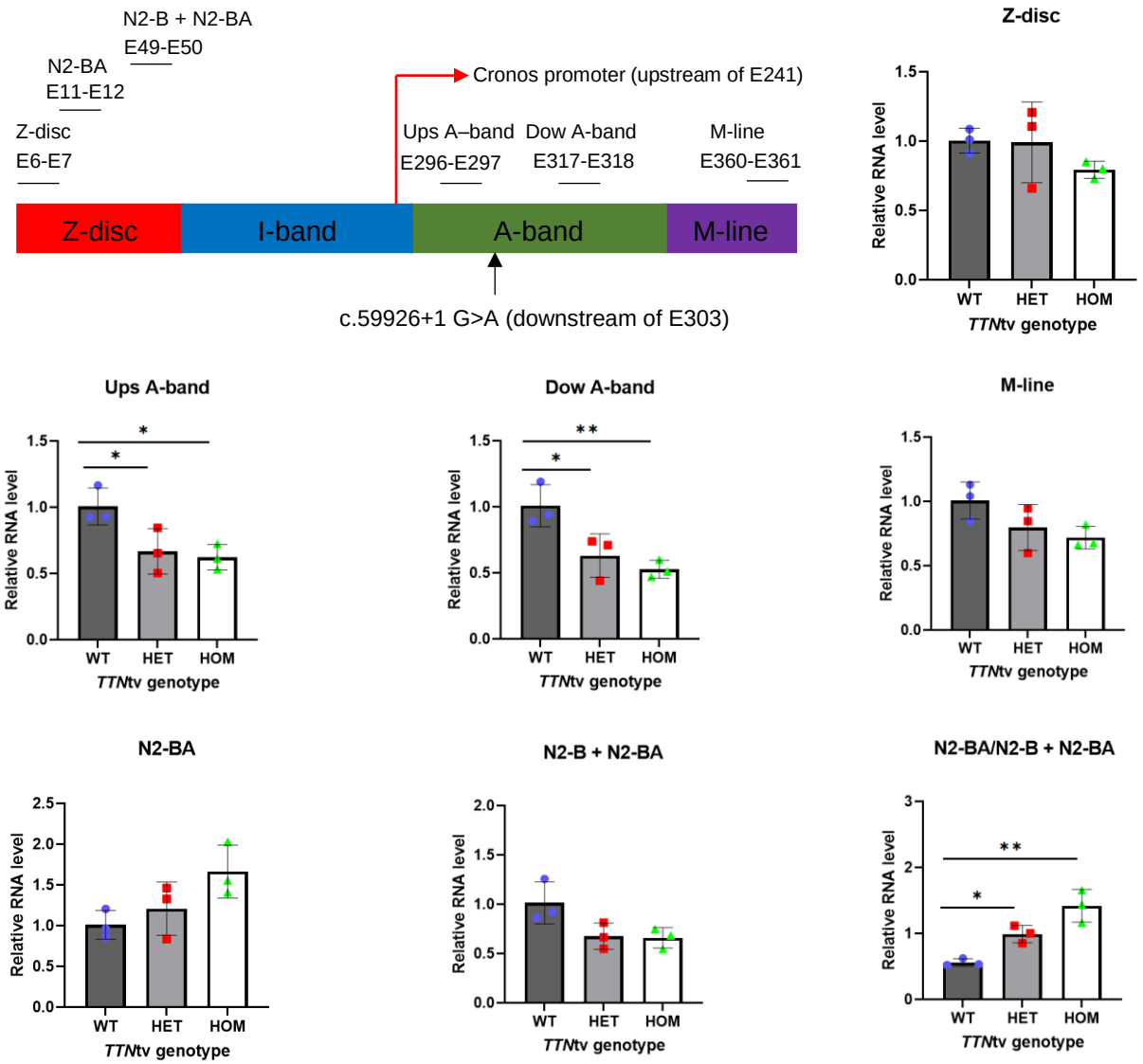


Figure 6-17 *TTN* transcriptional levels of day 31 atrial-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A. PCR probes were designed to be specifically spanning different pairs of exons responsible to the *TTN* regions; Z-disc (exons 6-7), A-band upstream of the *TTN*tv site (exons 296-297), A-band downstream of the *TTN*tv site (exons 317-318), and M-line (exons 360-361). The N2-BA probes (exons 11-12) were localised within the Z-disc region. The probes spanning exons 49-50 were designed to target both the cardiac specific N2-B and N2-BA isoforms. Each dataset was obtained from different wells of differentiation using a single iPSC batch. All quantifications were made relative to the WT control, except for the N2-BA/N-2B + N2-BA ratio. n=3, one-way ANOVA with Turkey's post-hoc. HET, heterozygous *TTNtv*; HOM, homozygous *TTNtv*.

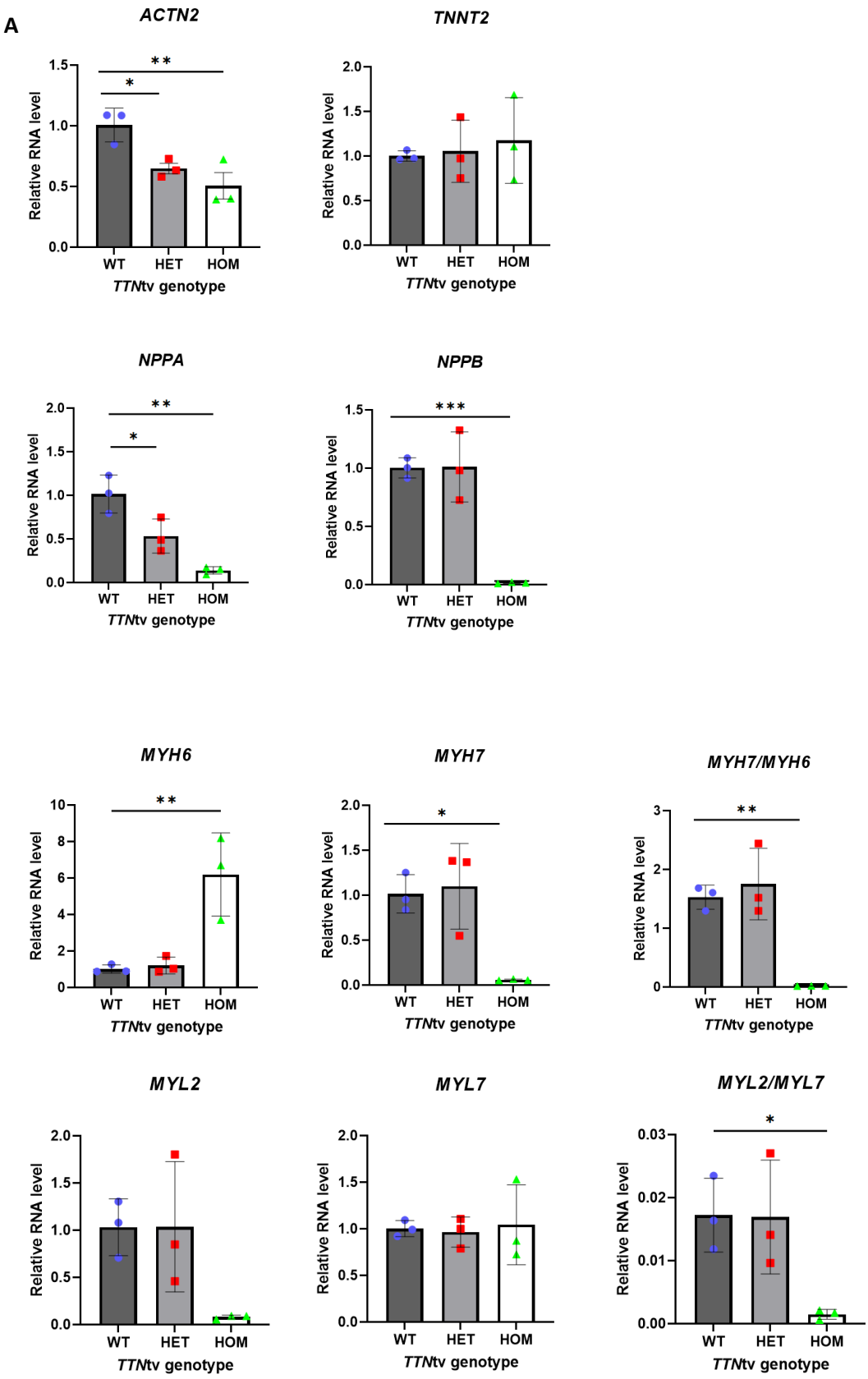
There were no dramatic changes in expression of pan-cardiac markers (*ACTN2* and *TNNT2*) between the *TTNtv* genotypes (Figure 6-18A), suggesting that the cardiac differentiation potential was fully retained. The *ACTN2* downregulation in the homozygous line was not considered as a marker for defective cardiac differentiation because it was not complimented by the*TNNT2* expression level.

The heterozygous line did not show an increase in expression of heart failure markers (*NPPA*, *NPPB*; Figure 6-18B). The modest *NPPA* downregulation in the heterozygous line (Figure 6-18B) was considered to be a result of iPSC-CM phenotypic variability and immaturity. The homozygous line showed an obvious decrease in *NPPA* and *NPPB*, suggesting an underdeveloped CM phenotype.

There was no modification of expression of myofibrillar protein isoforms observed in the heterozygous line as shown by the normal expression of *MYH6*, *MYH7*, *MYL7* and *MYL2* (Figure 6-18C). In contrast, the homozygous line showed a downregulation of expression of *MYL2*, a downregulation of expression of *MYH7* and an upregulation of expression of *MYH6*, indicating a delay in CM development. The absence of the statistical significance in *MYL2* expression in the homozygous line should be explained by the limit of one-way ANOVA analysis power.

Similar to *MYL7*, there was no dramatic change in the *mClover* expression between the *TTNtv* genotypes (Figure 6-18D), highlighting that the *mClover* expression mirrored the *MYL7* expression and *TTN* c.59926+1 G>A had no effect on *mClove*r expression.

Lastly, *T-CAP* expression was normal in both the heterozygous and homozygous lines (Figure 6-18E). There may be a reduction in *T-CAP* expression in the homozygous line, but there was no statistical significance suggested by one-way ANOVA.


A
ACTN2
TNNT2
Relative RNA level
WT
HET
HOM
TTNtv genotype
B
NPPA
NPPB
C
MYH6
MYH7
MYH7/MYH6
MYL2
MYL7
MYL2/MYL7

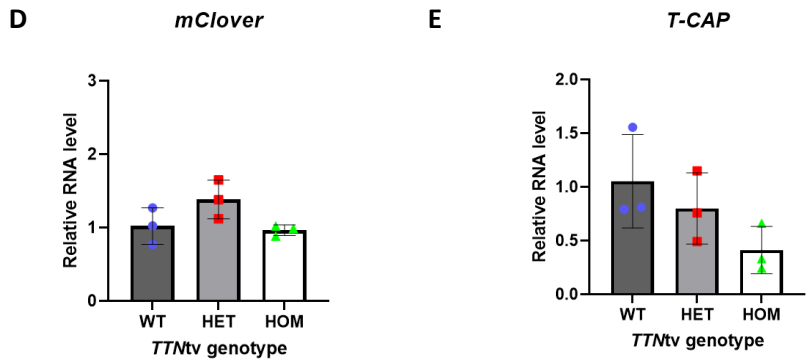


Figure 6-18 Transcriptional expression profile of day 31 atrial-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A. A) Pan-cardiac markers: *ACTN2* and *TNNT2*. B) Heart failure markers *NPPA* and *NPPB*. C) Myosin heavy chain (*MYH6* and *MYH7)* and light chain (*MYL2* and *MYL7*) isoforms. D) *mClover*. Each dataset was obtained from different wells of differentiation using a single iPSC batch. All quantifications were made relative to the WT control, except for the *MYH6/MYH7* and *MYL2/MYL7* ratios. n=3, mean ± S.D., one-way ANOVA with Turkey's post-hoc. HET, heterozygous *TTNtv*; HOM, homozygous *TTNtv*.

6.4.1.2 Effect of *TTNtv* on *TTN* splicing

In agreement with the observation in the ventricular-like iPSC-CMs, *TTN* c.59926+1 G>A caused aborted *TTN* splicing between exon 303 and 304 (Figure 6-19A & B) in the atrial-like iPSC-CMs. There was no change in total *TTN* expression between the *TTNtv* genotypes (Figure 6-19C). Only the mutant *TTN* species (>552 bp) was detected in the homozygous line while the heterozygous line expressed an equivalent level of the WT *TTN* (552 bp) and the mutant *TTN* species. Taken together, there was no difference in *TTN* transcriptional expression between ventricular and atrial-like iPSC-CMs with respect to *TTNtv* c.59926+1 G>A.

In addition, cDNA sequencing revealed that the mutant *TTN* transcript contained the additional 85 bp immediately downstream of exon 303 (Figure 6-20), again repeating the finding in the ventricular-like iPSC-CMs.

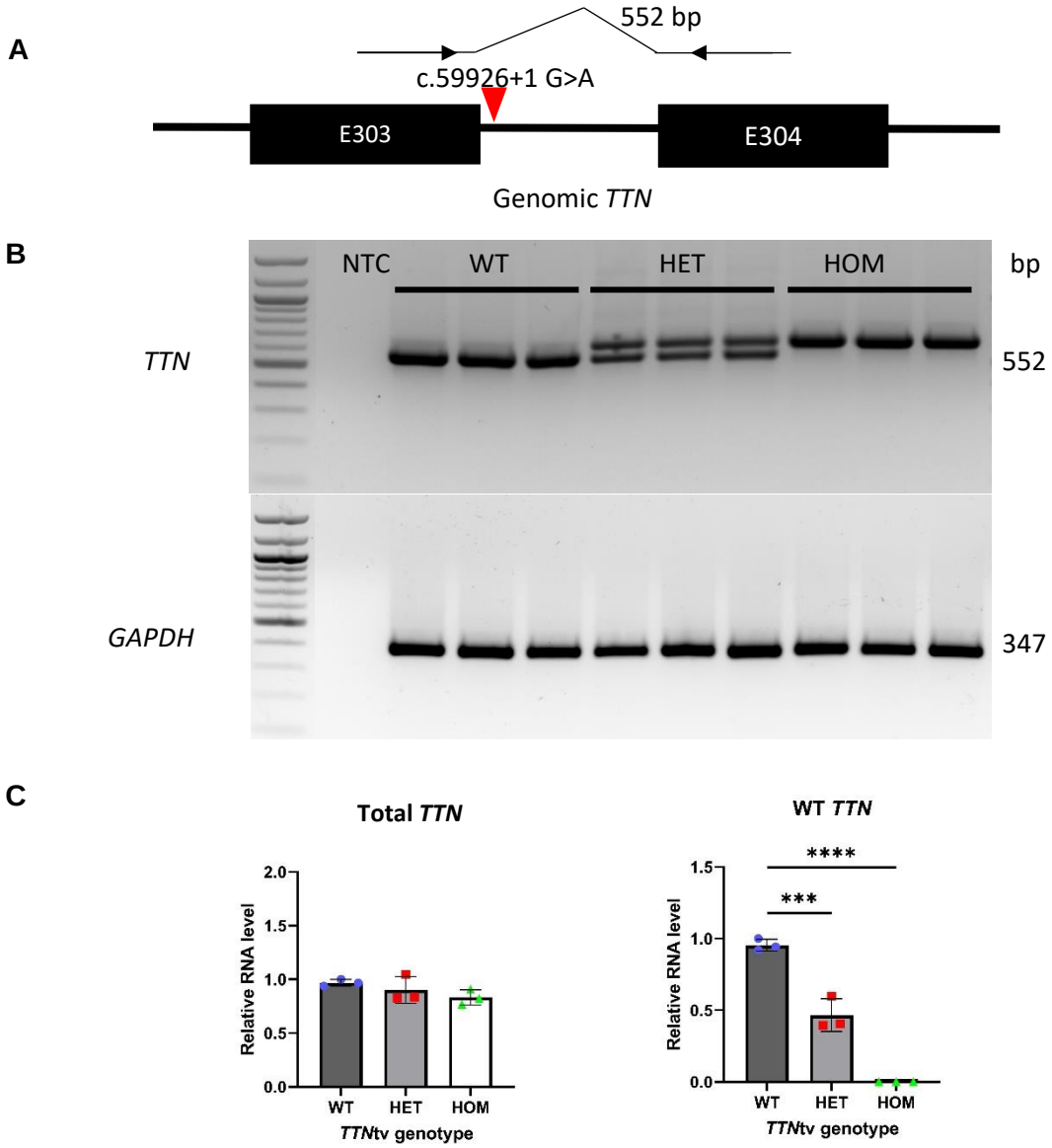


Figure 6-19 *TTN* c.59926+1 G>A caused changed *TTN* transcript species in atrial-like *MYL7-mClover* iPSC-CMs. A) Simplified genomic map over the introduced *TTN*tv site and PCR probes spanning exon 303-304. B) Agarose gels showing PCR products of day 31 iPSC-CMs. The homozygous *TTNtv* (HOM) line produced a higher molecular weight *TTN* transcript (mutant, top band) than those of the WT species (552 bp, lower band) whilst the heterozygous *TTNtv* (HET) line expressed both the WT and the mutant *TTN* transcripts. C) Quantification of *TTN* transcriptional levels based on B as normalised to *GAPDH* and as relative to the WT control. Total *TTN* was a summation of both the WT and the mutant species. Each dataset was obtained from different wells of differentiation using a single iPSC batch. n=3, mean ± S.D., one-way ANOVA with Tukey's post-hoc.

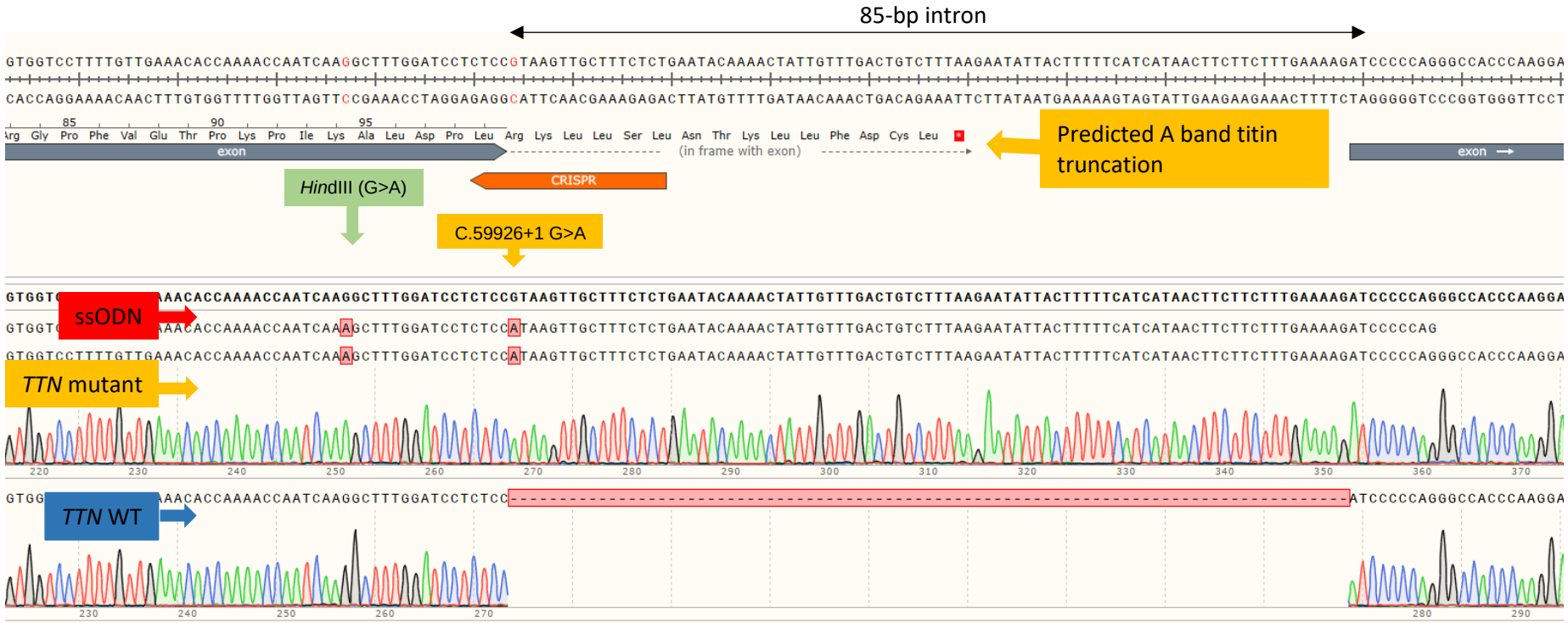


Figure 6-20 *TTN* c.59926+1 G>A caused intronic retention in atrial-like *MYL7-mClover* iPSC-CMs. cDNA sequence from the mutant titin transcript contained additional 85 nt immediately downstream of exon 303, which was predicted to cause titin A-band truncation via a premature stop codon following additional 15 amino acids (R-K-L-L-S-L-N-T-K-L-L-F-D-C-L). The sequence was aligned against the WT *TTN* transcript and ssODN highlighting the additional G>A silent mutation creating the *Hind*III site.

6.4.2 Functional analysis

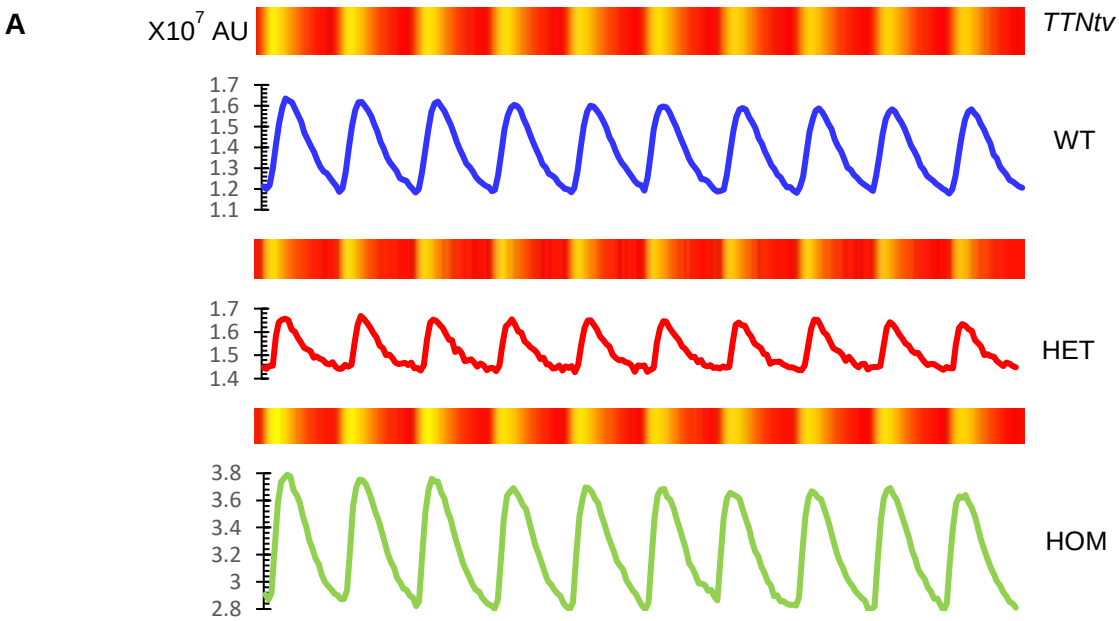
6.4.2.1 Effect of *TTNtv* on spontaneous contraction

The *TTNtv* heterozygous line developed a normal spontaneous contraction resembling to the beating characteristics of the WT line. However, it is worth noting that the heterozygous line was found to occasionally spontaneously beat faster than the WT line. Both the WT and heterozygous lines had an onset of spontaneous contraction at approximately day 9-11. Meanwhile, the spontaneous contraction was not detected in the homozygous line.

6.4.2.2 Effect of *TTNtv* on Ca^{2+} transient dynamics

Day 33 atrial-like iPSC-CMs were transduced with adenoviruses harbouring *RGECO* and their Ca^{2+} transient dynamics was analysed by CalTrack 2 days post transfection.

Interestingly, the heterozygous line exhibited a shortened Ca^{2+} transient morphology compared to the WT control (Figure 6-21A & B) accompanied by the computed reduced CD (Figure 6-21C), which was further supported by the decrease in both time to the Ca^{2+} peak (Figure 6-21D) and time to the Ca^{2+} decay (Figure 6-21E). In contrast, a slightly lengthened Ca^{2+} transient was observed in the homozygous line as illustrated by the extended CD, devoid of prominent changes in the computed time to the Ca^{2+} peak and time to the Ca^{2+} decay.



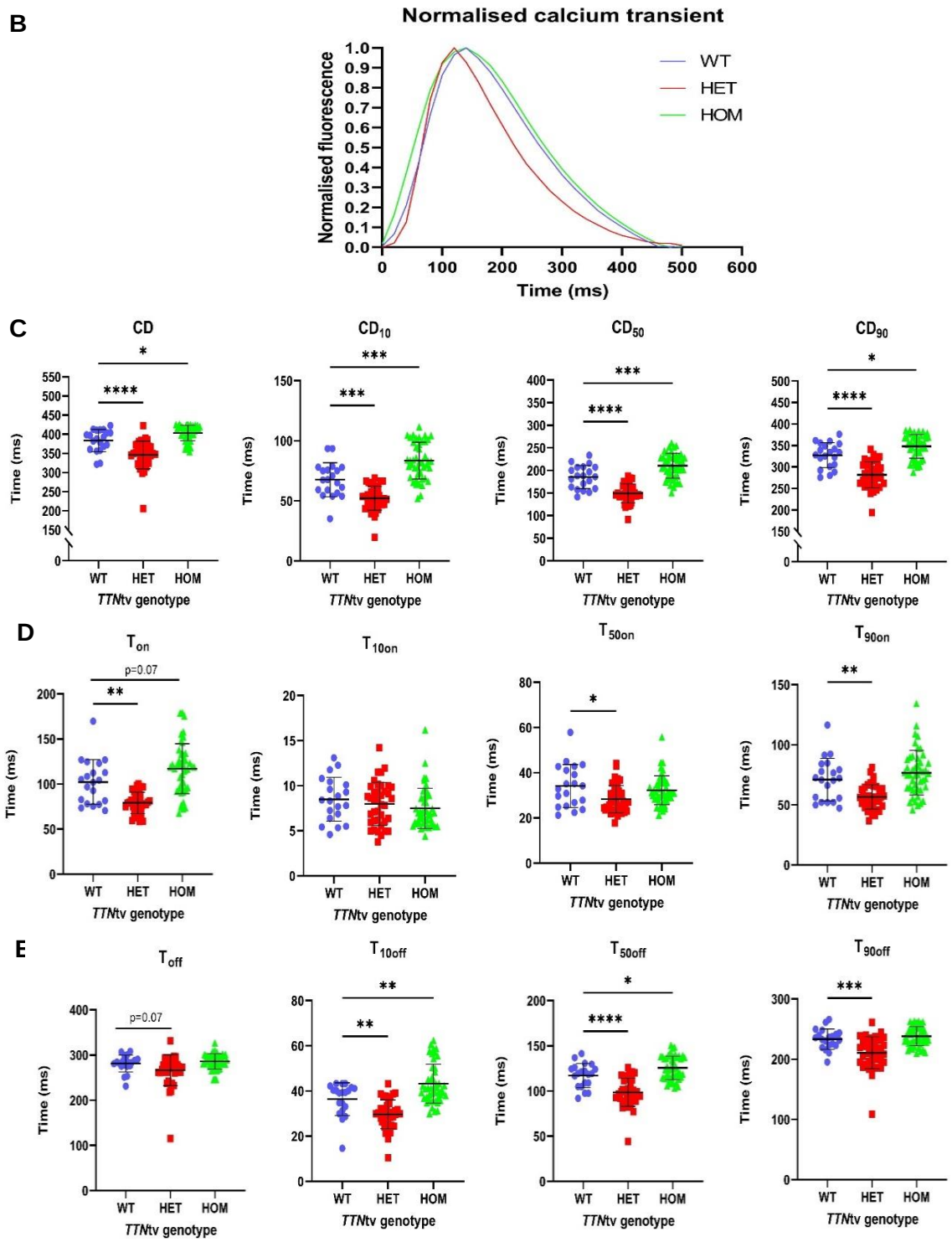


Figure 6-21 Ca^{2+} transient analysis of day 35 atrial-like *MYL7-mClover* iPSC-CMs carrying *TTN* c.59926+1 G>A. A) Representative kymograph and intracellular, raw fluorescent Ca^{2+} traces. B) Normalised Ca^{2+} traces. C) Ca^{2+} transient duration (CD), CD at 10% decay (CD_{10}), CD at 50% decay (CD_{50}) and CD at 90% decay (CD_{90}). D) Time to Ca^{2+} peak (T_{on}), time to 10% Ca^{2+} peak (T_{10on}), time to 50% Ca^{2+} peak (T_{50on}) and time to 90% Ca^{2+} peak (T_{90on}). E) Time to Ca^{2+} decay (T_{off}), time to 10% Ca^{2+} decay (T_{10off}), time to 50% Ca^{2+} decay (T_{50off}) and time to 90% Ca^{2+} decay (T_{90off}). Each dataset was obtained from different cells from a single batch of differentiation. Comparative analysis was made relative to the *TTN* WT control line. HET, homozygous *TTNtv*; HOM, homozygous *TTNtv*. Mean ± S.D.; n WT = 30, HET = 27, HOM = 24; one-way ANOVA with Tukey's post hoc (CD, CD_{50}, CD_{90}, T_{10on}, T_{10off} and T_{50off}) or Kruskal-Wallis with Dunn correction post-hoc (CD_{10}, T_{on}, T_{50on}, T_{90on}, T_{off} and T_{90off}).

6.4.3 Effect of *TTN*tv on sarcomeric protein localisation

There was no difference in ACTN2-marking Z-disc striation detected between the day 31 atrial-like *TTNtv* heterozygous and WT lines (Figure 6-22A), indicating relatively normal sarcomere formation. Only the *TTNtv* homozygous line displayed no development of sarcomere structure. In addition, there was no change in mClover expression between the *TTNtv* genotypes. Morphometric-wise, the homozygous line had a decreased cell area while the heterozygous line maintained a normal cell size (Figure 6-22B).

Further titin staining revealed that the homozygous line stably produced exclusively truncated titin proteins devoid of the M-line region i.e. undetectable M8 epitope (Figure 6-23). Meanwhile, the heterozygous line showed a relatively normal striation of T12 and M8 alternating pattern.

TNNT2 and F-actin double staining confirmed a normal thin filament organisation in the heterozygous line (Figure 6-24). In contrast, the investigation in the homozygous line suggested disrupted myofibrillogenesis as shown by the fragmented TNNT2/F-actin positive elements.

As confirmed by the MYOM1 and T-CAP staining (Figure 6-25), no sarcomere structure formed in the homozygous line. In the heterozygous line, both T-CAP and MYOM1 were normally detected in the heterozygous line.

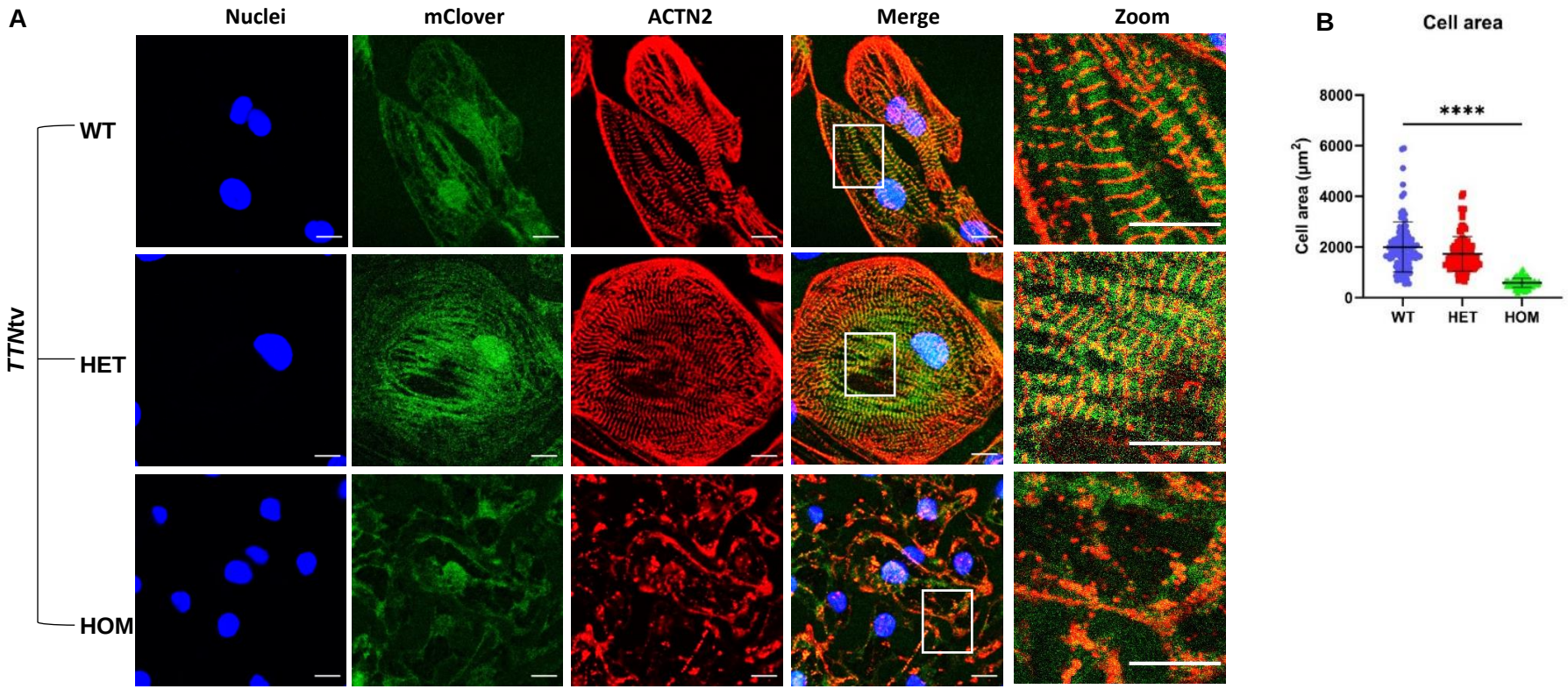


Figure 6-22 *TTN* c.59926+1 G>A abolished sarcomerogeneis in atrial-like *MYL7-mClover* iPSC-CMs. A) Representative ACTN2 IF images of day 31 *MYL7-mClover* iPSC-CMs harbouring WT *TTN*, heterozygous *TTNtv* (HET) and homozygous *TTNtv* (HOM) counter-stained with nuclear-specific DAPI. mClover fluorescence was acquired via a fluorescein-specific filter set. Merge is an overlay of ACTN2, mClover and nuclei images. Zoom images are digital zoom images from the square area destined in Merge. Scale bar 10 µm. B) Quantification of cell area of ACTN2-positive iPSC-CMs. Each dataset was obtained from different cells from a single batch of differentiation. Mean ± S.D., n WT=80, n HET=80, n HOM=80, Kruskal-Wallis with Dunn correction post-hoc.

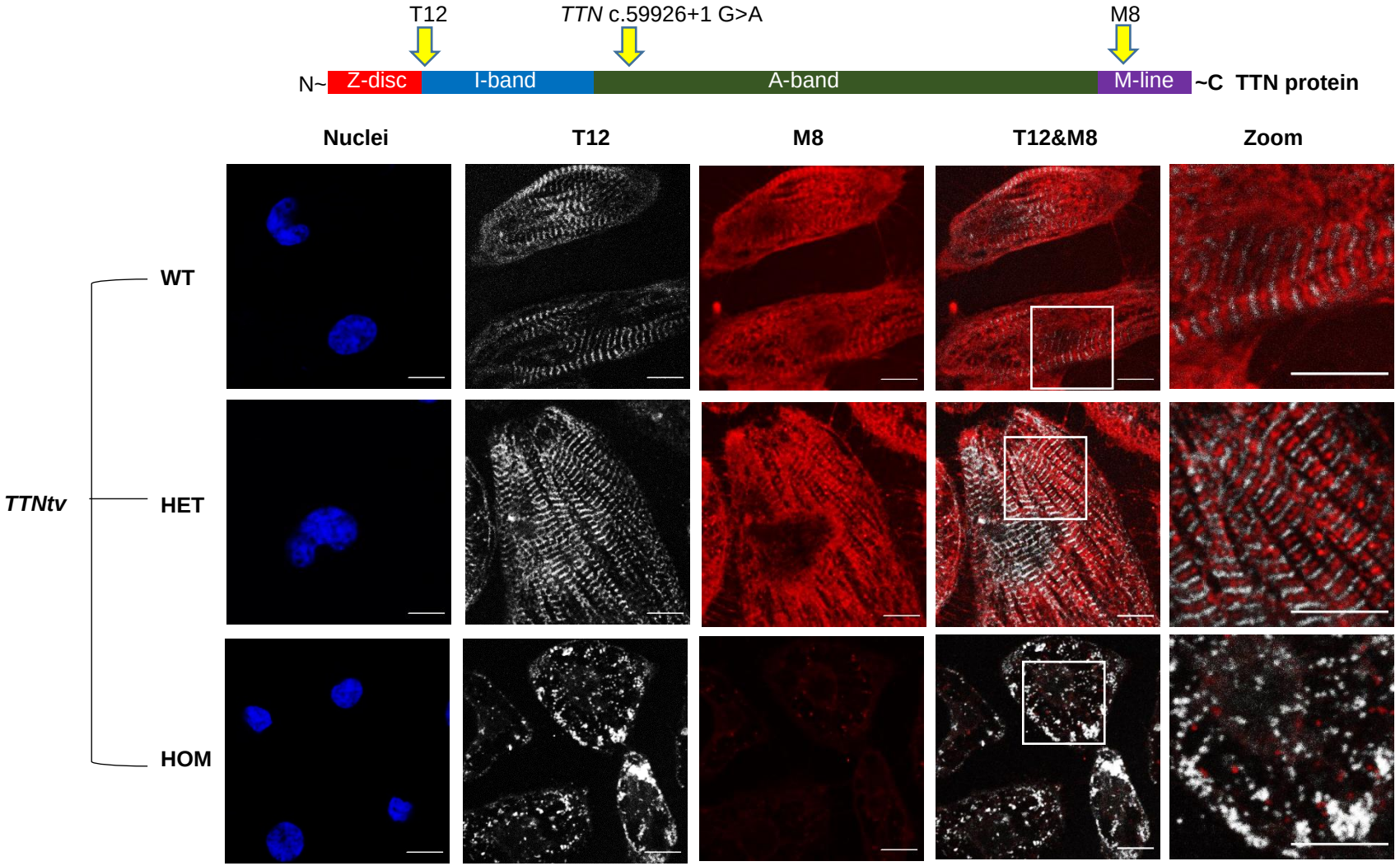


Figure 6-23 *TTN* c.59926+1 G>A truncated titin in atrial-like *MYL7-mClover* iPSC-CMs. IF staining for TTN T12 (recognising the early I-band region), TTN M8 (recognising the M-band titin) and nuclei (DAPI) of day 31 iPSC-CMs. Merge is an overlay of T12 and M8 images. Zoom images are digital zoom images from the square area destined in Merge. Scale bar 10 μm.

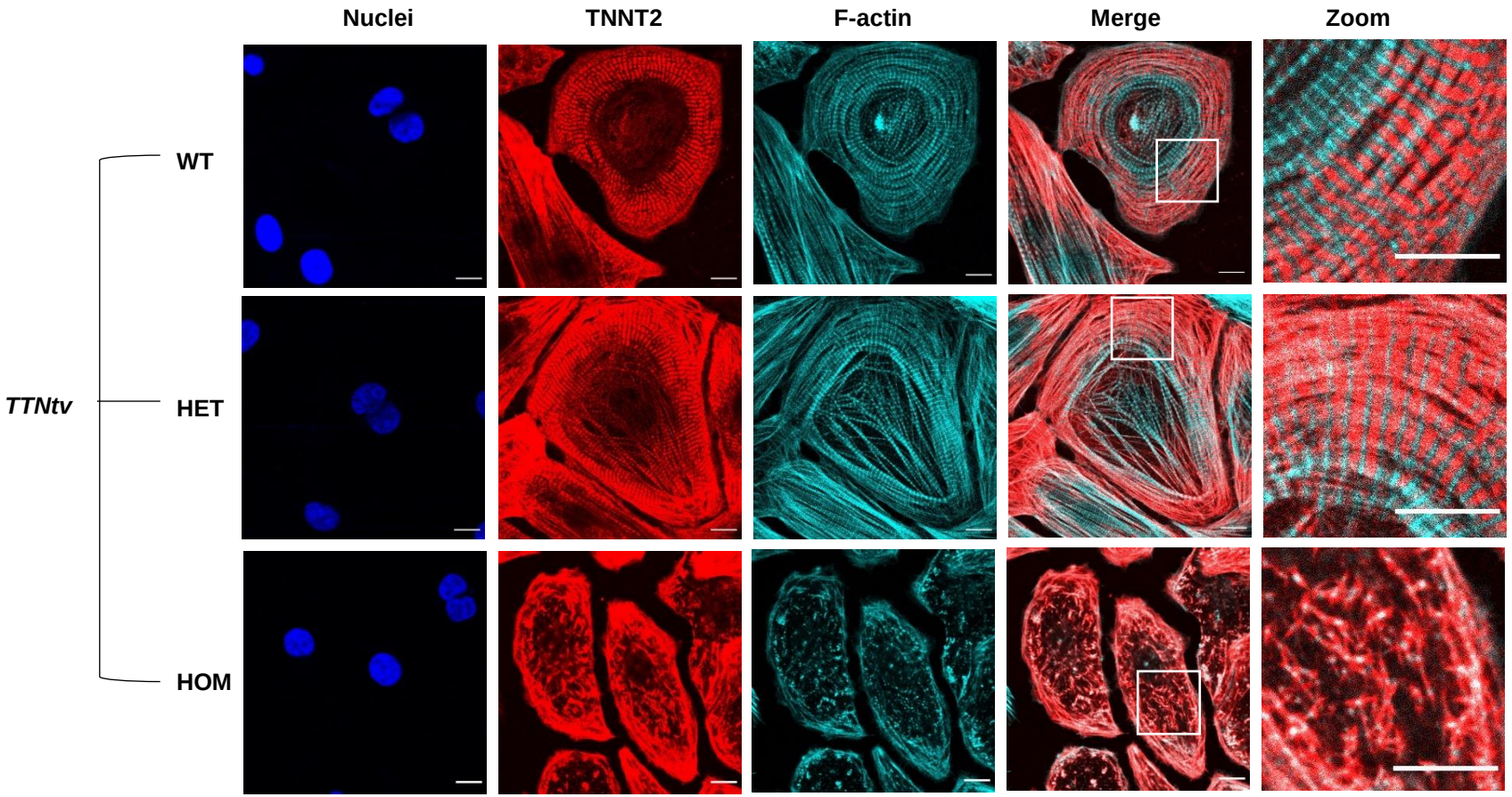


Figure 6-24 *TTN* c.59926+1 G>A impeded myofibrillogenesis in atrial-like *MYL7-mClover* iPSC-CMs carrying *TTNtv* c.59926+1 G>A. IF staining for TNNT2, F-actin and nuclei (DAPI) of day 31 iPSC-CMs. Merge is an overlay between TNNT2 and F-actin images. Zoom images are digital zoom images from the square area destined in Merge. Scale bar 10 μm.

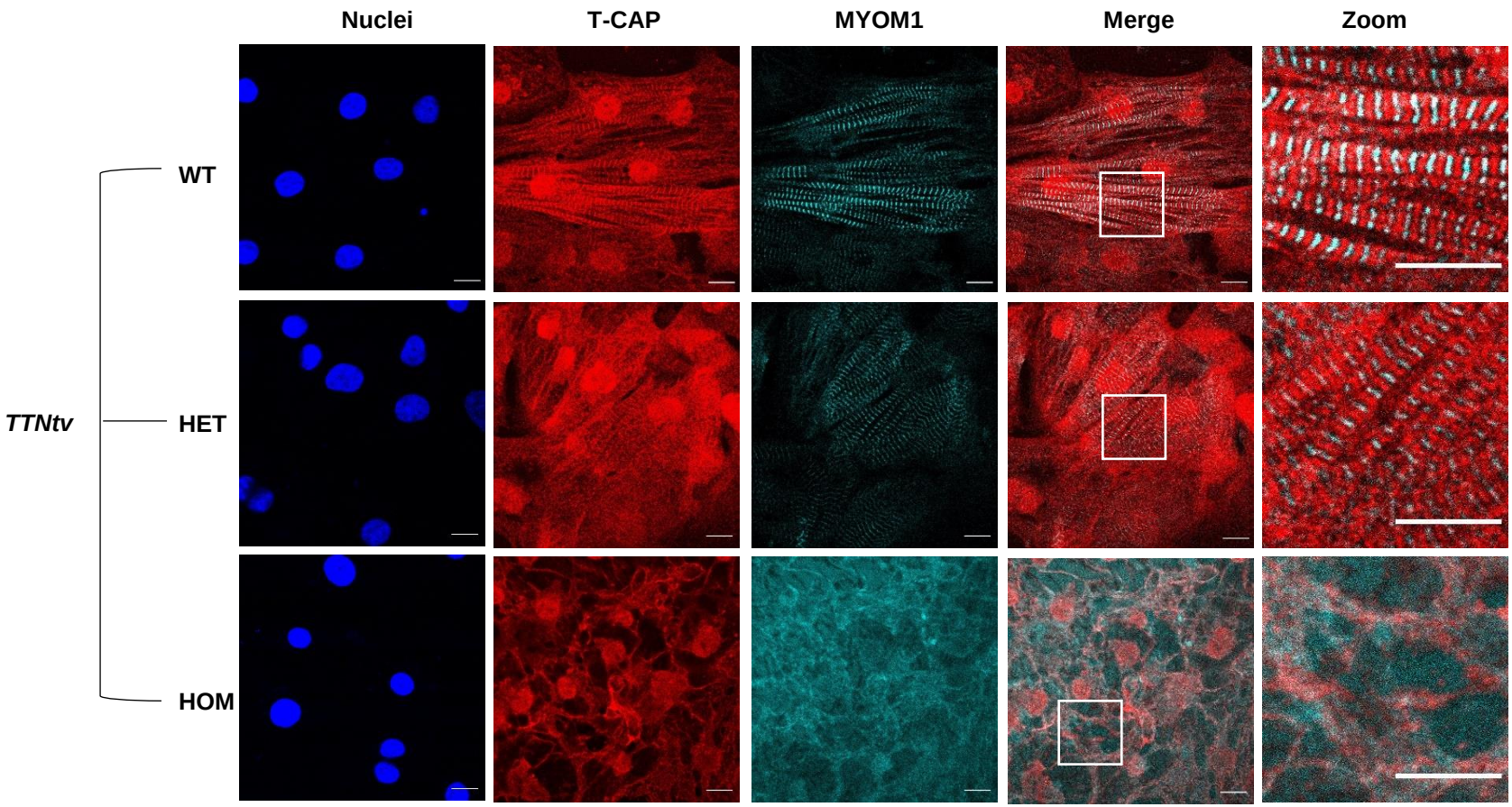


Figure 6-25 Normal Z-disc and M-line protein localisation in atrial-like *MYL7-mClover* iPSC-CMs harbouring *TTN* c.59926+1 G>A. IF staining for T-CAP, MYOM1 and nuclei (DAPI) of day 31 iPSC-CMs. Merge is an overlay between T-CAP and MYOM1 images. Zoom images are digital zoom images from the square area destined in Merge. Scale bar 10 µm.

6.5 Discussion

The *MYL7-mClover* line was initially designed as an atrial-specific fluorescent reporter line which was intended to be used in atrial-specific studies. The potential applications of the *MYL7-mClover* reporter systems include isolation of atrial-like cells for use in modelling atrial diseases; this is especially relevant as *TTNtv* has been associated with lone atrial fibrillation (Ahlberg et al., 2018, Andreasen et al., 2020, Olesen et al., 2018) and ventricular arrhythmia in DCM patients (Ahlberg et al., 2018, Corden et al., 2019, Tayal et al., 2017). Therefore, *TTNtv* is a potential candidate for exploring the utility of subtype-specific fluorescent reporter iPSC-CMs. This chapter specifically focused on characterisation of *TTN* c.59926+1 G>A consequence(s) with regards to atrial-like and ventricular-like *MYL7-mClover* iPSC-CMs. However, the characterisation of the *MYL7-mClover* line was incomplete, where the ability of this system to facilitate atrial-like iPSC-CM purification remains to be investigated (Chapter 4). Thus, the subtype-specific *TTN* c.59926+1 G>A characterisation relied on using the atrial (Chapter 4) and ventricular (Chapter 2) differentiation protocols.

I) Generation of TTNtv iPSC line

Using CRISPR/Cas9, *TTN* c.59926+1 G>A was successfully integrated into the *MYL7-mClover* iPSCs. Both *TTNtv* heterozygous and homozygous lines were produced. The presence of this *TTNtv* inhibited the normal RNA splicing event between exon 303 and 304, thereby contributing to the intron retention and A-band titin truncation. This was in agreement with the observations in the KOLF2-C1 line (Chapter 5).

II) Characterisation of homozygous TTNtv

To further investigate *TTN* c.59926+1 G>A consequences with regards to CM subtype, both the *TTNtv* heterozygous and homozygous lines were differentiated towards the atrial or ventricular phenotype. At certain levels, it appeared that the phenotypes in the homozygous line resembled immature CMs with low *MYH7/MYH6* and *MYL2/MYL7* ratios and prolonged

Ca^{2+} transient characteristic in both the atrial and ventricular-like settings (Karbassi et al., 2020).

However, at the structural level, homozygous *TTN* c.59926+1 G>A inhibited sarcomerogenesis, even though the truncated titin lacking the M-line was likely able to form a complex with presumed nascent myofibrils containing ACTN2, F-actin and TNNT2. The M-line region of titin contains a kinase domain, phosphorylation sites and various binding sites for its partner proteins e.g. MYOM1 and MURF1 (Weinert et al., 2006, Mayans et al., 1998). Consistent with the observation in this chapter, targeted biallelic deletion of the whole titin M-line domain was showed to arrest myofibrillogenesis at the early stage in a mouse ESC-CM model (Musa et al., 2006). Nevertheless, the observation in the homozygous setting alone does not directly give a clinical insight as the patients present heterozygous *TTNtv* mutations.

III) Characterisation of heterozygous TTNtv in ventricular-like iPSC-CMs

The ventricular-like, heterozygous *TTNtv* iPSC-CMs presented no obvious molecular remodelling (no indication of upregulation of heart failure marker expression, myofibrillar isoform switching and changes in *TTN* expression) and Ca^{2+} transient alteration. Interestingly, the sarcomere contractility analysis suggested a delay in sarcomere maturation according to the shortened sarcomere length and the increased sarcomere length variation. However, there were no changes in sarcomere fractional shortening and Ca^{2+} transient kinetics at optimal culture conditions.

At the structural level, T-CAP and MYOM-1 were normally detected in the ventricular-like cells. However, the T-CAP immunofluorescence study in iPSC-CMs may need a careful interpretation as this protein is poorly expressed in immature CMs (Di Baldassarre et al., 2018). Investigation in mature iPSC-CMs will confirm this finding. With regards to the truncated titin stability, there were no myofibrillar fragments (or aggregates) similar to those found in the homozygous line identified in the heterozygous line, likely suggesting that they may be effectively removed by the PQC system. However, the possibility that the truncated titin inserts

itself into the sarcomere in the heterozygous line is not totally excluded as such investigations were not initiated. Further protein quantification assays e.g. WB for titin in the myofibrillar fraction versus the cytoplasmic fraction would allow clarification of truncated titin's ability to incorporate into the sarcomere.

IV) Characterisation of heterozygous TTNtv in atrial-like iPSC-CMs

In agreement with the observation in the ventricular-like iPSC-CMs, there was no evidence of molecular remodelling (regarding induction of heart failure marker expression, myofibrillar isoform switching and *TTN* expression) and sarcomere structure alteration in the atrial-like iPSC-CMs containing the heterozygous *TTNtv*.

However, the Ca^{2+} transient kinetics appeared to be faster as compared to the WT control, suggesting changes in either sarcomeric Ca^{2+} sensitivity or Ca^{2+} handling. Notably, more biological replicates for this measurement will be required to confirm this finding as the data presented in this thesis was from only one batch of iPSC-CMs. It is important to note that the Ca^{2+} transient analysis in this study relied on a cytoplasmic Ca^{2+} indicator RGECO. Using e.g. troponin I/ troponin T-conjugated RGECO (Sparrow et al., 2019) would provide a more detailed insight into the involvement of myofilament in Ca^{2+} sensitivity, which could better elucidate whether *TTN* c.59926+1 G>A mediates changes in the myofilament Ca^{2+} sensitivity. The observed acceleration in the Ca^{2+} transient dynamics may be alternatively explained by increased Ca^{2+} uptake and release activity. This notion needs a further molecular confirmation, especially protein quantification for Ca^{2+} regulatory proteins e.g. RYR, SERCA, PLN and SLN. Noteworthy, SLN is specifically expressed in atrial CMs (Ng et al., 2010) and any increased SLN activity would explain the observed change in the Ca^{2+} decay kinetics specific to the atrial-like cells.

Whether the altered Ca^{2+} transient dynamics could impair sarcomere contractility also needs further investigations. However, following a couple of attempts, the atrial-like cells, both those carrying WT *TTN* and *TTNtv*, failed to be absolutely paced at 2 Hz after a few minutes, making

SarcTrack analysis impossible. This anomalous behaviour may be associated with the intrinsic arrhythmogenicity of atrial-like iPSC-CMs (Goldfracht et al., 2020). It is uncertain whether c.59926+1 G>A increases the arrhythmogenic risk in the atrial-like iPSC-CMs since there was occasionally a faster spontaneous contraction rate observed. A detailed electrophysiological characterisation would clarify the association of the *TTNtv* and arrhythmia in iPSC-CMs.

V) Potential application of MYL7-mClover reporter line in disease studies

Although the alteration of *TTNtv*-associated Ca^{2+} transient dynamics in the atrial-like cells needs further validation e.g. by investigating their electrophysiology and biomechanical stress response, it highlights that atrial-specific fluorescent reporter systems such as the *MYL7-mClover* line may be useful for a selective disease study. Notably, the presence of the *TTNtv* did not interfere with the *MYL7*/*mClover* expression; nor did the presence of the fluorescent reporter alter the *TTNtv*-associated phenotype. Once the *MYL7-mClover* reporter line is confirmed as a reliable atrial reporter system, the mClover fluorescence would allow FACS-assisted atrial-like cell purification. The "purer" atrial-like iPSC-CMs would provide more accurate insights into an association between the *TTNtv* and atrial diseases e.g. atrial fibrillation.

6.5.1 *TTNtv* delays sarcomere maturation

Though it was suggested by semi-quantitative PCR that there was no allelic imbalance of *TTN* transcriptional expression in the atrial-like and ventricular-like, heterozygous *TTNtv* models, the situation at the protein level is unexplored. As discussed in Chapter 4, titin protein gel electrophoresis is technically challenging due to the gigantic protein size. Nevertheless, quantification of the titin protein expression would allow an understanding of the stability of the truncated titin in the heterozygous setting.

The co-localisation of ACTN2 and the truncated titin protein observed in the homozygous *TTN* c.59926+1 G>A line suggested that the truncated titin may be able to bind to ACTN2. This situation may also present in the heterozygous line although no direct protein-protein

interaction studies e.g. co-immunoprecipitation were initiated. At the early stage of myofibrillogenesis, both the WT and the truncated titin may form a complex with nascent myofibrils; while only those myofibrils with the full-length titin could initiate sarcomerogenesis (Chopra et al., 2018), the presence of the unusual nascent myofibrils with the truncated titin may hinder such a process, causing a delay in sarcomere maturation as depicted by the observed shortened sarcomere length accompanied by the increased variance of the sarcomere length.

6.5.2 *TTNtv* and dysregulated protein quality control (PQC)

The sarcomere is a dynamic structure where its protein components undergo constant turnover. Maintenance of this contractile unit relies on the balance between sarcomere protein synthesis and degradation. Impaired PQC has been implicated in common cardiac diseases including DCM, HCM and heart failure (Willis et al., 2014).

UPS and autophagy function as the interconnected central PQC pathways in eukaryotes, where a downregulation of one system activates the other (Willis et al., 2014). While the autophagy (macroautophagy in this context) catalytically hydrolyses long-lived proteins, insoluble protein aggregates and unused organelles encapsulated in a double-membrane structure called the autophagosome, the UPS targets ubiquitinated short-lived proteins and misfolded protein for degradation at the so-called proteasome organelle.

Several myofibrillar proteins were shown to be targeted redundantly by E3-ubiquitin ligases called “MURF1/2” (Figure 2-26) for degradation by the UPS e.g. troponin I, troponin T, titin and T-CAP (Witt et al., 2005). It was highlighted by (Fomin et al., 2021) that MUFR1 protein expression was depleted in *TTNtv*-containing, end-stage DCM myocardial tissues as compared to either idiopathic DCM or healthy indivi.l,dual samples. This study also illustrated an impairment in ubiquitination of the truncated titin, which likely supported the increased titin-

positive aggregate formation, whereas hyper-ubiquitination was presented in the WT titin, which may explain titin haploinsufficiency and compromised contractility.

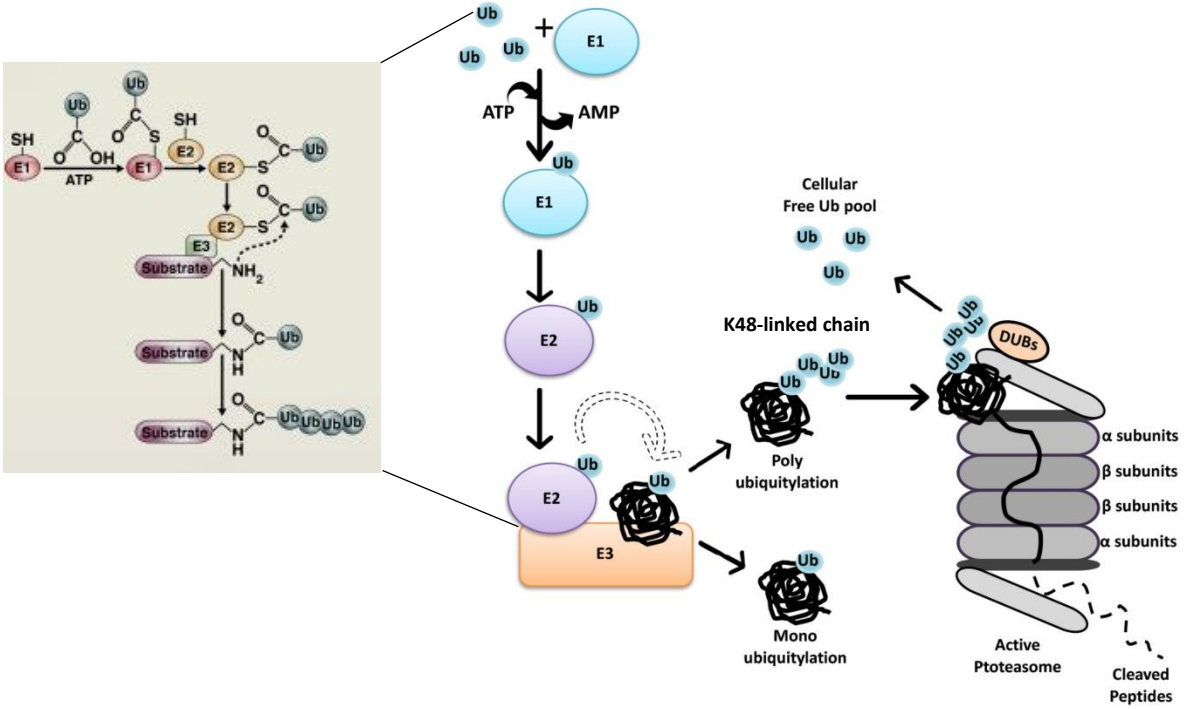


Figure 6-26 UPS mechanism. UPS relies on an ubiquitin signal attached to a target protein by a process known as ubiquitination, which is reversible by deubiquitination enzymes (DUBs). Ubiquitination involves 3 cascade steps catalysed by E1 (ubiquitin activating), E2 (ubiquitin-conjugating) and E3 (ubiquitin ligase) enzymes. Ubiquitin is firstly activated via ATP hydrolysis and covalently bound to E1. The active ubiquitin molecule is then transferred to E2. Finally, the ubiquitin is attached to specific target proteins. The ubiquitin on the target protein itself also serves as a substrate for ubiquitination, making the targeted protein become polyubiquitinated. Lysine 48-linked ubiquitin chain is important for targeting the protein for proteasomal degradation whereas monoubiquitination involves in other functions e.g. protein transport, DNA repair and histone regulation. The 26S proteasome is an ATP-dependent protease comprising of the 20S and 19S components. The 19S cap involves recognising and unfolding of misfolded substrate before passing it into the 20S protease catalytic core formed by α and β subunits. Modified from (Kocaturk and Gozuacik, 2018) and (Adhikari and Chen, 2009).

MURF1 has been implicated in both skeletal and cardiac diseases. MURF1 mediated skeletal muscle atrophy where KO of the gene encoding MURF1 protected rat and mouse models against atrophy (Bodine et al., 2001, Labeit et al., 2010). In the human heart, loss of function mutations (p.A48V, p.I130M and p.Q247*) in MURF1 are associated with HCM, where an apparent impairment in myofibrillar protein ubiquitination and UPS-mediated degradation were shown in mice bearing the mutations (Chen et al., 2012). Meanwhile, overexpression of the gene encoding MURF1 destabilised the sarcomeric M-line structure in chick CMs (McElhinny et al., 2002) and caused left ventricular wall thinning and increased susceptibility to heart failure upon pressure overload in mice (Willis et al., 2009).

6.5.3 Length-dependent Ca^{2+} sensitivity of tension/activation

The Frank-Starling mechanism (Frank, 1895, Knowlton and Starling, 1912) describes the situation where cardiac force of contraction and stroke volume (the amount of blood ejected during each contractile cycle) increase in response to increased ventricular volume during diastole. The elevated ventricular filling augments preload, the initial CM stretching before contraction. Here, myofilament Ca^{2+} sensitivity is enhanced at a longer sarcomere length, possibly through modified myofilament lattice spacing and phosphorylation status of e.g. troponin I and MYBPC3, thereby improving force development – this is known as length dependent Ca^{2+} sensitivity of tension/activation (Chung et al., 2016).

Titin is another protein that plays a role in length-dependent activation. A DCM rat model carrying a homozygous *Rbm20* mutation (loss-of-function) presented impaired cardiac *TTN* splicing with exclusive expression of an unusually large N2-BA isoform, depressed passive tension, abolished Frank-Starling response and reduced length dependent activation (Ait-Mou et al., 2016). (Ait-Mou et al., 2016) reasoned that the titin strain mediated the length dependent activation via stretch-induced structural rearrangement of thick and thin filaments which promoted actomyosin interaction and hence, force production at the greater sarcomere length.

It is intriguing to study whether *TTN* c.59926+1 G>A can also modify length dependence of activation which might subsequently trigger DCM. Analysis of isolated cardiac myofibrils from end stage DCM patients carrying A-band *TTNtv* (c.70387_70394 del(GAAA)2insGAAA and c.56769 T>A) showed no alterations in sarcomere length, sarcomere relaxation kinetics and length dependence of maximal force development as compared to DCM samples without *TTNtv* or heathy donors (Vikhorev et al., 2017). However, relying on the same samples, (Vikhorev et al., 2022) identified enhanced Ca^{2+} sensitivity of force and loss of the length dependence of Ca^{2+} sensitivity attributable to depleted phosphorylation of troponin I and MYBPC3.

In this chapter, the sarcomere length appeared to be slightly shorter in the ventricular-like iPSC-CMs carrying heterozygous *TTN* c.59926+1 G>A without a prominent contribution to the sarcomere relaxation and the Ca^{2+} transient kinetics, likely suggesting unaltered myofilament Ca^{2+} sensitivity. It was unfortunate that the contractility data of the atrial-like iPSC-CMs was unavailable. Assuming a similar sarcomere length decrease in the atrial-like iPSC-CMs, this may suggest an implication of the *TTNtv* in reduced Ca^{2+} sensitivity, which translates into the observed acceleration of the Ca^{2+} decay kinetics.

6.5.4 Arrhythmogenic potential of RA-induced atrial-like iPSC-CMs

Atrial-like iPSC-CMs generated by RA induction were shown to be highly susceptible to spontaneous arrhythmia though reversible by electrical field potential stimulation or using antiarrhythmic drugs (Goldfracht et al., 2020). This presents a challenge for studies of atrial fibrillation associated with genetic mutations; whether the mutations directly initiate arrhythmogenesis or whether they are secondary factors that exacerbate the intrinsic phenotype of the atrial-like cells is uncertain.

The observed faster Ca^{2+} transient kinetics may be indicative of increased arrhythmogenic potential in the atrial-like cells carrying heterozygous c.59926+1 G>A although electrophysiological studies would be required for validation. The fact that the altered Ca^{2+} transient phenotype was absent in the ventricular-like cells may suggest that the Ca^{2+} handling aberration was caused by an altered activity of atrial-specific component(s).

Both PLN and SLN function as negative regulators of the SERCA activity via modulating the apparent Ca^{2+} affinity and Ca^{2+} transport velocity (Asahi et al., 2002). While PLN is highly expressed in ventricular CMs, SLN is specifically expressed in atrial CMs throughout development (Ng et al., 2010), which makes it a potential candidate for developing a future atrial-specific fluorescent reporter (Chen et al., 2016, Josowitz et al., 2014, Chirikian et al., 2021). A reduction in the SLN activity increases Ca^{2+} reuptake from the cytosol into the SR and hence accelerates the Ca^{2+} decay kinetics. Therefore, the shortened time to Ca^{2+} decay

detected in the atrial-like cells carrying heterozygous *TTN* c.59926+1 G>A may be attributable to reduced SLN expression, which remains to be confirmed by e.g. WB.

Enhanced Ca^{2+} cycling can drive atrial fibrillation development (Yeh et al., 2008). *Sln*-KO mice showed enhancement in SERCA Ca^{2+} pump activity, Ca^{2+} transient amplitude and contractility complemented with blunt β-adrenergic response in the atria but not in the ventricles (Babu et al., 2007). *Sln* null mice were also susceptible to atrial fibrillation when aged, accompanied by increased spontaneous Ca^{2+} oscillation incidents, triggered activity frequency and tissue fibrosis in the atria (Xie et al., 2012). In human, SLN deficiency in the atria was identified in patients suffering from atrial fibrillation and heart failure independent from changes in expression of other Ca^{2+} handling proteins (Uemura et al., 2004, Shanmugam et al., 2011).

6.5.5 Concluding remarks

TTN c.59926+1 G>A was successfully engineered in the *MYL7-mClover* reporter iPSC. This *TTNtv* caused retention of the intron immediately downstream of exon 303 and generated a premature stop codon, leading to truncation of titin from the early A-band region.

The homozygous *TTNtv* arrested myofibrillogenesis resulted in immature CM phenotypes with reduced *MYH7/MYH6* and *MYL2/MYL7* expression ratios and lengthened Ca^{2+} transient duration in both the ventricular-like and atrial-like iPSC-CMs.

The heterozygous *TTNtv* induced no apparent molecular remodelling (overall *TTN* expression, heart failure marker expression and myofibrillar isoform switching) and sarcomere disorganisation in both the ventricular-like and atrial-like iPSC-CMs.

The ventricular-like cells carrying the heterozygous *TTNtv* showed a modest reduction of sarcomere length and increased variation in the sarcomere length, suggesting delayed sarcomere maturation. Although the cells showed slightly faster contraction kinetics, there was no change in the relaxation kinetics or the fractional shortening. In addition, no aberrant Ca^{2+} transient dynamics were observed in these cells. Thus, overall, no obvious effect of the heterozygous *TTNtv* on functional behaviour was detected in the ventricular-like iPSC-CMs.

However, an acceleration of Ca^{2+} transient was present in the atrial-like cells carrying the heterozygous *TTNtv*, likely suggesting changes in atrial-specific Ca^{2+} handling and/or myofibrillar Ca^{2+} sensitivity.

In summary, *TTN* c.59926+1 G>A was characterised in a subtype-specific manner in the *MYL7-mClover* iPSC-CMs by relying on the atrial and ventricular differentiation protocols. The presence of the *TTNtv* did not interfere with the *MYL7/mClover* expression. As the aberrant Ca^{2+} transient kinetics were observed only in the atrial-like cells carrying the heterozygous *TTNtv*, the atrial-specific *MYL7-mClover* reporter could be used in future to generate purer atrial-like cell populations, allowing a detailed characterisation of this *TTNtv*.

Chapter 7

Discussion

7.1 Generation of subtype-specific fluorescent reporter lines

Advances in iPSC-CM differentiation have provided an infinite source of cells for cardiovascular research. However, the iPSC-CM population's heterogeneity challenges selective studies. This thesis aims at addressing this issue by designing fluorescent reporter systems that rely on the endogenous promoter activity present in the different parts of the adult heart – *MYL7* and *PITX2* in the atria, *MYL2* in the ventricles and *HCN4* in the nodal structures (Ng et al., 2010). The fluorescent expression from these systems could allow identification and isolation of the respective CM subtypes for subsequent uses.

The complementarity between gRNA and target sequences, together with the homology between the fluorophore donor template and sequences surrounding the target site allows CRISPR/Cas9-assisted generation of fluorescent reporter lines. The *MYL7-mClover*, *PITX2-mClover* and *HCN4-mSarlet* reporter lines were successfully generated by relying on the targeting plasmids carrying ~1000-bp HAs. The high HDR efficiency (>60%) of these 3 targeting projects were explained by using the most efficient CRISPR/Cas9 delivery method "RNP electroporation" (Burger et al., 2016) and antibiotic selection. However, repetitive motifs within the homology region may introduce errors during homologous recombination and reduce the targeting efficiency (Pascarella et al., 2020, George and Alani, 2012). Due to this concern, an lssDNA with relatively short HAs (~50-nt) was used as a donor template for the *MYL2-mClover* line. The complication of using the lssDNA is its incompatibility with RNP electroporation, as electroporation has proven an inefficient method of delivery of these lssDNA molecules. Chemical transfection of RNP was also not particularly efficient in our experience. Therefore, the lssDNA was chemically transfected into iPSCs together with a CRISPR/Cas9 expression plasmid. However, this method resulted in a poor HDR efficiency (~0.004%).

ssDNAs are advantageous in terms of HDR efficiency, cellular toxicity and off-target event when compared to dsDNA donors (Ranawakage et al., 2021, Canaj et al., 2019). However, this likely applies only to short ssDNAs e.g. ssODN. Recent work has shown that dsDNA (plasmid) presented a more efficient template for insertion of a large fluorophore construct in iPSCs (~700 bp, close to the size of *mClover* and *mScarlet*) compared to its equivalent lssDNA (~10% versus ~50% HDR) while the opposite outcome was reported for a relatively small construct (~90 bp) (~85% versus ~25% HDR) (Canaj et al., 2019). In addition, the longer life time of the CRISPR/Cas9 plasmid relative to RNP preparations of CRISPR/Cas9 also results in a prolonged expression, allowing continue cleaving within the *MYL2* locus, and thus potentially introducing deletion mutations at the non-targeted allele.

An alternative to lssDNA, targeting plasmids with short HAs allowing targeted insertion of fluorophores at the *MYL2* locus have been published recently using plasmid donors with ~500-bp HAs (Chirikian et al., 2021, Galdos et al., 2021, Luo et al., 2021). These works relied on a similar targeting strategy described in this thesis – using CRISPR/Cas9 to replace the original stop codon with a P2A-*fluorophore* construct. Hence, this method could be used to improve the efficiency for generation of the *MYL2-mClover* reporter line.

For all reporter lines, the editing efficiency may be further enhanced by using chemically modified gRNAs (Riesenberg et al., 2022). The chemical modification ensures gRNA stability, allowing more time for CRISPR/Cas9 to access its target sequences and generate DSBs, a prerequisite for HDR-mediated insertion. Increased cutting efficiency may also eliminate the necessity of including a selection cassette in a targeting vector (and antibiotic use), avoiding a need of subsequent excision of the cassette.

7.2 Faithful reporter systems: do they enable specific iPSC-CM subtype identification?

All of the reporter lines were differentiated into CMs by WNT signalling modulation. The *PITX2-mClover* line lost the cardiac differentiation possibly due to poor iPSC handling. Two clones of the *MYL2-mClover* line contained deletion mutations on the non-targeted alleles and did

not express both MYL2 and mClover proteins. One clone displayed a delay in the differentiation, resulting in no transcriptional expression of *MYL2* and *mClover*. The other clone showed normal differentiation and transcriptional expression of *MYL2* and *mClover*; therefore, the loss of MYL2 may be explained by protein instability associated with the P2A sequence and the deletion mutation (Manivannan et al., 2020), while the loss of mClover may be associated with the translational fall-off mechanism (Liu et al., 2017).

Time course analysis in both the *MYL7-mClover* and *HCN4-Scarlet* lines suggested that the fluorophores recapitulated the *MYL7* and *HCN4* expression, respectively. Therefore, the fluorescence signals could be used to track expression of *MYL7* or *HCN4* in individual cells. RA was shown to promote an atrial differentiation with atrial-type Ca^{2+} transient, contractility and molecular features in the *MYL7* reporter line. Though RA is known to upregulate *MYL7* expression (Lemme et al., 2018, Miao et al., 2020, Cyganek et al., 2018), this effect was not clearly observed in the *MYL7* reporter iPSC-CMs, possibly indicating that the cells had an immature phenotype (Gunawan et al., 2021). Maturation of iPSC-CMs by 3-dimentional culture was previously illustrated to drive differential *MYL7* expression between RA-treated and non-treated condition (Goldfracht et al., 2020). Fine tuning the RA concentration (Devalla et al., 2015) may allow further development of the atrial differentiation protocol. In future work, (*MYL7*-driven) mClover fluorescence quantification e.g. by a simple plate reader or flow analysis could guide which conditions better promote the atrial pathway. Further enrichment of atrial-like population could be accomplished by FACS-sorting for cells which highly express the mClover marker. The mClover expression also showed a well-to-well variation when induced by RA even though the cells were differentiated from a single batch of iPSCs, which can hinder a reproducibility of iPSC-CM phenotype. To overcome this issue, iPSC-CM differentiation and phenotyping would need to be performed using an automated platform (Paull et al., 2015) or by several experimental replicates.

Despite being a specific marker for adult nodal CMs, *HCN4* is generally expressed in iPSC-CMs irrespective of subtype (Yechikov et al., 2016). The expression of *HCN4* partly explains

automaticity though the majority of iPSC-CMs are ventricular-like populations, where adult ventricular cells do not express this gene and do not beat spontaneously. I therefore believe that the *HCN4*-driven *mScarlet* expression may not faithfully indicate nodal-like population in immature iPSC-CMs. Previous studies have shown that upon iPSC-CM maturation, *HCN4* expression was downregulated and became more restricted to nodal-like CMs, corresponding to a reduction in spontaneous contractile activity (Ichimura et al., 2020, Bosman et al., 2013, Yechikov et al., 2016). Further FACS analysis of expression of nodal markers e.g. ISLET1 SHOX2 (Ichimura et al., 2020) in HCN4/*HCN4*-driven mScarlet positive populations between cells differentiated by the conventional method and cells differentiated by a maturation-promoting protocol e.g. 3-dimentional culture would clarify whether *HCN4* is a specific nodal marker in mature iPSC-CMs. Alternatively, investigating how the *HCN4*/*HCN4*-driven *mScarlet* expression responds to a nodal differentiation protocol e.g. by modulating NODAL signalling (Yechikov et al., 2020) would confirm if the *HCN4-mScarlet* reporter allows nodal-like cell identification. Nevertheless, these future experiments may require a new clonal reporter line because the clone analysed in this thesis showed a defect in cardiac differentiation potential, which is likely associated with poor handling or prolongued culture of the iPSCs. Since the differentiation potential was retained in some cells, subcloning might be a way forward to acquire a better performing clone.

7.3 Subtype-specific TTNtv modelling

TTNtv is the most common genetic cause of atrial fibrillation (Choi et al., 2018) and DCM (Roberts et al., 2015). However, their underlying mechanism are poorly understood. iPSC-CMs have been increasingly used to study the disease mechanisms of *TTNtv* in DCM (Schick et al., 2018, Chopra et al., 2018, Fomin et al., 2021, Gramlich et al., 2015, Hinson et al., 2015) but to my current knowledge, no mutations have been reported for atrial fibrillation, possibly due to the lack of an atrial-like cell model. Due to its contribution in both atrial and ventricular CMs, this thesis characterised *TTNtv* in a subtype-specific manner.

TTN c.59926+1 G>A, a splice variant located in the A-band, is associated with atrial fibrillation and ventricular tachycardia in DCM patients (Hoorntje et al., 2018). This mutation was introduced into the parental KOLF2-C1 iPSCs for baseline characterisation before being incorporated into the *MYL7-mClover* iPSCs. Since the question of whether the *MYL7-mClover* reporter allows atrial-like cell purification was not sufficiently addressed, a subtype-specific *TTNtv* characterisation had to rely on using the ventricular and RA-induced atrial differentiation protocols.

The mutation caused titin A-band truncation, though a clear phenotype was only detected in the homozygous model. Homozygous *TTNtv* completely disrupted sarcomerogenesis and restricted differentiation to more embryonic-like stages. However, such phenotypes were absent in the heterozygous model. This contradicts previous studies which described heterozygous A-band *TTNtv*-associated sarcomere disorganisation in iPSC-CMs (Chopra et al., 2018, Hinson et al., 2015). However, my model supports previous investigations in human biopsies (McAfee et al., 2021, Fomin et al., 2021, Vikhorev et al., 2017) and a zebrafish model (Huttner et al., 2018), which showed normal sarcomere alignment/striation in the heterozygous state. The difference may arise from different sites of the *TTNtv* mutation or differences in experimental approaches.

In addition, the heterozygous *TTN* c.59926+1 G>A mutation did not result in transcriptional changes in both the atrial-like and ventricular-like iPSC-CMs with regard to myofilament isoforms, *TTN* expression levels and markers for heart failure. In the ventricular-like model, the mutation led to a decrease in sarcomere length and an increase in variation of this parameter, suggesting delayed sarcomere maturation. I believe that truncated titin may trigger protein degradation as the mutant protein impeded sarcomerogenesis. At early stages, truncated titin proteins may be efficiently removed by protein degradation pathways as suggested by the fact that there were no myofibrillar aggregates detected. It is worth reiterating that iPSC-CMs displayed immature phenotypes, thus better recapitulating early development and pathogenesis. Meanwhile, continuous expression of truncated titin proteins may

aggravate PQC function, leading to accumulation of such mutant proteins, increased myofibrillar protein degradation (i.e. sarcomere loss) and impaired myocardial function as seen in end stage DCM (Fomin et al., 2021). Time course characterisation of *TTNtv*-carrying iPSC-CMs would clarify this notion, considering that prolonged culture promotes CM maturation (Ahmed et al., 2020).

The only different feature between the atrial-like and ventricular-like iPSC-CMs with respect to heterozygous *TTN* c.59926+1 G>A consequence was the Ca^{2+} signalling. While normal Ca^{2+} transient dynamics were observed in the ventricular-like model, it appeared that the atrial-like cells displayed acceleration in both Ca^{2+} upstroke and decay velocity, representing increased Ca^{2+} release and reuptake. Increase Ca^{2+} transport activity, both by SERCA and RYR, leads to delay after depolarisation i.e. spontaneous depolarisation following a complete cycle of action potential (Landstrom et al., 2017). If this unusual phenomenon brings the membrane potential beyond the threshold, it can result in so-called triggered activity, generating abnormal, fast atrial rhythm, a characteristic of atrial fibrillation.

7.4 Future directions

7.4.1 Validation of subtype-specific fluorescent reporter lines

To validate the subtype-specific property, the *MYL7-mClover* and *HCN4-mScarlet* iPSC-CMs will need to undergo cell isolation by FACS. Electrophysiological study of the mClover and mScarlet positive fractions will indicate whether they represent atrial-like and nodal-like populations, respectively. Further, comparative expression profiling for the sorted fractions and pre-sorting sample would help to indicate whether the fluorescent reporter systems are useful tools for enrichment of their respective CM subtype. They could then be applied to disease modelling questions, such as the role of *TTNtv* in atrial cells.

7.4.2 Dissecting *TTNtv* pathogenesis

The atrial-specific *TTNtv* phenotype presents an opportunity for use of the *MYL7-mClover* reporter line. Though atrial-like iPSC-CMs can be obtained by RA induction, isolation of these

subtypes from pools of cells may further enrich the atrial-specific phenotype. This could be achieved by FACS-aided purification of cells expressing high levels of mClover.

Work being carried out by Max Cumberland (University of Birmingham) is focussing on the electrophysiology of *TTN* c.59926+1 G>A bearing atrial-like and ventricular iPSC-CMs. This will clarify the consequence of the altered atrial-selective Ca^{2+} signalling in action potential generation. A combination of the Ca^{2+} data and electrophysiology will lead to confirmation of the intrinsic arrhythmogenic potential of the *TTNtv*.

To further understand the molecular pathway(s) underlying the Ca^{2+} transient changes, key Ca^{2+} handling proteins need to be analysed, both quantitatively (levels of expression and post-translational modification) and functionally. With respect to the atrial-selective phenotype, SLN is a top candidate, in that reduction in SLN expression enhances SERCA activity (Babu et al., 2007) and predisposes to atrial fibrillation (Xie et al., 2012).

As titin protein turnover is regulated by the UPS (Witt et al., 2005), quantification of titin protein level, titin ubiquitination and ubiquitin ligase MURF1/2 level is required to determine if *TTNtv* alters the UPS activity and mediates sarcomere insufficiency, thereby disrupting CM's function. All together, these studies would benefit our understanding of *TTNtv* pathogenesis.

7.5 Concluding remarks

This thesis explored the application of a fluorescent reporter in discriminating iPSC-CM subtypes. *MYL7-mClover*, *HCN4-mScarlet*, *PITX2-Clover* and *MYL2-mClover* reporter lines were established using CRISPR/Cas9. Unfortunately, the *PITX2-mClover* line lost its differentiation potential and the *MYL2-mClover* line did not express MYL2 and mClover proteins. The *MYL7-mClover* and *HCN4-mScarlet* lines, however, constitute faithful reporter systems where the expression of the fluorophores correlate with the expression pattern of their target markers. Whether the *MYL7* and *HCN4* reporters could enable FACS-assisted purification of atrial-like and nodal-like cells, respectively, remains to be confirmed.

This thesis also sought to identify the subtype-specific contribution to the *TTNtv* disease mechanism. Heterozygous *TTN* c.59926+1 G>A did not induce myofibrillar protein isoform switching, expression of heart failure markers and overall *TTN* transcriptional changes in both atrial-like and ventricular-like iPSC-CMs. In the homozygous model, the truncated titin arrested myofribrillogenesis, leading to accumulation of myofibril aggregates. However, normal sarcomere structure was preserved in the heterozygous model irrespective of the CM subtype, likely suggesting effective removal of the truncated titin proteins. The *TTNtv*-associated increased protein degradation activity may also increase myofilament turnover and sarcomere insufficiency, which could underlie the myocardial dysfunction in DCM. The only atrial-specific phenotype observed was an abbreviated Ca^{2+} transient. Though awaiting electrophysiology confirmation, this aberration could influence the risk of atrial fibrillation.

Further exploration of the phenotype and the utility of the engineered *MYL7-mClover* line for atrial CM purification may present an opportunity to study atrial-specific diseases such as atrial fibrillation caused by *TTNtv* or other mutations. It could also be used to study the effects of pharmacological interventions in enriched atrial-like cell populations.

References

ADHIKARI, A. & CHEN, Z. J. 2009. Diversity of polyubiquitin chains. *Developmental cell,* 16**,** 485-486.

ADIKUSUMA, F., PILTZ, S., CORBETT, M. A., TURVEY, M., MCCOLL, S. R., HELBIG, K. J., BEARD, M. R., HUGHES, J., POMERANTZ, R. T. & THOMAS, P. Q. 2018. Large deletions induced by Cas9 cleavage. *Nature,* 560**,** E8-E9.

AGARKOVA, I. & PERRIARD, J.-C. 2005. The M-band: an elastic web that crosslinks thick filaments in the center of the sarcomere. *Trends in cell biology,* 15**,** 477-485.

AHLBERG, G., REFSGAARD, L., LUNDEGAARD, P. R., ANDREASEN, L., RANTHE, M. F., LINSCHEID, N., NIELSEN, J. B., MELBYE, M., HAUNSØ, S. & SAJADIEH, A. 2018. Rare truncating variants in the sarcomeric protein titin associate with familial and early-onset atrial fibrillation. *Nature communications,* 9**,** 1-11.

AHMED, R. E., ANZAI, T., CHANTHRA, N. & UOSAKI, H. 2020. A brief review of current maturation methods for human induced pluripotent stem cells-derived cardiomyocytes. *Frontiers in cell and developmental biology,* 8**,** 178.

AIRD, E. J., LOVENDAHL, K. N., ST MARTIN, A., HARRIS, R. S. & GORDON, W. R. 2018. Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. *Communications biology,* 1**,** 1-6.

AIT-MOU, Y., HSU, K., FARMAN, G. P., KUMAR, M., GREASER, M. L., IRVING, T. C. & DE TOMBE, P. P. 2016. Titin strain contributes to the Frank–Starling law of the heart by structural rearrangements of both thin-and thick-filament proteins. *Proceedings of the National Academy of Sciences,* 113**,** 2306-2311.

ALANIS-LOBATO, G., ZOHREN, J., MCCARTHY, A., FOGARTY, N. M., KUBIKOVA, N., HARDMAN, E., GRECO, M., WELLS, D., TURNER, J. M. & NIAKAN, K. K. 2021. Frequent loss of heterozygosity in CRISPR-Cas9–edited early human embryos. *Proceedings of the National Academy of Sciences,* 118.

ANDREASEN, L., BERTELSEN, L., GHOUSE, J., LUNDEGAARD, P. R., AHLBERG, G., REFSGAARD, L., RASMUSSEN, T. B., EISKJÆR, H., HAUNSØ, S. & VEJLSTRUP, N. 2020. Early-onset atrial fibrillation patients show reduced left ventricular ejection fraction and increased atrial fibrosis. *Scientific reports,* 10**,** 1-8.

ASAHI, M., KURZYDLOWSKI, K., TADA, M. & MACLENNAN, D. H. 2002. Sarcolipin inhibits polymerization of phospholamban to induce superinhibition of sarco (endo) plasmic reticulum Ca2+-ATPases (SERCAs). *Journal of Biological Chemistry,* 277**,** 26725-26728.

BABU, G. J., BHUPATHY, P., TIMOFEYEV, V., PETRASHEVSKAYA, N. N., REISER, P. J., CHIAMVIMONVAT, N. & PERIASAMY, M. 2007. Ablation of sarcolipin enhances sarcoplasmic reticulum calcium transport and atrial contractility. *Proceedings of the National Academy of Sciences,* 104**,** 17867-17872.

BANG, M.-L., CENTNER, T., FORNOFF, F., GEACH, A. J., GOTTHARDT, M., MCNABB, M., WITT, C. C., LABEIT, D., GREGORIO, C. C. & GRANZIER, H. 2001. The complete gene sequence of titin, expression of an unusual≈ 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. *Circulation research,* 89**,** 1065-1072.

BARRANGOU, R., FREMAUX, C., DEVEAU, H., RICHARDS, M., BOYAVAL, P., MOINEAU, S., ROMERO, D. A. & HORVATH, P. J. S. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science,* 315**,** 1709-1712.

BEDADA, F. B., CHAN, S. S., METZGER, S. K., ZHANG, L., ZHANG, J., GARRY, D. J., KAMP, T. J., KYBA, M. & METZGER, J. M. 2014. Acquisition of a quantitative, stoichiometrically conserved ratiometric marker of maturation status in stem cell-derived cardiac myocytes. *Stem cell reports,* 3**,** 594-605.

BELUS, A. & WHITE, E. 2001. Effects of antibiotics on the contractility and Ca2+ transients of rat cardiac myocytes. *Eur. J. Pharmacol.,* 412**,** 121-6.

BERS, D. M. 2002. Cardiac excitation–contraction coupling. *Nature,* 415**,** 198-205.

BIRD, S., DOEVENDANS, P., VAN ROOIJEN, M., BRUTEL DE LA RIVIERE, A., HASSINK, R., PASSIER, R. & MUMMERY, C. 2003. The human adult cardiomyocyte phenotype. *Cardiovascular research,* 58**,** 423-434.

BIZY, A., GUERRERO-SERNA, G., HU, B., PONCE-BALBUENA, D., WILLIS, B. C., ZARZOSO, M., RAMIREZ, R. J., SENER, M. F., MUNDADA, L. V., KLOS, M., DEVANEY, E. J., VIKSTROM, K. L., HERRON, T. J. & JALIFE, J. 2013. Myosin light chain 2-based selection of human iPSC-derived early ventricular cardiac myocytes. *Stem Cell Res,* 11**,** 1335-47.

BLACKWOOD, E. A., BILAL, A. S., AZIZI, K., SARAKKI, A. & GLEMBOTSKI, C. C. 2020. Simultaneous isolation and culture of atrial myocytes, ventricular myocytes, and non-myocytes from an adult mouse heart. *JoVE (Journal of Visualized Experiments)***,** e61224.

BLOCH, K. D., SEIDMAN, J., NAFTILAN, J. D., FALLON, J. T. & SEIDMAN, C. E. 1986. Neonatal atria and ventricles secrete atrial natriuretic factor via tissue-specific secretory pathways. *Cell,* 47**,** 695-702.

BODINE, S. C., LATRES, E., BAUMHUETER, S., LAI, V. K.-M., NUNEZ, L., CLARKE, B. A., POUEYMIROU, W. T., PANARO, F. J., NA, E. & DHARMARAJAN, K. 2001. Identification of ubiquitin ligases required for skeletal muscle atrophy. *Science,* 294**,** 1704-1708.

BOGDANOV, K. Y., MALTSEV, V. A., VINOGRADOVA, T. M., LYASHKOV, A. E., SPURGEON, H. A., STERN, M. D. & LAKATTA, E. G. 2006. Membrane potential fluctuations resulting from submembrane Ca2+ releases in rabbit sinoatrial nodal cells impart an exponential phase to the late diastolic depolarization that controls their chronotropic state. *Circulation research,* 99**,** 979-987.

BÖGEHOLZ, N., PAULS, P., DECHERING, D. G., FROMMEYER, G., GOLDHABER, J. I., POTT, C., ECKARDT, L., MUELLER, F. U. & SCHULTE, J. S. 2018. Distinct occurrence of proarrhythmic afterdepolarizations in atrial versus ventricular cardiomyocytes: implications for translational research on atrial arrhythmia. *Frontiers in Pharmacology,* 9**,** 933.

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**,** 69-84.

BOOTMAN, M. D., HIGAZI, D. R., COOMBES, S. & RODERICK, H. L. 2006. Calcium signalling during excitation-contraction coupling in mammalian atrial myocytes. *Journal of cell science,* 119**,** 3915-3925.

BOSMAN, A., SARTIANI, L., SPINELLI, V., DEL LUNGO, M., STILLITANO, F., NOSI, D., MUGELLI, A., CERBAI, E. & JACONI, M. 2013. Molecular and functional evidence of HCN4 and caveolin-3 interaction during cardiomyocyte differentiation from human embryonic stem cells. *Stem cells and development,* 22**,** 1717-1727.

BRINKMAN, E. K., CHEN, T., AMENDOLA, M. & VAN STEENSEL, B. 2014. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Research,* 42**,** e168-e168.

BROGNA, S. & WEN, J. 2009. Nonsense-mediated mRNA decay (NMD) mechanisms. *Nature structural & molecular biology,* 16**,** 107-113.

BURGER, A., LINDSAY, H., FELKER, A., HESS, C., ANDERS, C., CHIAVACCI, E., ZAUGG, J., WEBER, L. M., CATENA, R. & JINEK, M. 2016. Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development,* 143**,** 2025-2037.

BURRIDGE, P. W., MATSA, E., SHUKLA, P., LIN, Z. C., CHURKO, J. M., EBERT, A. D., LAN, F., DIECKE, S., HUBER, B. & MORDWINKIN, N. M. 2014. Chemically defined generation of human cardiomyocytes. *Nature methods,* 11**,** 855-860.

CALERO-NIETO, F. J., BERT, A. G. & COCKERILL, P. N. 2010. Transcription-dependent silencing of inducible convergent transgenes in transgenic mice. *Epigenetics Chromatin,* 3**,** 3.

CANAJ, H., HUSSMANN, J. A., LI, H., BECKMAN, K. A., GOODRICH, L., CHO, N. H., LI, Y. J., SANTOS, D. A., MCGEEVER, A. & STEWART, E. M. 2019. Deep profiling reveals substantial heterogeneity of integration outcomes in CRISPR knock-in experiments. *BioRxiv*, 841098.

CHEN, J., KUBALAK, S. W., MINAMISAWA, S., PRICE, R. L., BECKER, K. D., HICKEY, R., ROSS, J. & CHIEN, K. R. 1998. Selective requirement of myosin light chain 2v in embryonic heart function. *Journal of Biological Chemistry,* 273**,** 1252-1256.

CHEN, S. N., CZERNUSZEWICZ, G., TAN, Y., LOMBARDI, R., JIN, J., WILLERSON, J. T. & MARIAN, A. J. 2012. Human molecular genetic and functional studies identify TRIM63, encoding Muscle RING Finger Protein 1, as a novel gene for human hypertrophic cardiomyopathy. *Circulation research,* 111**,** 907-919.

CHEN, Z., XIAN, W., BELLIN, M., DORN, T., TIAN, Q., GOEDEL, A., DREIZEHNTER, L., SCHNEIDER, C. M., WARD-VAN OOSTWAARD, D. & NG, J. K. M. J. E. H. J. 2016. Subtype-specific promoter-driven action potential imaging for precise disease modelling and drug testing in hiPSC-derived cardiomyocytes. *Eur. Heart J.,* 38**,** 292-301.

CHIEN, K. R., KNOWLTON, K. U. & CHIEN, S. 1991. Regulation of cardiac gene expression during myocardial growth and hypertrophy: molecular studies of an adaptive physiologic response. *The FASEB Journal,* 5**,** 3037-3064.

CHIRIKIAN, O., GOODYER, W. R., DZILIC, E., SERPOOSHAN, V., BUIKEMA, J. W., MCKEITHAN, W., WU, H., LI, G., LEE, S. & MERK, M. 2021. CRISPR/Cas9-based targeting of fluorescent reporters to human iPSCs to isolate atrial and ventricular-specific cardiomyocytes. *Scientific reports,* 11**,** 1-10.

CHOI, S. H., WENG, L.-C., ROSELLI, C., LIN, H., HAGGERTY, C. M., SHOEMAKER, M. B., BARNARD, J., ARKING, D. E., CHASMAN, D. I. & ALBERT, C. M. 2018. Association between titin loss-of-function variants and early-onset atrial fibrillation. *Jama,* 320**,** 2354-2364.

CHOPRA, A., KUTYS, M. L., ZHANG, K., POLACHECK, W. J., SHENG, C. C., LUU, R. J., EYCKMANS, J., HINSON, J. T., SEIDMAN, J. G. & SEIDMAN, C. E. 2018. Force generation via β-cardiac myosin, titin, and α-actinin drives cardiac sarcomere assembly from cell-matrix adhesions. *Developmental cell,* 44**,** 87-96. e5.

CHRISTOFFELS, V. M., HABETS, P. E., FRANCO, D., CAMPIONE, M., DE JONG, F., LAMERS, W. H., BAO, Z.-Z., PALMER, S., BIBEN, C. & HARVEY, R. P. 2000. Chamber formation and morphogenesis in the developing mammalian heart. *Developmental biology,* 223**,** 266-278.

CHRISTOFOROU, N., LIAU, B., CHAKRABORTY, S., CHELLAPAN, M., BURSAC, N. & LEONG, K. W. 2013. Induced pluripotent stem cell-derived cardiac progenitors differentiate to cardiomyocytes and form biosynthetic tissues. *PloS one,* 8**,** e65963.

CHU, G., LESTER, J. W., YOUNG, K. B., LUO, W., ZHAI, J. & KRANIAS, E. G. 2000. A single site (Ser16) phosphorylation in phospholamban is sufficient in mediating its maximal cardiac responses to β-agonists. *Journal of Biological Chemistry,* 275**,** 38938-38943.

CHUNG, J.-H., BIESIADECKI, B. J., ZIOLO, M. T., DAVIS, J. P. & JANSSEN, P. M. 2016. Myofilament calcium sensitivity: role in regulation of in vivo cardiac contraction and relaxation. *Frontiers in physiology,* 7**,** 562.

CHURKO, J. M., GARG, P., TREUTLEIN, B., VENKATASUBRAMANIAN, M., WU, H., LEE, J., WESSELLS, Q. N., CHEN, S.-Y., CHEN, W.-Y. & CHETAL, K. 2018. Defining human cardiac transcription factor hierarchies using integrated single-cell heterogeneity analysis. *Nature communications,* 9**,** 1-14.

CHUVA DE SOUSA LOPES, S. M., HASSINK, R. J., FEIJEN, A., VAN ROOIJEN, M. A., DOEVENDANS, P. A., TERTOOLEN, L., BRUTEL DE LA RIVIÈRE, A. & MUMMERY, C. L. 2006. Patterning the heart, a template for human cardiomyocyte development. *Developmental dynamics: an official publication of the American Association of Anatomists,* 235**,** 1994-2002.

CLAUSS, S. & KÄÄB, S. 2011. Is Pitx2 growing up? : Am Heart Assoc.

CLOWES, C., BOYLAN, M. G., RIDGE, L. A., BARNES, E., WRIGHT, J. A. & HENTGES, K. E. 2014. The functional diversity of essential genes required for mammalian cardiac development. *genesis,* 52**,** 713-737.

CORDEN, B., JARMAN, J., WHIFFIN, N., TAYAL, U., BUCHAN, R., SEHMI, J., HARPER, A., MIDWINTER, W., LASCELLES, K. & MARKIDES, V. 2019. Association of titin-truncating genetic variants with life-threatening cardiac arrhythmias in patients with dilated cardiomyopathy and implanted defibrillators. *JAMA network open,* 2**,** e196520-e196520.

CYGANEK, L., TIBURCY, M., SEKERES, K., GERSTENBERG, K., BOHNENBERGER, H., LENZ, C., HENZE, S., STAUSKE, M., SALINAS, G. & ZIMMERMANN, W.-H. J. J. I. 2018. Deep phenotyping of human induced pluripotent stem cell–derived atrial and ventricular cardiomyocytes. *JCI Insight,* 3.

DE VRIES, J. E., VORK, M. M., ROEMEN, T., DE JONG, Y. F., CLEUTJENS, J., VAN DER VUSSE, G. & VAN BILSEN, M. 1997. Saturated but not mono-unsaturated fatty acids induce apoptotic cell death in neonatal rat ventricular myocytes. *Journal of lipid research,* 38**,** 1384-1394.

DEN HARTOGH, S. C. & PASSIER, R. J. S. C. 2016. Concise review: fluorescent reporters in human pluripotent stem cells: contributions to cardiac differentiation and their applications in cardiac disease and toxicity. *Stem Cells,* 34**,** 13-26.

DENNING, C., BORGDORFF, V., CRUTCHLEY, J., FIRTH, K. S., GEORGE, V., KALRA, S., KONDRASHOV, A., HOANG, M. D., MOSQUEIRA, D. & PATEL, A. 2016. Cardiomyocytes from human pluripotent stem cells: from laboratory curiosity to industrial biomedical platform. *Biochim. Biophys. Acta,* 1863**,** 1728-48.

DEUTSCHBAUER, A. M., JARAMILLO, D. F., PROCTOR, M., KUMM, J., HILLENMEYER, M. E., DAVIS, R. W., NISLOW, C. & GIAEVER, G. 2005. Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics,* 169**,** 1915-1925.

DEVALLA, H. D., SCHWACH, V., FORD, J. W., MILNES, J. T., EL-HAOU, S., JACKSON, C., GKATZIS, K., ELLIOTT, D. A., CHUVA DE SOUSA LOPES, S. M., MUMMERY, C. L., VERKERK, A. O. & PASSIER, R. 2015. Atrial-like cardiomyocytes from human pluripotent stem cells are a robust preclinical model for assessing atrial-selective pharmacology. *EMBO Mol Med,* 7**,** 394-410.

DEVINE, W. P., WYTHE, J. D., GEORGE, M., KOSHIBA-TAKEUCHI, K. & BRUNEAU, B. G. 2014. Early patterning and specification of cardiac progenitors in gastrulating mesoderm. *Elife,* 3**,** e03848.

DI BALDASSARRE, A., CIMETTA, E., BOLLINI, S., GAGGI, G. & GHINASSI, B. 2018. Human-induced pluripotent stem cell technology and cardiomyocyte generation: progress and clinical applications. *Cells,* 7**,** 48.

DUBOIS, N. C., CRAFT, A. M., SHARMA, P., ELLIOTT, D. A., STANLEY, E. G., ELEFANTY, A. G., GRAMOLINI, A. & KELLER, G. 2011. SIRPA is a specific cell-surface marker for isolating cardiomyocytes derived from human pluripotent stem cells. *Nature biotechnology,* 29**,** 1011-1018.

EL-ARMOUCHE, A. & ESCHENHAGEN, T. 2009. β-Adrenergic stimulation and myocardial function in the failing heart. *Heart failure reviews,* 14**,** 225-241.

ELLIOTT, D. A., BRAAM, S. R., KOUTSIS, K., NG, E. S., JENNY, R., LAGERQVIST, E. L., BIBEN, C., HATZISTAVROU, T., HIRST, C. E., YU, Q. C., SKELTON, R. J., WARD-VAN OOSTWAARD, D., LIM, S. M., KHAMMY, O., LI, X., HAWES, S. M., DAVIS, R. P., GOULBURN, A. L., PASSIER, R., PRALL, O. W., HAYNES, J. M., POUTON, C. W., KAYE, D. M., MUMMERY, C. L., ELEFANTY, A. G. & STANLEY, E. G. 2011. NKX2-5(eGFP/w) hESCs for isolation of human cardiac progenitors and cardiomyocytes. *Nat Methods,* 8**,** 1037-1040.

ENGLAND, J. & LOUGHNA, S. 2013. Heavy and light roles: myosin in the morphogenesis of the heart. *Cellular and Molecular Life Sciences,* 70**,** 1221-1239.

ESCHENHAGEN, T. & CARRIER, L. 2019. Cardiomyopathy phenotypes in human-induced pluripotent stem cell-derived cardiomyocytes—a systematic review. *Pflügers Archiv-European Journal of Physiology,* 471**,** 755-768.

EUSTICE, D. C. & WILHELM, J. M. 1984. Mechanisms of action of aminoglycoside antibiotics in eucaryotic protein synthesis. *Antimicrobial agents and chemotherapy,* 26**,** 53-60.

EVANS, S. M., YELON, D., CONLON, F. L. & KIRBY, M. L. 2010. Myocardial lineage development. *Circulation research,* 107**,** 1428-1444.

FEYEN, D. A., MCKEITHAN, W. L., BRUYNEEL, A. A., SPIERING, S., HÖRMANN, L., ULMER, B., ZHANG, H., BRIGANTI, F., SCHWEIZER, M. & HEGYI, B. 2020. Metabolic maturation media improve physiological function of human iPSC-derived cardiomyocytes. *Cell reports,* 32**,** 107925.

FIEDOROWICZ, K., ROZWADOWSKA, N., ZIMNA, A., MALCHER, A., TUTAK, K., SZCZERBAL, I., NOWICKA-BAUER, K., NOWACZYK, M., KOLANOWSKI, T. J. & ŁABĘDŹ, W. 2020. Tissue-specific promoter-based reporter system for monitoring cell differentiation from iPSCs to cardiomyocytes. *Scientific reports,* 10**,** 1-13.

FOMIN, A., GÄRTNER, A., CYGANEK, L., TIBURCY, M., TULETA, I., WELLERS, L., FOLSCHE, L., HOBBACH, A. J., VON FRIELING-SALEWSKY, M. & UNGER, A. 2021. Truncated titin proteins and titin haploinsufficiency are targets for functional recovery in human cardiomyopathy due to TTN mutations. *Science translational medicine,* 13**,** eabd3079.

FRANCO, D., LAMERS, W. H. & MOORMAN, A. F. 1998. Patterns of expression in the developing myocardium: towards a morphologically integrated transcriptional model. *Cardiovasc Res,* 38**,** 25-53.

FRANK, O. 1895. Zur dynamik des herzmuskels. *Z Biol,* 32**,** 370-447.

FREIBURG, A., TROMBITAS, K., HELL, W., CAZORLA, O., FOUGEROUSSE, F., CENTNER, T., KOLMERER, B., WITT, C., BECKMANN, J. S. & GREGORIO, C. C. 2000. Series of exon-skipping events in the elastic spring region of titin as the structural basis for myofibrillar elastic diversity. *Circulation research,* 86**,** 1114-1121.

FU, J.-D., JIANG, P., RUSHING, S., LIU, J., CHIAMVIMONVAT, N. & LI, R. A. 2010. Na+/Ca2+ exchanger is a determinant of excitation–contraction coupling in human embryonic stem cell–derived ventricular cardiomyocytes. *Stem Cells Dev,* 19**,** 773-782.

FUKUSHIMA, H., YOSHIOKA, M., KAWATOU, M., LÓPEZ-DÁVILA, V., TAKEDA, M., KANDA, Y., SEKINO, Y., YOSHIDA, Y. & YAMASHITA, J. K. 2020. Specific induction and long-term maintenance of high purity ventricular cardiomyocytes from human induced pluripotent stem cells. *PloS one,* 15**,** e0241287.

FUNAKOSHI, S., FERNANDES, I., MASTIKHINA, O., WILKINSON, D., TRAN, T., DHAHRI, W., MAZINE, A., YANG, D., BURNETT, B. & LEE, J. 2021. Generation of mature compact ventricular cardiomyocytes from human pluripotent stem cells. *Nature communications,* 12**,** 1-23.

FÜRST, D. O., OSBORN, M., NAVE, R. & WEBER, K. 1988. The organization of titin filaments in the half-sarcomere revealed by monoclonal antibodies in immunoelectron microscopy: a map of ten nonrepetitive epitopes starting at the Z line extends close to the M line. *The Journal of cell biology,* 106**,** 1563-1572.

FURUKAWA, T., ONO, Y., TSUCHIYA, H., KATAYAMA, Y., BANG, M.-L., LABEIT, D., LABEIT, S., INAGAKI, N. & GREGORIO, C. C. 2001. Specific interaction of the potassium channel β-subunit minK with the sarcomeric protein T-cap suggests a T-tubule-myofibril linking system. *Journal of molecular biology,* 313**,** 775-784.

GACITA, A. M., FULLENKAMP, D. E., OHIRI, J., POTTINGER, T., PUCKELWARTZ, M. J., NOBREGA, M. A. & MCNALLY, E. M. 2021. Genetic variation in enhancers modifies cardiomyopathy gene expression and progression. *Circulation,* 143**,** 1302-1316.

GALDOS, F. X., LEE, C., LEE, S., GOODYER, W., PAIGE, S., ESCOBAR, G., XU, S., BAK, R. O., PORTEUS, M. & WU, S. M. 2021. Dual TBX5-Lineage and MYL2 Reporter System For Identification of Left Ventricular Cardiomyocytes During Human Induced Pluripotent Stem Cell Differentiation. *bioRxiv*.

GAO, Y. & PU, J. 2021. Differentiation and application of human pluripotent stem cells derived cardiovascular cells for treatment of heart diseases: Promises and challenges. *Frontiers in Cell and Developmental Biology,* 9.

GEORGE, C. M. & ALANI, E. 2012. Multiple cellular mechanisms prevent chromosomal rearrangements involving repetitive DNA. *Critical reviews in biochemistry and molecular biology,* 47**,** 297-313.

GHAZIZADEH, Z., ZHU, J., FATTAHI, F., TANG, A., SUN, X., AMIN, S., TSAI, S.-Y., KHALAJ, M., ZHOU, T. & SAMUEL, R. M. 2022. A dual SHOX2: GFP; MYH6: mCherry knockin hESC reporter line for derivation of human SAN-like cells. *Iscience,* 25**,** 104153.

GIBSON, D. G., YOUNG, L., CHUANG, R.-Y., VENTER, J. C., HUTCHISON, C. A. & SMITH, H. O. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods,* 6**,** 343-345.

GIGANTI, D., YAN, K., BADILLA, C. L., FERNANDEZ, J. M. & ALEGRE-CEBOLLADA, J. 2018. Disulfide isomerization reactions in titin immunoglobulin domains enable a mode of protein elasticity. *Nature communications,* 9**,** 1-11.

GILDA, J. E. & GOMES, A. V. 2017. Proteasome dysfunction in cardiomyopathies. *The Journal of physiology,* 595**,** 4051-4071.

GOLBUS, J. R., PUCKELWARTZ, M. J., FAHRENBACH, J. P., DELLEFAVE-CASTILLO, L. M., WOLFGEHER, D. & MCNALLY, E. M. 2012. Population-based variation in cardiomyopathy genes. *Circulation: Cardiovascular Genetics,* 5**,** 391-399.

GOLDFRACHT, I., PROTZE, S., SHITI, A., SETTER, N., GRUBER, A., SHAHEEN, N., NARTISS, Y., KELLER, G. & GEPSTEIN, L. 2020. Generating ring-shaped engineered heart tissues from ventricular and atrial human pluripotent stem cell-derived cardiomyocytes. *Nat. Commun.,* 11**,** 1-15.

GOODYER, W. R., BEYERSDORF, B. M., PAIK, D. T., TIAN, L., LI, G., BUIKEMA, J. W., CHIRIKIAN, O., CHOI, S., VENKATRAMAN, S. & ADAMS, E. L. 2019. Transcriptomic profiling of the developing cardiac conduction system at single-cell resolution. *Circulation research,* 125**,** 379-397.

GORDAN, R., GWATHMEY, J. K. & XIE, L.-H. 2015. Autonomic and endocrine control of cardiovascular function. *World journal of cardiology,* 7**,** 204.

GOVERSEN, B., VAN DER HEYDEN, M. A., VAN VEEN, T. A. & DE BOER, T. P. 2018. The immature electrophysiological phenotype of iPSC-CMs still hampers in vitro drug screening: Special focus on IK1. *Pharmacology & therapeutics,* 183**,** 127-136.

GRAMLICH, M., PANE, L. S., ZHOU, Q., CHEN, Z., MURGIA, M., SCHÖTTERL, S., GOEDEL, A., METZGER, K., BRADE, T. & PARROTTA, E. 2015. Antisense-mediated exon skipping: a therapeutic strategy for titin-based dilated cardiomyopathy. *EMBO molecular medicine,* 7**,** 562-576.

GRANT, A. O. 2009. Cardiac ion channels. *Circulation: Arrhythmia and Electrophysiology,* 2**,** 185-194.

GRUEN, M. & GAUTEL, M. 1999. Mutations in β-myosin S2 that cause familial hypertrophic cardiomyopathy (FHC) abolish the interaction with the regulatory domain of myosin-binding protein-C. *Journal of molecular biology,* 286**,** 933-949.

GUNAWAN, M. G., SANGHA, S. S., SHAFAATTALAB, S., LIN, E., HEIMS-WALDRON, D. A., BEZZERIDES, V. J., LAKSMAN, Z. & TIBBITS, G. F. 2021. Drug screening platform using human induced pluripotent stem cell-derived atrial cardiomyocytes and optical mapping. *Stem cells translational medicine,* 10**,** 68-82.

GUO, W., SCHAFER, S., GREASER, M. L., RADKE, M. H., LISS, M., GOVINDARAJAN, T., MAATZ, H., SCHULZ, H., LI, S. & PARRISH, A. M. 2012. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nature medicine,* 18**,** 766-773.

GUO, W., ZHU, C., YIN, Z., WANG, Q., SUN, M., CAO, H. & GREASER, M. L. 2018. Splicing factor RBM20 regulates transcriptional network of titin associated and calcium handling genes in the heart. *International journal of biological sciences,* 14**,** 369.

GUO, Y. & PU, W. T. 2020. Cardiomyocyte maturation: new phase in development. *Circulation research,* 126**,** 1086-1106.

GUYTON, A. & HALL, J. 2000. Textbook of medical physiology. 10. *WB Sanders Co*.

HAGGERTY, C. M., DAMRAUER, S. M., LEVIN, M. G., BIRTWELL, D., CAREY, D. J., GOLDEN, A. M., HARTZEL, D. N., HU, Y., JUDY, R. & KELLY, M. A. J. C. 2019. Genomics-First Evaluation of Heart Disease Associated With Titin-Truncating Variants. *Circulation*
140**,** 42-54.
HANG, C., SONG, Y., LI, Y. N., ZHANG, S., CHANG, Y., BAI, R., SALEEM, A., JIANG, M., LU, W. & LAN, F. 2021. Knockout of MYOM1 in human cardiomyocytes leads to myocardial atrophy via impairing calcium homeostasis. *Journal of cellular and molecular medicine,* 25**,** 1661-1676.
HANNES, T., WOLFF, M., DOSS, M. X., PFANNKUCHE, K., HAUSTEIN, M., MÜLLER-EHMSEN, J., SACHINIDIS, A., HESCHELER, J., KHALIL, M. & HALBACH, M. 2015. Electrophysiological characteristics of embryonic stem cell-derived cardiomyocytes are cell line-dependent. *Cell Physio. Biochem.,* 35**,** 305-14.
HATTORI, F., CHEN, H., YAMASHITA, H., TOHYAMA, S., SATOH, Y.-S., YUASA, S., LI, W., YAMAKAWA, H., TANAKA, T. & ONITSUKA, T. 2010. Nongenetic method for purifying stem cell–derived cardiomyocytes. *Nature methods,* 7**,** 61-66.
HAYASHI, T., ARIMURA, T., ITOH-SATOH, M., UEDA, K., HOHDA, S., INAGAKI, N., TAKAHASHI, M., HORI, H., YASUNAMI, M. & NISHI, H. 2004. Tcap gene mutations in hypertrophic cardiomyopathy and dilated cardiomyopathy. *Journal of the American College of Cardiology,* 44**,** 2192-2201.
HERMAN, D. S., LAM, L., TAYLOR, M. R., WANG, L., TEEKAKIRIKUL, P., CHRISTODOULOU, D., CONNER, L., DEPALMA, S. R., MCDONOUGH, B. & SPARKS, E. 2012. Truncations of titin causing dilated cardiomyopathy. *New England Journal of Medicine,* 366**,** 619-628.
HERSHBERGER, R. E., HEDGES, D. J. & MORALES, A. 2013. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nature Reviews Cardiology,* 10**,** 531-547.
HERZOG, W. 2018. The multiple roles of titin in muscle contraction and force production. *Biophysical reviews,* 10**,** 1187-1199.
HINSON, J. T., CHOPRA, A., NAFISSI, N., POLACHECK, W. J., BENSON, C. C., SWIST, S., GORHAM, J., YANG, L., SCHAFER, S. & SHENG, C. C. 2015. Titin mutations in iPS cells define sarcomere insufficiency as a cause of dilated cardiomyopathy. *Science,* 349**,** 982-986.
HOEKSTRA, M., MUMMERY, C. L., WILDE, A. A., BEZZINA, C. R. & VERKERK, A. O. 2012. Induced pluripotent stem cell derived cardiomyocytes as models for cardiac arrhythmias. *Frontiers in physiology,* 3**,** 346.
HOORNTJE, E. T., VAN SPAENDONCK-ZWARTS, K. Y., TE RIJDT, W. P., BOVEN, L., VINK, A., VAN DER SMAGT, J. J., ASSELBERGS, F. W., VAN WIJNGAARDEN, J., HENNEKAM, E. A. & PINTO, Y. M. 2018. The first titin (c. 59926+ 1G> A) founder mutation associated with dilated cardiomyopathy. *European journal of heart failure,* 20**,** 803.
HORIKOSHI, Y., YAN, Y., TERASHVILI, M., WELLS, C., HORIKOSHI, H., FUJITA, S., BOSNJAK, Z. J. & BAI, X. 2019. Fatty acid-treated induced pluripotent stem cell-derived human cardiomyocytes exhibit adult cardiomyocyte-like energy metabolism phenotypes. *Cells,* 8**,** 1095.
HUBER, I., ITZHAKI, I., CASPI, O., ARBEL, G., TZUKERMAN, M., GEPSTEIN, A., HABIB, M., YANKELSON, L., KEHAT, I. & GEPSTEIN, L. 2007. Identification and selection of cardiomyocytes during human embryonic stem cell differentiation. *FASEB J.,* 21**,** 2551-63.
HUTTNER, I. G., WANG, L. W., SANTIAGO, C. F., HORVAT, C., JOHNSON, R., CHENG, D., VON FRIELING-SALEWSKY, M., HILLCOAT, K., BEMAND, T. J. & TRIVEDI, G. 2018. A-band titin truncation in zebrafish causes dilated cardiomyopathy and hemodynamic stress intolerance. *Circulation: Genomic and Precision Medicine,* 11**,** e002135.
IBRAHIM, M., SIEDLECKA, U., BUYANDELGER, B., HARADA, M., RAO, C., MOSHKOV, A., BHARGAVA, A., SCHNEIDER, M., YACOUB, M. H. & GORELIK, J. 2013. A critical

role for Telethonin in regulating t-tubule structure and function in the mammalian heart. *Human molecular genetics,* 22**,** 372-383.
ICHIMURA, H., KADOTA, S., KASHIHARA, T., YAMADA, M., ITO, K., KOBAYASHI, H., TANAKA, Y., SHIBA, N., CHUMA, S. & TOHYAMA, S. 2020. Increased predominance of the matured ventricular subtype in embryonic stem cell-derived cardiomyocytes in vivo. *Scientific reports,* 10**,** 1-12.
IKONNIKOV, G., YELLE, D., WONG, E. & CHAUDHRY, S. 2013. Physiology of cardiac conduction and contractility. *McMaster Pathophysiology Review*.
INAGAKI, N., HAYASHI, T., ARIMURA, T., KOGA, Y., TAKAHASHI, M., SHIBATA, H., TERAOKA, K., CHIKAMORI, T., YAMASHINA, A. & KIMURA, A. 2006. αB-crystallin mutation in dilated cardiomyopathy. *Biochemical and biophysical research communications,* 342**,** 379-386.
IONTA, V., LIANG, W., KIM, E. H., RAFIE, R., GIACOMELLO, A., MARBÁN, E. & CHO, H. C. 2015. SHOX2 overexpression favors differentiation of embryonic stem cells into cardiac pacemaker cells, improving biological pacing ability. *Stem cell reports,* 4**,** 129-142.
IRION, S., LUCHE, H., GADUE, P., FEHLING, H. J., KENNEDY, M. & KELLER, G. 2007. Identification and targeting of the ROSA26 locus in human embryonic stem cells. *Nat Biotechnol,* 25**,** 1477-1482.
JIANG, F. & DOUDNA, J. A. J. A. R. O. B. 2017. CRISPR–Cas9 structures and mechanisms. *Annu Rev Biophys,* 46**,** 505-529.
JIANG, H., HOOPER, C., KELLY, M., STEEPLES, V., SIMON, J. N., BEGLOV, J., AZAD, A. J., LEINHOS, L., BENNETT, P. & EHLER, E. 2021. Functional analysis of a gene-edited mouse model to gain insights into the disease mechanisms of a titin missense variant. *Basic research in cardiology,* 116**,** 1-18.
JIANG, Y., PARK, P., HONG, S.-M. & BAN, K. 2018. Maturation of cardiomyocytes derived from human pluripotent stem cells: current strategies and limitations. *Molecules and cells,* 41**,** 613.
JOSOWITZ, R., LU, J., FALCE, C., D'SOUZA, S. L., WU, M., COHEN, N., DUBOIS, N. C., ZHAO, Y., SOBIE, E. A., FISHMAN, G. I. & GELB, B. D. 2014. Identification and purification of human induced pluripotent stem cell-derived atrial-like cardiomyocytes based on sarcolipin expression. *PLoS One,* 9**,** e101316.
JOST, N. 2009. Transmembrane ionic currents underlying cardiac action potential in mammalian hearts. *Advances in cardiomyocyte research. Ed. Péter P. Nánási. Transworld Research Network, Tivandrum, Kerala, India***,** 1-45.
KAMAKURA, T., MAKIYAMA, T., SASAKI, K., YOSHIDA, Y., WURIYANGHAI, Y., CHEN, J., HATTORI, T., OHNO, S., KITA, T. & HORIE, M. 2013. Ultrastructural maturation of human-induced pluripotent stem cell-derived cardiomyocytes in a long-term culture. *Circulation Journal,* 77**,** 1307-1314.
KAMP, T. J. & HELL, J. W. 2000. Regulation of cardiac L-type calcium channels by protein kinase A and protein kinase C. *Circulation research,* 87**,** 1095-1102.
KANE, C., COUCH, L. & TERRACCIANO, C. M. 2015. Excitation–contraction coupling of human induced pluripotent stem cell-derived cardiomyocytes. *Frontiers in cell and developmental biology,* 3**,** 59.
KARBASSI, E., FENIX, A., MARCHIANO, S., MURAOKA, N., NAKAMURA, K., YANG, X. & MURRY, C. E. 2020. Cardiomyocyte maturation: advances in knowledge and implications for regenerative medicine. *Nature Reviews Cardiology,* 17**,** 341-359.
KATTMAN, S. J., WITTY, A. D., GAGLIARDI, M., DUBOIS, N. C., NIAPOUR, M., HOTTA, A., ELLIS, J. & KELLER, G. 2011. Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell stem cell,* 8**,** 228-240.
KIM, H. & KIM, J. S. 2014. A guide to genome engineering with programmable nucleases. *Nat Rev Genet,* 15**,** 321-234.
KIM, H. S., YOON, J. W., LI, H., JEONG, G. O., PARK, J. J., SHIN, S. E., JANG, I. H., KIM, J. H. & PARK, W. S. 2017. Functional expression and pharmaceutical efficacy of

cardiac-specific ion channels in human embryonic stem cell-derived cardiomyocytes. *Scientific reports,* 7**,** 1-11.
KIM, J. H., LEE, S.-R., LI, L.-H., PARK, H.-J., PARK, J.-H., LEE, K. Y., KIM, M.-K., SHIN, B. A. & CHOI, S.-Y. 2011. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PloS one,* 6**,** e18556.
KLABUNDE, R. 2011. *Cardiovascular physiology concepts*, Lippincott Williams & Wilkins.
KLATT, D., CHENG, E., HOFFMANN, D., SANTILLI, G., THRASHER, A. J., BRENDEL, C. & SCHAMBACH, A. 2020. Differential Transgene Silencing of Myeloid-Specific Promoters in the AAVS1 Safe Harbor Locus of Induced Pluripotent Stem Cell-Derived Myeloid Cells. *Hum. Gene Ther.,* 31**,** 199-210.
KLAUS, A., MÜLLER, M., SCHULZ, H., SAGA, Y., MARTIN, J. F. & BIRCHMEIER, W. 2012. Wnt/β-catenin and Bmp signals control distinct sets of transcription factors in cardiac progenitor cells. *Proceedings of the National Academy of Sciences,* 109**,** 10921-10926.
KLEBER, A. G. & SAFFITZ, J. E. 2014. Role of the intercalated disc in cardiac propagation and arrhythmogenesis. *Frontiers in physiology,* 5**,** 404.
KM, S. B., CIRULLI, E., BOLZE, A., ROWAN, C., ELHANAN, G., GRZYMSKI, J., LEE, W. & WASHINGTON, N. 2022. TTN truncating variants in hiPSI exons show high penetrance for cardiomyopathy in carriers with atrial fibrillation.
KNOWLTON, F. & STARLING, E. 1912. The influence of variations in temperature and blood-pressure on the performance of the isolated mammalian heart. *The Journal of physiology,* 44**,** 206.
KOBAYASHI, T. & SOLARO, R. J. 2005. Calcium, thin filaments, and the integrative biology of cardiac contractility. *Annual review of physiology,* 67**,** 39.
KOCATURK, N. M. & GOZUACIK, D. 2018. Crosstalk between mammalian autophagy and the ubiquitin-proteasome system. *Frontiers in cell and developmental biology***,** 128.
KOIVUMÄKI, J. T., NAUMENKO, N., TUOMAINEN, T., TAKALO, J., OKSANEN, M., PUTTONEN, K. A., LEHTONEN, Š., KUUSISTO, J., LAAKSO, M. & KOISTINAHO, J. 2018. Structural immaturity of human iPSC-derived cardiomyocytes: in silico investigation of effects on function and disease modeling. *Frontiers in physiology,* 9**,** 80.
KOSICKI, M., TOMBERG, K. & BRADLEY, A. 2018. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nature biotechnology,* 36**,** 765-771.
KOSS, K., GRUPP, I. & KRANIAS, E. G. 1998. The relative phospholamban and SERCA2 ratio: a critical determinant of myocardial contractility. *Alterations of Excitation-Contraction Coupling in the Failing Human Heart.* Springer.
KÖTTER, S., UNGER, A., HAMDANI, N., LANG, P., VORGERD, M., NAGEL-STEGER, L. & LINKE, W. A. 2014. Human myocytes are protected from titin aggregation-induced stiffening by small heat shock proteins. *Journal of Cell Biology,* 204**,** 187-202.
KRUSE, S. W., SUINO-POWELL, K., ZHOU, X. E., KRETSCHMAN, J. E., REYNOLDS, R., VONRHEIN, C., XU, Y., WANG, L., TSAI, S. Y. & TSAI, M.-J. 2008. Identification of COUP-TFII orphan nuclear receptor as a retinoic acid–activated receptor. *PLoS biology,* 6**,** e227.
KUBALAK, S. W., MILLER-HANCE, W. C., O'BRIEN, T. X., DYSON, E. & CHIEN, K. R. 1994. Chamber specification of atrial myosin light chain-2 expression precedes septation during murine cardiogenesis. *Journal of Biological Chemistry,* 269**,** 16961-16970.
KUSSAUER, S., DAVID, R. & LEMCKE, H. 2019. hiPSCs derived cardiac cells for drug and toxicity screening and disease modeling: what micro-electrode-array analyses can tell us. *Cells,* 8**,** 1331.
LABEIT, S., KOHL, C. H., WITT, C. C., LABEIT, D., JUNG, J. & GRANZIER, H. 2010. Modulation of muscle atrophy, fatigue and MLC phosphorylation by MuRF1 as indicated by hindlimb suspension studies on MuRF1-KO mice. *Journal of Biomedicine and Biotechnology,* 2010.

LAFLAMME, M. A., CHEN, K. Y., NAUMOVA, A. V., MUSKHELI, V., FUGATE, J. A., DUPRAS, S. K., REINECKE, H., XU, C., HASSANIPOUR, M. & POLICE, S. 2007. Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. *Nature biotechnology,* 25**,** 1015-1024.

LAHMERS, S., WU, Y., CALL, D. R., LABEIT, S. & GRANZIER, H. 2004. Developmental control of titin isoform expression and passive stiffness in fetal and neonatal myocardium. *Circulation research,* 94**,** 505-513.

LAKATTA, E. G., MALTSEV, V. A. & VINOGRADOVA, T. M. 2010. A coupled SYSTEM of intracellular Ca2+ clocks and surface membrane voltage clocks controls the timekeeping mechanism of the heart's pacemaker. *Circulation research,* 106**,** 659-673.

LAKSMAN, Z., WAUCHOP, M., LIN, E., PROTZE, S., LEE, J., YANG, W., IZADDOUSTDAR, F., SHAFAATTALAB, S., GEPSTEIN, L. & TIBBITS, G. F. 2017. Modeling atrial fibrillation using human embryonic stem cell-derived atrial tissue. *Scientific reports,* 7**,** 1-11.

LANDSTROM, A. P., DOBREV, D. & WEHRENS, X. H. 2017. Calcium signaling and cardiac arrhythmias. *Circulation research,* 120**,** 1969-1993.

LANGE, M. J., LYDDON, T. D. & JOHNSON, M. C. 2019. Diphtheria toxin A-resistant cell lines enable robust production and evaluation of DTA-encoding lentiviruses. *Scientific reports,* 9**,** 1-11.

LEAVITT, A. D. & HAMLETT, I. 2011. Homologous recombination in human embryonic stem cells: a tool for advancing cell therapy and understanding and treating human disease. *Clin Transl Sci,* 4**,** 298-305.

LEE, J. H., PROTZE, S. I., LAKSMAN, Z., BACKX, P. H. & KELLER, G. M. 2017. Human pluripotent stem cell-derived atrial and ventricular cardiomyocytes develop from distinct mesoderm populations. *Cell stem cell,* 21**,** 179-194. e4.

LEMME, M., ULMER, B. M., LEMOINE, M. D., ZECH, A. T., FLENNER, F., RAVENS, U., REICHENSPURNER, H., ROL-GARCIA, M., SMITH, G. & HANSEN, A. J. S. C. R. 2018. Atrial-like engineered heart tissue: an in vitro model of the human atrium. *Stem Cell Reports,* 11**,** 1378-1390.

LESCROART, F., CHABAB, S., LIN, X., RULANDS, S., PAULISSEN, C., RODOLOSSE, A., AUER, H., ACHOURI, Y., DUBOIS, C. & BONDUE, A. 2014. Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development. *Nature cell biology,* 16**,** 829-840.

LEWINTER, M. M. & GRANZIER, H. L. 2013. Titin is a major human disease gene. *Circulation,* 127**,** 938-944.

LI, W., LUO, X., ULBRICHT, Y. & GUAN, K. 2021. Blebbistatin protects iPSC-CMs from hypercontraction and facilitates automated patch-clamp based electrophysiological study. *Stem Cell Research,* 56**,** 102565.

LIAN, X., HSIAO, C., WILSON, G., ZHU, K., HAZELTINE, L. B., AZARIN, S. M., RAVAL, K. K., ZHANG, J., KAMP, T. J. & PALECEK, S. P. J. P. O. T. N. A. O. S. 2012. Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. *Proc Natl Acad Sci U S A,* 109**,** E1848-E1857.

LIN, B., LIN, X., STACHEL, M., WANG, E., LUO, Y., LADER, J., SUN, X., DELMAR, M. & BU, L. 2017. Culture in glucose-depleted medium supplemented with fatty acid and 3, 3′, 5-triiodo-l-thyronine facilitates purification and maturation of human pluripotent stem cell-derived cardiomyocytes. *Frontiers in endocrinology,* 8**,** 253.

LIN, F.-J., YOU, L.-R., YU, C.-T., HSU, W.-H., TSAI, M.-J. & TSAI, S. Y. 2012. Endocardial cushion morphogenesis and coronary vessel development require chicken ovalbumin upstream promoter-transcription factor II. *Arteriosclerosis, thrombosis, and vascular biology,* 32**,** e135-e146.

LINO, C. A., HARPER, J. C., CARNEY, J. P. & TIMLIN, J. A. 2018. Delivering CRISPR: a review of the challenges and approaches. *Drug delivery,* 25**,** 1234-1257.

LIP, G., FAUCHIER, L., FREEDMAN, S. B., VAN GELDER, I., NATALE, A., GIANNI, C., NATTEL, S., POTPARA, T., RIENSTRA, M. & TSE, H.-F. 2016. Atrial fibrillation. *Nature reviews. Disease primers,* 2**,** 16016-16016.

LIU, F., FANG, Y., HOU, X., YAN, Y., XIAO, H., ZUO, D., WEN, J., WANG, L., ZHOU, Z. & DANG, X. 2020. Enrichment differentiation of human induced pluripotent stem cells into sinoatrial node-like cells by combined modulation of BMP, FGF, and RA signaling pathways. *Stem Cell Research & Therapy,* 11**,** 1-15.

LIU, Z., CHEN, O., WALL, J., ZHENG, M., ZHOU, Y., WANG, L., RUTH VASEGHI, H., QIAN, L. & LIU, J. 2017. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Scientific reports,* 7**,** 1-9.

LOPEZ, C. A., AL-SIDDIQI, H. H. A., PURNAMA, U., IFTEKHAR, S., BRUYNEEL, A. A., KERR, M., NAZIR, R., DA LUZ SOUSA FIALHO, M., MALANDRAKI-MILLER, S. & ALONAIZAN, R. 2021. Physiological and pharmacological stimulation for in vitro maturation of substrate metabolism in human induced pluripotent stem cell-derived cardiomyocytes. *Scientific reports,* 11**,** 1-13.

LUO, X.-L., ZHANG, P., LIU, X., HUANG, S., RAO, S.-L., DING, Q. & YANG, H.-T. 2021. Myosin light chain 2 marks differentiating ventricular cardiomyocytes derived from human embryonic stem cells. *Pflügers Archiv-European Journal of Physiology,* 473**,** 991-1007.

LYKKE-ANDERSEN, S. & JENSEN, T. H. 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nature reviews Molecular cell biology,* 16**,** 665-677.

MAKARENKO, I., OPITZ, C., LEAKE, M., NEAGOE, C., KULKE, M., GWATHMEY, J., DEL MONTE, F., HAJJAR, R. & LINKE, W. 2004. Passive stiffness changes caused by upregulation of compliant titin isoforms in human dilated cardiomyopathy hearts. *Circulation research,* 95**,** 708-716.

MANFRA, O., FRISK, M. & LOUCH, W. E. 2017. Regulation of cardiomyocyte T-tubular structure: opportunities for therapy. *Current Heart Failure Reports,* 14**,** 167-178.

MANIVANNAN, S. N., DAROUICH, S., MASMOUDI, A., GORDON, D., ZENDER, G., HAN, Z., FITZGERALD-BUTT, S., WHITE, P., MCBRIDE, K. L. & KHARRAT, M. 2020. Novel frameshift variant in MYL2 reveals molecular differences between dominant and recessive forms of hypertrophic cardiomyopathy. *PLoS genetics,* 16**,** e1008639.

MARCZENKE, M., FELL, J., PICCINI, I., ROEPKE, A., SEEBOHM, G. & GREBER, B. 2017. Generation and cardiac subtype-specific differentiation of PITX2-deficient human iPS cell lines for exploring familial atrial fibrillation. *Stem Cell Research,* 21**,** 26-28.

MARON, B. J., GARDIN, J. M., FLACK, J. M., GIDDING, S. S., KUROSAKI, T. T. & BILD, D. E. 1995. Prevalence of hypertrophic cardiomyopathy in a general population of young adults: echocardiographic analysis of 4111 subjects in the CARDIA study. *Circulation,* 92**,** 785-789.

MAYANS, O., VAN DER VEN, P. F., WILM, M., MUES, A., YOUNG, P., FÜRST, D. O., WILMANNS, M. & GAUTEL, M. 1998. Structural basis for activation of the titin kinase domain during myofibrillogenesis. *Nature,* 395**,** 863-869.

MAZZONE, A., STREGE, P. R., TESTER, D. J., BERNARD, C. E., FAULKNER, G., DE GIORGIO, R., MAKIELSKI, J. C., STANGHELLINI, V., GIBBONS, S. J. & ACKERMAN, M. J. 2008. A mutation in telethonin alters Nav1. 5 function. *Journal of Biological Chemistry,* 283**,** 16537-16544.

MCAFEE, Q., CHEN, C. Y., YANG, Y., CAPORIZZO, M. A., MORLEY, M., BABU, A., JEONG, S., BRANDIMARTO, J., BEDI JR, K. C. & FLAM, E. 2021. Truncated titin proteins in dilated cardiomyopathy. *Science translational medicine,* 13**,** eabd7287.

MCELHINNY, A. S., KAKINUMA, K., SORIMACHI, H., LABEIT, S. & GREGORIO, C. C. 2002. Muscle-specific RING finger-1 interacts with titin to regulate sarcomeric M-line and thick filament structure and may have nuclear functions via its interaction with glucocorticoid modulatory element binding protein-1. *The Journal of cell biology,* 157**,** 125-136.

MCINTOSH, D. B., DUGGAN, G., GOUIL, Q. & SALEH, O. A. 2014. Sequence-dependent elasticity and electrostatics of single-stranded DNA: signatures of base-stacking. *Biophysical journal,* 106**,** 659-666.

MCMURRAY, J. J., ADAMOPOULOS, S., ANKER, S. D., AURICCHIO, A., BÖHM, M., DICKSTEIN, K., FALK, V., FILIPPATOS, G. & FONSECA, C. 2012. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *European heart journal,* 33**,** 1787-1847.

MEILHAC, S. M., ESNER, M., KELLY, R. G., NICOLAS, J.-F. & BUCKINGHAM, M. E. 2004. The clonal origin of myocardial cells in different regions of the embryonic mouse heart. *Developmental cell,* 6**,** 685-698.

MIAO, S., ZHAO, D., WANG, X., NI, X., FANG, X., YU, M., YE, L., YANG, J., WU, H. & HAN, X. 2020. Retinoic acid promotes metabolic maturation of human Embryonic Stem Cell-derived Cardiomyocytes. *Theranostics,* 10**,** 9686.

MINAMISAWA, S., WANG, Y., CHEN, J., ISHIKAWA, Y., CHIEN, K. R. & MATSUOKA, R. 2003. Atrial chamber-specific expression of sarcolipin is regulated during development and hypertrophic remodeling. *J Biol Chem,* 278**,** 9570-9575.

MIURA, H., QUADROS, R. M., GURUMURTHY, C. B. & OHTSUKA, M. J. N. P. 2018. Easi-CRISPR for creating knock-in and conditional knockout mouse models using long ssDNA donors. *Nat Protoc,* 13**,** 195.

MOHAPATRA, B., JIMENEZ, S., LIN, J. H., BOWLES, K. R., COVELER, K. J., MARX, J. G., CHRISCO, M. A., MURPHY, R. T., LURIE, P. R. & SCHWARTZ, R. J. 2003. Mutations in the muscle LIM protein and α-actinin-2 genes in dilated cardiomyopathy and endocardial fibroelastosis. *Molecular genetics and metabolism,* 80**,** 207-215.

MONTEIRO, L. M., VASQUES-NÓVOA, F., FERREIRA, L., PINTO-DO-Ó, P. & NASCIMENTO, D. S. 2017. Restoring heart function and electrical integrity: closing the circuit. *NPJ Regenerative medicine,* 2**,** 1-13.

MONTERO, P., FLANDES-IPARRAGUIRRE, M., MUSQUIZ, S., PÉREZ ARALUCE, M., PLANO, D., SANMARTÍN, C., ORIVE, G., GAVIRA, J. J., PROSPER, F. & MAZO, M. M. 2020. Cells, materials, and fabrication processes for cardiac tissue engineering. *Frontiers in Bioengineering and Biotechnology,* 8**,** 955.

MOORE, J. C., FU, J., CHAN, Y.-C., LIN, D., TRAN, H., TSE, H.-F. & LI, R. A. 2008. Distinct cardiogenic preferences of two human embryonic stem cell (hESC) lines are imprinted in their proteomes in the pluripotent state. *Biochemical and biophysical research communications,* 372**,** 553-558.

MURPHY, M., RASNIK, I., CHENG, W., LOHMAN, T. M. & HA, T. 2004. Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophysical journal,* 86**,** 2530-2537.

MUSA, H., MEEK, S., GAUTEL, M., PEDDIE, D., SMITH, A. J. & PECKHAM, M. 2006. Targeted homozygous deletion of M-band titin in cardiomyocytes prevents sarcomere formation. *Journal of cell science,* 119**,** 4322-4331.

MYERS, F. B., ZARINS, C. K., ABILEZ, O. J. & LEE, L. P. 2013. Label-free electrophysiological cytometry for stem cell-derived cardiomyocyte clusters. *Lab on a Chip,* 13**,** 220-228.

NAGARAJAN, V. D., HO, S. Y. & ERNST, S. 2019. Anatomical considerations for His bundle pacing. *Circulation: Arrhythmia and Electrophysiology,* 12**,** e006897.

NAGUEH, S. F., SHAH, G., WU, Y., TORRE-AMIONE, G., KING, N. M., LAHMERS, S., WITT, C. C., BECKER, K., LABEIT, S. & GRANZIER, H. L. 2004. Altered titin expression, myocardial stiffness, and left ventricular function in patients with dilated cardiomyopathy. *Circulation,* 110**,** 155-162.

NAYEROSSADAT, N., MAEDEH, T. & ALI, P. A. 2012. Viral and nonviral delivery systems for gene delivery. *Adv Biomed Res,* 1**,** 27.

NEAGOE, C., KULKE, M., DEL MONTE, F., GWATHMEY, J. K., DE TOMBE, P. P., HAJJAR, R. J. & LINKE, W. A. 2002. Titin isoform switch in ischemic human heart disease. *Circulation,* 106**,** 1333-1341.

NEUBAUER, S. 2007. The failing heart—an engine out of fuel. *New England Journal of Medicine,* 356**,** 1140-1151.

NG, S. Y., WONG, C. K. & TSANG, S. Y. 2010. Differential gene expressions in atrial and ventricular myocytes: insights into the road of applying embryonic stem cell-derived cardiomyocytes for future therapies. *American journal of physiology-Cell physiology,* 299**,** C1234-C1249.

NIELSEN, J. B., FRITSCHE, L. G., ZHOU, W., TESLOVICH, T. M., HOLMEN, O. L., GUSTAFSSON, S., GABRIELSEN, M. E., SCHMIDT, E. M., BEAUMONT, R. & WOLFORD, B. N. 2018. Genome-wide study of atrial fibrillation identifies seven risk loci and highlights biological pathways and regulatory elements involved in cardiac development. *The American Journal of Human Genetics,* 102**,** 103-115.

OBERMANN, W., GAUTEL, M., STEINER, F., VAN DER VEN, P., WEBER, K. & FÜRST, D. O. 1996. The structure of the sarcomeric M band: localization of defined domains of myomesin, M-protein, and the 250-kD carboxy-terminal region of titin by immunoelectron microscopy. *The Journal of cell biology,* 134**,** 1441-1453.

OCEGUERA-YANEZ, F., KIM, S.-I., MATSUMOTO, T., TAN, G. W., XIANG, L., HATANI, T., KONDO, T., IKEYA, M., YOSHIDA, Y. & INOUE, H. 2016. Engineering the AAVS1 locus for consistent and scalable transgene expression in human iPSCs and their differentiated derivatives. *Methods,* 101**,** 43-55.

OLESEN, M., LUNDEGAARD, P., AHLBERG, G., REFSGAARD, L., ANDREASEN, L., RANTHE, M., LINSCHEID, N., NIELSEN, J., MELBYE, M. & HAUNSOE, S. J. E. H. J. 2018. Titin-truncating variants associates with atrial fibrillation, compromises assembly of the sarcomere. *Eur. Heart J.,* 39**,** ehy564. 195.

OSBORN, L., HESSION, C., TIZARD, R., VASSALLO, C., LUHOWSKYJ, S., CHI-ROSSO, G. & LOBB, R. 1989. Direct expression cloning of vascular cell adhesion molecule 1, a cytokine-induced endothelial protein that binds to lymphocytes. *Cell,* 59**,** 1203-1211.

PAIGE, S. L., PLONOWSKA, K., XU, A. & WU, S. M. 2015. Molecular regulation of cardiomyocyte differentiation. *Circulation research,* 116**,** 341-353.

PAPAPETROU, E. P. & SCHAMBACH, A. J. M. T. 2016. Gene insertion into genomic safe harbors for human gene therapy. *Mol Ther,* 24**,** 678-684.

PAPATHANASIOU, S., MARKOULAKI, S., BLAINE, L. J., LEIBOWITZ, M. L., ZHANG, C.-Z., JAENISCH, R. & PELLMAN, D. 2021. Whole chromosome loss and genomic instability in mouse embryos after CRISPR-Cas9 genome editing. *Nature communications,* 12**,** 1-7.

PARK, Y.-G., SOHN, C. H., CHEN, R., MCCUE, M., YUN, D. H., DRUMMOND, G. T., KU, T., EVANS, N. B., OAK, H. C. & TRIEU, W. 2019. Protection of tissue physicochemical properties using polyfunctional crosslinkers. *Nature biotechnology,* 37**,** 73-83.

PASCARELLA, G., HASHIMOTO, K., BUSCH, A., LUGINBÜHL, J., PARR, C., HON, C., YIP, W., KRATZ, A., BONETTI, A. & AGOSTINI, F. 2020. Non-allelic homologous recombination of Alu and LINE-1 elements generates somatic complexity in human genomes. *BioRxiv*.

PAULL, D., SEVILLA, A., ZHOU, H., HAHN, A. K., KIM, H., NAPOLITANO, C., TSANKOV, A., SHANG, L., KRUMHOLZ, K. & JAGADEESAN, P. 2015. Automated, high-throughput derivation, characterization and differentiation of induced pluripotent stem cells. *Nature methods,* 12**,** 885-892.

PEI, Y., SIERRA, G., SIVAPATHAM, R., SWISTOWSKI, A., RAO, M. S. & ZENG, X. 2015. A platform for rapid generation of single and multiplexed reporters in human iPSC lines. *Sci Rep,* 5**,** 9205.

PEREA-GOMEZ, A., VELLA, F. D., SHAWLOT, W., OULAD-ABDELGHANI, M., CHAZAUD, C., MENO, C., PFISTER, V., CHEN, L., ROBERTSON, E. & HAMADA, H. 2002. Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks. *Developmental cell,* 3**,** 745-756.

PETKOVA, M. A. & DOBRZYNSKI, H. 2021. Do human sinoatrial node cells have t-tubules? *Translational Research in Anatomy,* 25**,** 100131.

PICCINI, I., RAO, J., SEEBOHM, G. & GREBER, B. 2015. Human pluripotent stem cell-derived cardiomyocytes: Genome-wide expression profiling of long-term in vitro maturation in comparison to human heart tissue. *Genom. Data,* 4**,** 69-72.

PIONER, J. M., SANTINI, L., PALANDRI, C., MARTELLA, D., LUPI, F., LANGIONE, M., QUERCETO, S., GRANDINETTI, B., BALDUCCI, V. & BENZONI, P. 2019. Optical investigation of action potential and calcium handling maturation of hiPSC-cardiomyocytes on biomimetic substrates. *International journal of molecular sciences,* 20**,** 3799.

PLAČKIĆ, J. & KOCKSKÄMPER, J. 2018. Isolation of atrial and ventricular cardiomyocytes for in vitro studies. *Experimental Models of Cardiovascular Diseases.* Springer.

POULIN, H., MERCIER, A., DJEMAI, M., POULIOT, V., DESCHENES, I., BOUTJDIR, M., PUYMIRAT, J. & CHAHINE, M. 2021. iPSC-derived cardiomyocytes from patients with myotonic dystrophy type 1 have abnormal ion channel functions and slower conduction velocities. *Scientific reports,* 11**,** 1-13.

PREDMORE, J. M., WANG, P., DAVIS, F., BARTOLONE, S., WESTFALL, M. V., DYKE, D. B., PAGANI, F., POWELL, S. R. & DAY, S. M. 2010. Ubiquitin proteasome dysfunction in human hypertrophic and dilated cardiomyopathies. *Circulation,* 121**,** 997-1004.

PSARAS, Y., MARGARA, F., CICCONET, M., SPARROW, A. J., REPETTI, G. G., SCHMID, M., STEEPLES, V., WILCOX, J. A., BUENO-OROVIO, A. & REDWOOD, C. S. 2021. CalTrack: High-throughput automated calcium transient analysis in cardiomyocytes. *Circulation research,* 129**,** 326-341.

RADEMAKER, M. T. & RICHARDS, A. M. 2005. Cardiac natriuretic peptides for cardiac health. *Clinical Science,* 108**,** 23-36.

RANAWAKAGE, D. C., OKADA, K., SUGIO, K., KAWAGUCHI, Y., KUNINOBU-BONKOHARA, Y., TAKADA, T. & KAMACHI, Y. 2021. Efficient CRISPR-Cas9-mediated knock-in of composite tags in zebrafish using long ssDNA as a donor. *Frontiers in cell and developmental biology***,** 1926.

REINHARDT, A., KAGAWA, H. & WOLTJEN, K. 2020. N-terminal amino acids determine KLF4 protein stability in 2A peptide-linked polycistronic reprogramming constructs. *Stem cell reports,* 14**,** 520-527.

RIESENBERG, S., HELMBRECHT, N., KANIS, P., MARICIC, T. & PÄÄBO, S. 2022. Improved gRNA secondary structures allow editing of target sites resistant to CRISPR-Cas9 cleavage. *Nature communications,* 13**,** 1-8.

RIVERA, T., ZHAO, Y., NI, Y. & WANG, J. 2020. Human-induced pluripotent stem cell culture methods under cGMP conditions. *Current Protocols in Stem Cell Biology,* 54**,** e117.

ROBERTS, A. M., WARE, J. S., HERMAN, D. S., SCHAFER, S., BAKSI, J., BICK, A. G., BUCHAN, R. J., WALSH, R., JOHN, S. & WILKINSON, S. 2015. Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. *Science translational medicine,* 7**,** 270ra6-270ra6.

ROBERTSON, J. A. 2001. Human embryonic stem cell research: ethical and legal issues. *Nat Rev Genet,* 2**,** 74-8.

ROUET, P., SMIH, F. & JASIN, M. 1994. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol,* 14**,** 8096-8106.

RYCKEBUSCH, L., WANG, Z., BERTRAND, N., LIN, S.-C., CHI, X., SCHWARTZ, R., ZAFFRAN, S. & NIEDERREITHER, K. 2008. Retinoic acid deficiency alters second heart field formation. *Proceedings of the National Academy of Sciences,* 105**,** 2913-2918.

SADAYAPPAN, S., GULICK, J., OSINSKA, H., MARTIN, L. A., HAHN, H. S., DORN, G. W., KLEVITSKY, R., SEIDMAN, C. E., SEIDMAN, J. G. & ROBBINS, J. 2005. Cardiac myosin-binding protein-C phosphorylation and cardiac function. *Circulation research,* 97**,** 1156-1163.

SÁNCHEZ-QUINTANA, D. 2003. Anatomy of cardiac nodes and atrioventricular specialized conduction system. *Revista espanola de cardiologia,* 56**,** 1085-1092.

SANDER, J. D. & JOUNG, J. K. J. N. B. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol,* 32**,** 347.

SARTIANI, L., BETTIOL, E., STILLITANO, F., MUGELLI, A., CERBAI, E. & JACONI, M. E. 2007. Developmental changes in cardiomyocytes differentiated from human

embryonic stem cells: a molecular and electrophysiological approach. *Stem cells,* 25**,** 1136-1144.

SCHAFER, S., DE MARVAO, A., ADAMI, E., FIEDLER, L. R., NG, B., KHIN, E., RACKHAM, O. J., VAN HEESCH, S., PUA, C. J. & KUI, M. 2017. Titin-truncating variants affect heart function in disease cohorts and the general population. *Nature genetics,* 49**,** 46-53.

SCHEEL, O., FRECH, S., AMUZESCU, B., EISFELD, J., LIN, K.-H. & KNOTT, T. 2014. Action potential characterization of human induced pluripotent stem cell–derived cardiomyocytes using automated patch-clamp technology. *Assay and drug development technologies,* 12**,** 457-469.

SCHICK, R., MEKIES, L. N., SHEMER, Y., EISEN, B., HALLAS, T., JEHUDA, R. B., BEN-ARI, M., SZANTAI, A., WILLI, L. & SHULMAN, R. J. P. O. 2018. Functional abnormalities in induced Pluripotent Stem Cell-derived cardiomyocytes generated from titin-mutated patients with dilated cardiomyopathy. *Plos One,* 13**,** e0205719.

SCHWACH, V., COFIÑO-FABRES, C., TEN DEN, S. A. & PASSIER, R. 2022. Improved Atrial Differentiation of Human Pluripotent Stem Cells by Activation of Retinoic Acid Receptor Alpha (RARα). *Journal of personalized medicine,* 12**,** 628.

SCHWACH, V., VERKERK, A. O., MOL, M., MONSHOUWER-KLOOTS, J. J., DEVALLA, H. D., ORLOVA, V. V., ANASTASSIADIS, K., MUMMERY, C. L., DAVIS, R. P. & PASSIER, R. 2017. A COUP-TFII human embryonic stem cell reporter line to identify and select atrial cardiomyocytes. *Stem cell reports,* 9**,** 1765-1779.

SCHWEIZER, P. A., DARCHE, F. F., ULLRICH, N. D., GESCHWILL, P., GREBER, B., RIVINIUS, R., SEYLER, C., MÜLLER-DECKER, K., DRAGUHN, A., UTIKAL, J. J. S. C. R. & THERAPY 2017. Subtype-specific differentiation of cardiac pacemaker cell clusters from human induced pluripotent stem cells. *Stem Cell Res Ther,* 8**,** 229.

SHANMUGAM, M., MOLINA, C. E., GAO, S., SEVERAC-BASTIDE, R., FISCHMEISTER, R. & BABU, G. J. 2011. Decreased sarcolipin protein expression and enhanced sarco (endo) plasmic reticulum Ca2+ uptake in human atrial fibrillation. *Biochemical and biophysical research communications,* 410**,** 97-101.

SHARMA, A., TOEPFER, C. N., SCHMID, M., GARFINKEL, A. C. & SEIDMAN, C. E. 2018. Differentiation and Contractile Analysis of GFP-Sarcomere Reporter hiPSC-Cardiomyocytes. *Curr Protoc Hum Genet,* 96**,** 21 12 1-21 12 12.

SHAVE, R., BAGGISH, A., GEORGE, K., WOOD, M., SCHARHAG, J., WHYTE, G., GAZE, D. & THOMPSON, P. D. 2010. Exercise-induced cardiac troponin elevation: evidence, mechanisms, and implications. *Journal of the American College of Cardiology,* 56**,** 169-176.

SHEIKH, F., LYON, R. C. & CHEN, J. 2015. Functions of myosin light chain-2 (MYL2) in cardiac muscle and disease. *Gene,* 569**,** 14-20.

SIEGERT, R., PERROT, A., KELLER, S., BEHLKE, J., MICHALEWSKA-WŁUDARCZYK, A., WYCISK, A., TENDERA, M., MORANO, I. & ÖZCELIK, C. 2011. A myomesin mutation associated with hypertrophic cardiomyopathy deteriorates dimerisation properties. *Biochemical and biophysical research communications,* 405**,** 473-479.

SINGH, V. P., RUBINSTEIN, J., ARVANITIS, D. A., REN, X., GAO, X., HAGHIGHI, K., GILBERT, M., IYER, V. R., KIM, D. H. & CHO, C. 2013. Abnormal calcium cycling and cardiac arrhythmias associated with the human Ser96Ala genetic variant of histidine-rich calcium-binding protein. *Journal of the American Heart Association,* 2**,** e000460.

SMITH, D. L. & FERNHALL, B. 2022. *Advanced cardiovascular exercise physiology*, Human Kinetics.

SOMA, Y., MORITA, Y., KISHINO, Y., KANAZAWA, H., FUKUDA, K. & TOHYAMA, S. 2021. The present state and future perspectives of cardiac regenerative therapy using human pluripotent stem cells. *Frontiers in Cardiovascular Medicine,* 8.

SONTAYANANON, N., REDWOOD, C., DAVIES, B. & GEHMLICH, K. 2020. Fluorescent PSC-derived cardiomyocyte reporter lines: generation approaches and their applications in cardiovascular medicine. *Biology,* 9**,** 402.

SOUTHERN, P. & BERG, P. 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *Journal of molecular and applied genetics,* 1**,** 327-341.
SPARROW, A. J., SIEVERT, K., PATEL, S., CHANG, Y.-F., BROYLES, C. N., BROOK, F. A., WATKINS, H., GEEVES, M. A., REDWOOD, C. S. & ROBINSON, P. 2019. Measurement of myofilament-localized calcium dynamics in adult cardiomyocytes and the effect of hypertrophic cardiomyopathy mutations. *Circulation research,* 124**,** 1228-1239.
SPÄTER, D., ABRAMCZUK, M. K., BUAC, K., ZANGI, L., STACHEL, M. W., CLARKE, J., SAHARA, M., LUDWIG, A. & CHIEN, K. R. 2013. A HCN4+ cardiomyogenic progenitor derived from the first heart field and human pluripotent stem cells. *Nature cell biology,* 15**,** 1098-1106.
STANLEY, W. C., RECCHIA, F. A. & LOPASCHUK, G. D. 2005. Myocardial substrate metabolism in the normal and failing heart. *Physiological reviews,* 85**,** 1093-1129.
SU, M., WANG, J., KANG, L., WANG, Y., ZOU, Y., FENG, X., WANG, D., AHMAD, F., ZHOU, X. & HUI, R. 2014. Rare variants in genes encoding MuRF1 and MuRF2 are modifiers of hypertrophic cardiomyopathy. *International journal of molecular sciences,* 15**,** 9302-9313.
SUNG, P. & KLEIN, H. 2006. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol,* 7**,** 739-50.
TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* 126**,** 663-76.
TANIGUCHI, M., SANBO, M., WATANABE, S., NARUSE, I., MISHINA, M. & YAGI, T. 1998. Efficient production of Cre-mediated site-directed recombinants through the utilization of the puromycin resistance gene, pac: a transient gene-integration marker for ES cells. *Nucleic acids research,* 26**,** 679-680.
TAYAL, U., NEWSOME, S., BUCHAN, R., WHIFFIN, N., WALSH, R., BARTON, P. J., WARE, J. S., COOK, S. A. & PRASAD, S. K. J. J. O. T. A. C. O. C. 2017. Truncating variants in titin independently predict early arrhythmias in patients with dilated cardiomyopathy. *J. Am. Coll. Cardiol.,* 69**,** 2466-2468.
TESSARI, A., PIETROBON, M., NOTTE, A., CIFELLI, G., GAGE, P. J., SCHNEIDER, M. D., LEMBO, G. & CAMPIONE, M. 2008. Myocardial Pitx2 differentially regulates the left atrial identity and ventricular asymmetric remodeling programs. *Circulation research,* 102**,** 813-822.
TOEPFER, C. N., SHARMA, A., CICCONET, M., GARFINKEL, A. C., MÜCKE, M., NEYAZI, M., WILLCOX, J. A., AGARWAL, R., SCHMID, M. & RAO, J. 2019. SarcTrack: an adaptable software tool for efficient large-scale analysis of sarcomere function in hiPSC-cardiomyocytes. *Circulation research,* 124**,** 1172-1183.
TOHYAMA, S., HATTORI, F., SANO, M., HISHIKI, T., NAGAHATA, Y., MATSUURA, T., HASHIMOTO, H., SUZUKI, T., YAMASHITA, H. & SATOH, Y. 2013. Distinct metabolic flow enables large-scale purification of mouse and human pluripotent stem cell-derived cardiomyocytes. *Cell stem cell,* 12**,** 127-137.
TRAN, D. B., WEBER, C. & LOPEZ, R. A. 2020. Anatomy, Thorax, Heart Muscles. *StatPearls [Internet].* StatPearls Publishing.
TROMBITÁS, K., REDKAR, A., CENTNER, T., WU, Y., LABEIT, S. & GRANZIER, H. 2000. Extensibility of isoforms of cardiac titin: variation in contour length of molecular subsegments provides a basis for cellular passive stiffness diversity. *Biophysical journal,* 79**,** 3226-3234.
UEMURA, N., OHKUSA, T., HAMANO, K., NAKAGOME, M., HORI, H., SHIMIZU, M., MATSUZAKI, M., MOCHIZUKI, S., MINAMISAWA, S. & ISHIKAWA, Y. 2004. Down-regulation of sarcolipin mRNA expression in chronic atrial fibrillation. *European journal of clinical investigation,* 34**,** 723-730.
VERMIJ, S. H., ABRIEL, H. & VAN VEEN, T. A. 2017. Refining the molecular organization of the cardiac intercalated disc. *Cardiovascular research,* 113**,** 259-275.

VIKHOREV, P. G., SMOKTUNOWICZ, N., MUNSTER, A. B., COPELAND, O. N., KOSTIN, S., MONTGIRAUD, C., MESSER, A. E., TOLIAT, M. R., LI, A. & DOS REMEDIOS, C. G. 2017. Abnormal contractility in human heart myofibrils from patients with dilated cardiomyopathy due to mutations in TTN and contractile protein genes. *Scientific reports,* 7**,** 1-11.

VIKHOREV, P. G., VIKHOREVA, N. N., YEUNG, W., LI, A., LAL, S., DOS REMEDIOS, C. G., BLAIR, C. A., GUGLIN, M., CAMPBELL, K. S. & YACOUB, M. H. 2022. Titin-truncating mutations associated with dilated cardiomyopathy alter length-dependent activation and its modulation via phosphorylation. *Cardiovascular research,* 118**,** 241-253.

VOLPATO, V., SMITH, J., SANDOR, C., RIED, J. S., BAUD, A., HANDEL, A., NEWEY, S. E., WESSELY, F., ATTAR, M. & WHITELEY, E. 2018. Reproducibility of molecular phenotypes after long-term differentiation to human iPSC-derived neurons: a multi-site omics study. *Stem cell reports,* 11**,** 897-911.

WALKLATE, J., FERRANTINI, C., JOHNSON, C. A., TESI, C., POGGESI, C. & GEEVES, M. A. 2021. Alpha and beta myosin isoforms and human atrial and ventricular contraction. *Cellular and Molecular Life Sciences,* 78**,** 7309-7337.

WANG, J., SCHWINGER, R., FRANK, K., MÜLLER-EHMSEN, J., MARTIN-VASALLO, P., PRESSLEY, T. A., XIANG, A., ERDMANN, E. & MCDONOUGH, A. A. 1996. Regional expression of sodium pump subunits isoforms and Na+-Ca++ exchanger in the human heart. *The Journal of clinical investigation,* 98**,** 1650-1658.

WANG, Y., WANG, F., WANG, R., ZHAO, P. & XIA, Q. 2015. 2A self-cleaving peptide-based multi-gene expression system in the silkworm Bombyx mori. *Scientific reports,* 5**,** 1-10.

WATKINS, H., ASHRAFIAN, H. & REDWOOD, C. 2011. Inherited cardiomyopathies. *New England Journal of Medicine,* 364**,** 1643-1656.

WEINERT, S., BERGMANN, N., LUO, X., ERDMANN, B. & GOTTHARDT, M. 2006. M line–deficient titin causes cardiac lethality through impaired maturation of the sarcomere. *The Journal of cell biology,* 173**,** 559-570.

WENG, L.-C., HALL, A. W., CHOI, S. H., JURGENS, S. J., HAESSLER, J., BIHLMEYER, N. A., GRARUP, N., LIN, H., TEUMER, A. & LI-GAO, R. 2020. Genetic determinants of electrocardiographic P-wave duration and relation to atrial fibrillation. *Circulation: Genomic and Precision Medicine,* 13**,** 387-395.

WILLEMS, E., SPIERING, S., DAVIDOVICS, H., LANIER, M., XIA, Z., DAWSON, M., CASHMAN, J. & MERCOLA, M. 2011. Small-molecule inhibitors of the Wnt pathway potently promote cardiomyocytes from human embryonic stem cell–derived mesoderm. *Circulation research,* 109**,** 360-364.

WILLIS, M. S., BEVILACQUA, A., PULINILKUNNIL, T., KIENESBERGER, P., TANNU, M. & PATTERSON, C. 2014. The role of ubiquitin ligases in cardiac disease. *Journal of molecular and cellular cardiology,* 71**,** 43-53.

WILLIS, M. S., IKE, C., LI, L., WANG, D.-Z., GLASS, D. J. & PATTERSON, C. 2007. Muscle ring finger 1, but not muscle ring finger 2, regulates cardiac hypertrophy in vivo. *Circulation research,* 100**,** 456-459.

WILLIS, M. S., SCHISLER, J. C., LI, L., RODRÍGUEZ, J. E., HILLIARD, E. G., CHARLES, P. C. & PATTERSON, C. 2009. Cardiac muscle ring finger-1 increases susceptibility to heart failure in vivo. *Circulation research,* 105**,** 80-88.

WILSON, C., BELLEN, H. J. & GEHRING, W. J. 1990. Position effects on eukaryotic gene expression. *Annu Rev Cell Biol,* 6**,** 679-714.

WITT, S. H., GRANZIER, H., WITT, C. C. & LABEIT, S. 2005. MURF-1 and MURF-2 target a specific subset of myofibrillar proteins redundantly: towards understanding MURF-dependent muscle ubiquitination. *Journal of molecular biology,* 350**,** 713-722.

WU, S.-P., CHENG, C.-M., LANZ, R. B., WANG, T., RESPRESS, J. L., ATHER, S., CHEN, W., TSAI, S.-J., WEHRENS, X. H. & TSAI, M.-J. 2013. Atrial identity is determined by a COUP-TFII regulatory network. *Developmental cell,* 25**,** 417-426.

WU, X., KRIZ, A. J. & SHARP, P. A. 2014. Target specificity of the CRISPR-Cas9 system. *Quantitative biology,* 2**,** 59-70.

XIAO, B., ZHONG, G., OBAYASHI, M., YANG, D., CHEN, K., WALSH, M. P., SHIMONI, Y., CHENG, H., TER KEURS, H. & CHEN, S. W. 2006. Ser-2030, but not Ser-2808, is the major phosphorylation site in cardiac ryanodine receptors responding to protein kinase A activation upon β-adrenergic stimulation in normal and failing hearts. *Biochemical Journal,* 396**,** 7-16.

XIE, L.-H., SHANMUGAM, M., PARK, J. Y., ZHAO, Z., WEN, H., TIAN, B., PERIASAMY, M. & BABU, G. J. 2012. Ablation of sarcolipin results in atrial remodeling. *American Journal of Physiology-Cell Physiology,* 302**,** C1762-C1771.

YANG, Q., HEWETT, T. E., KLEVITSKY, R., SANBE, A., WANG, X. & ROBBINS, J. 2001. PKA-dependent phosphorylation of cardiac myosin binding protein C in transgenic mice. *Cardiovascular research,* 51**,** 80-88.

YANG, X., PABON, L. & MURRY, C. E. 2014. Engineering adolescence: maturation of human pluripotent stem cell–derived cardiomyocytes. *Circulation research,* 114**,** 511-523.

YANG, X., RODRIGUEZ, M. L., LEONARD, A., SUN, L., FISCHER, K. A., WANG, Y., RITTERHOFF, J., ZHAO, L., KOLWICZ JR, S. C. & PABON, L. 2019. Fatty acids enhance the maturation of cardiomyocytes derived from human pluripotent stem cells. *Stem Cell Reports,* 13**,** 657-668.

YECHIKOV, S., COPACIU, R., GLUCK, J. M., DENG, W., CHIAMVIMONVAT, N., CHAN, J. W. & LIEU, D. K. 2016. Same-Single-Cell Analysis of Pacemaker-Specific Markers in Human Induced Pluripotent Stem Cell-Derived Cardiomyocyte Subtypes Classified by Electrophysiology. *Stem Cells,* 34**,** 2670-2680.

YECHIKOV, S., KAO, H. K., CHANG, C.-W., PRETTO, D., ZHANG, X.-D., SUN, Y.-H., SMITHERS, R., SIRISH, P., NOLTA, J. A. & CHAN, J. W. 2020. NODAL inhibition promotes differentiation of pacemaker-like cardiomyocytes from human induced pluripotent stem cells. *Stem cell research,* 49**,** 102043.

YEH, Y.-H., WAKILI, R., QI, X.-Y., CHARTIER, D., BOKNIK, P., KÄÄB, S., RAVENS, U., COUTU, P., DOBREV, D. & NATTEL, S. 2008. Calcium-handling abnormalities underlying atrial arrhythmogenesis and contractile dysfunction in dogs with congestive heart failure. *Circulation: Arrhythmia and Electrophysiology,* 1**,** 93-102.

YIP, B. H. 2020. Recent advances in CRISPR/Cas9 delivery strategies. *Biomolecules,* 10**,** 839.

ZAMMIT, P. S., KELLY, R. G., FRANCO, D., BROWN, N., MOORMAN, A. F. & BUCKINGHAM, M. E. 2000. Suppression of atrial myosin gene expression occurs independently in the left and right ventricles of the developing mouse heart. *Developmental dynamics: an official publication of the American Association of Anatomists,* 217**,** 75-85.

ZAUNBRECHER, R. J., ABEL, A. N., BEUSSMAN, K., LEONARD, A., VON FRIELING-SALEWSKY, M., FIELDS, P. A., PABON, L., REINECKE, H., YANG, X. & MACADANGDANG, J. 2019. Cronos titin is expressed in human cardiomyocytes and necessary for normal sarcomere function. *Circulation,* 140**,** 1647-1660.

ZHANG, J., TAO, R., CAMPBELL, K. F., CARVALHO, J. L., RUIZ, E. C., KIM, G. C., SCHMUCK, E. G., RAVAL, A. N., DA ROCHA, A. M. & HERRON, T. J. 2019a. Functional cardiac fibroblasts derived from human pluripotent stem cells via second heart field progenitors. *Nat. Commun.,* 10**,** 1-15.

ZHANG, J., WILSON, G. F., SOERENS, A. G., KOONCE, C. H., YU, J., PALECEK, S. P., THOMSON, J. A. & KAMP, T. J. 2009. Functional cardiomyocytes derived from human induced pluripotent stem cells. *Circulation research,* 104**,** e30-e41.

ZHANG, J. Z., TERMGLINCHAN, V., SHAO, N. Y., ITZHAKI, I., LIU, C., MA, N., TIAN, L., WANG, V. Y., CHANG, A. C. Y., GUO, H., KITANI, T., WU, H., LAM, C. K., KODO, K., SAYED, N., BLAU, H. M. & WU, J. C. 2019b. A Human iPSC Double-Reporter System Enables Purification of Cardiac Lineage Subpopulations with Distinct Function and Drug Response Profiles. *Cell Stem Cell,* 24**,** 802-811 e5.

ZHANG, P. 2017. CaMKII: The molecular villain that aggravates cardiovascular disease. *Experimental and therapeutic medicine,* 13**,** 815-820.

ZHANG, Q., JIANG, J., HAN, P., YUAN, Q., ZHANG, J., ZHANG, X., XU, Y., CAO, H., MENG, Q., CHEN, L., TIAN, T., WANG, X., LI, P., HESCHELER, J., JI, G. & MA, Y. 2011. Direct differentiation of atrial and ventricular myocytes from human embryonic stem cells by alternating retinoid signals. *Cell Res.,* 21**,** 579-87.

ZHANG, W., KAI, K., CHOI, D. S., IWAMOTO, T., NGUYEN, Y. H., WONG, H., LANDIS, M. D., UENO, N. T., CHANG, J. & QIN, L. 2012. Microfluidics separation reveals the stem-cell–like deformability of tumor-initiating cells. *Proceedings of the National Academy of Sciences,* 109**,** 18707-18712.

ZHAO, M.-T., SHAO, N.-Y. & GARG, V. 2020. Subtype-specific cardiomyocytes for precision medicine: Where are we now? *Stem Cells,* 38**,** 822-833.

ZHU, S., GULERIA, R. S., THOMAS, C. M., ROTH, A., GERILECHAOGETU, F., KUMAR, R., DOSTAL, D. E., BAKER, K. M. & PAN, J. 2016. Loss of myocardial retinoic acid receptor α induces diastolic dysfunction by promoting intracellular oxidative stress and calcium mishandling in adult mice. *Journal of molecular and cellular cardiology,* 99**,** 100-112.

ZOU, J., TRAN, D., BAALBAKI, M., TANG, L. F., POON, A., PELONERO, A., TITUS, E. W., YUAN, C., SHI, C. & PATCHAVA, S. 2015. An internal promoter underlies the difference in disease severity between N-and C-terminal truncation mutations of Titin in zebrafish. *elife,* 4**,** e09406.

Appendix

Table S1 Recipes for buffers and solutions

Buffer	Composition
2X Binding buffer for lssDNA purification	10 mM Tris pH8, 1mM EDTA, 2M NaCl
Cell lysis buffer for gDNA extraction	0.1% SDS, 400 µg/ml proteinase K and 40 µg/ml RNAse A
Hybridisation Buffer for lssDNA purification	1x CutSmart buffer and 20 µM hybridisation primer (Table S4)
Neutralising Solution for lssDNA purification	40 mM HCl and 0.1 M Tris pH 8, 0.02 M EDTA, 0.2% w/v
Orange-G DNA loading buffer	50% v/v glycerol, 0.1 M Tris-HCl pH 8, 0.2% w/v Orange G
Tween buffer for lssDNA purification	5 mM Tris pH8, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween
Tyrode's-HEPES solution	141.4 mM NaCl, 5 mM HEPES, 1 mM $MgCl_2$, 1.8 mM KCl, 0.35 mM NaH_2PO_4, 420 µM $CaCl_2$, and 5.56 mM glucose pH 7.4

Table S2 List of commercial kits and other commercial products

Kits/products	Company and catalogue number
10X TGS running buffer	Biorad #1610732
4X Laemmli Sample Buffer	Biorad #1610747
50X TAE	Severn Biotech #20-60001-50
Agarose	Sigma #A9539-500G
All-trans-retinoic acid	Sigma #R2625-50MG
Alt-R® S.p. HiFi Cas9 Nuclease V3	IDT #1081060
AnykD™ Mini-PROTEAN®-TGX™ Stain-free Precast Protein Gel	Biorad #456-8129
ATP-dependent Plasmid Safe Nuclease, including 25 mM ATP	Epicentre #E3101K
B27™ Supplement	Gibco #17504-044
B27™ Supplement, minus insulin	Gibco #A18956-01
BD Stemflow™ Human and Mouse Pluripotent Stem Cell Analysis Kit	BD Biosciences #560477
CHIR-99021	Selleck #S2924
CloneJET PCR Cloning Kit	ThermoFisher #D1232
CloneR™	STEMCELL #05888
cOmplete™, EDTA-free Protease Inhibitor Cocktail	Roche #4693132001
CutSmart buffer	NEB # B7204
DNAse I	NEB #M0303S
Dulbecco's Modified Eagle Medium/F12	Gibco #11320033
Dynabeads™ M280 Streptavidin	ThermoFisher #11205D
ElectroSHOX competent cells	Bioline #BIO-85038
EnGen® sgRNA Synthesis Kit, *S. pyogenes*	NEB #E3322
EnGen® Spy Cas9 NLS	NEB# M0646T
Essential 8™ medium	Gibco #A1517001
Fish gelatin	Sigma #G7041-100G
Fluoroshield™ with DAPI	Sigma #F6057-20ML
Foetal bovine serum	Gibco #A4766801

FuGENE6®	Promega #E2691
GeneJet Gel Extraction Kit	ThermoFisher #K0692
GeneJET Plasmid Midiprep Kit	ThermoFisher #K0482
GeneJET Plasmid Miniprep Kit	ThermoFisher #K0503
IWR1	Selleck #S7086
KnockOut™ Serum Replacement	Gibco #10828010
Monarch® DNA Elution Buffer	NEB # T1016L
Monarch® PCR & DNA Cleanup Kit	NEB #T1030S
NEBuilder® HiFi DNA Assembly Cloning Kit containing NEB® 5-alpha Competent *E. coli*	NEB #E5520S
Neon™ Transfection System 10 µl kit	Invitrogen #MPK1096
Opti-MEM™ I Reduced Serum Medium	Gibco #31985062
PBS, pH7.4	Gibco #10010023
Pierce™ BCA Protein Assay Kit	ThermoFisher #23227
PowerUp™ SYBR™ Green Master Mix	Applied Biosystems #A25742
qPCRBIO cDNA Synthesis Kit	PCRBiosystems #PB30.11-10
Quick Calf Intestinal Alkaline Phosphatase	NEB #M0525S
Reduced Growth Factor, Phenol Red-free Matrigel	Corning #356231
ReLeaSR™	STEMCELL #05872
RevitaCell™	Gibco #A2644501
RIPA Lysis and Extraction Buffer	ThermoFisher #89900
RPMI-1640	Gibco #11875-093
RPMI-1640, no glucose	Gibco #11879020
siR-DNA	Universal Biologicals #CY-SC007
T4 DNA ligase	NEB #M0202S
TaqMan™ Fast Advanced Master Mix, no UNG	Applied Biosystems #AA359
Thiazovivin	Selleck #S1456
Trans-Blot Turbo Mini 0.2 µm PVDF Transfer Packs	Biorad #1704156
Tris-EDTA	Invitrogen #12090015
TRIzol™	Invitrogen #15596026
TrypLE™ Select Enzyme (1X), no phenol red	Gibco #12563011
Trypsin-EDTA (0.25%), phenol red	Gibco #25200056
Tween-20 (Polyoxyethylene-20-Sorbitan Monolaurate)	VWR #9005-64-5
Vitronectin (VTN-N) Recombinant Human Protein, Truncated	Gibco #A14700
Y-27632 (Dihydrochloride)	STEMCELL #72302

Table S3 List of instruments

Equipment/ instrument	**Company**
CFX96 Touch Real-Time PCR Detection System	Biorad
ChemiDoc™ MP Imaging System	Biorad
FX500 cell sorter	Sony
Gene Pulser Xcell	Biorad
IX81 microscope	Olympus
LSRFORTESSA™ X20 CELL ANALYZER	BD Biosciences
Neon™ Transfection System	Invitrogen

SP8 lightning confocal microscope	Leica
T100™ Thermal Cycler	Biorad
Trans-Blot Turbo transfer system	Biorad

Table S4 List of primers and DNA oligos

DNA oligo	Sequence (5' to 3')
bGH-F2	CAGCCTCGACTGTGCCTTCTAG
Exneo2	GTTGTGCCCAGTCATAGCCGAATAG
HCN4-3HR-F2	CTGTACAAGTAGGCTGGGCCCTTCCTTCCC
HCN4-3HR-R2	CCCCGGGTCTAGGGTACCTCTAGGAAGCTTTCACCTCCCACTGTACAGATGG
HCN4-5HR-F2	GGAATTCTCTAGGAGATCTTCTAGCTCGAGCCTTTATGCCTAAGCCAG
HCN4-5HR-R2	TGGCTCCGCTTCCTAGATTGGATGGCAGTTTG
HCN4-F	TCCTCAGGTTCTTTGCCACCC
HCN4-F2	AGGAGTACCCCATGATGCGA
HCN4-F3	GACACCTACTGCCGCCTCTA
HCN4-R4	GCGTGGACTCGAACCTCAAA
HCN4-mScarlet-F2	CCATCCAATCTAGGAAGCGGAGCCACGAAC
HCN4-mScarlet-R2	GGAAGGGCCCAGCCTACTTGTACAGCTCGTCCATG
HCN4-R	CCATCTGCCTTTCTCTGGCT
hTTN-F5	GCTTTGCAGGGGGTATCAGT
hTTN-F6	AACCTGGGCCACCTAGAGAT
hTTN-qF1	GTCCCCATCGCCCATAAGAC
hTTN-qR1	GAGGCACTTCAGGACCTGTG
hTTN-qF3	TGCTACTGTCCAGGGTCTCA
hTTN-qR3	CTTCCAGGTCTTCAGGACGC
hTTN-qF4	ACCCGCTTTGAGGTTACTGG
hTTN-qR4	CCGCATCAAGTTCTCCCTCA
hTTN-qF6	GGATCTGTGTCCTCTAGCTGC
hTTN-qR6	GAAGCTTCCTCTTTGAGGGCT
hTTN-qF7	CCCGGGAGCTCAAGAAGAAA
hTTN-qR7	TGACTTTGGGTGTGGCAACT
hTTN-qF9	TGAGGCCTTGAATGACAGCG
hTTN-qR9	GATCACTGGGGCAGCTCTTT
mClover5'-R	TTCAGGGTCAGCTTGCCGTT
mClover-F	TCTTTCAAGGACGACGGGAC
mClover-F2	CACAACGTTGAGGACGGCAG
mClover-qF*	GACGACGGGACCTACAAGAC
mClover-qR*	GATGTTGCCGTCCTCCTTGA
mClover-R	CGCTGAACTTGTGGCCGT
mScarlet-F	TACAAGGCCAAGAAGCCCGT

mScarlet-qF*	CGCGTGATGAACTTCGAGGA
mScarlet-qR*	GCGGAGCTTCACCTTGTAGA
mScarlet-R	ACCTTCAGCTTGGCGGTCTG
MYL2-A1	CACCGCATCACCCACGGAGAAGAGA
MYL2-A2	AAACTCTCTTCTCCGTGGGTGATGC
MYL2-F	CTGCAAGGCTGTCACTAGCT
MYL2-F4	TACCCAGGGAGGGAGCTTAT
MYL2-F5	TGGACCCTGTGGTCACTTCAT
MYL2-R	CCTCGTCCTTAGCACGTGTT
MYL2-R3	CTACCATGTGTCAGTCCGCT
MYL2-R5	GCACGTGTTGCTGGCTCATT
MYL7-3HR-F	CCTGCATGATAAGGGGCAGGGCCAGGCCCA
MYL7-3HR-R2	CCCCGGGTCTAGGGTACCTCTAGGACACTGTTCCAGCATGGTC
MYL7-5HR-F	GGAATTCTCTAGGAGATCTTCTAGCTTCCCAGCCCCACTGATC
MYL7-5HR-R	GGCCCCGCTTCCTTCCTCTTTCTCGTCTCCATGAGTAATAATA
MYL7-F2	CAGACAGAAAGGCCACAGGA
MYL7-F3	CACTGGGGTCCCTACTCTGA
MYL7-mClover-F	CGAGAAAGAGGAAGGAAGCGGGGCCACGAAC
MYL7-mClover-R	CTGGCCCTGCCCCTTATCATGCAGGTGCTACCTGAGAC
MYL7-R2	CCAGCTCCCAGGAGGTAGAT
MYL7-R3	CAAAAGGTTCAGACACCCCATCAA
NS-TV-link1	GATCCATTGTACAGGGCCCCTCGAGA
NS-TV-link2	GATCTCTCGAGGGGCCCTGTACAATG
NS-TV-link3	GTACAAGTAGCAATTGTTACTAGTGGGCCCA
NS-TV-link4	AGCTTGGGCCCACTAGTAACAATTGCTACTT
PITX2-3HR-F	CCTGCATGATAAGCCGCAACCACAGCGCCG
PITX2-3HR-R2	CCCCGGGTCTAGGGTACCTCTAGGAGGTGTAGGTTTCCCCGTAAAGGAGT
PITX2-5HR-F2	GGAATTCTCTAGGAGATCTTCTAGCCAGGGAAGGAAGGGAGGAATTGCGT
PITX2-5HR-R	GGCCCCGCTTCCCACGGGCCGGTCCACTGC
PITX2-F	CGCCGACTCCTCCGTATGTT
PITX2-F2	ATCCTTCGAAAACGCCCTCC
PITX2-mClover-F	GGACCGGCCCGTGGGAAGCGGGGCCACGAAC
PITX2-mClover-R	CTGTGGTTGCGGCTTATCATGCAGGTGCTACCTGAGAC
PITX2-R	TGTCAGACACTGAGGACATCCC
PITX2-R2	GGGAGCATGACCTTTGCCTC

pJET1.2-F	CGACTCACTATAGGGAGAGCGGC (ThermoFisher #D1232)
pJET1.2-R	AAGAACATCGATTTTCCATGGCAG (ThermoFisher #D1232)
pJET-Btn-F	CGACTCACTATAGGGAGAGCGGC
pJET-Hybridization	GATATCTTGCTGAAAAACTCGAGCCATCCG
T7-HCN4-A	TTCTAATACGACTCACTATAG**GAAGGGCCCAGCTCATAGAT**GTTTTAGAGCTAGA
T7-HCN4-B	TTCTAATACGACTCACTATAGG**CTGCCATCCAATCTATGAGC**GTTTTAGAGCTAGA
T7-MYL2-A	TTCTAATACGACTCACTATAGG**CCACGGAGAAGAGAAGGACT**GTTTTAGAGCTAGA
T7-MYL2-B	TTCTAATACGACTCACTATAGG**CATCACCCACGGAGAAGAGA**GTTTTAGAGCTAGA
T7-MYL7-A	TTCTAATACGACTCACTATAGG**CACCCATGGAGACGAGAAAG**GTTTTAGAGCTAGA
T7-MYL7-B	TTCTAATACGACTCACTATAGG**ATTCCTCTTTCTCGTCTCCA**GTTTTAGAGCTAGA
T7-MYL7-C	TTCTAATACGACTCACTATAGG**CTGTGCTACATCATCACCCA**GTTTTAGAGCTAGA
T7-PITX2-A	TTCTAATACGACTCACTATAG**GTCCTAGGATCCCGGCGCTGT**GTTTTAGAGCTAGA
T7-PITX2-B	TTCTAATACGACTCACTATAGG**TGGGTGCGGCTCACACGGGC**GTTTTAGAGCTAGA

Bold, 20-nt CRISPR sequences; *, used in SYBR green assay.

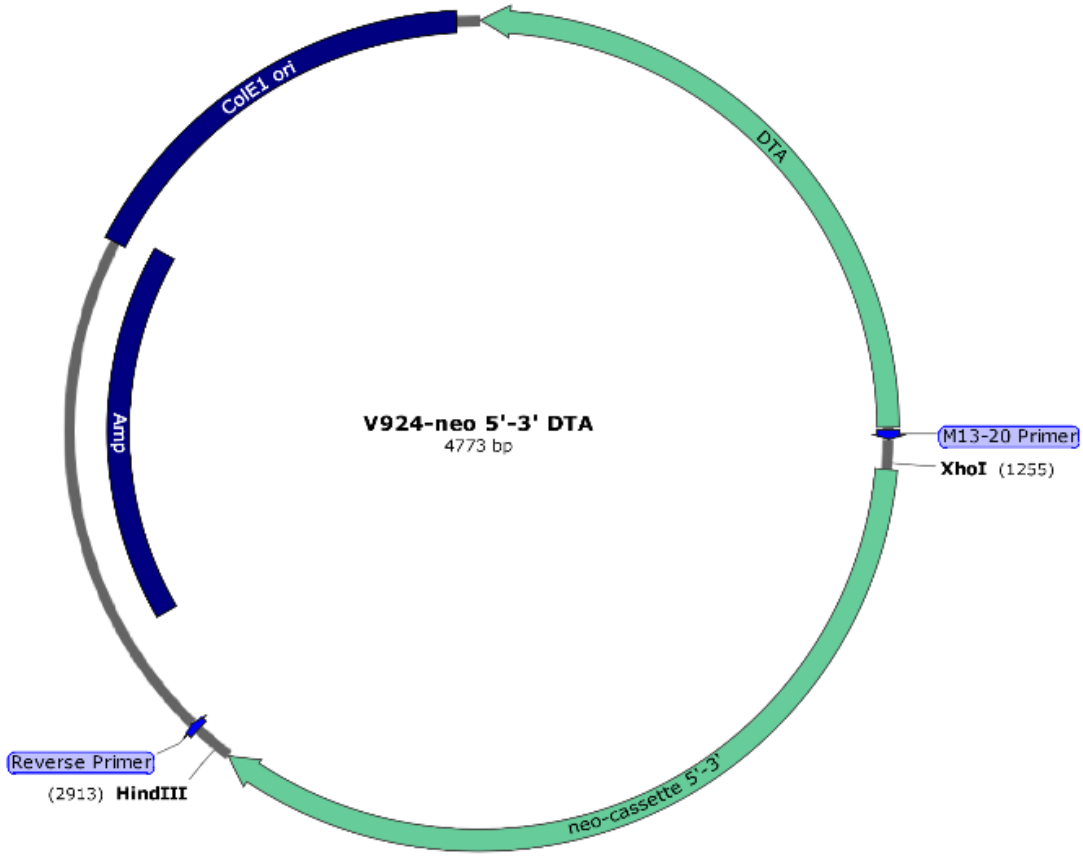


Figure S1 Plasmid map for V294-neo-dTA. This vector was used as a backbone for generating the targeting vectors described in Chapter 3

Figure S2 Plasmid map for PL452. This plasmid was used to obtain the *Neo* cassette fragment in Chapter 3.

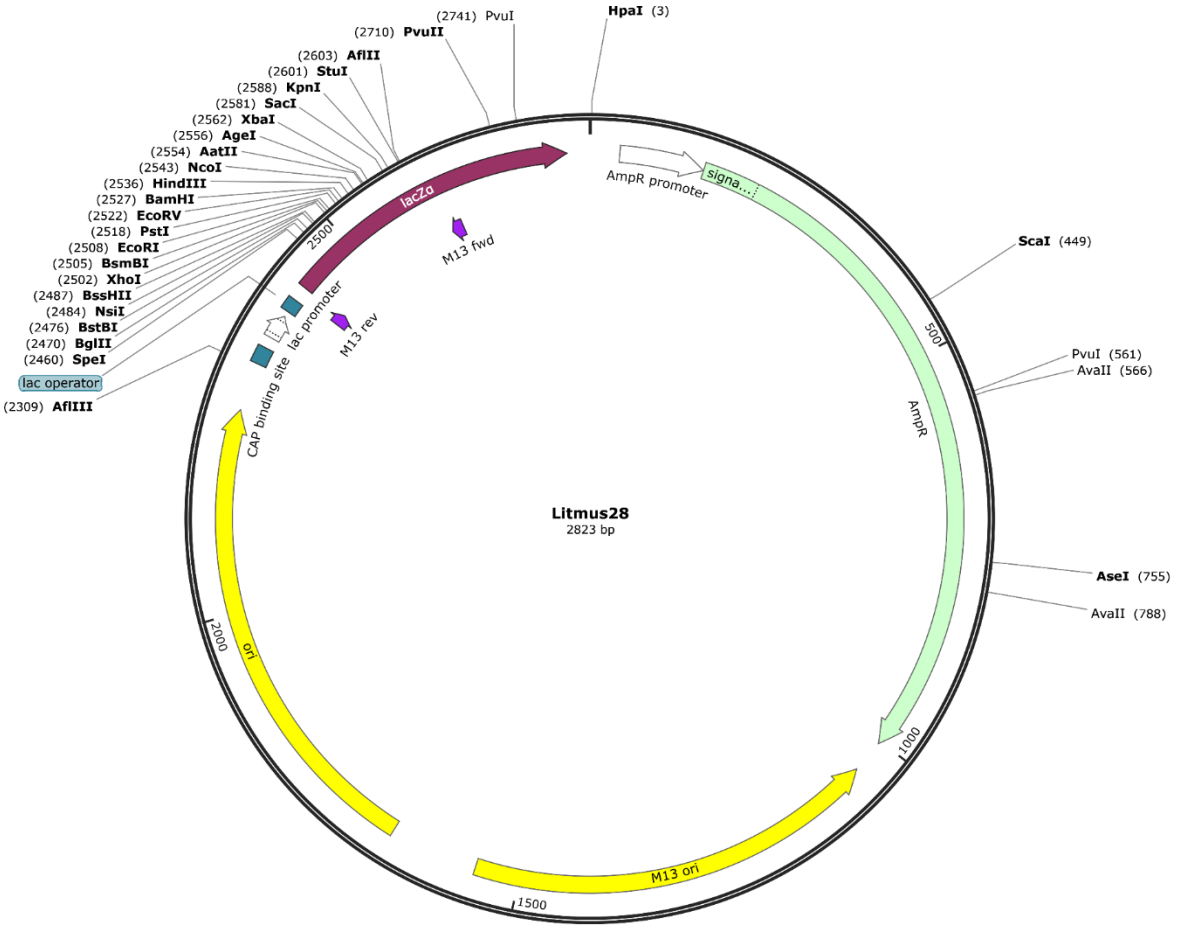


Figure S3 Plasmid map for Litmus28. This plasmid was used for cloning the PCR products in Chapter 3 at the multiple cloning site (*LacZα*).

Figure S4 Plasmid map for pX330-puro. This plasmid was used for Golden Gate cloning of gRNA sequence at the *Bbs*I sites in Chapter 3.

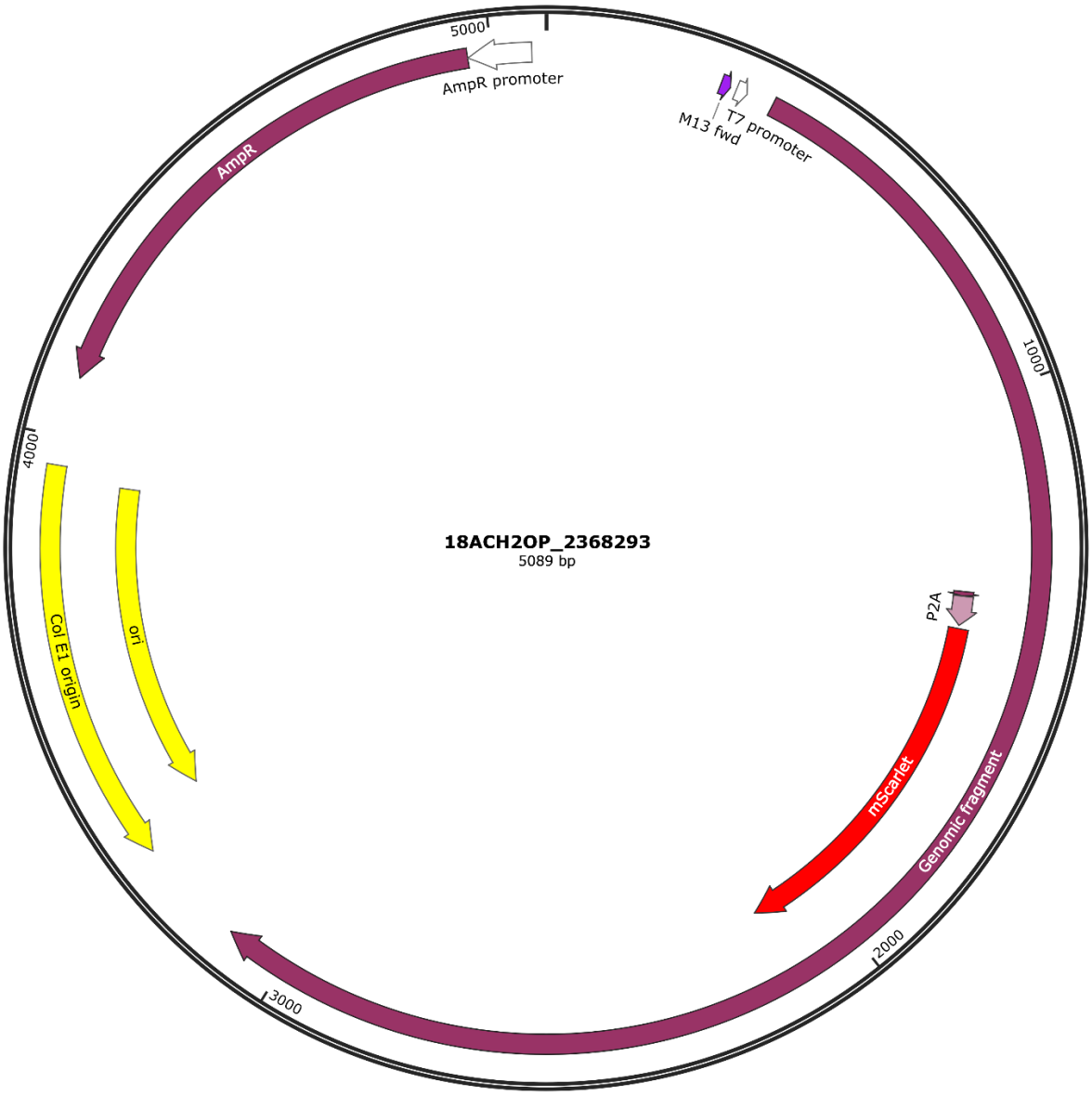


Figure S5 Plasmid map for 18ACH20P_2368293. This plasmid contains a *mSarlet* open reading frame required for generation of the fluorescent reporter in Chapter 3.

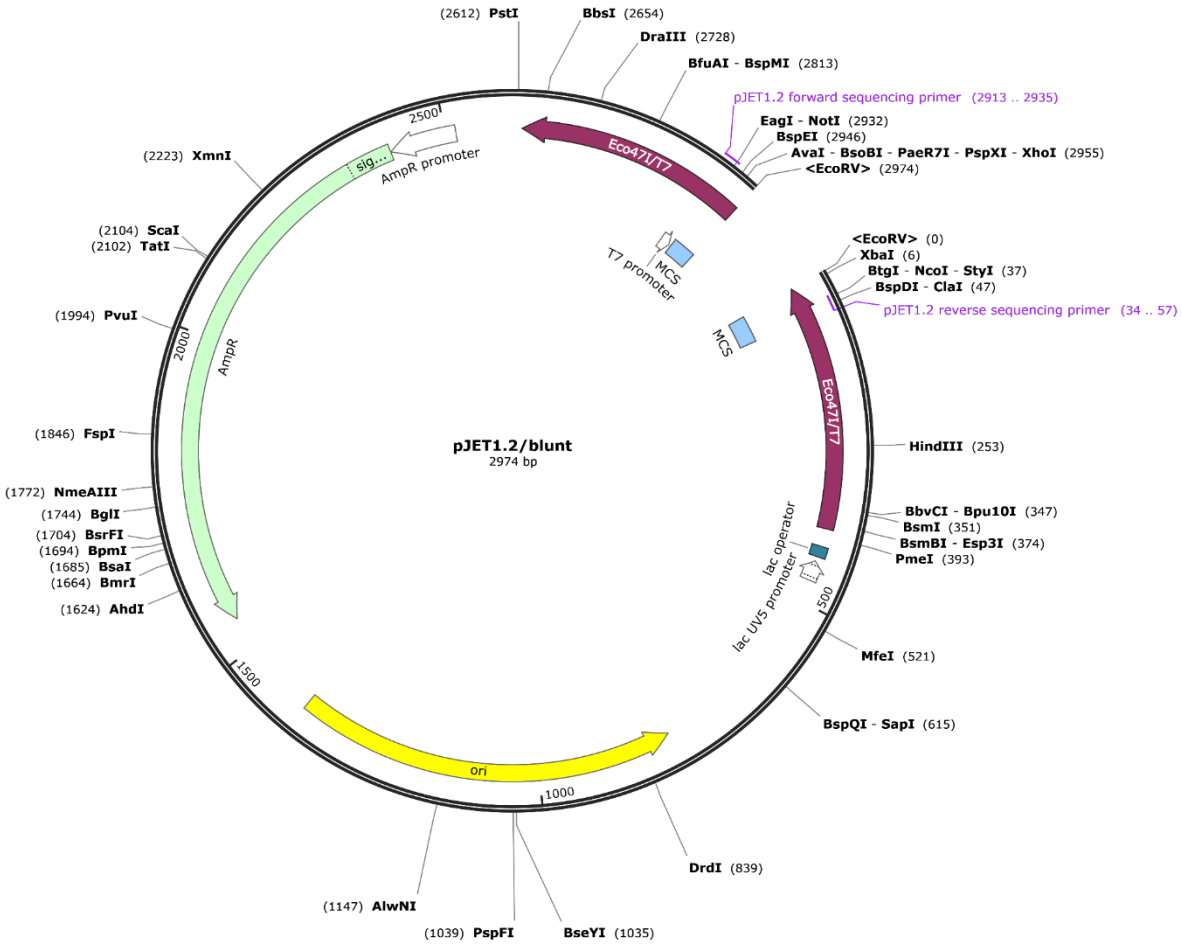


Figure S6 Plasmid map for pJET1.2/blunt. This plasmid was used for cloning the blunt end PCR products in Chapter 3.

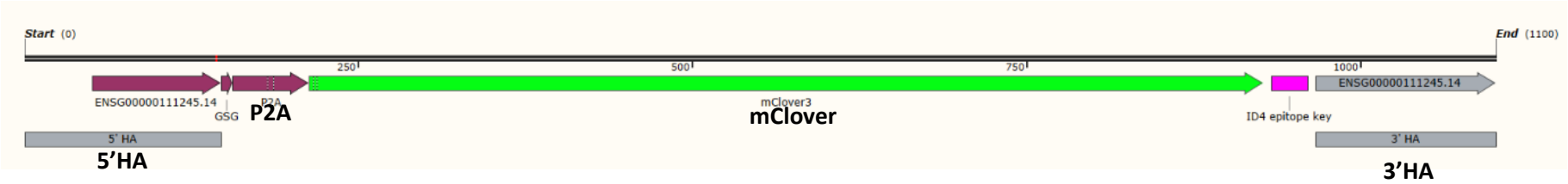


Figure S7 *MYL2*-P2A-*mClover* gBlock template.