The effects of kirrel1 isoform expression on C2C12 differentiation and fusion *in vitro*

by

Mr Rhys Stewart McColl

Dissertation presented for the degree of

Doctor of Philosophy in the Faculty of Science at

Stellenbosch University

Supervisor: **Professor Kathryn H. Myburgh**

Co-supervisor: **Dr Peter Durcan**

Department of Physiological Sciences

March 2023

Declaration

"By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third-party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification."

March 2023

Verklaring

"Deur hierdie tesis elektronies in te lewer, verklaar ek dat die geheel van die werk hierin vervat, my eie, oorspronklike werk is, dat ek die alleen outeur daarvan is (behalwe in die mate uitdruklik anders aangedui), dat reproduksie en publikasie daarvan deur die Universiteit van Stellenbosch nie derdepartyregte sal skend nie en dat ek dit nie vantevore, in die geheel of gedeeltelik, ter verkryging van enige kwalifikasie aangebied het nie."

Maart 2023

Abstract

Adult skeletal muscle myogenesis involves the fusion of muscle progenitor cells into multi-nucleated myofibers, a process crucial for the growth and repair of muscle tissue. Vertebrate myoblast fusion is a relatively poorly understood process that involves a multitude of cell adhesion molecules, actin regulators and fusion proteins. A more comprehensive understanding of myogenesis is essential to better assess muscle myopathies and for the development of improved interventions. The kirrel family of mammalian cell adhesion molecules are highly involved in the production and maintenance of complex tissue structures such as the slit diaphragm in the kidney. The *Drosophila* paralogs of the kirrel proteins are known to be vital for actin regulation during myoblast fusion with the mechanisms of this regulation being mostly understood. However, these same findings have not been confirmed with regards to mammalian myoblast fusion; an arguably more complex process than that in the fly. It has been demonstrated that kirrel1A and its associated splice variant, kirrel1B, are differentially expressed in regenerating mouse muscle tissue; although, the exact roles of these molecules during this process are not clear. More recently, kirrel3 has been shown to be required for the successful fusion of mouse myoblasts. The aim of this study was to determine the effects of *kirrel1A* and *kirrel1B* expression levels on C2C12 differentiation and fusion *in vitro*. Three genetic strategies were employed to assess *kirrel1* activity during C2C12 myogenesis, these being; CRISPR/Cas9 modification, shRNA knockdown and retroviral overexpression. CRISPR/Cas9 was used to disrupt *kirrel1* expression by modifying genomic regulatory regions between exons 1 and 2 of the gene. The individual knockdown/inhibition of *kirrel1A* and *kirrel1B* mRNA activity was achieved using shRNAs. Overexpression was carried out by wild-type *kirrel1A* and *kirrel1B* gene-cloning followed by retroviral transduction. Additionally, a kirrel1A-mCherry mutant was overexpressed in the C2C12s. The differentiation of the various cell lines was assessed via western blotting, PCR analysis and phase-contrast microscopy. The experiments suggest that although the moderate overexpression of *kirrel1A* or *kirrel1B* has little effect on myotube production, the gross overexpression of kirrel1 variants leads to a drastic reduction in myogenesis, potentially due to increased steric hindrance at the cell surface. Moreover, our findings demonstrate the requirement for *kirrel1A* during myotube formation as no tubes were seen in *kirrel1A*-knockdown myoblasts. This inhibition appeared to be unrelated to the expression of the myogenic regulatory factors. However, it is still unclear whether there is a similar requirement for kirrel1B during fusion. The expression of a mutant form of *kirrel1A* with an mCherry tag inserted close to an intracellular cleavage site resulted in a complete lack of myotubes; seemingly due to altered early MRF expression. In each case where cell lines produced myotubes, non-reducing western blotting revealed large kirrel1-containing complexes that accumulated as fusion progressed. These complexes were not seen in any of the non-fusing cell lines. From the results it is apparent that kirrel1 is required for healthy myogenesis and that further research is required to fully understand the mechanisms of this regulation.

Opsomming

Volwasse skeletspier miogenese behels die samesmelting van spiervoorloperselle in multi-kern miovesels, 'n proses wat noodsaaklik is vir die groei en herstel van spierweefsel. Vertebrate mioblaste samesmelting is 'n relatief swak verstaande proses wat 'n menigte sel adhesie molekules, aktien reguleerders en samesmeltings proteïene behels. 'n Meer omvattende begrip van miogenese is noodsaaklik om spiermiopatologie beter te assesseer en vir die ontwikkeling van verbeterde intervensies. Die kirrelfamilie van soogdier seladhesiemolekules is hoogs betrokke by die produksie en instandhouding van komplekse weefselstrukture soos die spleetdiafragma in die nier. Dit is bekend dat die Drosophila-paraloë van die kirrel-proteïene noodsaaklik is vir aktienregulering tydens mioblastesamesmelting, met die meganismes van hierdie regulering wat meestal verstaan word. Hierdie bevindinge is egter nie bevestig met betrekking tot soogdier mioblastfusie nie; 'n waarskynlik meer komplekse proses as dié in die vlieg. Dit is gedemonstreer dat kirrel1A en sy geassosieerde splitsingsvariant, kirrel1B, differensieel uitgedruk word in regenererende muisspierweefsel; alhoewel, die presiese rolle van hierdie molekules tydens hierdie proses is nie duidelik nie. Meer onlangs is getoon dat kirrel3 benodig word vir die suksesvolle samesmelting van muismioblaste. Die doel van hierdie studie was om die effekte van kirrel1A en kirrel1B uitdrukkingsvlakke op C2C12 differensiasie en samesmelting *in vitro* te bepaal. Drie genetiese strategieë is aangewend om kirrel1-aktiwiteit tydens C2C12-miogenese te bepaal, hierdie is; CRISPR/Cas9-modifikasie, shRNA-afslaan en retrovirale ooruitdrukking. CRISPR/Cas9 is gebruik om kirrel1-uitdrukking te ontwrig deur genomiese regulatoriese streke tussen eksons 1 en 2 van die geen te wysig. Die individuele knockdown / inhibisie van kirrel1A en kirrel1B mRNA aktiwiteit is bereik met behulp van shRNAs. Ooruitdrukking is uitgevoer deur wilde-tipe kirrel1A en kirrel1B geenkloning gevolg deur retrovirale transfeksie. Daarbenewens is 'n kirrel1A-mCherry mutant ooruitgedruk in die C2C12s. Die differensiasie van die verskillende sellyne is beoordeel deur middel van Westerse kladding, PCR-analise en fasekontrasmikroskopie. Ons eksperimente dui daarop dat alhoewel die matige ooruitdrukking van kirrel1A of kirrel1B min effek op miobuisproduksie het, die growwe ooruitdrukking van kirrel1 variante lei tot 'n drastiese vermindering in miogenese, moontlik as gevolg van verhoogde steriese hindering by die seloppervlak. Boonop demonstreer ons bevindinge die vereiste vir kirrel1A tydens miobuisvorming aangesien geen buise in kirrel1A-afslaan-mioblaste gesien is nie. Hierdie inhibisie blyk nie verband te hou met die uitdrukking van die miogeniese regulerende faktore nie. Dit is egter nog onduidelik of daar 'n soortgelyke vereiste vir kirrel1B tydens samesmelting is. Die uitdrukking van 'n mutante vorm van kirrel1A met 'n mCherry-merker wat naby 'n intrasellulêre splitsingsplek ingevoeg is, het gelei tot 'n volledige gebrek aan miobuise; skynbaar as gevolg van veranderde vroeë MRF-uitdrukking. In elke geval waar sellyne miobuise geproduseer het, het nie-reduserende Westerse kladding groot kirrel1-bevattende komplekse aan die lig gebring wat opgehoop het soos die samesmelting gevorder het. Hierdie komplekse is nie in enige van die nie-samesmeltende sellyne gesien nie. Uit die resultate is dit duidelik dat kirrel1 nodig is vir gesonde miogenese en dat verdere navorsing nodig is om die meganismes van hierdie regulasie ten volle te verstaan.

In loving memory of Stewart McColl

Acknowledgements

Firstly, I would like to thank the National Research Foundation of South Africa for their continued financial support throughout my degree.

To Prof Kathy Myburgh, thank you for taking me on as a student all those years back. I have learnt so much since joining the muscle group and have obtained valuable skills that will put me in good stead for the future. You have always been patient with us and willing to sit down for hours to iron out the details. I have had a great time doing this research and will forever be grateful for the opportunity.

To Dr. Peter Durcan, this project would have been impossible without your planning, help in the lab and guidance. I only hope that the results of this thesis are as interesting to you as they are to me.

To the various muscle group members, past and present, I appreciate all your valuable insights. Having a fun and strong research group has made my time in the department very enjoyable and the time has flown.

To Dr Isabelle Richard, thank you for hosting Dr Durcan and for your help in designing the CRISPR/Cas9 guide sequences that were used in this study.

To Professor Ehtesham Arif, thank you for the donation of your *kirrel1A*-mCherry plasmid.

To my family, Mom, Dad, Evan and Wendy. Even though we live on other sides of the country, you are all always on my mind. Mom and Dad, the selfless sacrifices you have made for your children have not gone unnoticed. Love you all.

To Kelby, my strongest supporter. Thank you for your unwavering commitment. To the other Englishes, you have been there for Kelby and I throughout this process and we can't thank you enough.

Finally, to the boys, I would have been done in half the time if it weren't for you all.

Table of contents

List of tables and figures

List of acronyms and abbreviations

Arp2/3	Actin related protein 2/3 complex
ADAM12	ADAM Metallopeptidase Domain 12
BFP	blue fluorescent protein
bHLH	basic helix-hoop-helix proteins
Blow	Blown Fuse
BOC	Brother of Cdo
BSA	bovine serum albumin
CAF	Central Analytical Facilities
CAM	cell adhesion molecule
cAMP	cyclic adenosine monophosphate
CD	cluster of differentiation
CDK	cyclin-dependent-kinase
CDS	coding DNA sequence
CIP	calf intestinal alkaline phosphatase
circRNAs	circular RNAs
CREB	cAMP-response element binding protein
Crk	CT10 regulator of kinase
CTCF	CCCTC-binding factor
Dck	dreadlocks
Dia	Diaphanous
DLL	Delta-like ligand
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	dimethylsulfoxide
DPak1 & DPak3	*Drosophila* p21-activated kinase 1
DSwip-1	*Drosophila* Swiprosin-1
D-titin	*Drosophila*-titin
Duf	Dumbfounded
D-WIP/Sltr	*Drosophila* WASP-interacting protein/Solitary
ECM	extracellular matrix
Eff-1	epithelial fusion failure 1
Elmo	engulfment and cell motility
ERK	extracellular signal-regulated kinase
FACS	fluorescence-activated cell sorting
FAK	focal adhesion kinase

FBS	fetal bovine serum
FC	muscle founder cell
FCM	fusion competent myoblast
FGF	fibroblast growth factors
FGFR	fibroblast growth factor receptor
FM	freezing medium
FNIII	fibronectin type III
FuRMAS	fusion-restricted myogenic-adhesive structure
Fz	frizzled receptors
Gas1	growth arrest-specific protein 1
GEFs	guanine nucleotide exchange factors
GFP	green fluorescent protein
GM	growth medium
GPI	glycosylphosphatidylinositol
Grb2	Growth factor receptor-bound protein 2
GTP	guanosine triphosphate
H3K4me3	histone H3 lysine K4 trimethylated
Hbs	Hibris
HGF	hepatocyte growth factor
ICAM-1	intercellular adhesion molecule 1
ID	inhibitor of DNA binding
IDT	Integrated DNA Technologies
Ig	immunoglobulin
IGF	insulin-like growth factor
IGF-1	insulin-like growth factor-1
IGF-2	insulin-like growth factor-2
IGFR	IGF receptor
IL-4	interleukin-4
IL-6	interleukin-6
IR	insulin receptor
IRM	irre recognition module
IrreC	Irregular Chiasm
IRS	insulin receptor substrate
JNK	Jun N-terminal kinases
K1Ash	kirrel1A short hairpin
K1Bsh	kirrel1B short hairpin
Kirre	Kin-of-IrreC

kirrel	Kin of IRRE-like proteins
MAPK	mitogen-activated protein kinase
Mbc	myoblast city
mdx	X-linked muscular dystrophy
MEF2	myocyte enhancer factor 2
MRF	myogenic regulatory factor
MRF4	myogenin and muscle-specific regulatory factor 4
mTOR	mammalian target of rapamycin
Myf5	myogenic factor 5
MyHC	myosin heavy chain
Myo1c	myosin 1C
MyoII	myosin II
MyoD	myoblast determination protein
MYONAP	myogenesis-related and NCAM-associated protein
Nap1	Nck-associated protein 1
NCAM	neural cell adhesion molecule (also known as CD56)
NCBI	National Centre for Biotechnology Information
NO	intracellular nitric oxide
NPFs	actin nucleation-promoting factors
PI3K	lipid kinase phosphatidylinositol 3-kinase
PAK1/2	p21 activated kinases 1 and 2
PBS	phosphate buffered saline
PDGF	platelet-derived growth factor
PIP2	phospholipid Phosphatidylinositol-(4,5)-biphosphate
PIP3 PH	phosphatidylinositol-3,4,5-triphosphate pleckstrin homology
PKA	protein kinase A
PLS	podosome-like structure
PTM	post-translational modification
Rac	Ras-related C3 botulinum toxin substrate
RhoA	Ras homolog family member A
RNAi	RNA interference
ROCK2	activated Rho-associated, coiled-coil containing protein kinase 2
Rok	Rho-associated protein kinase
Rols7/Ants	Rolling Pebbles 7/Antisocial
Rst	Roughest
RTK	receptor tyrosine kinase
SC	Satellite cell

Scar	Suppressor of cAMP receptor
shRNAs	short hairpin RNAs
SINE	short interspersed nuclear element
Sing	transmembrane protein Singles bar
SkM	Skeletal muscle
Sltr	Solitary
SMA	spinal muscular atrophy
Sns	Sticks and Stones
SRF	serum response factor
TGF-β	transforming growth factor-β
TNF-α	tumour necrosis factor-α
TSS	Transcription Start Site
VCAM-1	vascular cell adhesion protein 1
WASP	Wiskott-Aldrich syndrome protein
WAVE	WASP family verprolin homologs

CHAPTER ONE

LITERATURE REVIEW

1.1 Background and rationale

Skeletal muscle (SkM) constitutes 30-40% of the total lean mass of the average, healthy adult and is essential for controlled movement and structural support (Huard *et al.*, 2002). Distinct from both cardiac and smooth muscle, skeletal muscle (SkM) is made up of bundles of striated myofibers consisting of multiple elongated, multi-nucleated muscle cells called myocytes (Chal and Pourquié, 2017).

During embryogenesis, the SkM of the limbs and torso is formed through the development of the dorso-lateral region of somites derived from the paraxial mesoderm (Yin *et al.*, 2013; Blake and Ziman, 2014; Chal and Pourquié, 2017). Pre-myogenic muscle progenitor cells within the somites undergo numerous rounds of proliferation with a majority of these cells then progressing through lineage commitment, maturing into fusion capable SkM myoblasts (Yin *et al.*, 2013; Chal and Pourquié, 2017). These myoblasts fuse with each other and surrounding multi-nucleated muscle cells to produce the multinucleated myotubes that make up the myofibers (Yin *et al.*, 2013). The SkM found in the head and neck develop in a similar fashion but from cells originating in the neural crest and anterior paraxial mesoderm (Chal and Pourquié, 2017).

In adult tissue, the growth and repair of the post-mitotic myofibers is facilitated by a population of mono-nucleated muscle progenitor cells called satellite cells (SCs) (Chal and Pourquié, 2017). These cells typically lie dormant along the exterior of the cell membrane of the myocytes where they, upon activation, enter the cell cycle and migrate along the myofibers to the site of growth/repair (Yin *et al.*, 2013; Chal and Pourquié, 2017). Once at the target location, proliferating and myogenically committed satellite cells (SCs) become terminally differentiated and fuse with each other and existing myofibers; a process that is highly dependent on a variety of cell adhesion molecules (CAMs) (Yin *et al.*, 2013; Chal and Pourquié, 2017; Krauss *et al.*, 2017; Rout *et al.*, 2022).

In both adult and embryonic tissues, myogenic programs are largely controlled by a number of transcription factors including the paired-box proteins Pax3 and Pax7 as well as 4 muscle specific, basic helix-hoop-helix proteins (bHLH) known collectively as the myogenic regulatory factors (MRFs) (Singh and Dilworth, 2013; Zammit, 2017). The 4 myogenic regulatory factors include; myogenic factor 5 (Myf5), myoblast determination protein (MyoD), myogenin and muscle-specific regulatory factor 4 (MRF4) (Singh and Dilworth, 2013; Zammit, 2017). These MRFs are temporally expressed during myogenesis, promoting the proliferation of activated SCs and initiating myogenic lineage commitment and subsequent myoblast

differentiation and fusion (Singh and Dilworth, 2013; Zammit, 2017). The expression of a number of essential myogenic proteins are initiated by MRF-mediated transcriptional programmes and importantly include fusion proteins as well the sarcomere and contractile proteins, such as desmin and myosin (Singh and Dilworth, 2013; Zammit, 2017). The expression of the MRFs is in turn regulated by a range of secreted extracellular signalling molecules as well as cell-cell interactions that involve an array of cell-surface complexes and signalling cascades (Singh and Dilworth, 2013; Zammit, 2017). Some of the cell-surface proteins such as members of the cadherin and Ig superfamilies play indispensable roles in facilitating cell-cell adhesion and alignment before fusion as well regulating adherens junction morphology, actin polymerization and pore formation during fusion (Krauss, 2010; Knight and Kothary, 2011). In spite of current insights, the mechanisms driving the complex process of myoblast differentiation and fusion are not fully understood with Myomaker and Myomixer being the only vertebrate, muscle-specific fusion proteins identified to date (Singh and Dilworth, 2013; Zammit, 2017).

Notwithstanding the plasticity of SCs and overall regenerative capability of SkM, defective muscle function is a symptom of a number of developmental disorders and muscular dystrophies (Knight and Kothary, 2011; Chal and Pourquié, 2017). In adult tissue, severe muscle injuries, neuromuscular diseases, late-onset dystrophies, cachexia and aging lead to a dramatic loss in SkM volume and a drastically decreased quality of life (Knight and Kothary, 2011; Chal and Pourquié, 2017). For example, rhabdomyosarcomas are aggressive and highly malignant cancerous tissue derived from SkM progenitor cells that fail to fully differentiate and fuse (Knight and Kothary, 2011; Chal and Pourquié, 2017). In order to combat these numerous and complex pathologies, a more comprehensive understanding of SkM regeneration will be required for the development of the next generation of muscle disease and injury treatments; these including cell, RNA and gene therapies as well as production of *in vitro* tissue for grafting.

The vertebrate Kin of IRRE-like proteins (kirrel) proteins are evolutionarily conserved members of the immunoglobulin (Ig) superfamily of cell adhesion molecules (CAMs) and include kirrel1, kirrel2 and kirrel3 (Durcan *et al.,* 2013; Tamir-Livne *et al.,* 2017). Trans interactions between cell-surface kirrel proteins and their binding partners promote cell-recognition and adhesion as well as regulating actin dynamics and cytoskeletal reorganization (Krauss, 2010; Önel *et al.,* 2014; Tamir-Livne *et al.,* 2017; Duan *et al.,* 2018). The *Drosophila* homologues of the vertebrate kirrel proteins, Dumbfounded (Duf) and Roughest (Rst), have been extensively studied in the context of myogenesis and are known to be essential for myoblast fusion with double-knockout of the genes resulting in a complete arrest of the process (Strünkelnberg *et al.,* 2001; Rochlin *et al.,* 2010; Shilagardi *et al.,* 2013; Kim *et al*., 2015). However, the requirement for kirrel activity during vertebrate cell fusion has not been fully explored with very little being known about which kirrel variants are involved, their temporal expression patterns or how their signalling

events are transduced. Recently, it has been found that kirrel1 is temporally expressed during mammalian myogenesis while kirrel3 regulates critical pre-fusion events (Durcan *et al.*, 2013; Tamir-Livne *et al.*, 2017). Tamir-Livine *et al.*, (2017) demonstrated that Kirrel3 is required for the fusion of C2 myoblasts cultured *in vitro* with the silencing of kirrel3 translation or the deletion of regions of the intracellular domain resulting in a significant reduction in fusion capacity. Furthermore, in zebrafish, the knockdown of *kirrel3l* translation was seen to interfere with myoblast fusion with phylogenetic analysis revealing the zebrafish protein to be most similar to mammalian kirrel3 (Srinivas *et al.*, 2007). Kirrel1 and its truncated splice variant, kirrel1B, have also been seen to be differentially expressed immediately preceding and during myoblast fusion in injured, regenerating mouse SkM (Durcan *et al.*, 2013). In light of these findings, as well as those in the fly, further research is needed to better understand the function of the kirrel family of adhesion molecules with regard to myoblast fusion.

1.2 Myofiber structure

As mentioned above, myofibers are elongated, multi-nucleated, tube-shaped muscle cells that are formed by the fusion of multiple mono-nucleated myoblasts (Chal and Pourquié, 2017; Zammit, 2017). Bundles of striated myofibers run parallel to one another along the length of SkM with interactions between actin and myosin proteins within these structures facilitating muscle contractions (figure 1.1) (Chal and Pourquié, 2017; Zammit, 2017). The extent of fusion events required to form a mature myofiber becomes evident when it is considered that in humans, myofibers can range from 2 to 50 cm in length and a thickness of between 10 and 100 μm, containing up to 200 nuclei (Chal and Pourquié, 2017; Zammit, 2017). Myofibers contain long, thin cylindrical contractile organelles called myofibrils of between 1-2 μm in diameter which run through the sarcoplasm i.e., the muscle cell cytoplasm (Chal and Pourquié, 2017; Zammit, 2017).

Myofibrils span the length of the elongated myofiber structure and are anchored to the myotendinous junctions at the poles of the cell (Chal and Pourquié, 2017). Myofibrils comprise of thousands of longitudinally arranged contractile sarcomere units of between 1 and 3.5 μm in length in humans (Moo *et al.*, 2016; Chal and Pourquié, 2017). In SkM, roughly 2500 myofibrils are contained within each adult myofiber, occupying upwards of 80% of the cell volume, pushing the majority of organelles and nuclei to the periphery of the cell (Cretoiu *et al.*, 2018). In healthy SkM, the numerous myonuclei are evenly distributed along the length of the myofiber and are anchored to the sarcolemma by an astral microtubular network (Liu *et al.*, 2020). However, clustering of myonuclei does occur at neuromuscular and myotendinous junctions (Liu *et al.*, 2020).

The specialized myofiber sarcolemma provides structural stability and allows for neuronal signal transduction (Cretoiu *et al.*, 2018). The sarcolemma is anchored to the basal lamina via the transmembrane

dystrophin-associated glycoprotein complex, a connection that also links the cell cytoskeleton to extracellular matrix (ECM) proteins surrounding the fibers (Cretoiu *et al.*, 2018).

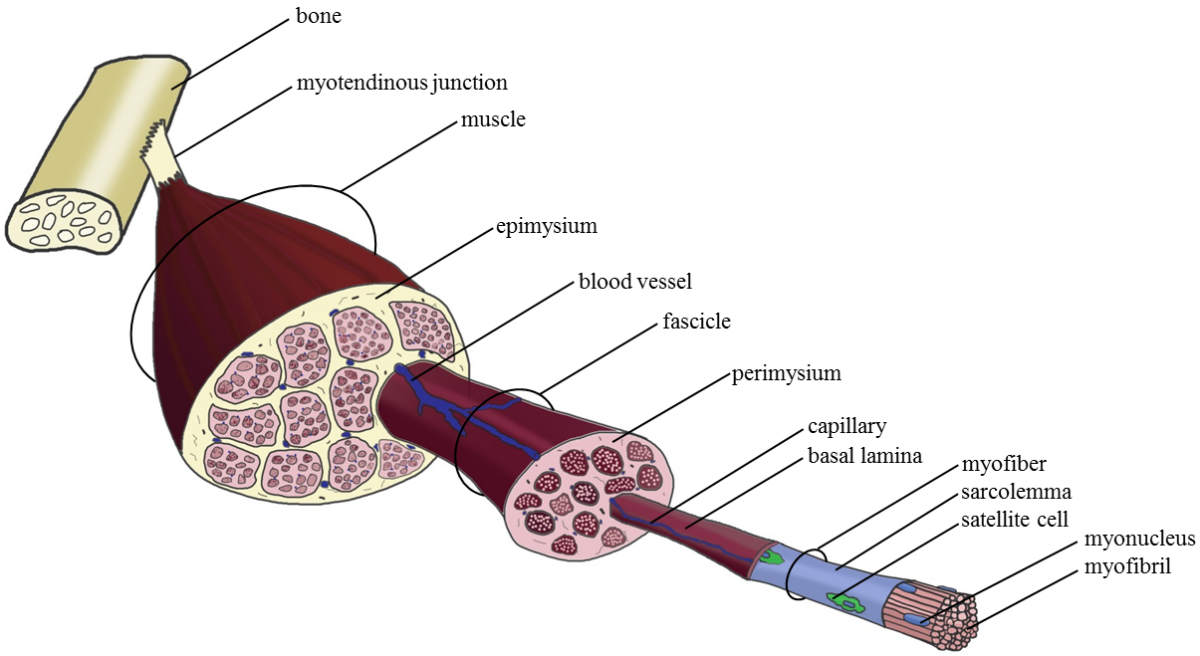


Figure 1.1 Mammalian skeletal muscle structure, myonucleus and satellite cell localization. Contractile myofibrils make up the bulk of the myofiber volume. Myonuclei are positioned away from the central myofibrils and can be seen just under the sarcolemma. Quiescent satellite cells typically reside in their niche between the basal lamina and cell-membrane/sarcolemma of the myofiber structure. Stability for the myofibers and the satellite cell niches are indirectly provided by the further layering of the structure: bundles/fascicles of striated myofibers are embedded in a matrix of connective tissue that runs along the length of the tissue. Constructed using references: Charge and Rudnicki, 2004; Chal and Pourquié, 2017; Zammit 2017.

1.3 Embryonic development of vertebrate skeletal muscle

Skeletal muscle is formed by cells originating in the paraxial mesoderm during embryogenesis (Chal and Pourquié, 2017). The SkM of the torso and limbs is derived from tissue within the somites, masses of bilaterally paired paraxial mesoderm that run along the head-to-tail axis of the embryo (Chal and Pourquié, 2017). During early development, the somites become compartmentalized into the dorsal epithelial dermomyotome and ventral mesenchymal sclerotome regions (Chal and Pourquié, 2017). The dermomyotome forms SkM, brown adipose tissue and the dermis of the back while the sclerotome gives rise to the axial skeleton (Chal and Pourquié, 2017). The SkM of the head and the neck develop from the anterior paraxial mesoderm (Chal and Pourquié, 2017). Muscle progenitor cells in mesoderm tissue are characterized by their expression of the Pax3 and/or Pax7 transcription factors (Le Grand and Rudnicki, 2007). As somites mature, SkM myogenesis is initiated by the downregulation of Pax3 and the expression of the MRF proteins in cells in the dorsal and ventral lips of the dermomyotome (Chal and Pourquié, 2017). A population of progenitor cells preferentially expressing Pax7 over Pax3, but do not express the MRF proteins, are found in the central portion of the dermomyotome (Kassar-Duchossoy *et al.*, 2004; Relaix *et al.*, 2005). These cells are kept in this pre-myogenic state throughout embryogenesis and migrate to and

populate the developing SkM, providing myogenic precursors for future myogenesis (Le Grand and Rudnicki, 2007). These precursors are maintained in adult SkM and termed satellite cells (SCs) (Yin *et al.*, 2013; Chal and Pourquié, 2017).

1.4 Satellite cells

The myonuclei contained within SkM myocytes are terminally differentiated and unable to undergo-mitotic division (Charge and Rudnicki, 2004; Le Grand and Rudnicki, 2007). The replenishment of myonuclei and the repair of damaged myofibers is instead facilitated by a local population of postnatal, muscle-specific progenitors called satellite cells (SC) (Le Grand and Rudnicki, 2007). Satellite cells typically reside in a dormant/quiescent state in their niche situated between the basal lamina and sarcolemma of the muscle fibers (figure 1.1) and are activated by stress as a result of strain or trauma to the muscle tissue (Charge and Rudnicki, 2004). Activated SCs and progenitors derived from SC, known as skeletal myoblasts, then migrate to the site of the muscle trauma where they undergo numerous rounds of replication before differentiation and fusion to produce multinucleated myofibers (Le Grand and Rudnicki, 2007). This process is analogous to that occurring during embryonic and fetal myogenesis. Activated SCs that do not become terminally differentiated, return to quiescence to replenish the pool of SCs (Le Grand and Rudnicki, 2007; Yin *et al.*, 2013). Hence a population of mono-nucleated muscle-specific progenitor cells with the potential to become fusion competent are maintained.

Satellite cells constitute the bulk of the resident adult stem cells found in SkM and can be traced back to the $Pax7^{+}$ progenitors of the central domain of the dermomyotome (Yin *et al.*, 2013; Chal and Pourquié, 2017). The continued production of the paired-box transcription factors is essential for the maintenance and self-renewal of the SC population and is regulated in large by Notch signalling with inhibition of this pathway resulting in the loss of Pax3 and Pax7 expression in SCs and ultimately the progressive loss of muscle progenitors in SkM (Charge and Rudnicki, 2004; van der Velden *et al.*, 2008; Vasyutina *et al.*, 2009). Likewise, the knockout of the *pax7* gene results in the depletion of the SC population (Le Grand and Rudnicki, 2007).

The expansion of SkM peaks during the fetal and perinatal stages of development with SC progenitors accounting for roughly 30% of the mono-nucleated cells in the tissue during this time (Hellmuth and Allbrook, 1971). After this peak, the abundance of SCs steadily declines as the SkM develops during growth until the progenitors account for roughly 5% of the mono-nucleated cells found in the adult tissue (Cardasis and Cooper, 1975; Schmalbruch and Hellhammer, 1976; Asfour *et al.*, 2018). During the regeneration of adult SkM, a number of extracellular signalling molecules are involved in the activation of SCs and include fibroblast growth factors (FGF), insulin-like growth factors (IGF), hepatocyte growth

factor (HGF) and several of the secreted Wnt proteins (Charge and Rudnicki, 2004; van der Velden *et al.*, 2008). The activation of quiescent SCs and the initiation of myogenesis is characterized by the downregulation of Pax3/Pax7 and expression of the MRFs (Charge and Rudnicki, 2004; van der Velden *et al.*, 2008; Chal and Pourquié, 2017).

1.5 The myogenic regulatory factors

Myf5, MyoD, Mrf4 and myogenin are all members of the basic helix-loop-helix (bHLH) family of transcription factors and are critical for the regulation of SkM progenitor cell activation, replication and differentiation during vertebrate embryogenesis and postnatal myogenesis (Joulia *et al,* 2003; Holterman and Rudnicki 2005; Hernández-Hernández *et al.*, 2017). They form homodimers and heterodimers with a wide range of ubiquitous transcriptional regulators including the helix-loop-helix E-proteins, the inhibitor of DNA binding (ID) proteins and members of the MADS box transcription factors (Moran *et al.*, 2002; Sartorelli and Caretti, 2005; Wang and Baker, 2015). The dimerization of the MRFs is required for DNA-binding and the subsequent activation of target genes (Sartorelli and Caretti, 2005; Wang and Baker, 2015). Following dimerization with E proteins, the MRFs, via their basic helix-loop-helix (bHLH) domains, bind to E-box DNA sequences found in the promoter and enhancer regions of a number of downstream muscle and non-muscle-specific genes (Hernández-Hernández *et al.*, 2017; Asfour *et al.*, 2018). The MRFs preferentially form heterodimers with the pro-myogenic E2A proteins, including E12 and E47, initiating myogenesis (Hernández-Hernández *et al.*, 2017; Asfour *et al.*, 2018).

Myf5 and MyoD are the first of the MRF proteins to be up-regulated during the myogenic process with an increase in their expression characterizing the entry of the progenitors into the cell cycle and the initiation of the myogenic process (Charge and Rudnicki, 2004; Le Grand and Rudnicki, 2007). The levels of the Myf5 and MyoD proteins are in part regulated during the cell cycle by mRNA targeting microRNAs (Zammit, 2017). Myf5 expression is initially upregulated during the G_0 phase of the cell cycle, before dropping to a minimum during mid G_1 and then rising through G_1/S and S to reach a maximal at G_2 before declining again slightly at the G_2/M transition (Singh and Dilworth, 2013; Yin *et al.*, 2013; Zammit, 2017). MyoD expression is absent during G_0 and peaks in mid G_1, before reaching its lowest levels at the G_1/S boundary before rising again to a similar level as Myf5 at G_2/M (Singh and Dilworth, 2013; Yin *et al.*, 2013; Zammit, 2017). Gene knockout studies carried out in mice show that up-regulation of Myf5 promotes the proliferation of SCs and myoblasts while increased MyoD activity stimulates myogenin production and an exit from the cell cycle (figure 1.2) (Kuang *et al.*, 2007, Ustanina *et al.*, 2007). Additional knockout studies have also shown that a degree of functional redundancy exists within the MRF signalling pathways (Chanoine *et al.*, 2004). During embryogenesis, mouse embryos lacking a functional *MyoD* gene show no physical abnormalities but express four-fold higher levels of *Myf5* (Rudnicki *et al.*, 1993; Megeney *et al.*,

1996). Similarly, embryos lacking *Myf5* express the MyoD protein at much higher levels and produce functional SkM, although the initiation of the myogenic process is delayed (Rudnicki *et al.*, 1993; Megeney *et al.*, 1996). However, mice with defective *Myf5* and *MyoD* genes display severely impaired muscle development at birth with *Mrf4* only able to partially compensate for their absence, initiating limited myogenesis (Kassar-Duchossoy *et al.*, 2004). Additionally, MyoD is able to initiate myogenic progression in *Myf5/Mrf4*-null mouse embryos, albeit this activation is delayed (Hernández-Hernández *et al.*, 2017). The effectiveness of the redundancy between Myf5 and MyoD is drastically reduced in adult SkM with the crossing of *MyoD*-null or *Myf5*-null mice with X-linked muscular dystrophy (*mdx*) mutants resulting in the exacerbation of the dystrophic condition in both of the compound types (Megeney *et al.*, 1996; Gayraud-Morel *et al.*, 2007; Le Grand and Rudnicki, 2007; Ustanina *et al.*, 2007). Knockout of *Myf5*, *MyoD* and *Mrf4* results in the complete failure of myoblasts to undergo lineage specification and the failure of myotubes to form (Le Grand and Rudnicki, 2007; Ustanina *et al.*, 2007).

Following progression through the cell cycle and the subsequent upregulation of MyoD; myogenin is expressed, resulting in myoblasts exiting the cell cycle and undergoing terminal differentiation (Charge and Rudnicki, 2004; Ustanina *et al.*, 2007). The expression of myogenin results in the rapid exit of myoblasts from the cell cycle through pathways that inhibit the E2F family of cell cycle promoting transcription factors (Faralli and Dilworth, 2012; Yin *et al.*, 2013; Zammit, 2017). While Myf5, MyoD and Mrf4 display a degree of functional redundancy during foetal and adult SkM regeneration, myogenin is indispensable for this process (Le Grand and Rudnicki, 2007; Zammit, 2017). The knockout of myogenin prevents terminal differentiation of myoblasts and ultimately prevents the production of SkM (Zammit, 2017; Ganassi *et al.*, 2018). After differentiation has progressed, the expression of late-stage myogenic proteins is initiated and includes proteins such Mrf4, the fusion proteins Myomaker and Myomixer as well as the large contractile protein myosin (Charge and Rudnicki, 2004; Ustanina *et al.*, 2007; Luo *et al.*, 2015; Sampath *et al.*, 2018). The myocyte enhancer factor 2 (MEF2) family of proteins positively regulate MyoD, myogenin and Mrf4, thus promoting myogenic progression (Asfour *et al.*, 2018).

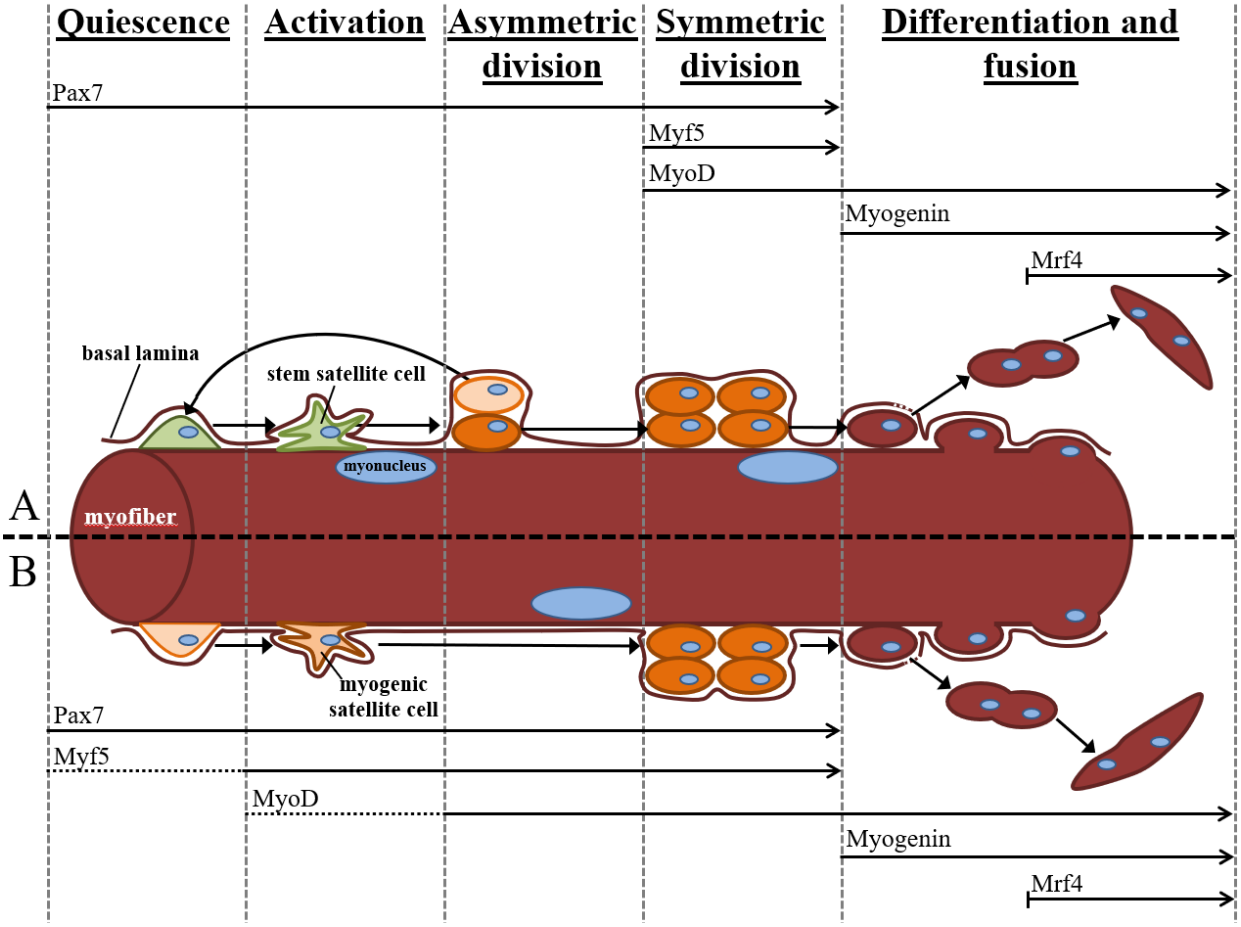


Figure 1.2 Generalized myogenic regulatory factor expression patterns during adult mammalian muscle progenitor cell quiescence, activation, proliferation, differentiation and fusion. Satellite cells can be divided into two groups based on their propensity to undergo either asymmetric or symmetric division. (A) More stem-like satellite cells (green) ($Pax7^{+}/Myf5^{-}/MyoD^{-}$) make up about 10% of the progenitor cell pool and undergo asymmetric division. This type of division results in two functionally different daughter cells with one daughter cell returning to quiescence while the other goes on to divide symmetrically. (B) Roughly 90% of activated satellite cells ($Pax7^{+}/Myf5^{+}/MyoD^{+}$) in SkM undergo symmetric cell division, resulting in functionally identical daughter cells that then go on to differentiate and fuse to from multinucleated myofibers. (A+B) The differentiation of muscle progenitors is initiated by the upregulation of MyoD and myogenin. Myogenic regulatory factor 4 (Mrf4) is only expressed in differentiated myocytes and regulates muscle hypertrophy. Constructed using references: Charge and Rudnicki, 2004; Ustanina *et al.*, 2007; Zammit, 2017; Asfour *et al.*, 2018.

1.6 Satellite cell activation

In adult SkM muscle, SCs are typically mitotically quiescent and display limited gene expression and protein synthesis (Cornelison and Wold, 1997; Tajbakhsh, 2005; Le Grand and Rudnicki, 2007). Satellite cells are activated by signalling cascades initiated by stress to the SkM induced by mechanical stretch, injury or myodegenerative diseases (Le Grand and Rudnicki, 2007). The mechanisms controlling the activation of these myogenic precursors are complex and not fully elucidated but are known to involve a variety of membrane proteins, extracellular signalling molecules, microRNAs and signalling pathways (Le Grand and Rudnicki, 2007). Once activated, SCs are primed for myogenic differentiation, however it is the expression of the MRFs and the paired-box transcription factors that determines their commitment towards full differentiation or their return to quiescence (Kuang *et al.*, 2007; Le Grand and Rudnicki, 2007). The

propensity of muscle precursors to remain quiescent or become differentiated is controlled in large by cross-talk between the Notch and Wnt/β-catenin signalling cascades (Brack *et al.*, 2008; Le Grand *et al.,* 2009). Notch signalling is highly nuanced and heavily influenced by both the density of Notch receptors on the cell surface as well as the distribution of its extracellular, membrane-bound binding partner Delta-like ligand (DLL) (Brack *et al.*, 2008; Le Grand *et al.,* 2009). Notch signalling is essential for both the return of myoblasts to quiescence as well as the proliferation of activated myoblasts; in quiescent SCs, the pathway is essential for the maintenance of Pax3 and Pax7 expression while also inhibiting the expression of the MRFs (Brack *et al.,* 2008; Le Grand *et al.,* 2009). However, during proliferation, Notch signalling prevents pre-mature differentiation by inhibiting MRF activity (Le Grand and Rudnicki, 2007; Brack *et al.*, 2008). Signalling via the canonical Wnt/β-catenin pathway is a major trigger of MRF production and during the early stages of myogenesis, includes the expression of Myf5 and MyoD (Brack *et al.*, 2008; Le Grand *et al.*, 2009).

1.7 Myoblast migration

Cellular migration is essential for the successful growth and repair of tissues within an organism. Once activated, myoblasts move to the outside of the basal lamina where they start to proliferate and migrate as a swam along the length of the myofibers to the site of repair where along with Pax7 and Myf5, the cells begin to express MyoD (Le Grand and Rudnicki, 2007; Griffin *et al.,* 2010). Chemokines released by injured myofibers, myoblasts and inflammatory cells at the injury site play an important role in controlling myoblast swarm migration (Lafreniere *et al.* 2006; Griffin *et al.,* 2010). The primary chemokines involved in the regulation of myoblast migration include the FGFs, HGF, platelet-derived growth factor (PDGF) and interleukin-4 (IL-4) (Bischoff, 1997; Hawke and Garry, 2001; Lafreniere *et al*., 2006). Besides the above-mentioned canonical chemokines, cleaved extracellular domains from cell surface proteins such as members of the Ig superfamily of CAMs act as chemoattractants during myoblast migration (Strünkelnberg *et al.*, 2001; Menon *et al.*, 2005; Johnson *et al.*, 2017).

The integrin family of cell surface proteins is essential for cell migration and facilitates the linkage of contractile, cellular actin and extracellular matrix (ECM) ligands (Friedl and Brocker, 2000; Hynes, 2002; Gallant *et al.*,2005). During the initial stages of cell movement, hand-like structures called lamellipodia are formed via the distortion of the leading edge of the cell (Humphries *et al.,* 2006; Huttenlocher and Horwitz, 2011). The lamellipodia reach forward in the direction of movement and attach to ECM proteins via the integrin-based adhesion points (Friedl and Wolf, 2010; Huttenlocher and Horwitz, 2011). Cytoplasmic actin fibers running in the direction of movement then contract while at the same time integrins at the trailing edge of the cell detach from the ECM and the cell pulls itself forward (Friedl and Wolf, 2010; Huttenlocher and Horwitz, 2011).

1.8 Myoblast proliferation and satellite cell population replenishment

The stimulation of myoblast proliferation is primarily achieved by cell-cell contact and by extracellular cytokines and growth factors released by macrophages, myotubes and other proliferating myoblasts within the local tissue microenvironment (Hawke and Garry, 2001). After injury or weight-bearing, macrophages are drawn to the affected areas where they envelope debris and along with the local myoblasts and myotubes, they secrete a number of proliferation-promoting signalling molecules (Hawke and Garry, 2001). These molecules include the FGFs, HGF, insulin-like growth factor-1 (IGF-1), insulin-like growth factor-2 (IGF-2), a number of Wnt proteins and to a lesser extent, interleukin-4 (IL-4), interleukin-6 (IL-6), leukaemia inhibitory factor (LIF) and transforming growth factor-β (TGF-β) (Hawke and Garry, 2001; Le Grand *et al.,* 2009; Von Maltzahn *et al.,* 2012). Additionally, and as with all other cell types, cell cycle progression in SC and myoblasts is reliant on cytoplasmic cyclin-dependent-kinases (CDKs) (Knight and Kothary, 2011; Singh and Dilworth, 2013).

Activated muscle progenitors display both asymmetric and symmetric division with the nature of division thought to be largely dependent on the position of the daughter cells relative to the myofiber and epigenetic differences between sister chromatids (Yin *et al.,* 2013). These epigenetic and spatial differences are thought to affect the balance between notch and Wnt/β-catenin signalling as well as the levels of Myf5 and MyoD expression, resulting in a heterogeneous population of progenitors in the SkM (Kuang *et al.,* 2007). Less than 10% of activated SCs are $Pax7^{+}/Myf5^{-}/MyoD^{-}$ and show a propensity to undergo asymmetrical division but are also capable of symmetric replication (Kuang *et al.,* 2007; Asfour *et al.,* 2018). Following proliferation, a small percentage of undifferentiated myoblasts upregulate notch signalling while inhibiting Wnt/β-catenin signalling pathways, resulting in increased Pax7 expression and reduced MRF levels, all leading to their exit from the cell cycle and re-entry into the quiescent state as dormant SCs (Charge and Rudnicki, 2004; Relaix *et al.,* 2005; Brack *et al.,* 2008). It is this process that drives the self-renewal of the SC population in SkM (Hernández-Hernández *et al.,* 2017).

1.9 Myoblast differentiation

The differentiation of SkM myoblasts is a complex process that involves a number of steps including; the expression of muscle specific genes, withdrawal from the cell cycle, cell-cell recognition and adhesion, terminal differentiation and alignment, all followed by cell-cell fusion (Moran *et al.,* 2002; Krauss *et al.,* 2017). In addition to the widely studied secreted chemokines and growth factors, the coordinated changes in gene expression and cell morphology during myogenesis are regulated by a number of relatively poorly understood adhesion molecules (Chan and Hiiragi, 2017; Krauss *et al.,* 2017).

The adoption of myogenic transcriptional programs is primarily mediated by the activity of the MRFs (Holterman and Rudnicki 2005; Zammit, 2017). A number of non-muscle-specific factors within the cell work in conjunction with the MRFs to promote myogenic progression and include the MRFs' primary heterodimeric E protein partners, members of the MEF2 family of transcription factors, transcriptional coactivators, and chromatin remodelling factors (Moran *et al.,* 2002; Comai and Tajbakhsh, 2014). Extracellular signalling molecules are the earliest initiators of MRF activity and include IGF-1, IGF-2, FGF-6, nitric oxide (NO), tumour necrosis factor-α (TNF-α) and members of the Wnt protein family (Armand *et al.,* 2006; Chen *et al.,* 2007; Tanaka *et al.,* 2011; Cazzato *et al.,* 2014). These factors are released by a number of cell and tissue types including migrating myoblasts, myotubes and immune cells (Hawke and Garry, 2001; Jones *et al.,* 2005; Karalaki *et al.,* 2009). These extracellular molecules initiate signalling cascades that involve a wide range of intracellular kinases that are crucial for differentiation, directly regulating the expression and activity of the MRFs as well as a number of other non-muscle specific genes that are required for the process (Knight and Kothary 2011; Asfour *et al.,* 2018). The later stages of differentiation i.e., terminal differentiation, as well the process of fusion, are largely driven by cell-cell/myotube contacts (Gildor *et al.,* 2012). These cell surface interactions will be covered in more detail later in the thesis.

1.9.1 Regulation of differentiation: canonical early-stage signalling

PKA: Protein kinase A (PKA) is a cyclic adenosine monophosphate (cAMP)-dependent kinase and initiator of satellite cell activation, proliferation and depending on localization, promoter of differentiation (Knight and Kothary, 2011; Abmayr and Pavlath, 2012). Activation of PKA is achieved following the binding of extracellular Wnt1 and/or Wnt7a to frizzled receptors (Fz) (Chen *et al.,* 2005; Knight and Kothary, 2011). Active PKA subunits move into the nucleus where they phosphorylate and activate the CREB transcription factor, inducing the early expression of Myf5 and MyoD, resulting in SC activation (Knight and Kothary, 2011; Abmayr and Pavlath, 2012). Nuclear PKA prevents the pre-mature differentiation of myoblasts by inhibiting the activity of Myf5 and MyoD, albeit via unknown mechanisms; and by phosphorylating and deactivating the MEF2D transcription factor (Li *et al.,* 1992; Knight and Kothary, 2011). Accordingly, a reduction in nuclear PKA levels is necessary for myogenesis to proceed unhindered (Knight and Kothary, 2011).

The CDKs: The cyclin-dependent kinases (CDKs) are major regulators of cell cycle progression and play active roles in the modulation of myogenesis (Knight and Kothary, 2011; Abmayr and Pavlath, 2012). The cyclins and the CDKs can be separated into 3 main groups according to their functions, these being: the G_1 cyclins; the mitotic cyclins; and the non-cell cycle cyclins (Knight and Kothary, 2011; Abmayr and Pavlath, 2012). In order for differentiation to occur, the cell cycle must be exited and accordingly, there is a decrease

in G_1 and mitotic CDK activity and expression (Knight and Kothary, 2011; Abmayr and Pavlath, 2012). Once the G_1 and mitotic CDKs have been silenced and the cell cycle exited, the expression of the non-cell cycle CDKs is initiated, and includes the production of the differentiation-promoting CDK5 and CDK9 (Knight and Kothary, 2011; Singh and Dilworth, 2013).

The Ras-Raf-MEK-ERK cascade: The extracellular signal-regulated kinases (ERKs) include the ERK1 and ERK2 isoforms and are part of the mitogen-activated protein kinase (MAPK) family which also includes the Jun N-terminal kinases (JNK) and the p38 mitogen-activated protein kinases (Knight and Kothary, 2011; Wang and Almazan, 2016). The extracellular signal-regulated kinases (ERKs) are involved in signalling pathways initiated by the binding of extracellular ligands to receptor tyrosine kinase (RTK) family members such as the fibroblast growth factor receptor (FGFR) (Knight and Kothary, 2011; Wang and Almazan, 2016). The Ras-Raf-MEK-ERK cascade plays a critical role in cell-proliferation and survival and can be initiated by a number of mitogens including the FGFs, HGF, IGFs, leukaemia-inhibitory factor (LIF) and platelet-derived growth factor (PDGF); although not all of these factors elicit the same response (Kook *et al.*, 2008; Knight and Kothary, 2011).

The PI3K-Akt1/2 cascade: The Akt family of serine/threonine protein kinases consists of 3 mammalian homologues, these being Akt1, Akt2 and Akt3 (Fayard *et al.*, 2005; Knight and Kothary, 2011). The activation of Akt signalling cascades is initiated by the binding of IGF to the IGF receptor (IGFR) or of insulin to the insulin receptor (IR), this resulting in the autoactivation of the intercellular tyrosine kinase domain of the receptors (Dong and Liu, 2005). Both Akt1 and Akt2 are involved in myogenic signalling with studies showing that the 2 isoforms are required at different stages of the process; however, how their activities are differentially controlled is not known with both enzymes being stimulated by IGF (Knight and Kothary, 2011). It is thought that the nuanced nature of Akt signalling is determined by extracellular IGF concentrations in conjunction with cell-cell contacts (Knight and Kothary, 2011).

1.9.2 Regulation of differentiation: canonical mid- to late-stage signalling

The p38α MAPK cascade: The p38 subfamily of mitogen-activated kinases (MAPKs) are powerful regulators of differentiation in a number of cell types, including myoblasts (Stanton *et al.*, 2004; Bhat *et al.*, 2007). There are 4 p38 isoforms and include the p38α, p38β, p38γ and p38δ MAPKs; however, only the p38α isoform has been shown to be critical for mammalian myoblast differentiation both *in vivo* and *in vitro* (Perdiguero *et al.*, 2007a; 2007b; Ruiz-Bonilla *et al.*, 2008; Knight and Kothary, 2011). Furthermore, it has been found that the γ isoform is required for ideal myoblast fusion (Perdiguero *et al.*, 2007a; 2007b; Ruiz-Bonilla *et al.*, 2008).

Homophilic interactions between cell-surface N-cadherin clusters on adjacent myoblasts are a major stimulus of p38 MAPK activity (Knight and Kothary, 2011). A level of functional redundancy exists between members of the cadherin superfamily with M-cadherin being able to substitute for the loss of N-cadherin in myoblasts (Knight and Kothary, 2011). Once activated, p38α/β MAPKs promote myogenic differentiation via the activation of proteins that intensify the transcriptional activity of MyoD and by inhibiting the activities of quiescence and cell cycle related transcription factors (Knight and Kothary, 2011; Krauss *et al.*, 2017).

Wnt and β-catenin: As differentiation progresses and prior to fusion, Wnt signalling via the frizzled receptor and its Lrp5/6 co-receptor, leads to the unbinding of β-catenin from the intracellular domain of N-cadherin (Mrozik *et al.*, 2018). The cytoplasmic β-catenin is then free to translocate into the nucleus and initiate myogenic transcriptional programs (Mrozik *et al.*, 2018).

The canonical extracellular signalling molecules and the cascades they initiate in mammalian muscle precursor cells are shown in figure 1.3.

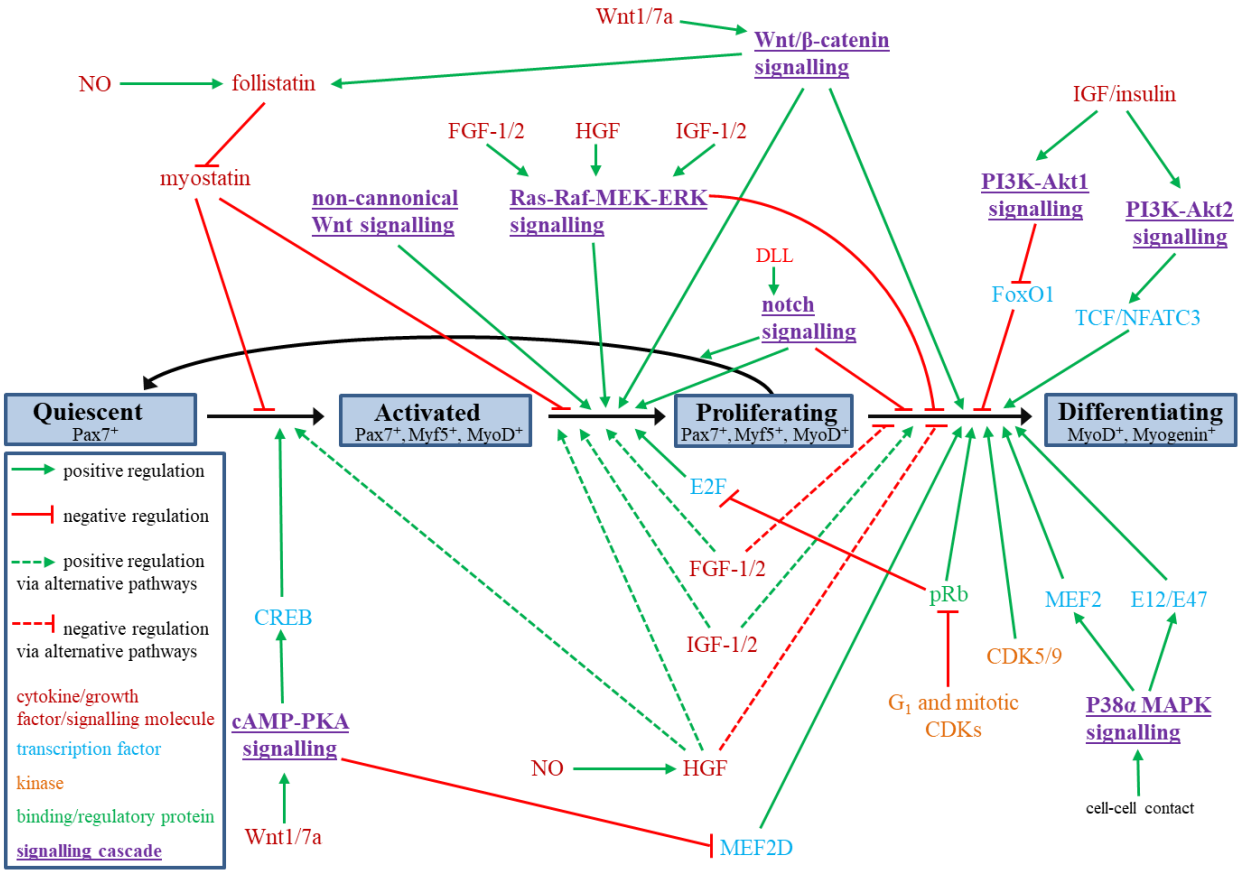


Figure 1.3 The major binding proteins, cytokines, signalling cascades and transcription factors involved in the regulation of mammalian satellite cells and myoblasts during myogenesis. The effects of signalling molecules and their resulting cascades on the activation, proliferation and differentiation of mammalian SkM progenitor cells as described in the text. Alternative cascades and pathways that are not named in this figure and are represented by dashed lines. Constructed using references: Charge and Rudnicki, 2004; Le Grand *et al.*, 2009; Knight and Kothary, 2011; Zammit, 2017; Asfour *et al.*, 2018.

1.9.3 Regulation of differentiation: cell adhesion molecules and terminal differentiation

The final steps of myoblast differentiation are highly dependent on cell-cell contacts that involve a variety of CAMs including the cadherins, Ig-domain adhesion molecules, integrins, neogenin and the cluster of differentiation (CD) molecules (Krauss, 2010; Knight and Kothary 2011; Gildor *et al.*, 2012; Asfour *et al.*, 2018). In general, the expression of these CAMs is upregulated with increased proximity to and adhesion with the appropriate fusion partners with the molecules clustering at sites of cell-cell contact (Gildor *et al.*, 2012; Przewoźniak *et al.*, 2013). Groups of myoblasts migrating towards their myotube targets are maintained in a semi-differentiated state by Delta-like ligand (DLL) which is presented on the cell-membranes of fellow migrating myoblasts (Brack *et al.*, 2008; Gildor *et al.*, 2012). Continuous presentation of Delta-like ligand (DLL) stimulates Notch signalling which in-turn prevents the terminal differentiation of the migrating progenitors (Brack *et al.*, 2008; Gildor *et al.*, 2012). These semi-differentiated myoblasts are still capable of division and some continue to proliferate once the myoblasts reach the site of growth or repair (Brack *et al.*, 2008; Yin *et al.*, 2013). As the myoblasts reach their targets, cell-cell recognition and adhesion is primarily facilitated by the cadherins and Ig-domain containing CAMs; rapidly leading to the cessation of Notch signalling, allowing for terminal differentiation and fusion (Brack *et al.*, 2008; Gildor *et al.*, 2012). Vital for recognition and adhesion, actin-driven filopodia from myoblast membranes reach outwards to facilitate contacts with extracellular matrix (ECM) factors as well as with adjacent cells and myotubes (Segal *et al.*, 2016; Bischoff *et al.*, 2021). The filopodia facilitate these interactions via molecules embedded in their membranes with some of these proteins being clustered at the tips (Segal *et al.*, 2016; Bischoff *et al.*, 2021). The adhesion of myoblasts to myotubes increases the concentration of secreted pro-myogenic factors that reach the progenitors from the muscle fibres as well as allowing for signalling cascades that require direct cell-cell contact i.e., juxtracrine signalling (Momiji *et al.*, 2019). Cell-adhesion facilitates the activity of critical myogenic molecules such as the GTPase Rac and pathways such as the p38 MAPK and Akt/mTOR pathways (Nowak *et al.*, 2009; Cadot *et al.*, 2012). As the fusion process progresses, most CAMs are relocated away from the direct site of fusion to allow for closer contact between the fusion partners and fusion pore formation; these CAMs include the cadherins, integrins and CD molecules which remain critical for continued adhesion as well as regulating cell alignment during fusion (D'Souza-Schorey and Chavrier, 2006; Önel *et al.*, 2014). Adhesion molecules that regulate fusion via actin remodelling, such as the irre recognition proteins, remain close to the site of fusion and are essential to the fusion process (Durcan *et al.*, 2013; Tamir-Livne *et al.*, 2017).

Due to the intricate nature of cell-adhesion and differentiation in vertebrate systems, most of what is known about these mechanisms has been derived from *Drosophila* models. Furthermore, due to the overlapping

nature of adhesion, the final steps of differentiation and the initiation of fusion, it is difficult to determine exactly how these cell surface molecules influence these inseparable events.

The irre recognition molecules: The irre recognition module (IRM) is a small group of evolutionarily conserved proteins of the immunoglobulin superfamily that are involved in cell attraction, recognition, adhesion and signalling with pleiotropic functions in a wide range of tissue types (Fischbach *et al* 2009). In mammals, the IRM includes nephrin as well as the three kirrel molecules i.e., kirrel1, kirrel2 and kirrel3, also known as neph1, neph3 and neph2 respectively (Krauss, 2010; Durcan *et al.*, 2013). The kirrel proteins preferentially bind in *trans* to cell-surface nephrin, facilitating the initial recognition and adhesion between migrating myoblasts and myotubes (figure 1.4) (Galletta *et al.*, 2004; Krauss, 2010; Gildor *et al.*, 2012). The irre molecules have been shown to be present on exploratory myoblast filopodia; outward reaching extensions that increase the success rate of recognition and adhesion (Rochlin *et al.*, 2010; Segal *et al.*, 2016). Furthermore, it is thought that irre proteins on myotubes have their putative extracellular regions cleaved in a proteosome-dependent manner, a process that creates a myoblast-attracting chemogradient that gets stronger as the progenitors approach their fusion targets (Strunkelnberg *et al.*, 2003; Menon *et al.*, 2005; Tamir-Livne *et al.*, 2017).

Although the IRM molecules are known to be heavily involved in myoblast recognition and adhesion, it is unclear how these CAMs may modulate myogenic transcriptional programs i.e., MRF activity (Galletta *et al.*, 2004; Krauss, 2010; Durcan *et al.*, 2013; Tamir-Livne *et al.*, 2017). However, recently it has been shown that kirrel1 is a feedback regulator of the Hippo pathway in HEK293 cells (Paul *et al.*, 2022); the same has not been shown in myoblasts. The mechanisms of Hippo signalling are complex and relatively poorly understood but are known to affect cell proliferation and fate determination during the regeneration of a number of tissue types; in conjunction with Notch and Wnt/β-catenin signalling (Yin *et al.*, 2013; Liu *et al.*, 2021; Paul *et al.*, 2022). In myoblasts, Hippo signalling directly modulates the transcription of *pax7*, *myf5* and *myod* (Yin *et al.*, 2013; Liu *et al.*, 2021; Paul *et al.*, 2022); however, no research has shown that kirrel1 can regulate these MRFs via Hippo signalling. In C2C12s, silencing of the Dock family of Rac GEFs (specifically Dock1 and Dock5) results in the delayed expression of myogenin as well as a postponed exit from the cell cycle (Laurin *et al*, 2008). The IRM molecules are also known to regulate the localization of Dock proteins at the fusion synapse (Laurin *et al*, 2008; Rochlin *et al.*, 2010; Haralalka *et al.*, 2011). Once again, it is not known whether the IRM molecules can promote myogenin expression via Dock protein localization. Being the focus of this thesis, the nature of IRM-mediated adhesion, differentiation and fusion is more extensively covered later in the text.

The cadherins: As previously covered, the cadherin family of CAMs play a pivotal role in myoblast differentiation primarily through the initiation of the p38 kinase cascade and via the promotion of RhoA activity, all leading to increased MRF expression (Krauss, 2010; Knight and Kothary, 2011; Krauss *et al.*, 2017). Classical cadherins, such as N- and M-cadherin facilitate cell adhesion via the calcium dependent, homophilic binding of their ectodomains with cadherin molecules on adjacent myoblasts/myotubes (figure 1.4) (Knight and Kothary, 2011; Krauss *et al.*, 2017). Additionally, the intracellular regions of the cadherin molecules bind directly to β-catenin (and indirectly to α-catenin via β-catenin) which in turns tethers the cell-surface complex to the actin cytoskeleton, facilitating the spaciotemporal and cell morphology changes that are required for successful differentiation and fusion i.e., cell alignment and elongation (Knight and Kothary, 2011; Krauss *et al.*, 2017). It has been hypothesized that the cadherins form the base of a large, promyogenic cell surface complex that includes neogenin, Cdo, Brother of Cdo (BOC) and netrin (figure 1.4); a multifunctional group of proteins that promote adhesion, differentiation and fusion (Kang *et al.*, 2004; Krauss *et al.*, 2017).

Neogenin and netrin: Neogenin is an Ig and fibronectin type III (FNIII) repeat-containing cell-surface receptor which along with its ligand, netrin-3, is produced in myoblasts with interactions between the two occurring in an autocrine fashion (Kang *et al.*, 2004). Although netrins are generally considered secreted molecules, in myoblasts cultures netrins have been shown to become embedded within cell membranes via glycosylphosphatidylinositol (GPI) tail anchors (Kang *et al.*, 2004; Rajasekharan and Kennedy, 2009). The embedding of netrin within the myoblast cell membranes ensures that pro-myogenic neogenin-netrin signalling is only initiated following direct cell-cell contact (figure 1.4) (Kang *et al.*, 2004). In *trans* interactions between neogenin and netrin lead to the activation of the focal adhesion kinase (FAK), an enzyme that promotes myoblast differentiation and fusion via the upregulation of caveolin 3 and integrin β1 (Hindi *et al.*, 2013; Wang and Almazan, 2016). Furthermore, in non-myoblast mammalian cell culture, neogenin is known to promote the activities of both Dock1 and Trio, two GEFs that regulate actin dynamics during fusion (via activation of Rac1) (figure 1.4) (Kang *et al.*, 2004; Krauss *et al.*, 2017).

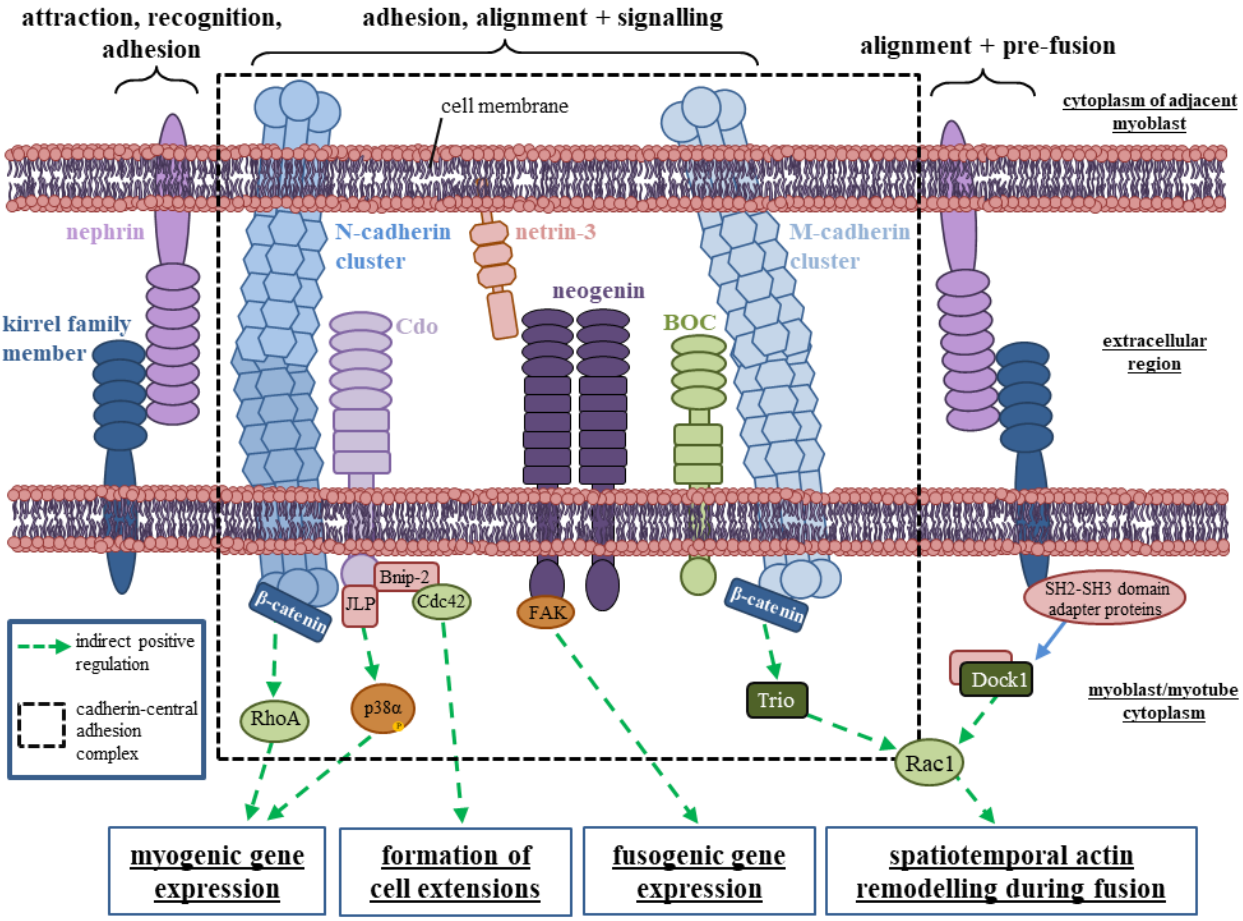


Figure 1.4 The IRM and cadherin-dependent recognition and adhesion between mammalian myoblasts and myoblasts/myotubes. The attraction to and initial recognition and adhesion of myoblasts to myotubes is thought to be in large facilitated by the irre recognition module molecules. The mammalian IRM molecules bind in *trans* with nephrin and also help regulate actin remodelling during fusion. A proposed myogenic complex consisting of cadherin clusters, Cdo, neogenin and BOC initiates a range of cascades that promote the formation of filopodia, myogenic gene transcription as well as participating in the fusion process. The components of this myogenic complex are expressed in both myoblasts and myotubes. These cadherin complexes are relocated away from the fusion site as the process progresses but continue to promote adhesion and alignment along the cell membrane. For simplicity, the cascades covered in-text are only represented in one of the fusion partners in the figure. Additionally, the regulation of actin remodelling by the IRM molecules will be covered in greater detail further along in the text. Constructed using references: Kang *et al.*, 2004; Hindi *et al.*, 2013; Wang and Almazan, 2016 and Krauss *et al.*, 2017.

Cell adhesion continued - Integrins, ADAM12 and the CD molecules: Also essential for myoblast adhesion and alignment during myogenesis are the integrins, the cluster of differentiation (CD) molecules and the ADAM Metallopeptidase Domain 12 (ADAM12) disintogren/metalloprotease (Przewoźniak *et al.*, 2013; Pizza *et al.*, 2017; Rout *et al.*, 2022). A number of integrins are known to be involved in myoblast adhesion and include integrin α3, integrin β1 and integrin α4β1 (Przewoźniak *et al.*, 2013; Pizza *et al.*, 2017). The canonical CD molecules involved in myogenesis include CD9, CD36, CD81, intercellular adhesion molecule 1 (ICAM-1) (also known as CD54), neural cell adhesion molecule (NCAM) (also known as CD56) and vascular cell adhesion protein 1 (VCAM-1) (also known as CD106) (Przewoźniak *et al.*, 2013; Pizza *et al.*, 2017; Rout *et al.*, 2022).

Prior to fusion, and in conjunction with the cadherins and irre molecules, homophilic *trans*-interactions between neural cell adhesion molecules (NCAMs) contribute towards the adhesion between myoblasts and myotubes (figure 1.5) (Przewoźniak *et al.,* 2013; Pizza *et al.,* 2017; Rout *et al.,* 2022). Bound to the intracellular tail of complexed NCAM molecules, the myogenesis-related and NCAM-associated protein (MYONAP) promotes the formation of filopodia, increasing the rate of new cell-cell contacts (figure 1.5) (Hirayama and Kim, 2008).

Heterophilic *trans*-interactions between vascular cell adhesion protein 1 (VCAM-1) and integrin α4β1 help facilitate the adhesion between myoblasts and myotubes as well as interactions between myoblasts and immune cells such as neutrophils and macrophages; cells critical for successful myogenesis *in vivo* (Choo *et al.,* 2017). Studies involving C2C12 myoblasts have shown that VCAM-1 is present on proliferating myoblasts while integrin α4β1 is expressed on the surface of multinucleated myotubes, in addition to the aforementioned immune cells (Rosen *et al.,* 1992; Choo *et al.,* 2017).

As differentiation progresses, the expression of the α3 and β1 integrin subunits is increased in myoblasts and myotubes with the two molecules complexing to form the integrin α3β1 dimer, a protein complex that co-localises in *cis* with the likewise upregulated ADAM12, CD9 and CD81 proteins (figure 1.5) (Przewoźniak *et al.,* 2013; Pizza *et al.,* 2017; Rout *et al.,* 2022). The integrin α3β1 dimer has also been shown to interact in *trans* with ADAM12 and other integrin α3β1 dimers on the surface of opposing fusion partners (Brzóska *et al.,* 2006; Przewoźniak *et al.,* 2013).

The homophilic in *trans* binding of intercellular adhesion molecule 1 (ICAM-1) proteins on myoblast/myotube partners facilitates adhesion as well as directly promoting fusion by stimulating the Rac1-mediated remodelling of the actin cytoskeleton (figure 1.5) (Przewoźniak *et al.,* 2013; Pizza *et al.,* 2017; Rout *et al.,* 2022).

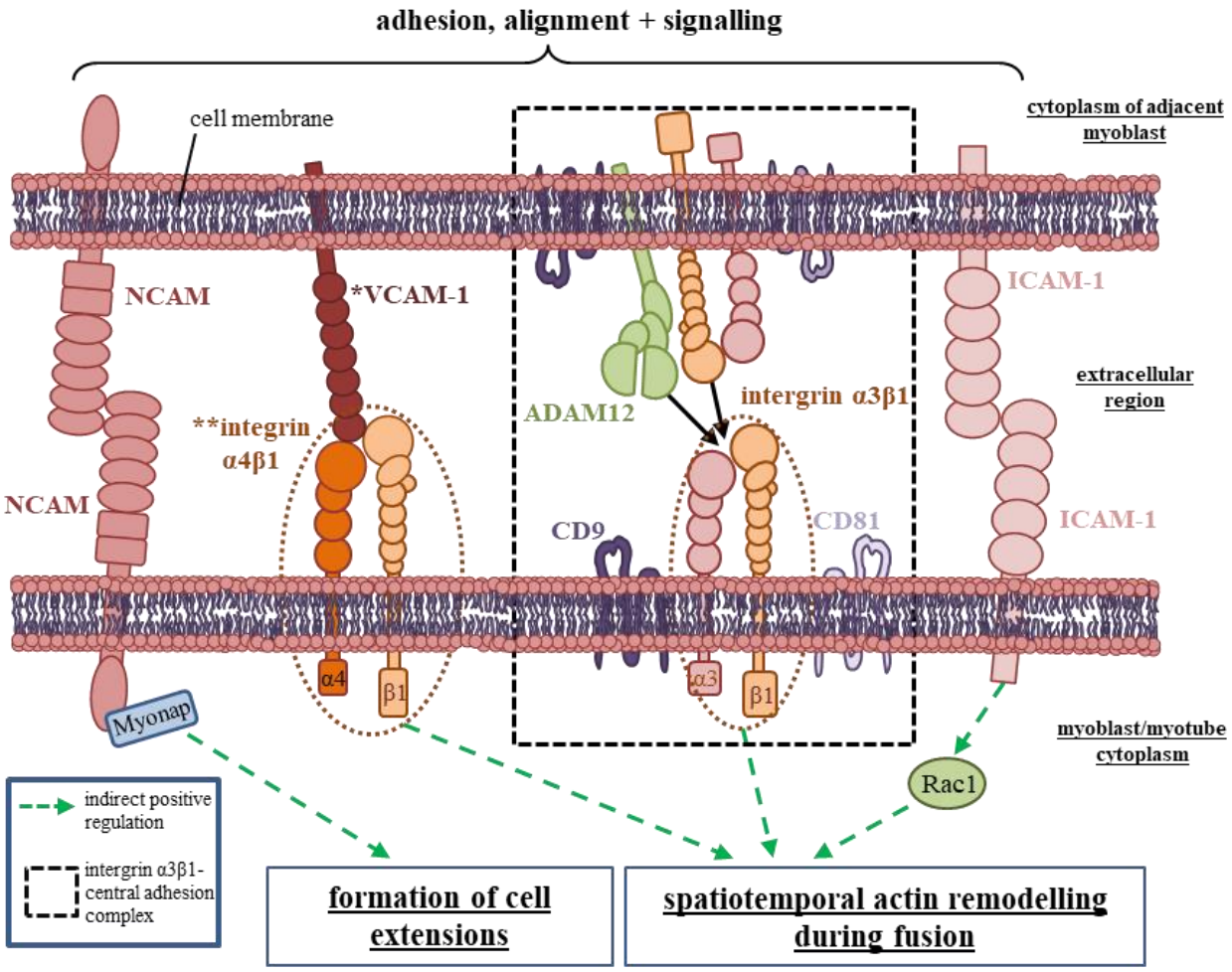


Figure 1.5 The integrins and CD molecules involved in the adhesion between myoblasts and myoblasts/myotubes. Further promoting adhesion, the integrins and CD molecules help facilitate a close association and alignment between the fusion partners. These molecules cluster in specific regions on the cell membranes; away from the immediate site of fusion pore formation. Homophilic binding of NCAM molecules promote filopodia formation, enhancing membrane closeness. In addition to facilitating adhesion, the heterophilic binding of VCAM-1 to integrin α4β1 and the homophilic biding of ICAM-1 molecules promote actin remodelling and fusion. The integrin α3β1 dimer co-localises in *cis* with the metalloprotease Adam12, CD9 and CD81. Integrin α3β1 interacts in *trans* with Adam12 and other α3β1 dimers on the opposing fusion partner and promote actin remodelling and fusion. For simplicity, the above-mentioned cascades are only represented in one of the fusion partners in the figure. Additionally, this figure does not account for the frequency/ratio of each receptor as clustering for CAMs is common. * indicates expression in myoblasts only. ** indicates expression in myotubes only. Constructed using references: Przewoźniak *et al.*, 2013; Pizza *et al.*, 2017; Rout *et al.*, 2022.

1.10 Myoblast fusion

Myoblast fusion is an intricate and relatively poorly understood process that is essential for the production of multi-nucleated myofibers (Krauss *et al.*, 2017). Genetic and mechanistic studies pioneered in *Drosophila* and subsequently expanded to zebrafish and mammalian models have shown that three consecutive steps are involved in myoblast fusion, these being; 1) cell recognition and adhesion, mediated by cell adhesion molecules (CAMs), 2) the enhancement of fusion partner closeness/integration, all followed by 3) the destabilization of the opposing lipid bilayers and the formation of a single-channel fusion

pore allowing for the transfer of cellular contents (figure 1.6) (Rochlin *et al.*, 2010; Kim *et al*., 2015; Krauss *et al.*, 2017).

Myoblast fusion has been extensively studied in *Drosophila* with many of the intracellular molecules regulating the process having been identified and where tested, shown to be evolutionarily conserved in vertebrates and in cultured cells (Krauss *et al.*, 2017). However, the mechanisms regulating vertebrate myoblast fusion (and cell fusion in general) are more complex than those in the fly and with many gaps in the knowledge (Krauss *et al.*, 2017). Despite the shortcomings of the current research, and due to the high level of evolutionary conservation of the core mechanisms controlling fusion, a number of valuable inferences can be obtained from the studies involving *Drosophila* myoblast fusion (Rochlin *et al.*, 2010; Krauss *et al.*, 2017). To better understand how the kirrel molecules may be involved in the differentiation and fusion of mammalian myoblasts, we will first look at these proteins with regard to *Drosophila* myogenesis.

1.11 Drosophila myogenesis: muscle founder cells, fusion competent myoblasts and the irre cell recognition module molecules

Two types of fusion muscle fusion cells: In *Drosophila* embryos, myoblast fusion takes place between two types of myoblasts: muscle founder cells (FCs), and fusion competent myoblasts (FCMs); with the type of myoblast being determined by a variety of transcription factors (Rochlin, 2010; Kim *et al*., 2015; Krauss, 2010). Each muscle in the fly is formed from a single muscle founder cell (FC) "seed" that attracts and fuses with surrounding fusion competent myoblasts (FCMs) and determines the orientation, position, size, nerve innervation pattern and muscle attachment site of the developing structure (Kim *et al*., 2015; Krauss *et al*, 2017). Muscle founder cells (FCs) represent a diverse population of myoblasts and can be distinguished by their expression of a number of transcription factors including Apterous, Even-skipped, Krüppel, Nautilus and Slouch; with these factors determining the characteristics of the FCs and accordingly, the morphology of the muscles produced (Kim *et al*., 2015). Conversely, the population of FCMs found in the developing embryo is larger and more uniform, characteristically expressing the transcription factor Lameduck (Rochlin *et al.*, 2010).

The irre recognition molecules in FCs and FCMs: In invertebrates, the IRM includes the two irre proteins Dumbfounded (Duf) and Roughest (Rst) as well as their binding partners, Sticks and Stones (Sns) and its paralog Hibris (Hbs) (Fischbach *et al* 2009; Krauss, 2010; Durcan *et al.*, 2013). Duf and Rst are homologous with the mammalian kirrel1, kirrel2 and kirrel3 proteins (Krauss, 2010; Durcan *et al.*, 2013). The vertebrate protein nephrin is homologous with both Sns and Hbs (Krauss, 2010; Durcan *et al.*, 2013). The Duf and Rst molecules share roughly 47% genetic identity with highly homologous cytoplasmic

regions and have seemingly overlapping functions during myoblast fusion (Rochlin *et al.*, 2010; Abmayr and Pavlath, 2012; Kim *et al*., 2015).

The recognition and adhesion between FCs and FCMs is heavily reliant on Ig and Ig-like domain-containing, type I transmembrane CAMs; most notably the IRM molecules (Rochlin *et al.*, 2010; Kim *et al*., 2015; Krauss *et al.*, 2017). In terms of IRM expression, FCs (and myotubes) exclusively express Duf and Rst whereas the FCMs typically express Sns and Hbs but have also been seen to produce Rst (Rochlin *et al.*, 2010; Önel *et al.*, 2014; Kim *et al*., 2015, Krauss *et al.*, 2017). The IRM molecules interact with their relevant binding partners via their multiple extracellular Ig and Ig-like domains and are essential regulators of fusion-driving actin polymerization in both FCs and FCMs (Fischbach *et al.*, 2009; Machado *et al.*, 2018). As with muscle development in the fly embryo, during the pupal stage, *Drosophila* muscles develop from individual Duf and Rst-expressing FCs/myotubes that attract Sns and Hbs-expressing FCMs (Kim *et al*., 2015). Both Duf and Rst can adhere directly to Sns or Hbs, facilitating the recognition and adhesion between FCs and FCMs (figures 1.6, 1.7 and 1.8) (Rochlin *et al.*, 2010; Kim *et al*., 2015). Duf and Rst are also capable of binding in *trans* to one another; however, it is unclear what cascades, if any, are initiated by these interactions (Galletta *et al.*, 2004). Sns and Hbs appear to be unable to bind to one another in *trans* (Galletta *et al.*, 2004). Despite the presence of homophilic interactions, Duf and Rst still have a strong preference for Sns (Abmayr and Pavlath, 2012).

Knockout and overexpression of the irre molecules in Drosophila: Although the single knockout of either *duf* or *rst* does not lead to impaired *Drosophila* myoblast fusion, the double knockout of these genes results in myoblasts being unable to properly recognise and adhere to their relevant fusion partners, thus preventing fusion and demonstrating the redundant functions of Duf and Rst (Ruiz-Gómez *et al.*, 2000; Strunkelnberg *et al.*, 2001; Rochlin *et al.*, 2010). The expression of Duf or Rst in these *duf;rst* double mutant embryos completely rescues the fusion process (Ruiz-Gómez *et al*., 2000; Strunkelnberg *et al.*, 2001). Further research has shown that in *duf*;*rst* double mutant embryos, Sns does not become enriched or localised at the FC-FCM contact site and is instead evenly distributed at the cell cortex (Galletta *et al.*, 2004). Ectopic expression studies in which Duf or Rst have been expressed on non-myoblast cells have also demonstrated the ability of these CAMs to attract FCMs towards heterologous cells (Ruiz-Gómez *et al.*, 2000; Strunkelnberg *et al.*, 2001; Kim *et al.*, 2015).

Hibris acts as a positive regulator of *Drosophila* myoblast fusion and can partially compensate for mutations to *sns* in cell culture models, however, the loss of both Sns and Hbs results in the complete inhibition of *Drosophila* myoblast fusion (Shelton *et al.*, 2009). Knockout studies carried out in fly embryos have demonstrated a critical function for Sns during embryogenesis, with the loss of Sns resulting in the complete

failure of the embryos to produce any muscle fibers, and accordingly, a failure to hatch; however, flies lacking the Sns paralog Hbs are still viable (Bour *et al.,* 2000). The expression of the differentiation-promoting MEF2 transcription factor is unchanged in these *sns* mutant *Drosophila* embryos, implying that the mutant muscle progenitor cells begin their differentiation programs but are unable to fuse (Bour *et al.*, 2000). Interestingly, the overexpression of Hbs in developing *Drosophila* embryos appears to have a dominant-negative effect on Sns and results in defective myoblast fusion, potentially due to increased Sns-Hbs heterodimerization (Lee and Chen, 2019).

Ectopic expression studies have also shown that the over-expression of Duf and Sns in the non-fusogenic Schneider's-line-2-derived S2R+ cell line (derived from *Drosophila* epithelial cells) was not enough to induce cell-fusion, although the over-expression of these 2 CAMs resulted in an extensive increase in the level of cell adhesion and F-actin enrichment at the sites of contact respectively (Shilagardi *et al.,* 2013). However, the overexpression of both Sns and the fusogenic *C.elegans*-derived epithelial fusion failure 1 (Eff-1) protein in the S2R+ cell line (*Drosophila*) resulted in a 7-fold increase in fusion relative to cells overexpressing Eff-1 alone (Shilagardi *et al.,* 2013). The overexpression of Duf and epithelial fusion failure 1 (Eff-1) did not result in increased fusion relative to Eff-1 overexpression alone, suggesting that membrane opposition brought about by increased cell adhesion is not enough to enhance Eff-1 mediated fusion (Shilagardi *et al.,* 2013). It is likely that the differences seen in the fusion of Eff-1+Duf and Eff-1+Sns overexpressing mutants is due to the different actin-regulating mechanisms of Sns vs Duf i.e., it appears that overexpressing Sns and promoting the production of the FCM protrusions is a more powerful driver of myoblast fusion than increasing Duf levels and leading to the formation of the actin sheath in FCs (Shilagardi *et al*., 2013).

1.12 Mechanisms regulating Drosophila myoblast fusion

1.12.1 Formation of the FuRMAS

Dumbfounded (Duf), Rst, Sns and Hbs contain 5, 5, 8 and 8 Ig and/or Ig-like domains in their extracellular regions respectively (Önel *et al.,* 2014; Kim *et al*., 2015). During *Drosophila* myoblast fusion, the IRM CAMs trigger distinct actin cytoskeletal reorganisation events that are essential for this process (Shilagardi *et al.*, 2013; Kim *et al*., 2015). Following adhesion, Duf or Rst in the FC stimulate the production of a thin, circular sheath of networked F-actin and non-muscle myosin II (MyoII) that runs parallel to the cell-membrane at the fusogenic synapse (figures 1.6, 1.7 and 1.8) (Shilagardi *et al.,* 2013; Kim *et al*., 2015). Conversely, Sns in the FCM promotes the formation of an F-actin-enriched podosome-like structure (PLS) at the site of fusion (Shilagardi *et al.,* 2013). The PLS facilitates the formation of finger-like protrusions which "invade" the FC and drive fusion (figures 1.6 and 1.9) (Kim *et al*., 2015; Chal and Pourquié, 2017).

As the early fusion process progresses, most bound IRM complexes relocate and arrange to form a ring-like structure at the focus of adhesion called the fusion-restricted myogenic-adhesive structure (FuRMAS) (figures 1.6 and 1.9) (Krauss, 2010; Rochlin *et al.,* 2010; Abmayr and Pavlath, 2012). F-actin and its associated regulators accumulate within the enclosed area of the FuRMAS structure in both FCs and FCMs (Kesper *et al.,* 2007; Richardson *et al.,* 2007). The FuRMAS gradually expands outwards, increasing its radius while continuing to promote the formation of the invasive PLS and actomyosin sheath in FCMs and FCs respectively (figure 1.6, 1.8 and 1.9) (Krauss, 2010; Rochlin *et al.,* 2010; Abmayr and Pavlath, 2012). Accordingly, the invasive protrusions and subsequent nascent fusion pores also develop within the area of the FuRMAS (Krauss, 2010; Rochlin *et al.,* 2010; Abmayr and Pavlath, 2012). On average, FCMs produce 4.3 of these actin-propelled protrusions which push against the FC membrane, resulting in an inward curvature of the membrane reaching depths of up to 1.9µm into the FC (Sens *et al.,* 2010). The mechanical force exerted by the invading protrusions on the FC sheath results in a MyoII-mediated increase in the cortical tension at the fusogenic synapse which in turn increases cell-cell closeness, restricting the boundary of the protrusions and encouraging pore formation (Kim *et al*., 2015; Deng *et al.,* 2017). Prior to fusion pore formation, the membranes of the opposing myoblasts become integrated at the tips of the invading protrusions i.e., hemifusion (figure 1.6) (Kim *et al*., 2015; Chal and Pourquié, 2017). Hemifusion and pore formation is largely controlled by specialized membrane-embedded fusion proteins (Krauss *et al.,* 2017). Nascent fusion pores form at the sites of hemifusion and gradually expand outwards by vesiculation of excess membrane and ultimately merge to form a single-channel fusion pore allowing for the movement of the FCM cytoplasm and nucleus into the FC (figure 1.6) (Kim *et al*., 2015; Chal and Pourquié, 2017; Deng *et al.,* 2017). As the fusion pore forms, the actin and MyoII that has accumulated at the fusion synapse dissolves and disperses from the area (Kim and Chen, 2019; Lee and Chen, 2019). The remaining FCM cell-membrane ultimately becomes integrated with that of the FC (Kim *et al*., 2015; Chal and Pourquié, 2017).

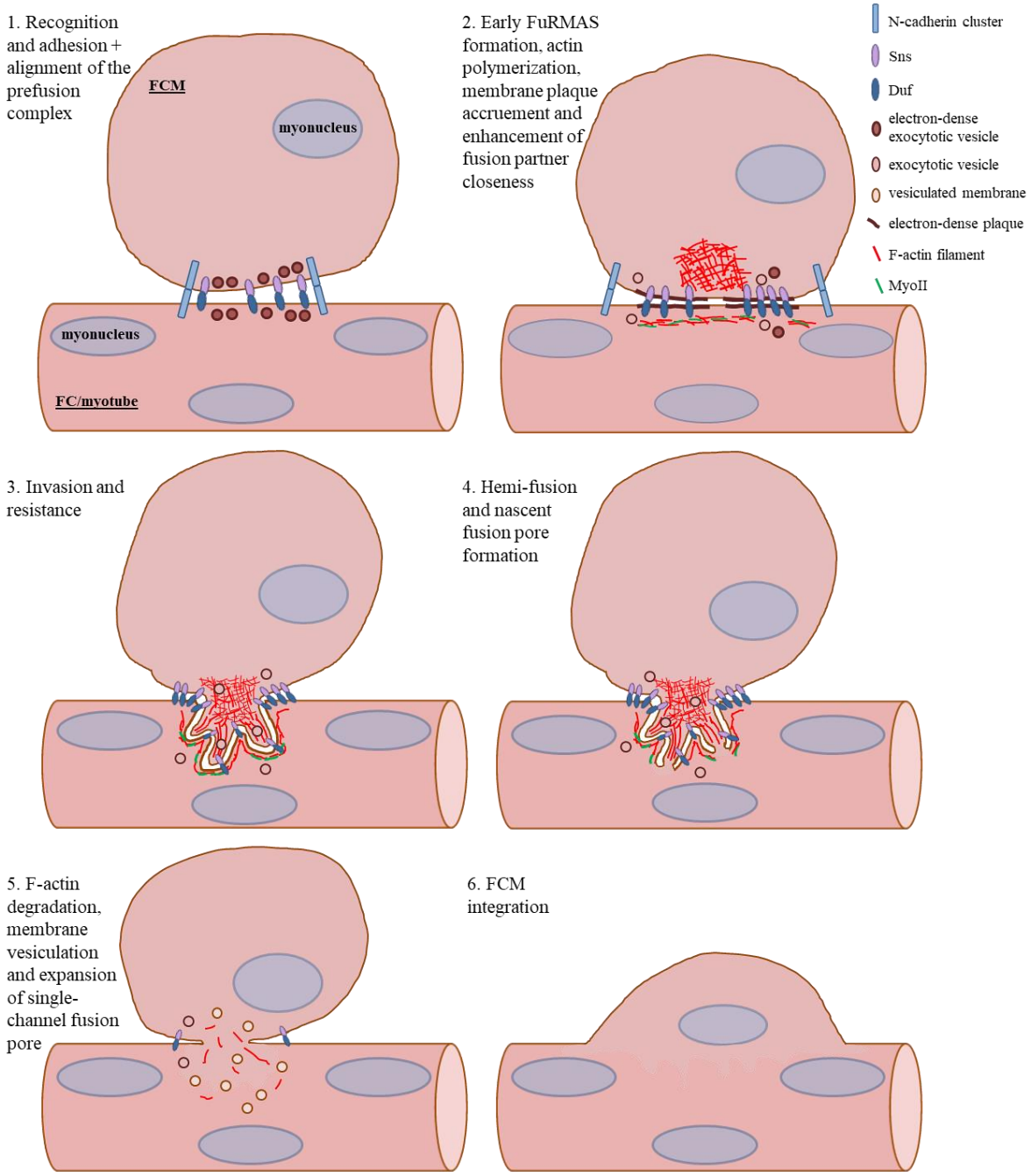


Figure 1.6 Schematic representation of the recognition, adhesion and fusion of *Drosophila* fusion competent myoblasts and founder cells. 1) The recognition, adhesion and alignment of FCs and FCMs is largely facilitated by interactions between IRM partners as well as the homophilic binding of N-cadherin clusters on opposing myoblasts. The electron-dense vesicles of the prefusion complex pair up and align themselves along the cell membrane of the adhered myoblasts. 2) The IRM molecules initiate F-actin polymerization in both fusion partners which then leads to closer cell-cell proximity. Sns in the FCM stimulates the production of an F-actin podosome-like structure while Duf and Rst in the FC initiate the production of a thin, circular sheath of F-actin and non-muscle MyoII. The electron dense vesicles of the prefusion complex fuse with the cell membrane to produce electron-dense plaques that are co-localised with the IRM and cadherin molecules in both myoblasts. The bound IRM complexes gradually arrange themselves into a ring-shaped formation referred to as the FuRMAS. 3) The F-actin podosome-like structure produces finger-like protrusions that extend from the FCM. As fusion progresses, the FuRMAS expands outwards allowing for invasion by FCM protrusions through the area enclosed by the ring and into the FC. 4) The membranes of the fusion partners become integrated at the tips of the protrusions which then become the sites of nascent fusion pore formation. 5) The nascent fusion pores expand and fuse to form a single fusion channel through which the cytoplasm and nucleus of the FCM moves into the FC. Depolymerization of F-actin occurs at the fusogenic synapse. 6) The cell membrane of the FCM becomes integrated with that of the FC during the final stage of fusion. Constructed using references: Krauss, 2010; Rochlin *et al.*, 2010; Abmayr and Pavlath, 2012; Önel *et al.*, 2014; Kim *et al*., 2015; Chal and Pourquié, 2017; Deng *et al.*, 2017.

1.12.2 Exocytic vesicles and the formation of electron-dense plaques

During the early fusion process, prior to and during the formation of the FuRMAS, electron-dense vesicles originating from the Golgi are trafficked along microtubules to the site of adhesion (Galletta *et al.*, 2004; Kim *et al*., 2007; Abmayr and Pavlath, 2012; Kim *et al*., 2015). These exocytic vesicles arrange themselves close to the cell membrane at the future fusogenic synapse in both FCs and FCMs and appear to pair up with vesicles in the opposing myoblast (figure 1.7) (Galletta *et al.*, 2004; Kim *et al*., 2007; Abmayr and Pavlath, 2012; Kim *et al*., 2015). These groups of paired electron-dense vesicles are collectively known as the prefusion complex (Galletta *et al.*, 2004; Kim *et al*., 2007; Abmayr and Pavlath, 2012; Kim *et al*., 2015). The paired vesicles then fuse with the cell-membranes to form electron-dense plaques that co-localise with the IRM CAMs and cadherin molecules and move outwards with the FuRMAS as it expands (Galletta *et al.*, 2004; Kim *et al*., 2007; Abmayr and Pavlath, 2012; Kim *et al*., 2015). It is thought that these plaques are formed and persist during an intermediate stage between fusion of the electron-dense vesicles and the formation of the nascent fusion pores (Önel *et al.*, 2014). Furthermore, exocytic vesicles from the Golgi shuttle IRM molecules to the cell membrane and may bind directly to IRM molecules already embedded in the membrane (Galletta *et al.*, 2004; Kim *et al*., 2007; Abmayr and Pavlath, 2012; Kim *et al*., 2015). As fusion progresses, exocytic vesicles shuttle fusion regulating proteins to the site of fusion (Galletta *et al.*, 2004; Kim *et al*., 2007; Abmayr and Pavlath, 2012; Kim *et al*., 2015).

The transmembrane protein Singles bar (Sing) is present in both FCs and FCMs and along with the Ca^{2+}-binding EF-hand protein *Drosophila* Swiprosin-1 (DSwip-1) (present only in FCMs), facilitates the exocytosis of the electron-dense vesicles of the prefusion complex (figures 1.7) (Önel *et al.*, 2014; Brunetti *et al.*, 2015; Kim *et al*., 2015). Sing is regulated by MEF2 and is required for adult *Drosophila* myoblast fusion with *sing* mutant myoblasts being able to adhere and form actin foci but are unable to fuse; thought to be due to the inhibited exocytosis of electron-dense vesicles (Estrada *et al.*, 2007; Önel *et al.*, 2014; Brunetti *et al.*, 2015). The exact localization of Sing during fusion is unclear but the protein is suspected to be trafficked in vesicles as well as being embedded in the cell membranes close to the fusogenic synapse (Estrada *et al.*, 2007; Önel *et al.*, 2014; Brunetti *et al.*, 2015; Rout *et al.*, 2022). Once bound to Duf or Rst, Sns in the FCM recruits *Drosophila* Swiprosin-1 (DSwip-1) to the same general region as the F-actin PLS (Önel *et al.*, 2014).

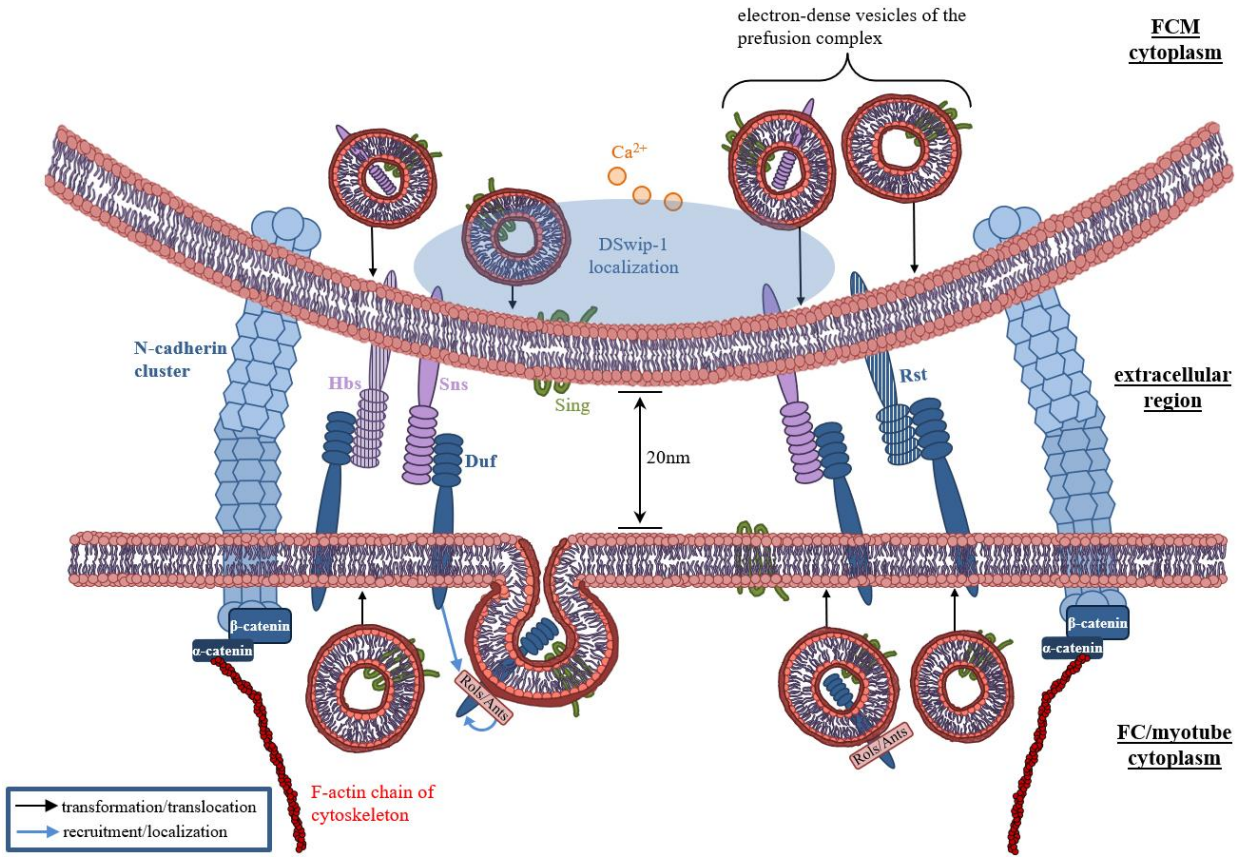


Figure 1.7 The irre-molecule dependent formation of the FuRMAS and build-up of the electron-dense plaques in *Drosophila* myoblasts/myotubes during the early fusion process. The IRM molecules facilitate the recognition and binding of fusion partners during myogenesis. During the early stages of fusion, the structure of the FuRMAS is not yet formed and the fusion partners are held roughly 20nm apart by CAMs at the site of adhesion. As fusion progresses, the electron dense vesicles are shuttled to and fuse with the cell membrane at the site of adhesion, beginning the build-up of the electron dense plaques. The sing protein is essential for the formation of these plaques, likely regulating the exocytosis of the electron-dense vesicles. The IRM molecules as well as the cadherins appear to localize with these plaques during the early stages of adhesion and fusion. The IRM molecules are also shuttled to the cell membrane in vesicles, a process regulated by other cell- surface IRM molecules, sing as well as Rols7. For greater simplicity, interactions between Rst and Hbs or Sns are not shown. The mechanisms activated by Rst binding to Hbs or Sns are consistent with those initiated by the binding of Duf and Hbs or Sns. Constructed using references: Galletta *et al.*, 2004; Kim *et al*., 2007; Abmayr and Pavlath, 2012; Önel *et al.*, 2014; Kim *et al*., 2015

1.12.3 Actin polymerization and production of the mechanosensitive actomyosin network

The degree of cell-cell closeness achieved by the initial binding of members of the IRM is not sufficient to facilitate myoblast fusion (Önel *et al.*, 2014; Kim *et al*, 2016). Instead, the binding of the IRM partners and the formation of the FuRMAS acts to localize and activate a number of adapter proteins, guanine nucleotide exchange factors (GEFs), GTPases and actin nucleation-promoting factors (NPFs) that together regulate the cytoskeletal reorganization required for fusion (Önel *et al.*, 2014; Kim *et al*., 2015; Lee and Chen, 2019). The accumulation and activation of these proteins results in the formation of the podosome-like structures (PLS) and actomyosin sheath in the myoblast fusion pair; structures essential for increased membrane proximity, protrusion invasion and fusion (Önel *et al.*, 2014; Kim *et al*., 2015; Lee and Chen, 2019). The nucleation of F-actin and the production of branching actin chains during polymerization is primarily

mediated by the Actin related protein 2/3 (Arp2/3) complex (Önel *et al.*, 2014; Kim *et al*., 2015; Lee and Chen, 2019).

Arp2/3: Genetic studies in *Drosophila* have highlighted the critical importance of actin cytoskeletal regulators during myoblast fusion with most of these regulators mediating their effects via the Actin related protein 2/3 (Arp2/3) complex; a seven-subunit, actin-nucleating complex/factor that is vital for the formation of branching actin networks (Abmayr and Pavlath, 2012; Önel *et al.*, 2014; Kim *et al*., 2015; Deng *et al.*, 2017; Lee and Chen, 2019). The Actin related protein 2/3 (Arp2/3) complex directly mediates the nucleation of the F-actin making up both the podosome-like PLS in FCMs and the actomyosin sheath in FCs (figures 1.8 and 1.9) (Abmayr and Pavlath, 2012; Önel *et al.*, 2014; Kim *et al*., 2015; Deng *et al.*, 2017; Lee and Chen, 2019). The activity of the Arp2/3 complex is primarily controlled by two actin nucleation-promoting factors (NPFs), these being the Suppressor of cAMP receptor (Scar) and the Wiskott-Aldrich syndrome protein (WASP) (Abmayr and Pavlath, 2012; Önel *et al.*, 2014; Kim *et al*., 2015; Deng *et al.*, 2017; Lee and Chen, 2019). Inhibition of Arp2/3 activity via RNA interference (RNAi) results in strong fusion-arrest phenotypes in developing *Drosophila* (Hudson and Cooley, 2002).

Scar and WASP: Both the Suppressor of cAMP receptor (Scar) and the Wiskott-Aldrich syndrome protein (WASP) are part of the WASP family of proteins and are critical for Arp2/3 activity; however, these two nucleation-promoting factors (NPFs) have distinct, non-interchangeable functions during the fusion process (Krauss *et al.*, 2017; Deng *et al.*, 2017). The Suppressor of cAMP receptor (Scar), also known as WASP family verprolin homologs (WAVE), is found in both FCMs and FCs and forms part of the Scar pentamer complex (Berger *et al.*, 2008; Lee and Chen, 2019). WASP is produced in FCMs but not in FCs and is recruited to the fusogenic synapse and stabilized in a tight, active complex by the *Drosophila* WASP-interacting protein/Solitary (D-WIP/Sltr) (figures 1.8 and 1.9) (Jin *et al.*, 2011; Lee and Chen, 2019). When associated with Solitary (Sltr), WASP binds to the barbed ends of actin filaments in the developing cytoskeleton while stimulating local Arp2/3 activity (Jin *et al.*, 2011; Lee and Chen, 2019). The cytoplasmic protein Blown Fuse (Blow) competes with WASP for Solitary (Sltr) binding, thus regulating the stability of the Sltr-WASP complex (figures 1.8 and 1.9) (Jin *et al.*, 2011). The destabilization of the Sltr-WASP complexes results in the unbinding of WASP from the actin filament ends and reduced Arp2/3 stimulation (Jin *et al.*, 2011). The removal of WASP from the actin filaments allows for end capping and the initiation of new branching filaments; processes essential for cytoskeletal organization (Jin *et al.*, 2011). Like WASP, neither Blow nor Sltr expression is detected in FCs during fusion (Sens *et al.*, 2010; Kim *et al*., 2015; Lee and Chen, 2019).

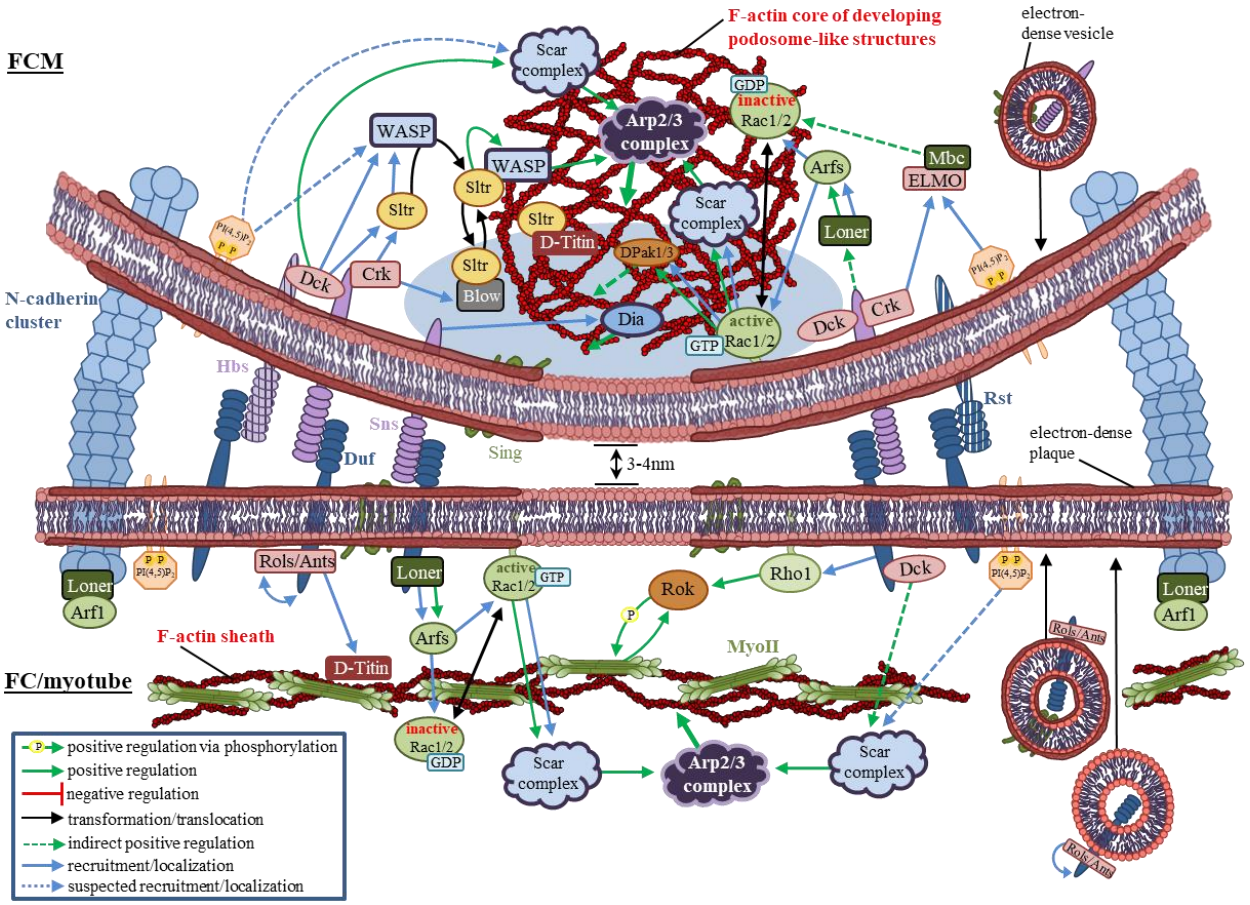


Figure 1.8 The irre molecule-mediated regulation of F-actin polymerization and actomyosin network formation in *Drosophila* myoblasts/myotubes during the early-mid fusion process. As fusion progresses, the IRM molecules accumulate at the site of fusion and the ring-shape of the FuRMAS begins to take shape. As the CAMs making up the FuRMAS expand outwards, the membranes of the FC and FCM fusion partners come into closer contact. In FCMs, Sns molecules promote pathways that initiate the development of the F-actin podosome-like structures and facilitating the invasion of the FC fusion partner. In FCs/myotubes, Duf and Rst promote the formation of the F-actin sheath as well as the mechanosensitive MyoII network, structures that detect and resist invasion by FCMs; processes that increases membrane intimacy and allow for fusion pore formation. For greater simplicity, interactions between Rst and Hbs or Sns are not shown. The mechanisms activated by Rst binding to Hbs or Sns are consistent with those initiated by the binding of Duf and Hbs or Sns. Constructed using references Menon and Chia, 2001; Galletta *et al.*, 2004; Bothe *et al.*, 2014; Deng *et al.*, 2015; Deng *et al.*, 2017; Lee and Chen, 2019.

Crk and Dck: The *Drosophila* SH2-SH3 domain-containing adapter proteins CT10 regulator of kinase (Crk) and dreadlocks (Dck) are heavily involved in the recruitment and activation of a number of proteins involved in cytoskeletal organization at the fusogenic synapse (Kaipa *et al.*, 2013; Kim and Chen, 2019; Lee and Chen, 2019). These two adapter proteins effectively act as a link between the irre recognition module (IRM) molecules and components of the actin polymerization machinery (Kim and Chen, 2019). The accumulated IRM molecules at the fusogenic synapse act to localize these SH2-SH3 adapters to the area (Kaipa *et al*., 2013; Lee and Chen, 2019). Dreadlocks (Dck) is expressed in both FCs and FCMs and has been shown to interact with Duf, Rst, Sns and Hbs via SH2 and/or SH3 domains within the adapter protein structure (Kaipa *et al*., 2013; Lee and Chen, 2019; Kim and Chen, 2019). In FCs, the SH3 domain of Dck binds to the proline-rich region within the cytoplasmic domain of Duf, promoting Scar complex

activity (figures 1.8 and 1.9) (Kaipa *et al*., 2013). In FCMs, Hbs and Sns bind to the SH2 and SH3 domains of Dck respectively and are thought to promote Scar complex activity and localize Sltr and WASP to the site of actin polymerization (figures 1.8 and 1.9) (Kaipa *et al*., 2013; Kim and Chen, 2019; Lee and Chen, 2019). Although it is unknown whether Crk is expressed in either FCs or FCMs alone or in both cell types, Crk has been demonstrated to bind the cytoplasmic domain of Sns resulting in the localization of Blow and Sltr to the synapse (Kaipa *et al*., 2013; Kim and Chen, 2019; Lee and Chen, 2019. Furthermore, Crk binding to Sns leads to the activation of the Rac family of GTPases which in turn promote Scar complex activity (figures 1.8 and 1.9) (Kaipa *et al*., 2013; Kim and Chen, 2019; Lee and Chen, 2019).

Loner, Mbc-Elmo, Rac and DPak3: In FCMs, Rac1 and Rac2 are vital for the activation of the Scar complex during myoblast fusion (Önel *et al.,* 2014; Deng *et al.,* 2017). The localization and activation of these GTPases is largely controlled by the IRM proteins at the fusion synapse. The GEF Loner is essential for fusion and is positively regulated by Sns which once promoted, Loner recruits and activates its effectors, the Arf small GTPases which then go on to localize Rac GTPases to the synapse (figures 1.8 and 1.9) (D'Souza-Schorey and Chavrier, 2006; Önel *et al.,* 2014; Mrozik *et al.,* 2018). Additionally, Loner and Arf1, control the translocation of N-cadherin molecules away from the centre of the synapse prior to fusion, most-likely to reduce the distance between the opposing membranes (D'Souza-Schorey and Chavrier, 2006; Önel *et al.,* 2014; Mrozik *et al.,* 2018). Sns-bound Crk localizes a bipartite Rac GEF complex made up of engulfment and cell motility (Elmo) and Myoblast city (Mbc), a complex essential for the activation of the Rac GTPases (Haralalka *et al.,* 2011) (figures 1.8 and 1.9). In addition to promoting Scar activity, Rac proteins in FCMs recruit and activate the *Drosophila* p21-activated kinase 1 (DPak1) and DPak3 (Kim *et al.,* 2015; Lee and Chen, 2019). The DPak proteins have partially redundant functions during myoblast fusion and play major roles in the organization of the actin filaments within the PLS into a densely packed structure (figures 1.8 and 1.9) (Lee and Chen, 2019). Current research appears inconclusive about the extent of Rac activity in FCs during fusion but it has been suggested that as with FCMs, Rac proteins are activated by Loner (bound to Duf) and are recruited to the site in part by Arf1 and Arf6 (D'Souza-Schorey and Chavrier, 2006; Önel *et al.,* 2014; Mrozik *et al.,* 2018). However, Elmo and Mbc expression is absent in FCs, preventing Rac activation via this pathway (Önel *et al.,* 2014; Kim *et al.,* 2015; Lee and Chen, 2019).

Rols7/Ants and D-titin: The FC-specific adapter protein Rolling Pebbles 7/Antisocial (Rols7/Ants) promotes the replenishment and enrichment of Duf and Rst at the adhesion focus and is thought to do so by co-translocating along with the IRM molecules in exocytic vesicles (figures 1.7 and 1.8) (Menon *et al.,* 2005; Kim *et al*., 2015). Along with Duf and Rst, Rols7 is expressed in FCs prior to fusion, however, as the fusion process progresses the IRM molecules and Rols7 are seen to translocate from the cytoplasm to

the cell membrane at sites of FC-FCM adhesion (Menon and Chia, 2001; Menon *et al.,* 2005; Kim *et al*., 2015). Furthermore, cell membrane-bound Duf and Rst recruit cytoplasmic Rols7 to the fusogenic synapse (Menon *et al.,* 2005). Through this Duf/Rst-Rols7 positive feedback loop, the IRM molecules accumulate at the FC synapse and in turn bind to and stabilize Sns proteins on the FCM surface (Galletta *et al.,* 2004; Menon *et al.,* 2005). In *rols7* mutants, Duf is not seen at the synapse and fusion is severely inhibited (Deng *et al.,* 2017; Lee and Chen, 2019). Additionally, via a region of ankyrin repeats in its structure, Rols7 attaches to the cytoplasmic tail of Duf and Rst and recruits the massive cytoskeletal protein *Drosophila*-titin (D-titin) to distinct points at the fusion site (Zhang *et al.,* 2000; Menon and Chia, 2001; Galletta *et al.*, 2004). *Drosophila*-titin (D-titin) is thought to be an essential structural component of the myoblast cytoskeleton during fusion and is highly involved in sarcomere assembly and function (Zhang *et al.,* 2000; Menon and Chia, 2001; Galletta *et al.*, 2004; Lee and Chen, 2019). In FCs, D-titin is recruited to the synapse by Sltr and is thought to be involved with actin filament organization within the PLS (Kim *et al.,* 2007).

Diaphanous: The *Drosophila* formin Diaphanous (Dia) colocalizes with the actin foci at the fusogenic synapse in FCMs and promotes the linear nucleation of actin (figures 1.8) (Goode and Eck, 2007). Diaphanous (Dia) is recruited to the synapse by Sns where its localization and activity is unaffected by the upstream regulators of Arp2/3 (Deng *et al.,* 2015). It has been shown that *dia* mutant *Drosophila* embryos and embryos expressing a dominant-negative *dia* develop thinner muscle fibers that contain fewer nuclei (Lee and Chen, 2019). However, the mechanism of this pathology is not fully understood and seems to involve more than actin dynamics as these *dia* mutant myoblasts do not appear to express myosin heavy chain (MyHC), suggesting a failure of the cells to fully differentiate and mature (Deng *et al.,* 2015). Additionally, it is unclear exactly how Dia is involved in actin polymerization as the protein still accumulates at FC-FCM synapse in *sltr* and *kette* mutants in which the actin foci do not form (Deng *et al.,* 2015).

Phosphatidylinositol-(4,5)-biphosphate: The phospholipid Phosphatidylinositol-(4,5)-biphosphate (PIP2) becomes enriched at the fusogenic synapse where it recruits Mbc and possibly Scar and WASP (figures 1.8) (Bothe *et al*., 2014; Deng *et al.,* 201). PIP2 localization at the synapse is promoted by CAMs such as the IRM proteins and it has been shown in fly embryos with mutations to *sns* that no PIP2 accumulates at the FC-FCM foci (Bothe *et al.,* 2014). By reducing the levels or availability of PIP2 *in vivo*, it has been demonstrated that the phospholipid is essential for fusion with lower PIP2 levels resulting is small, misshapen actin foci, likely due to misregulation of actin regulators (Bothe *et al.,* 2014).

The mechanosensitive actomyosin and spectrin networks - Rho1, Rok, myosin II and spectrin: The invasive finger-like protrusions extending from FCMs at the fusogenic synapse increase the cell surface contact area between the FC-FCM fusion partners; a hand-in-glove like confirmation that essential for successful myoblast fusion (figures 1.8 and 1.9) (Kim and Chen, 2019; Lee and Chen, 2019). In response to the mechanical force exerted by the protrusions, non-muscle myosin II (MyoII) and spectrin accumulate at the site of fusion in FCs where they restrict the boundary of the fusogenic synapse and push back against the invasive protrusions, ensuring a greater intimacy between the opposing lipid bilayers to better facilitate fusion (figure 1.9) (Kim and Chen, 2019; Lee and Chen, 2019).

The small GTPase Rho1 is recruited to the synapse by Duf and Rst in FCs where it activates Rho-associated protein kinase (Rok), which in turn phosphorylates the regulatory light chain of MyoII, activating and stabilizing MyoII (figures 1.8 and 1.9) (Kim and Chen, 2019; Lee and Chen 2019). The accumulation, activation and stabilization of MyoII increases the pressure applied by the FC membrane in response to the invasive protrusions, greatly increasing the FC-FCM contact area (Lee and Chen, 2019). Although single *rho1* or *rok* embryos display normal musculature, double *rho1*;*rok* mutant embryos display fusion defects (Kim *et al.*, 2015). The expression of the MyoII regulatory light chain in these double *rho1*;*rok* mutants partially rescues the fusion defects seen in these cells, highlighting the requirement of Rho-Rok for successful MyoII activation (Kim *et al.*, 2015). Despite the need for Rok during MyoII activation, MyoII is still enriched at the fusogenic synapse in the absence of Duf-mediated Rho1-Rok activity in fly embryos with MyoII accumulation being mechanically stimulated by the invasive protrusions (Kim *et al.*, 2015). Additionally, MyoII is seen to accumulate in FCs at the FC-FCM foci in response to mechanical stimuli prior to the buildup of Rho1 and Rok (Kim *et al.*, 2015).

The second mechanosensitive protein that accumulates at the fusion synapse in FCs is spectrin; a molecule generally considered a membrane scaffold component and is essential for cell shape maintenance and providing structural support to membranes (Kim and Chen, 2019; Lee and Chen, 2019). The chain-like spectrin tetramers that make up spectrin networks consist of two α-spectrin and two β-spectrin molecules and can bind actin at both ends of the tetramer structure (Kim and Chen, 2019; Lee and Chen, 2019). Of the two types of β-spectrin, these being β_{Heavy}-spectrin (β_H-spectrin) and β-spectrin, only β_H-spectrin has been shown to be involved with myoblast fusion (Duan *et al.*, 2018). β_{Heavy}-spectrin (β_H-spectrin) is recruited to the fusogenic synapse in FCs by Duf and Rst as well in response to mechanical stimuli from the FCM protrusions (Kim and Chen, 2019; Lee and Chen, 2019). The α/β_H-spectrin molecules accumulate and form a sieve-like physical barrier that only allows the formation and invasion of protrusions through areas in which spectrin is absent (Kim and Chen, 2019; Lee and Chen, 2019). As the FCM protrusions push through the holes in the α/β_H-spectrin barrier, the diameter of the invasive structures is constricted,

thus increasing the mechanical tension exerted by the narrower protrusions and improving cell fusion (Kim and Chen, 2019; Lee and Chen, 2019). Together with actin, spectrin at the foci dissolves following fusion pore formation (Kim and Chen, 2019; Lee and Chen, 2019).

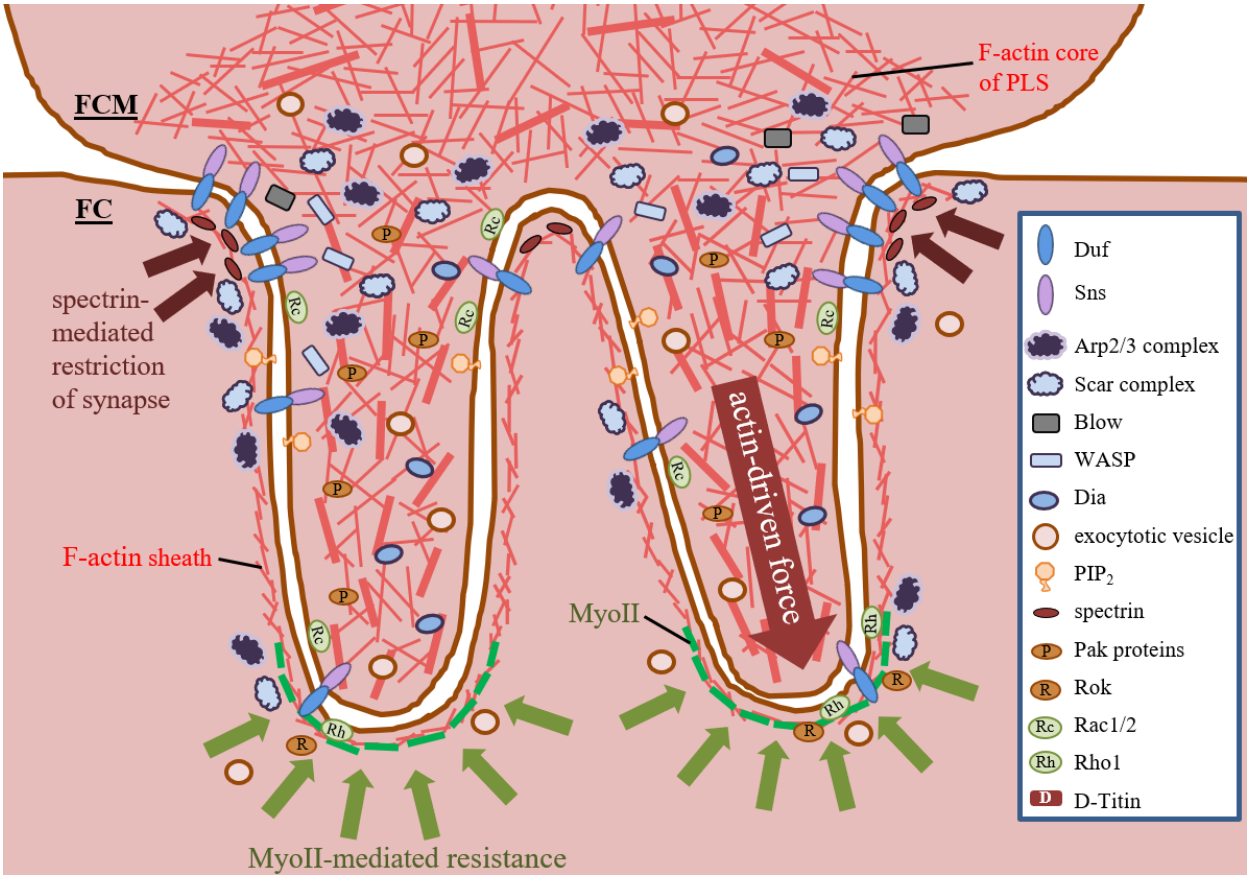


Figure 1.9 Localization of the IRM molecules and their effectors as they direct invasion and resistance during *Drosophila* myoblast fusion. The activity of Sns in FCMs promotes the development of the actin-driven protrusions that press into the FC body. This invasion is resisted by structures within the FC i.e., the β_H-spectrin network that constricts the base of the protrusions as well as by the MyoII sheets that form at the tips of the invasive structures. This process of invasion and resistance ensures a high level of intimacy between the two opposing cell membranes, allowing for fusion pore formation and easier integration of the two fusion partners. For simplicity, interactions between Rst and Hbs or Sns are not shown. Constructed using references: Kim *et al.*, 2015; Duan *et al.*, 2018; Kim and Chen, 2019; Lee and Chen 2019.

1.13 Vertebrate myoblast fusion

Myoblast fusion in invertebrates has been extensively researched with the mechanisms driving the process being well established. However, given the more complex nature of cell fusion in vertebrates, our understanding of this process in these systems is still relatively basic (Sohn *et al.*, 2009; Durcan *et al.*, 2013). Although valuable inferences can be obtained from research conducted in the fly and in other invertebrates, there is still the need to directly determine the functions of many of the evolutionarily conserved molecules found in vertebrates; these including the kirrel proteins. Other than the work done by Durcan *et al.*, (2013), no other assessments of kirrel1 activity during mammalian myogenesis have been made. Furthermore, no other publications have looked at the activity of the kirrel1B variant, in any context.

1.14 The kirrel cell adhesion molecules

The vertebrate kirrel proteins were first detected in kidney tissue and subsequently the majority of research into these molecules has focused on their involvement in the formation of structures in the brain and kidney (Gerke *et al.,* 2005; Sohn *et al.,* 2010; Durcan *et al.,* 2013; Yesildag *et al.,* 2015). The mammalian IRM molecules are known to be essential for the formation of the slit diaphragm in the kidney (making up ultrafiltration barriers) as well as the directing of neuronal migration, axonal fasciculation and the formation of synapses in the brain (Gerke *et al.,* 2003; Liu *et al.,* 2003; Nishida *et al.,* 2011; Martin *et al.,* 2015). However, little research has focused on these CAMs and vertebrate myoblast fusion with the few publications that have been released having mostly looked at kirrel3.

The earliest research regarding the IRM molecules and vertebrate myoblast fusion was carried out by Sohn *et al.,* (2009) who showed that nephrin, the vertebrate Sns paralog, is expressed during both embryonic SkM development and adult muscle tissue repair (in zebrafish and mice). Helmstädter *et al.,* (2012) showed that the mammalian kirrel1 protein is able to compensate for the deletion of the *duf* and *rst* genes in *Drosophila*; kirrel2 and kirrel3 were able to compensate for this loss. This both demonstrates the evolutionary conservation of these molecules between species as well as the functional diversity of mammalian kirrel proteins. Durcan *et al.,* (2013) showed that kirrel1A (the canonical kirrel1 variant) and its truncated splice variant, kirrel1B, are differentially expressed during the regeneration of mouse SkM tissue (Durcan *et al.,* 2013). Kirrel1A expression was seen to peak in conjunction with increased myogenin activity and myoblast fusion events, suggesting a role for kirrel1A in the production of myotubes (Durcan *et al.,* 2013). Conversely, kirrel1B expression was seen to decrease with the onset of muscle regeneration, although it is unclear whether this reduction is required for myogenesis or if kirrel1B is involved earlier on in the process (Durcan *et al.,* 2013). Tamir-Livne *et al.,* (2017) directly demonstrated the requirement for kirrel3 during mouse myoblast fusion *in vitro* with the knockdown of kirrel3 (by siRNAs) inhibiting C2 myoblast alignment and elongation and ultimately preventing fusion. Furthermore, Tamir-Livne *et al.,* (2017) showed that kirrel3 is transiently expressed at the tips of myocytes during early myoblast differentiation and that its expression is dependent on MyoD activity. This requirement for kirrel3 during vertebrate fusion was demonstrated by Srinivas *et al.,* (2007) who showed that the protein is required for the fusion of myoblasts during the production of fast-twitch muscles in zebrafish. The relatedness between the mammalian kirrel variants and their *Drosophila* paralogs can be seen in figure 1.10 below.

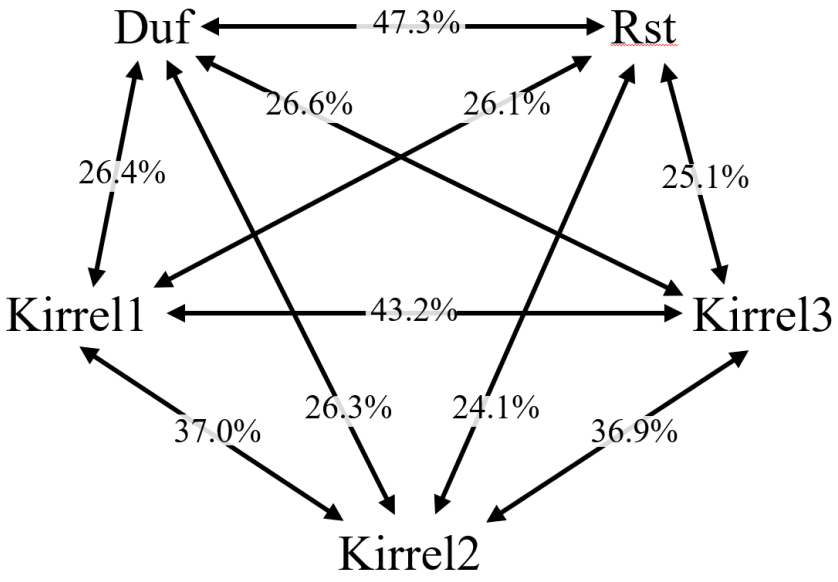


Figure 1.10 A similarity network showing the relatedness between the mammalian kirrel1, 2 and 3 proteins as well as the *Drosophila* Duf and Rst molecules. The above similarity network was generated with the T-Coffee multiple sequence alignment algorithm. The percentages displayed on the figure represent the relatedness between the full-length variants of the IRM molecules. Adapted from Helmstädter *et al.*, 2012.

1.14.1 Kirrel1A and kirrel1B protein structures

To date, two kirrel1 transcripts have been detected in mammalian SkM, these being kirrel1A and kirrel1B (Durcan *et al.*, 2013). The *kirrel1A* (NM_001170982.2) and *kirrel1B* (BC023765.1) mRNAs sequences contain 16 and 14 exons respectively with *kirrel1A* having two unique exons at its 3' end (figure 1.11) (Durcan *et al.*, 2013). The kirrel1A (NP_001164456.1) variant represents the full-length protein while kirrel1B (AAH23765) is a truncated isoform (Durcan *et al.*, 2013). Both proteins share the same ATG translation start codon which is located within exon 2. The predicted molecular weights of kirrel1A and kirrel1B are 87 and 70 kDa respectively with kirrel1A containing 789 aa and kirrel1B 634 aa (Durcan *et al.*, 2013). The first 605 aa of both kirrel1 isoforms are identical and encodes for a signal peptide (aa1-47), five extracellular Ig domains (aa54-151, aa151-243, aa256-339, aa340-422 and aa424-509), a transmembrane domain (aa529-551) as well as the first 53 aa of the cytoplasmic domain (figure 1.11) (Durcan *et al.*, 2013). However, the cytoplasmic domains of kirrel1A and kirrel1B differ significantly. The kirrel1A cytoplasmic tail contains a growth factor receptor bound 2 (Grb2) binding motif as well as a post synaptic density protein 95, *Drosophila* disc large tumour suppressor, zonula occludens 1 (PDZ) domain-binding motif at the c-terminus of the protein (aa787-789) (Sellin *et al.*, 2003; Harita *et al.*, 2008; Durcan *et al.*, 2013). The truncation of the kirrel1B cytoplasmic domain leads to the loss of both the Grb2 and PDZ binding motifs while a putative phosphatidylinositol-3,4,5-triphosphate pleckstrin homology (PIP3 PH) motif is predicted between aa607 and aa621 (Durcan *et al.*, 2013).

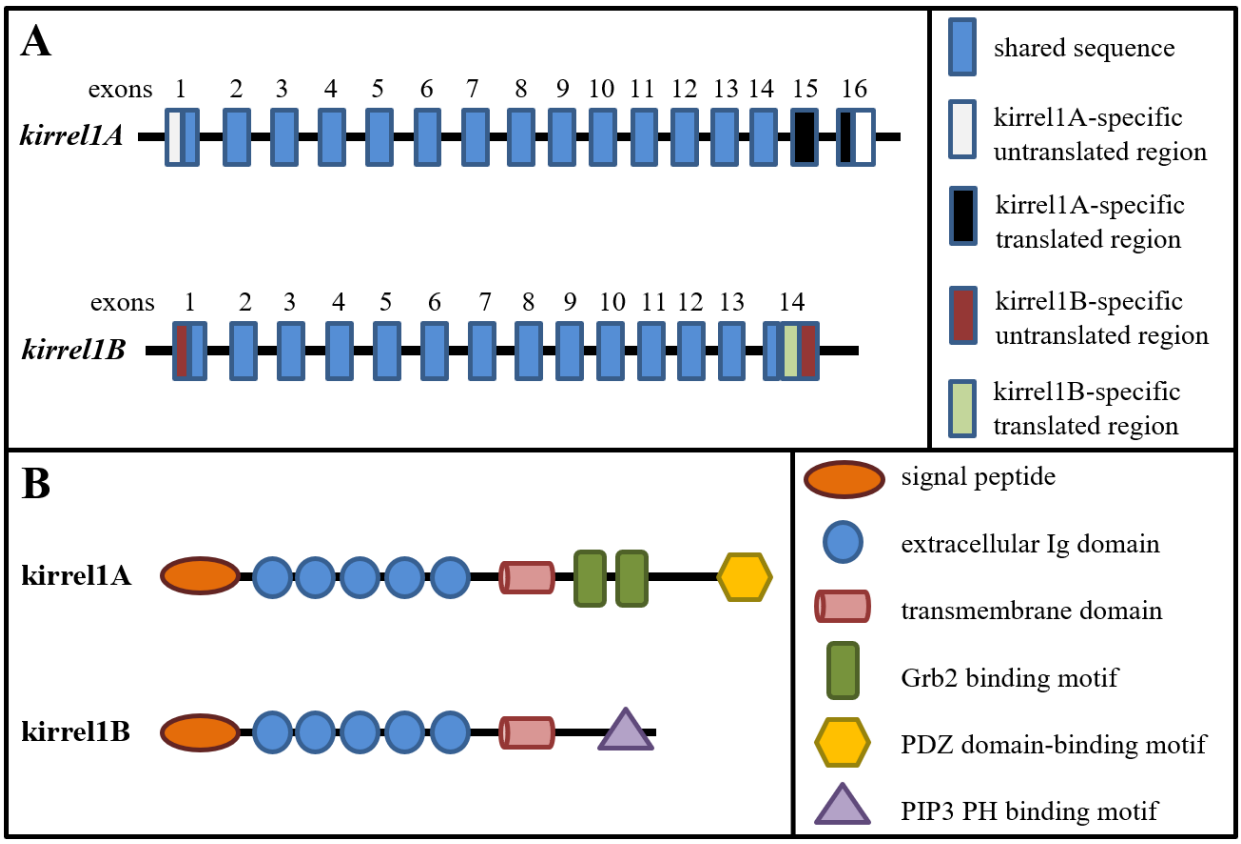


Figure 1.11 Schematic representation of the genetic and protein structures of kirrel1A and kirrel1B. A) The *kirrel1A* and *kirrel1B* genes contain 16 and 14 exons respectively. The first 13 exons of the two mRNA transcripts are identical; except for the unique untranslated regions at the 5' end of both *kirrel1A* and *kirrel1B*. The genes differ significantly after exon 13. B) The kirrel1A and kirrel1B proteins share the same first 605aa. The cytoplasmic domain of kirrel1A contains a Grb2 and PDZ domain-binding motifs while the intracellular tail of kirrel1B contains a PIP3 PH binding motif. Constructed using references: Sellin *et al.*, 2003; Harita *et al.*, 2008 and Durcan *et al.*, 2013.

1.14.2 Mechanisms of kirrel1 signaling

The mechanisms of kirrel1 signalling have not been fully explored, particularly with regards to myogenesis and its non-canonical splice variant, kirrel1B (Helmstädter *et al.*, 2012; Durcan *et al.*, 2013). The vertebrate kirrel1A protein has been shown to bind in *trans* to other kirrel1A molecules, kirrel2 as well to its preferred binding partner, nephrin, with the later interaction having been shown to induce actin nucleation at sites of podocyte cell-cell adhesion (Gerke *et al.*, 2005; Garg *et al.*, 2007; Yesildag *et al.*, 2015); an apparently similar process to the IRM-driven actin reorganization seen in *Drosophila* myoblasts prior to fusion (Kim *et al.*, 2015; Chal and Pourquié, 2017). The kirrel1A molecule initiates outside-in signalling following the in *trans* binding of nephrin to Ig domains within the kirrel1A extracellular region (Sohn *et al.*, 2009; Durcan *et al.*, 2013). This binding results in the phosphorylation of tyrosine residues (Tyr637 and Tyr638) within the cytoplasmic region of the protein by the Fyn kinase, leading to the recruitment of the SH2-SH3 domain-containing Grb2 protein (Drk in *Drosophila*) and resulting in actin nucleation (shown in podocytes) (Durcan *et al.*, 2013; Kawachi and Fukusumi, 2020). The nucleation factors that drive actin polymerization in *Drosophila* myoblasts (as previously covered) appear to be largely conserved in vertebrate systems e.g.,

Arp2/3, Scar/WAVE, WASP and D-WIP/WIP (Kim *et al.*, 2007; Durcan *et al.*, 2013; Önel *et al*., 2014; Kim and Chen, 2019). However, the roles of many of these factors during the process have not been directly determined. In the kidney, the PDZ domain-binding motif at the 3' end of the kirrel1A protein binds to the zonula ocludens-1 (ZO-1) scaffold protein, linking the membrane associated kirrel1A molecule to the cell zcytoskeleton; critical for the stability of the slit diaphragm (Durcan *et al.*, 2013; Kawachi and Fukusumi, 2020). Since the kirrel1B molecule is missing both the Grb2-binding and PDZ domain-binding regions, but maintains the same extracellular region as kirrel1A, it is thought that kirrel1B may act as a competitive inhibitor of kirrel1A and prevent actin polymerization (Durcan *et al.*, 2013).

1.14.3 Myomaker, Myomixer and the kirrel molecules

To date, no research appears to have assessed kirrel expression/signalling in relation to the activity of any known membrane fusion proteins e.g., Myomaker and Myomixer in myoblasts. Myomaker and Myomixer are the only known SkM-specific fusion proteins and have been shown to be essential for fusion in mammalian myoblasts (Singh and Dilworth, 2013; Zammit, 2017). For this reason, these two proteins have been the focus of a number of publications with regards to myogenesis. While the IRM molecules appear to be essential for the early to mid-stages of cell fusion in vertebrates (Shilagardi *et al.*, 2013; Lee and Chen, 2019), Myomaker and Myomixer are involved during the later stages of fusion and directly regulate lipid/membrane mixing (hemifusion) during pore formation (Singh and Dilworth, 2013; Zammit, 2017, Sampath *et al.*, 2018). In *Drosophila*, hemifusion starts at the tips of the actin-driven, invasive finger-like protrusions that develop in FCMs (see figure 1.6) (Kim *et al*., 2015; Chal and Pourquié, 2017). The mechanical force applied by these actin fingers increase cell-cell closeness (with the hand-in glove conformation) and allows fusogens to efficiently initiate hemifusion at the ends of the protrusions (Kim and Chen, 2019; Lee and Chen, 2019). The polymerization of actin in these protrusions is dependent on the build-up and activity of Duf/Rst and Sns (the *Drosophila* paralogs of the kirrel proteins and nephrin respectively) (Kim and Chen, 2019; Lee and Chen, 2019). The relationship between the actin regulating IRM molecules and fusion proteins has also been demonstrated by Shilargi *et al* (2013), who by modifying the expression of Duf, Rst, Sns as well as a hemifusion protein (Eff-1) in *Drosophila* myoblasts *in vitro*, highlighted their synergistic activities. Conversely, in vertebrates, these invasive protrusions are not well characterized (only one paper was found referencing these structures, Randrianarison-Huetz *et al.*, 2018) with any relationships between kirrel activity and that of hemifusion proteins such as Myomaker and Myomixer being unclear. However, with reference to findings in the fly, it is possible that kirrel-initiated structural changes in cell morphology and increased partner closeness may promote the activity of Myomaker and Myomixer. A list of the *Drosophila* molecules known to be required for healthy myoblast fusion, along with their vertebrate paralogs, are displayed in table 1.1.

Table 1.1 List of established molecules required for *Drosophila* myoblast fusion and their vertebrate counterparts

***Drosophila* fusion regulator**	**Vertebrate equivalents/ paralogs**	**Protein type**	**Proposed function during fusion (model tested) d=*Drosophila*, m=mouse, z=zebrafish**	**Localization (Drosophila)**	**References**
Sns	Nephrin	Ig domain-containing CAM	myoblast adhesion (d, m, z); actin regulation at synapse (d)	FCM	Srinivas *et al*., 2007; Shilagardi *et al*., 2013; Tamir-Livne *et al*., 2017
Hbs	Nephrin	Ig domain-containing CAM	myoblast adhesion (d, m, z); actin regulation at synapse (d)	FCM	Srinivas *et al*., 2007; Shilagardi *et al*., 2013; Tamir-Livne *et al*., 2017
Duf	kirrel family	Ig domain-containing CAM	myoblast adhesion (d, z); myoblast attraction (d, z); actin regulation at synapse (d, m (kirrel3), z (kirrel3l)	FC	Srinivas *et al*., 2007; Shilagardi *et al*., 2013; Tamir-Livne *et al*., 2017
Rst	kirrel family	Ig domain-containing CAM	myoblast adhesion (d, z); myoblast attraction (d, z); actin regulation at synapse (d, m (kirrel3), z (kirrel3l)	FC, FCM	Srinivas *et al*., 2007; Shilagardi *et al*., 2013; Tamir-Livne *et al*., 2017
Rols	TANC1	adapter protein	scaffold/adapter protein (d, m); Irre molecule localization and trafficking (d)	FC	Menon and Chia, 2001; Kim *et al*., 2015
sing	caveolin-3 (equivalent protein)	multipass transmembrane protein	vesicle trafficking and fusion (d, m)	ND	Estrada *et al*., 2007; Önel *et al*., 2014; Brunetti *et al*., 2015
Dck (dreadlocks)	Nck	SH2 and SH3 domain-containing adapter protein	links Irre molecules to actin regulating proteins (d); localization of actin regulating proteins (d); promotion of Scar complex activity (d)	ND	Kaipa *et al*., 2013; Kim and Chen, 2019
Crk	Crk/Crk-like	SH2 and SH3 domain-containing adapter protein	links Irre molecules to actin regulating proteins (d, z); localization of actin regulating proteins (d)	ND	Moore *et al*., 2007; Kaipa *et al*., 2013
Drk	Grb2	SH2 and SH3 domain-containing adapter protein	binds to Irre molecules (d)	ND	Kaipa *et al*., 2013; Kim and Chen, 2019
Mbc	Dock1, Dock5 (Dedicator of Cytokinesis)	bipartite Rac GEF	Rac activation and regulation (d, m, z)	FCM	Rochlin *et al*., 2010; Haralalka *et al*., 2011
Elmo	Elmo1	bipartite Rac GEF	Rac activation and regulation (d, m, z)	FCM	Rochlin *et al*., 2010; Haralalka *et al*., 2011
Loner	Brag2	Arf GEF	activates Arf proteins (d, m)	FCM	Pajcini *et al*., 2008; Önel *et al*., 2014
Arf1, Arf6	Arf1, Ar6	small GTPase	localization of Rac (d, m)	ND	Önel *et al*., 2014
Rac1, Rac2	Rac1, Rac2, Rac3, RhoG	Rho family GTPase	activates Scar complex and DPak proteins (d, m, z)	FCM	Rochlin *et al*., 2010; Önel *et al*., 2014
Kette	Nap1	component of Scar complex	regulation of Scar complex activity (d, m)	FC, FCM	Richardson *et al*., 2007; Abmayr and Pavlath, 2012
Scar/WAVE	WAVE	actin NPF and component of Scar complex	activation of the Arp2/3 complex (d)	FC, FCM	Richardson *et al*., 2007; Abmayr and Pavlath, 2012
Blow	none	PH domain-containing protein	regulation of WASP activity (d)	FCM	Sens *et al*., 2010; Jin *et al*., 2011
Sltr/D-WIP	WIP	WASP-interacting protein	localization of WASP (d); activation of WASP (d)	FCM	Kim *et al*., 2007; Jin *et al*., 2011
WASP	WASP	actin NPF	activation of the Arp2/3 complex (d)	FCM	Kim *et al*., 2007; Deng *et al*., 2017
Arp2/3 complex	Arp2/3 complex	actin nucleator	direct actin polymerization during fusion (d) formation of invasive PLS (d)	FC, FCM	Bothe *et al*., 2014; Deng *et al*., 2017
PIP2	PIP2	phospholipid	localization of actin regulators (d, m)	FC, FCM	Bothe *et al*., 2014; Deng *et al*., 2017
Rho1	RhoA	small GTPase	activates Rok (d)	FC	Kim and Chen, 2019; Lee and Chen, 2019
Rok	ROCK1, ROCK2	kinase	activation of MyoII (d)	FC	Kim *et al*., 2015; Lee and Chen, 2019
non-muscle MyoII	non-muscle MyoII	hexameric actin-binding protein	mechanosensor (d, m); increase cortical tension at fusion synapse (d, m)	FC	Duan *et al*., 2018; Shilagardi *et al*., 2013; Kim *et al*., 2015
DPak1, DPak3	Pak1, Pak3	serine/threonine kinase	organization of actin filaments in PLS (d)	FCM	Kim *et al*., 2015; Lee and Chen, 2019
D-Titin	Titin	filamentous protein	structural component of myoblast cytoskeleton during fusion (d)	FC, FCM	Menon and Chia, 2001; Galletta *et al*., 2004

1.15 Problem statement and aims

In light of the current research, it is clear that the mechanisms regulating vertebrate myoblast fusion are not fully understood with the kirrel proteins likely being critical for the process (Durcan *et al.,* 2013; Tamir-Livne *et al.,* 2017). Since the fusion of *Drosophila* myoblasts is highly dependent on the activity of these IRM molecules, it is likely that the same is true in vertebrates but further work looking at these paralogs is required to confirm conservation of function. Furthermore, research into the IRM molecules and vertebrate myogenesis has typically focused on the kirrel3 variant; the function of the kirrel1 variant remains relatively unexplored. However, it has been shown by Durcan *et al.,* (2013) that kirrel1A and kirrel1B are differentially expressed during myogenesis, although the functions of these proteins are not fully known.

For these reasons, the aim of this study was to determine the effects of kirrel1A and kirrel1B expression on the differentiation and fusion of murine skeletal myoblasts *in vitro*. In order to modify and assess the effects of kirrel1 expression on mammalian myoblast fusion *in vitro*, three strategies were implemented. The first strategy involved the CRISPR/Cas9 modification of the kirrel1 gene in an attempt to modify both kirrel1A and kirrel1B expression. The second strategy involved the individual knockdown of either kirrel1A or kirrel1B with shRNAs to better understand the unique roles of the two proteins. The final strategy involved the overexpression of kirrel1A, kirrel1B as well as a kirrel1 mutant. By looking at myotube formation and the expression of the MRFs during differentiation in the modified myoblasts, the functional activities of kirrel1A and kirrel1B can more clearly be defined.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Cell culture

Mammalian cell culture was carried out under sterile conditions in a class II type B2 biohazard safety cabinet (VividAir 2004 BioHazard Cabinet Class II B2). Cultures were incubated in an ESCO CellCulture® CO_2 incubator (CCL-170B-8) at 37°C and a 5% CO_2. All myoblast cell lines used to gather data during experimentation were between passages 13 and 19.

2.1.1 Culture media: growth medium

C2C12 (ATCC®, cat. CRL-1772™) and Phoenix-Ampho (ATCC®, cat. CRL-3213™) cells were maintained in growth medium (GM) consisting of Dulbecco's Modified Eagle's Medium (DMEM) (Gibco™, cat. 11960-044) supplemented with 10% (v/v) fetal bovine serum (FBS) (Gibco™, cat. 10493-106), 2 mM L-Glutamine (Gibco™, cat. 25030-081) and 2% (v/v) penicillin/streptomycin (Gibco™, cat. 15140-122).

2.1.2 Culture media: differentiation

C2C12s were differentiated in differentiation medium (DM) consisting of Dulbecco's Modified Eagle's Medium (DMEM) (Gibco™, cat. 11960-044) supplemented with 2% (v/v) horse serum (Gibco™, 26050088), 2 mM L-Glutamine (Gibco™, cat. 25030-081) and 2% (v/v) penicillin/streptomycin (Gibco™, cat. 15140-122).

2.1.3 Passaging, freezing and thawing

C212 and Phoenix-Ampho cells cultured to produce stocks or to be used in experiments were allowed to proliferate to a confluency of between 70-80% in T75 flasks (NEST Biotech, cat. 708001) before being passaged with 0.25% Trypsin-EDTA (Gibco™, cat. 25200-072). In order to passage the cultures, cell mono-layers were washed with phosphate buffered saline (PBS) (Sigma-Aldrich, cat. 79382) and lifted with trypsin. The resulting cell suspensions were then centrifuged to pellet. The pelleted cells were then resuspended in GM and cell counts performed. Cells to be frozen were centrifuged again and re-suspended at 1000000 cells/ml in 37°C freezing medium (FM) consisting of DMEM containing 30% (v/v) fetal bovine serum (FBS) and 8% (v/v) dimethylsulfoxide (DMSO) (Sigma-Aldrich, cat. D8418). The cells in freezing medium (FM) were aliquoted into cryopreservation vials (1 ml per vial) (Coring®, cat. 430489) and placed in a -80°C freezer for 12 hours after which time were moved to liquid nitrogen storage.

2.2 Cell-line preparation: CRISPR/Cas modification, shRNA knockdown and retroviral-overexpression

The permanent modification of *kirrel1* alleles in C2C12s was achieved using CRISPR/Cas9. Knockdown of wild-type kirrel1 mRNA activity was facilitated by short hairpin RNAs (shRNAs) while stable, kirrel1-overexpressing C2C12 cell-lines were produced using retroviruses.

2.2.1 CRISPR/Cas9 modification:

2.2.1.1 Guide sequence design and plasmid production

The CRISPR guide sequences used in this study were designed and produced in collaboration with Dr Isabelle Richard, the CNRS research director at Genethon in Evry, France. The first guide sequence (5'-TAAACGTGTAGGGTGTAGGC-3') targeted a region within intron 2 of the *kirrel1* gene while the second sequence (5'-CCTCCGGATAGATAAATGTC-3') targeted a region outside of the gene and downstream of exon 1; these guide sequences were ligated into the pU6-BFP (addgene; cat. 64323) and pU6-GFP plasmids respectively. Prior to ligation, the two plasmids were digested with BbsI (NEB, cat. R0539S), run on a 1% (m/v) agarose gel (Sigma-Aldrich, cat. 1168020250) containing 1/15000 SYBR Safe® (Invitrogen, cat. S33102) with the amplicons then being purified using a NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, cat. 740609.50). Following the gel clean-up, the digested plasmids were de-phosphorylated using the Quick CIP enzyme (NEB, cat.M0525S) in cutsmart buffer (NEB, cat. B6004). The complementary oligos (forward and reverse strands) used to produce the double-stranded CRISPR guide sequences were annealed in TE buffer (10 mM Tris, 0.1 mM EDTA; pH 9.0) containing 50 mM NaCl and was heated 100°C for 3 minutes and allowed to cool to room temperature for 20 minutes. The resulting double-stranded guide sequences were then phosphorylated using a T4 polynucleotide kinase (NEB, cat. M0201S) in cutsmart buffer and subsequently ligated into the digested plasmids by a T4 ligase (NEB, cat.M0202S). The plasmids were transformed into Stbl3 *E. coli* (Thermofisher Scientific, cat. A10469) made competent by the Mix & Go *E. coli* transformation kit (Zymo Research, cat. T3002). The transformed *E. coli* were plated onto nutrient agar (Sigma-Aldrich, cat. 70122 and A9799) containing 100 µg/ml ampicillin (Sigma-Aldrich, cat. A9518-5G) and incubated at 37°C for 16 hours. Single colonies were selected from these agar plates and individually expanded in 50 ml nutrient broth (Sigma-Aldrich, cat. 70122) containing 100 µg/ml ampicillin at 37°C for 24 hours on a shaking incubator. The plasmids were then purified using the Zymopure plasmid miniprep kit (Zymo Research, cat. D4210) and were sequenced at the Central Analytical Facilities' (CAF) sequencing unit at Stellenbosch University to determine ligation of the CRISPR guide sequences. A single forward primer was used for this sequencing (5'GAGGGCCTATTTCCCATGATT-3').

2.2.1.2 CRISPR plasmid transfection in C2C12s and cell sorting

C2C12s were plated into 6-well plates at 10000 cells/cm^2 and cultured until roughly 40% confluence. The two CRISPR plasmids (1.5 µg of each) were then transfected into the C2C12s using lipofectamine3000 (Thermofisher Scientific, cat. L3000001) over a 48-hour period. After 48 hours, the transfected C2C12s underwent fluorescence-activated cell sorting (FACS) in a BD Biosciences FACSMelody™ to isolate individual BFP and GFP positive cells. These dual-positive cells were individually sorted into the wells of a 96-well plate and were expanded to produce monoclonal stocks.

2.2.1.3 Genomic kirrel1 PCRs

To confirm genetic modification, the CRISPR clonal lines were cultured in 6-well plates to roughly 80% confluence and genomic DNA isolated from the myoblasts using a Macherey-Nagel Nucleospin® Mini kit for DNA - Cells and tissues (Macherey-Nagel, cat. 740952.50). DNA PCRs were carried out on 25 ng of the isolated genomic DNA samples using the Roche FastStart PCR Mastermix (Roche, cat. 04710444001) and the sequencing primers covered in figure 3.1A. The PCR thermal cycling conditions can be seen in appendix VIII, table I. These PCR reactions were then run on 1% agarose gels containing 1/15000 SYBR Safe® and using the peqGOLD 100 bp DNA Ladder Plus (VWR Peqlab, cat. 25-2020). The amplicons seen on the gels were cut out and purified using the NucleoSpin® Gel and PCR Clean-up kit. The purified amplicons were sequenced at the CAF sequencing unit using the same primers used in the PCRs.

2.2.2 shRNA knockdown:

2.2.2.1 Kirrel1 knockdown plasmid design and production

In order to individually knockdown *kirrel1A* and *kirrel1B* activity, 3 silencing plasmids were designed per mRNA variant; each plasmid producing a unique short interfering hairpin RNA (shRNA) sequence. The pSilencer™ 2.1-U6 hygro plasmid (Ambion, cat. AM5760) provided the backbone for the custom silencing plasmids. The shRNA oligos were purchased from Integrated DNA Technologies (IDT) and designed with sticky overhands to allow ligation between the HindIII and BamH1 digestion sites within the pSilencer plasmid (oligo sequences in appendix IV). The complementary strands making up the double-stranded shRNA sequences were annealed in TE buffer (10 mM Tris, 0.1 mM EDTA; pH 9.0) containing 50 mM NaCl and heated 100°C for 3 minutes and allowed to cool to room temperature for 20 minutes. The annealed, double-stranded oligos were then phosphorylated using a T4 polynucleotide kinase (NEB, cat. M0201S) in cutsmart buffer. The pSilencer plasmids were digested with the BamHI-high fidelity (NEB, cat. R3136S) and HindIII-high fidelity (NEB, cat. R3104S) enzymes in cutsmart buffer, run on a 1% agarose gel containing 1/15000 SYBR Safe® with the digested plasmid band being purified using a NucleoSpin®

Gel and PCR Clean-up kit. The isolated plasmids were de-phosphorylated using the Quick CIP enzyme in cutsmart buffer. The annealed shRNA sequences were ligated into the digested pSilencer plasmids using a T4 ligase. The ligated pSilencer plasmids were transfected into Stbl3 *E. coli*, expanded and isolated as described in section 2.2.1.1. The successful ligation of the pSilencer plasmids was confirmed via sequencing; plasmid-specific forward primer – 5'-GTTTTCCCAGTCACGACGTT-3', plasmid-specific reverse primer – 5' CTTTATGCTTCCGGCTCGTA-3'.

2.2.2.2 shRNA plasmid transfection into C2C12s and hygromycin selection

The *kirrel1A* and *kirrel1B*-silencing plasmid cocktails (3 plasmids per *kirrel1* variant, 1 µg of each plasmid), as well as an empty pSilencer control plasmid, were separately transfected into C2C12s using lipofectamine3000 as described in 2.2.1.2. After 48 hours of transfection, the resulting C2C12s were selected with 400 µg/ml hygromycin B (Invitrogen, cat.10687010) for 72 hours. Selection of the myoblasts with hygromycin at 400 µg/ml was continued until 100% cell death had occurred in a flask containing negative control myoblasts (roughly 72-96 hours). The surviving C2C12s were then expanded in GM containing half-strength hygromycin (200 µg/ml) to produce the 3 polyclonal cell line stocks i.e., pSilencer control, *kirrel1A* knockdown and *kirrel1B* knockdown. After thawing, and prior to plating out for experimentation, the modified cell lines were cultured in hygromycin at 200 µg/ml for a single passage. No hygromycin was used during experimentation.

2.2.3 Retroviral-overexpression:

2.2.3.1 Kirrel1 gene cloning and plasmid preparation

C2C12s were cultured to roughly 80% confluence in a T75 flask and RNA isolated from the cells using TRIzol reagent (Thermofisher Scientific, cat. 15596026) following to the manufacturer's protocol. cDNA was produced from the harvested RNA using the Superscript IV™ First-Strand Synthesis System (Invitrogen, cat. 18091050), making use of the included random hexamer primers. Custom cloning primers spanning the *kirrel1A* and *kirrel1B* coding DNA sequences (CDS) were purchased from Integrated DNA Technologies (IDT) and designed according to sequences obtained from the National Centre for Biotechnology Information (NCBI) database; the sequence IDs for these being NM_001170985.2 and BC023765.1 for *kirrel1A* and *kirrel1B* respectively. The primers used for the cloning of the *kirrel1A* mRNA include a forward primer containing an EcoRI digestion site and targeting the start of the *kirrel1A* CDS (5'-TAAGCAGAATTCATGACTCTGGAGAGCCCTAGC-3") and a reverse primer with a SalI digestion site and targeting the end of the CDS (5'-TAAGCAGTCGACCTACACATGAGTCTGCATGCG-3'). The *kirrel1B* cloning primers included a forward primer with a BamHI digestion site and matching the start of the *kirrel1B* CDS (5'-TAAGCAGGATCCATGACTCTGGAGAGCCCTAGC-3') as well as a reverse primer with an EcoRI

digestion site and matching the end of the CDS (5'-TAAGCAGAATTCTTAGAGGTAACGGAAATCATTAAAAGC-3'). The primer sets above were used in conjunction with the Roche FastStart PCR Mastermix (Roche, cat. 04710444001) to amplify the *kirrel1A* and *kirrel1B* CDS sequences from the prepared C2C12 cDNA (see appendix VIII for thermal cycling conditions, table II). The *kirrel1A* and *kirrel1B* PCR reactions were run on a 1% agarose gel and the CDS amplicons isolated using a NucleoSpin® Gel and PCR Clean-up kit. The *kirrel1A* amplicon (500 ng) was digested with EcoRI (NEB, cat. R0101S) and SalI (NEB, cat. R0138S) while the *kirrel1B* sequence (500 ng) was digested with BamHI (NEB, cat. R0136S) and EcoRI, both digestions taking place in cutsmart buffer at 37°C for 15 minutes. The two digested amplicons were cleaned-up by running on a 1% agarose gel and subsequently phosphorylated as described in 2.2.1.1. The pBABE retroviral plasmid was digested with either EcoRI and SalI or BamHI and EcoRI in order to ligate the *kirrel1A* and *kirrel1B* amplicons respectively. The digested pBABE plasmids were purified on a 1% agarose gel, dephosphorylated and the inserts (*kirrel1A* and *kirrel1B*) ligated as described in section 2.2.1.1. The prepared *kirrel1* retroviral plasmids were transfected into Stbl3 *E. coli*, expanded and isolated as described in section 2.2.1.1. The isolated plasmids were sequenced using the same gene cloning primers to confirm ligation (appendix VI).

2.2.3.2 Kirrel1A-mCherry mutant retroviral plasmid

A *kirrel1A-mCherry* mutant plasmid (also using the pBABE-puro retroviral plasmid backbone) was a kind gift from assistant Professor Ehtesham Arif working in the Dr Nihalani lab (renal disease biomarkers) at the Department of Medicine, Medical University of South Carolina. The *kirrel1A* in the mCherry mutant plasmid is of rat origin (95% homology with the mouse sequence). The mCherry fluorescent protein was inserted between Val575 and Asn576 of the kirrel1A molecule.

2.2.3.3 Retrovirus production, transduction into C2C12s and puromycin selection

Phoenix-Ampho cells (ATCC®, cat. CRL-3213™) were plated at 10000 cells/cm^2 in T25 flasks (NEST Biotech, cat. 707003) and cultured in GM till roughly 40% confluence. The control pBABE, pBABE-*kirrel1A*, pBABE-*kirrel1B* and the *kirrel1A*-mCherry mutant plasmids were transfected into the Phoenix-Ampho cells using lipofectamine3000 over 48 hours. After 24 hours of transfection, C2C12s were plated at 10000 cells/cm^2 in T25 flasks. After 48 hours of transfection, the culture media was removed from the Phoenix-Ampho cells, the cells washed twice with PBS and 2 ml of fresh GM added (to start production of viruses). After 15 hours, the first batch of retrovirus containing media was harvested from the Phoenix-Ampho flasks, centrifuged at 1500 g for 5 minutes and filtered through a 0.45 µm syringe filter unit (Merck-Millipore, cat. SLHV033RS) to remove loose cells. After removing the media from the C2C12s in the T25s, the filtered conditioned medium containing the retro viruses was then added to the myoblasts. This process of harvesting conditioned medium from the Phoenix-Ampho cells was repeated again 8 hours after

the first harvest and then again after another 4 hours. After adding the final batch of viruses, the C2C12s were cultured for 12 hours. The C2C12s were then moved into T75 flasks with 8 ml GM containing 5 µg/ml puromycin (Sigma-Aldrich, cat. P8833-25MG) to select for successfully modified myoblasts. Selection of the myoblasts with puromycin was continued until 100% cell death had occurred in a flask containing negative control myoblasts (roughly 48-72 hours). The surviving C2C12s were then expanded in GM containing half-strength puromycin (2.5 µg/ml) to produce the 4 stably-overexpressing polyclonal cell line stocks i.e., pBABE control, kirrel1A overexpressing, kirrel1B overexpressing and kirrel1A-mCherry mutant overexpressing line. After thawing, and prior to plating out for experimentation, the modified cell lines were cultured in puromycin at 2.5 µg/ml for a single passage. No puromycin was used during experimentation. All surfaces, glassware and plasticware coming into contact with retroviral particles were rinsed/washed with a 1/5 dilution of bleach. 5ml of this bleach solution was added to the Phoenix-Ampho T25 flasks prior to disposal.

2.3 C2C12 differentiation

C2C12s were plated at 10000 cells/cm^2 in 6-well plates and cultured in GM. After reaching roughly 80% confluence, the GM was removed, the cells washed with PBS and 2 ml of differentiation medium (DM) added to the wells of the plates. Cell lysates and total RNA were collected on each day of differentiation i.e., days 0 (prior to addition of DM), 1, 2, 3 and 4. An Olympus CKX31 phase contrast microscope was used to take pictures of the myoblasts on each day of differentiation. The DM in the plates was given a 50% refresh on day 3 of differentiation.

2.4 Analysis of differentiation: mRNA PCRs and western blotting

2.4.1 mRNA isolation, DNase treatment and reverse transcription

Total RNA was isolated from C2C12s differentiating in 6-well plates using 1 ml of TRIzol reagent and following to the manufacturer's protocol. Prior to reverse-transcription, 2 µg of isolated RNA was treated with a recombinant DNase I (Roche, cat. 4714728001). The Superscript IV™ First-Strand Synthesis System was then used to produce cDNA from the RNA samples, making use of the included random hexamer primers. A minus reverse transcription control was also included i.e., 2 µg of starting RNA in a RT reaction with no RT enzyme. The no RT control reaction was treated with DNase I and then used in PCRs to confirm complete digestion of any contaminating genomic DNA by the DNase I.

2.4.2 Semi-quantitative mRNA PCRs

Semi-quantitative *kirrel1* PCRs were carried out on the previously prepared cDNA using the Roche FastStart PCR Mastermix and the following two primer sets; 1) *kirrel1A* forward primer – 5'-GGCAGCTGGGTATCCTACAT-3', *kirrel1A* reverse primer – 5'-TGGCCATAGTCTGAGTGCTG-3'

and 2) *kirrel1B* forward primer - 5'-GAGGAAGCTGGACATCAAGG-3', *kirrel1B* reverse primer – 5'-CTAGGCCACGTTGAAAGGAG-3'. PCRs for *gapdh* were run alongside each *kirrel1* PCR with the *gapdh* amplicons being used as a loading controls; *gapdh* forward primer – 5'-GCAGTGGCAAAGTGGAGATT-3', *gapdh* reverse primer – 5'-GAATTTGCCGTGAGTGGAGT-3'. The thermal cycling conditions used in these reactions can be seen in appendix VIII, table III. The PCR reactions were run on 1% agarose gels containing 1/15000 SYBR Safe® and the intensities of the amplicons determined using the ImageJ analysis software (Fiji package). Relative *kirrel1* levels were calculated by dividing the *kirrel1* amplicon band intensity values by those of the corresponding *gapdh* bands.

2.4.3 Protein lysate preparation

Growth medium was removed from C2C12s differentiating in 6-well plate and the cells washed twice with PBS. Following removal of all liquid from the wells, 150µl of RIPA lysis buffer (100 mM Tris-HCl, 1mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 140mM NaCl, all in distilled H_2O, pH 7.5) containing 12µl protease inhibitor cocktail (Roche, cat. 04693116001) was added to each well while the plate was placed on ice. A cell scraper (NEST Biotech, cat. 710001) was used to detach the cell monolayers. The cell suspensions (in RIPA buffer) were then moved into Eppendorf tubes and left on ice for 60 minutes and subsequently sonicated (3x5second bursts at amplitude 3, Misonix Ultrasonic Liquid sonicator, model S-2000-010) and centrifuged at 16000 g for 15 minutes at 4°C. The protein containing supernatants were collected and the protein concentrations determined by use of a BCA assay (Pierce™, cat. 23227) and following the manufacturers protocol. The protein lysates were then stored at -80°C.

2.4.4 SDS-PAGE and western blotting

To denature and linearize protein samples prior to electrophoresis, lysates (20 µg total protein) were mixed at a ratio of 3:1 (v/v) with a reducing sample treatment buffer/loading dye mix (2mM EDTA, 20 mM tris pH 6.8, 0.1 mg/ml bromophenol blue, 2% SDS, 10% (v/v) mercaptoethanol, 20% (v/v) glycerol) and then heated to 95°C for 5 minutes on a heating block. The lysates were run on polyacrylamide gels consisting of a 4% acrylamide stacking stage and a 12.5% acrylamide running stage. A Precision Plus Protein™ Kaleidoscope™ ladder (Bio-Rad, cat.1610375) was used to determine protein sizes. Gels were run at 70 V, 400 mA in a running buffer consisting of 250 mM Tris-HCl, 14.4 mg/ml glycine and 0.1% SDS at pH 8.3. The proteins in the gels were then transferred to 0.2 µm nitrocellulose membranes (Bio-rad, cat. 162-0112) using a semi-dry Transblot® Turbo™ Transfer apparatus (Bio-Rad, cat. 1704150EDU). Transfers were carried out at 25 V, 2.5 A for 20 minutes in 30 ml of transfer buffer (25 mM Tris, 190 mM glycine, 20% (v/v) methanol, 0.1% (m/v) SDS, pH 8.3). Following transfer, the nitrocellulose membranes were blocked in 5% (v/v) fat-free milk in TBS-T buffer (0.15 M NaCl, 0.05 M Tris-HCl, 0.1% Tween® 20, pH 7.6) for 1 hour. The membranes were then washed 2x5 minutes in TBS-T on an orbital shaker, placed in

50 ml falcon tubes and treated with primary antibody in TBS-T containing 1% bovine serum albumin (BSA) (Roche, cat. 10735086001) for 4 hours on a roller mixer at room temperature. See appendix IX for antibody dilutions and catalogue numbers. Following the primary antibody treatment, the blots were washed for 4x5 minutes with TBS-T and treated with HRP-linked secondary antibodies in TBS-T with 1% BSA for 1 hour on a roller mixer at room temperature. The blots were then washed for 5x5 minutes with TBS-T. Immunoreactive bands were visualized using the SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Scientific, cat.34095) and a Bio-Rad Chemidoc MP gel system (Bio-Rad, cat 170-8280).

2.4.5 Pax7 hybridoma antibody harvesting

The Pax7 antibodies used in this study were produced in our lab using a mouse hybridoma line from the Developmental Studies hybridoma bank. The hybridomas were cultured in T75 flasks containing 25 ml of Advanced RPMI 1640 Medium (Gibco, cat. 12633012), 2mM L-Glutamine and 2% (v/v) penicillin/streptomycin and maintained until a cell density of between $5X10^5$ and $1X10^6$ cells/ml. Hybridomas were then moved into multiple T75 flasks (50 million cells per flask) containing 45 ml of medium. The vents on the lids of the flasks were sealed with parafilm and the cells cultured on a shaker (gentle shaking) at 37°C for 18 days. The cell suspensions were then centrifuged at 700 g for 10 minutes, the antibody-containing supernatant collected and filtered through a 0.22 µm syringe filter (Merck-Millipore, cat. SLGV033R). Per 200 ml of harvested supernatant, 6 ml of 1M Tris-HCl pH 7.5 was added to stabilize the pH. The antibody mix was then diluted with an equal volume of sterile glycerol to obtain the final working dilution and prevent freeze damage. Aliquots of the Pax7 antibodies were stored at -20°C.

2.5 Phase contrast image analysis

Average individual myotube areas and lengths were determined by manually outlining the tube structures on phase contrast microscopy pictures (40x magnification) using ImageJ. For each experimental repeat (n=3), three fields of view were assessed with five individual myotubes being measured per field of view. These 5 measurements were averaged, whereafter the 3 fields of view were averaged and statistical calculations were based on n=3 repeats. Total myotube coverage values were determined using custom ImageJ macros (three fields of view per experimental repeat, with statistical analysis based on the 3 repeats). The text commands making up the ImageJ macros as well as examples of analyzed pictures can be seen in appendix X.

2.6 Statistical analysis

Statistical analysis was carried out using the IBM SPSS Statistics 26 software. During differentiation analysis, statistically significant differences between control and modified myoblast values at individual

time points were determined by use of two-tailed t-tests. When more than two comparisons were being made, one-way analysis of variance (ANOVAs) along with the Tukey post-hoc test were performed. Statistical significance was taken as $p<0.05$.

CHAPTER THREE

RESULTS AND SPECIFIC DISCUSSIONS

To better understand the importance of kirrel1 protein expression during myogenesis, three separate *kirrel1* expression-modifying strategies were applied to C2C12s that were then differentiated *in vitro*. The first of these strategies involved the CRISPR/Cas9 modification of the regulatory regions of the *kirrel1* gene in the myoblasts; regions common to all *kirrel1* variants. Secondly, short interfering, hairpin RNAs (shRNAs) were used to knockdown the expression of either *kirrel1A* or *kirrel1B*, allowing for the assessment of the individual variants. The third and final approach involved the retroviral overexpression of either kirrel1A, kirrel1B or a kirrel1A-mCherry mutant. These three strategies will be covered separately in the ensuing results section. The assessment of each strategy included mRNA PCRs and western blotting to determine kirrel1A and kirrel1B levels during differentiation, myotube characterization as well as the analysis of transcription factor expression; specifically, the differentiation-driving MRFs.

3.1 CRISPR/Cas9 modification of kirrel1 gene integrity

The homozygous *kirrel1* gene produces a precursor mRNA which is known to be alternatively spliced to produce the *kirrel1A* and *kirrel1B* variants (Durcan *et al.*, 2013). To assess the role of the *kirrel1* gene in myoblast differentiation and fusion, the CRISPR/Cas9 system was used to modify the functionality of the gene.

3.1.1 The CRISPR/Cas9-mediated snipping of the homologous kirrel1 alleles in C2C12s leads to the disruption of gene integrity

Two custom CRISPR/Cas9 plasmids were designed to remove exons 1 and 2 of the *kirrel1* gene from the C2C12 genome (to render the gene inactive) with one plasmid snipping upstream of exon 2 and within intron 2 (cut site 1; GFP plasmid) and the other downstream of exon 1 on the C2C12 chromosome (cut site 2; BFP containing plasmid) (see figure 3.1A). These plasmids were designed for this project with the efficacy of the "snips" being tested for the first time during this study. Individual, dual-positive myoblasts (for BFP and GFP) were sorted via *fluorescence-activated cell sorting* (FACS) and expanded in order to produce stocks and to run DNA PCR reactions to confirm the modification of *kirrel1* genomic sequences (figure 3.1B).

Successful CRISPR/Cas9-mediated snipping was detected in roughly 1/15 of the clonal lines tested with two myoblast lines being produced that had at least one of their *kirrel1* alleles "inactivated" (by the removal of exons 1 and 2). One of these cell lines was viable under proliferating conditions but would completely die-off upon the initiation of differentiation and was not assessed further with regards to myotube

production in this study. Enhanced myoblast apoptosis is associated with and required for successful differentiation and fusion with typically between 30-35% of unmodified muscle progenitor cells undergoing this form of programmed cell death during myofiber production *in vitro* (Dee *et al.*, 2002; Schoneich *et al.*, 2014). This pattern of complete cell death seen in the excluded cell line has also been observed in modified C2C12 myoblasts expressing mutant α-actin proteins and there is evidence suggesting that the actin cytoskeleton acts as both a sensor of and mediator of apoptosis (Vandamme *et al.*, 2009; Desouza *et al.*, 2012). Although not specifically looked at in this study, it is possible that the knockout/modification of the *kirrel1* gene in this excluded, differentiation-incapable cell line resulted in a misregulation of apoptosis; potentially in an actin-dependent manner as the kirrel proteins are known to regulate actin dynamics during pre-fusion events (Durcan *et al.*, 2013; Tamir-Livne *et al.*, 2017). The second CRISPR/Cas9-modified monoclonal cell line isolated in this study did not exhibit the same die-off during differentiation and was used throughout this research project and will be referred to as the "K1CRISPR" cell line in-text. Three genomic DNA-specific primer sets (primer set 1, 2 and 3) were used to determine the knockout of both *kirrel1* alleles. The binding locations of these 3 primer sets can be seen in the figure below (figure 3.1A)

<u>Cut site 1:</u> intron 2 CRISPR/Cas9
CRISPR Guide sequence: 5'-TAAACGTGTAGGGTGTAGGC-3'
(mouse chr3:87139688-87139707)

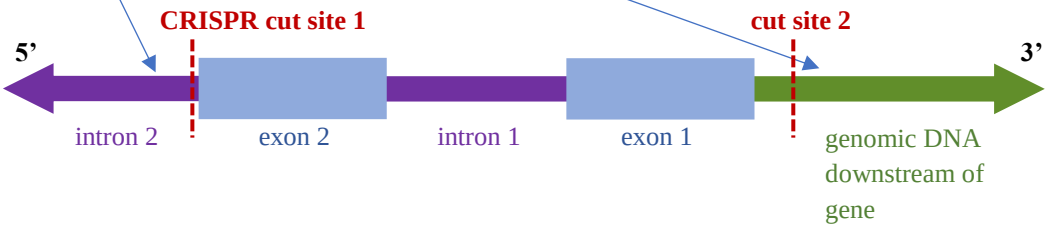
<u>Cut site 2:</u> downstream of exon 1 CRISPR/Cas9
CRISPR Guide sequence: 5'-CCTCCGGATAGATAAATGTC-3'
(chr3: 87175327-87175346)

<u>Primer set 1:</u> spanning both cut sites

Forward – 5'-CATCTCTTGCCCCTAACTGG-3' (chr3:87139401-87139420)

Reverse – 5'-CCACCCGAATGCACTTCTAT-3' (chr3: 87175475-87175494)

Wild-type amplicon: 36074bp. Double cut amplicon: 473bp

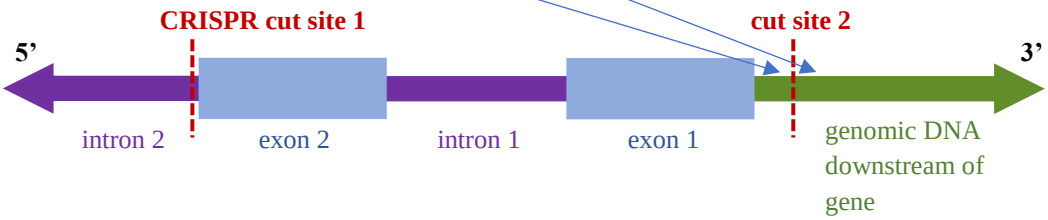


<u>Primer set 2:</u> Exon 1 cut-flanking primers for PCR and sequencing

Forward – 5'-GAAATCTTGACTGTCTTTG-3' (chr3:87175204–87175222)

Reverse – 5'-CACCTCAAGTTCCTGAGTGCTT-3' (chr3: 87175425-87175446)

Wild-type amplicon size: 224bp



<u>Primer set 3:</u> Exon 2 cut-flanking primers for PCR and sequencing

Forward – 5'-CATCTCTTGCCCCTAACTGG-3' (chr3:87139401–87139420)

Reverse – 5'-ATGACTCTGGAGAGCCCTAGC-3' (chr3:87141089-87141109)

Wild-type amplicon size: 1689bp

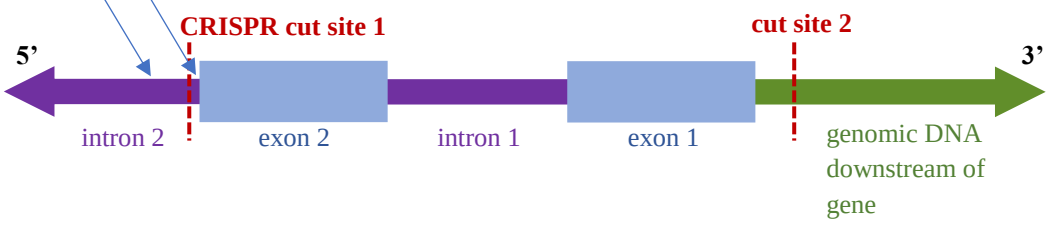


Figure 3.1A Genomic kirrel1 CRISPR/Cas9 cut sites and sequencing primers. The genomic sequence of kirrel1 was modified using CRISPR/Cas9 as described 3.1.1. The modification sites are represented by dotted red lines. Three sets of primers were designed to sequence the individually modified sites as well as the sequences between the two sites (Primer sets 1-3).

To assess whether the region of roughly 35600 bp and containing exons 1 and 2 of the *kirrel1* gene had been removed by the CRISPR/Cas9 treatment, genomic PCR reactions using primers spanning the two cut sites were carried out on C2C12 control and K1CRISPR genomic DNA (figure 3.1B, primer set 1). This deleted region contains the start codon of the *kirrel1* coding sequence (CDS) which is located within exon 2. A control C2C12 amplicon produced using this primer set was predicted to be 36074 bp long (too long to be amplified and visualized on gel) while a successfully cut allele would produce an amplicon of 473 bp (assuming no addition of random base pairs by the Cas9 enzyme at the cut site). A single K1CRISPR band was seen slightly below 500 bp when using primer set 1 (figure 3.1B, primer set 1), suggesting that the CRISPR/Cas9 treatment had removed the 35600 bp region from at least one of the homozygous *kirrel1* alleles. This amplicon was sequenced, confirming deletion of this region (appendix I).

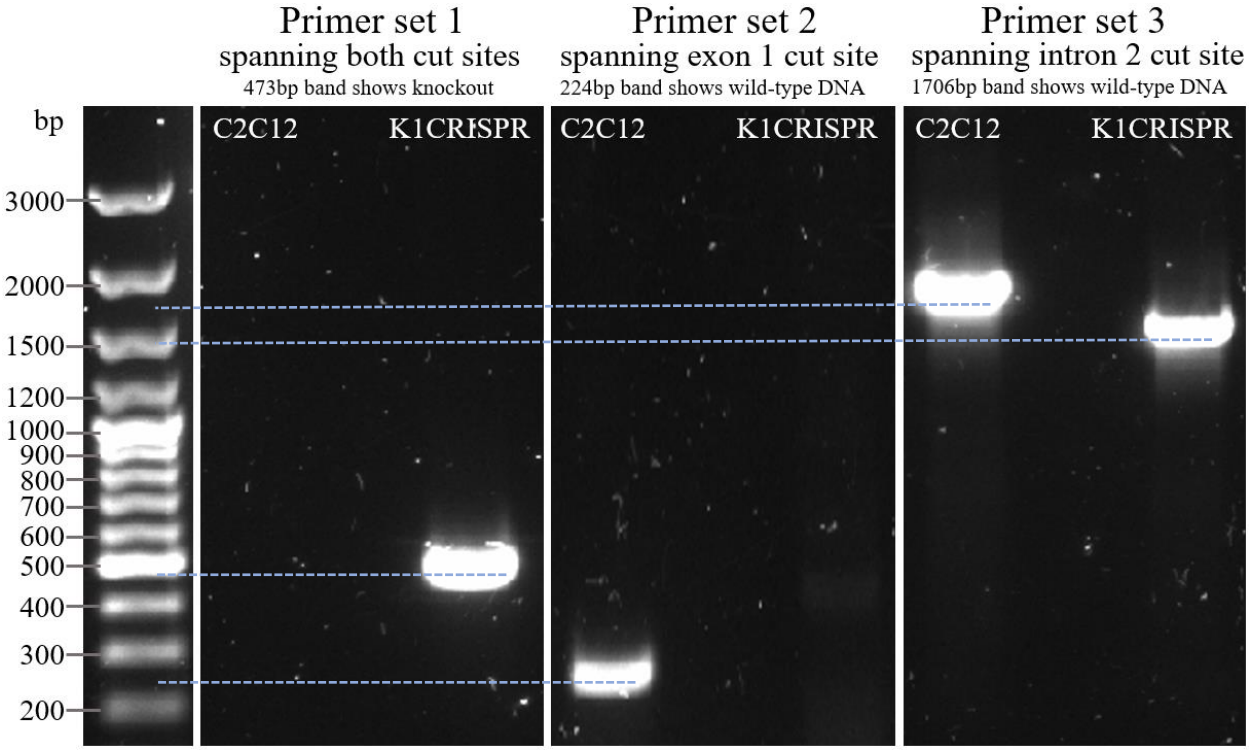


Figure 3.1B. *Kirrel1* amplicons from PCR reactions carried out on genomic DNA from control C2C12s or CRISPR/Cas9-modified C2C12s and using primers spanning snipped exons. Primer set 1 spanning both CRSIPR/Cas9 cut sites was used in genomic DNA PCRs to determine the success of the CRISPR/Cas9 mediated removal of a region of roughly 35600 bp between exons 1 and 2 of the *kirrel1* gene. An amplicon of 473 bp indicating the successful removal of this region. Primer set 2 spanning the cut site downstream of exon 1 was used to detect the presence of unmodified, wild-type *kirrel1* DNA; indicated by the presence of a 224 bp amplicon. Primer set 3 spanning the cut site within intron 2 of *kirrel1* was used to further show the presence of wild-type DNA with a predicted amplicon of 1706bp.

Primer set 1 was only able to confirm that at least one of the *kirrel1* alleles had been knocked-out, but not the knockout of both alleles. In order to confirm the knockout of both *kirrel1* alleles, additional primers (primer sets 2 and 3) were used to detect the presence of wild-type *kirrel1* DNA at the two CRISPR/Cas9 cut sites. Primers flanking the cut site downstream of exon 1 (figure 3.1B, primer set 2) of the *kirrel1* gene produced a control C2C12 PCR amplicon close to the predicted size of 224 bp and was sequenced and

successfully matched to the *kirrel1* reference sequence on the NCBI database (reference sequence AC102115.11, appendix I). As expected, no amplicon was obtained following PCRs using primer set 2 and K1CRISPR DNA, suggesting (along with primer set 1) that the targeted region was removed by the Cas9 enzyme; or that the DNA sequence close to the cut site had been modified and was not recognized by the primer set. Furthermore, a third primer set (figure 3.1B, primer set 3) flanking the CRISPR/Cas9 cut site within intron 2 was used to analyse modification/presence of this region. PCR reactions using this primer set were predicted to produce wild-type amplicons of 1706 bp which were subsequently seen when analysing the control C2C12 amplicon. This control amplicon was sequenced and successfully matched to the reference DNA sequence (appendix I). Unexpectedly, a K1CRISPR amplicon was seen at around 1500 bp (figure 3.1B, primer set 3). This indicates that one of the *kirrel1* alleles was not completely knocked-out by the Cas9 enzyme but instead any potential cut site was repaired with the smaller size of the amplicon indicating modifications to the DNA sequence in this region (figure 3.1B). Sequencing of this K1CRISPR amplicon confirmed a total size loss of roughly 270 bp. However, a sequence of around 281 bp close to the 3' end of the sequence could not be matched to any known reference sequences (appendix I). It appears that this random sequence of base pairs was added by the Cas9 enzyme prior to repair. Using this sequencing data, it was determined that the remaining K1CRISPR *kirrel1* allele was missing roughly 550 bp from its sequence when compared to the control with an additional 281 bp being added at the cut site accounting for the total size loss of around 270 bp. These lost and additional base pairs all fall within intron 2 and leaves unmodified the start of the coding sequence (CDS) of the gene, which lies within exon 2. Although poorly understood, numerous publications have highlighted the importance of intron integrity for healthy gene regulation (Jo and Choi, 2015; Shaul, 2017).

According to the UCSC Genome Browser and the ENCODE and Roadmap Epigenomic Consortia database, the CRISPR/Cas9 cut site within intron 2 intersects a short interspersed nuclear element (SINE), specifically a RSINE1 element (appendix II). Short interspaced nuclear elements (SINEs) are non-coding retrotransposons that are critical for the regulation of gene expression (Vassetzky and Kramerov, 2013; Zhang *et al.*, 2021). These transposons contribute towards chromatin folding, a complex, nuanced mechanism of gene regulation (Vassetzky and Kramerov, 2013; Zhang *et al.*, 2021). SINEs can be brought into the proximity of enhancer/promoter regions by chromatin folding (where they are thought to directly modulate gene expression) and have been shown to be involved in back-splicing (Vassetzky and Kramerov, 2013). Back-splicing produces circular RNAs (circRNAs); a poorly understood, relatively stable and generally non-coding type of single stranded RNA that regulates gene expression (via binding to miRNAs and RNA associated proteins); often modifying the expression of the genes from which the circRNAs themselves originate (Vassetzky and Kramerov, 2013; Nisar *et al.*, 2021; Zhang *et al.*, 2021). To date,

more than 50 disease phenotypes have been linked to the disruption of retrotransposon elements (Beauregard *et al.*, 2008). According to the work by Jokabi *et al*., (2018) using the Circtool software, RSINE1 elements, as found in intron 2 of *kirrel1*, are enriched in the early introns of murine genes that produce circRNAs. This raises the possibility of RSINE1 and circRNAs being involved in the regulation of *kirrel1* expression. Furthermore, after splicing, the production of proteins in the cytoplasm can be modulated by the SINE-mediated retention of mature mRNA sequences within the nucleus (Guo *et al.*, 2020). Modification of this RSINE1 element in the K1CRISPR line was confirmed by sequencing of the primer set 3 amplicon which showed missing and additional random base pairs at this site (appendix I). This RSINE1 sequence was seen to be intact in the control C2C12 DNA sequence. Additionally, this *kirrel1* RSINE1 lies downstream of and in direct contact with another SINE type termed PB1D10. Modification to the RSINE1 element in the K1CRISPR myoblasts, being adjacent to the PB1D10 sequence, could have affected functioning of this PB1D10 element; perhaps due to altered/faulty chromatin folding as a result of the added and missing base pairs.

Since an amplicon was seen for primer set 3 and not for primer set 2, it is reasonable to assume that the cut site flanked by primer set 2 in the remaining *kirrel1* allele was also cut, likely modified and repaired, resulting in the primers being unable to bind and produce an amplicon. Looking at the *kirrel1* gene on the Genome UCSC browser, it was determined that the CRISPR/Cas9 cut site downstream of exon 1 (flanked by primer set 2) is located within a 305 bp long regulatory region (ENCODE accession: EM10E0734476; appendix II). The proximity of this region to the *kirrel1* Transcription Start Site (TSS) heavily links this region to the regulation of the gene. Findings published on the ENCODE database indicate that in murine myoblasts, this regulatory region displays a proximal enhancer-like signal and is characterized by its proximity to the Transcription Start Site (TSS) (in this case within 500 bp of the TSS of the *kirrel1* gene) and its probable interactions with trimethylated form of the H3 histone (H3K4me3) as well as the CCCTC-binding factor (CTCF) (appendix II). H3K4me3 and CTCF are major components of epigenetic control, including the process of gene splicing (Plasschaert *et al.*, 2014; Kim *et al.*, 2015). Genetic fragment analysis has shown a strong association between H3K4me3 signals and SINE localization, suggesting a relationship between these regulatory elements (Ye *et al.*, 2020; Lambrot *et al.*, 2021).

In addition to the genomic *kirrel1* PCRs, sequencing PCR reactions covering the full-length of the *kirrel1A* and *kirrel1B* coding sequences (CDS) were performed. These sequencing reactions revealed that both the C2C12 control and K1CRISPR cell lines produced unmodified, mature *kirrel1A* and *kirrel1B* mRNAs that were all successfully matched to published reference sequences. This was expected as the cut downstream of exon 1 was far from the CDS region while the cut upstream of exon 2 was located within intron 2. However, with reference to the previously covered DNA PCRs (figure 3.1B), the pre-mature, unprocessed

kirrel1 mRNAs in the K1CRISPR line would have contained the modified intron 2. Upon further testing it would become apparent that the knockout of the single *kirrel1* allele coupled with the disruption of the regulatory regions within the remaining *kirrel1* allele in the K1CRISPR cell line resulted in the misregulation of *kirrel1* mRNA production and protein expression (seen section 3.1.2). Although both *kirrel1* alleles were not knocked-out in the K1CRISPR line, the resulting abnormal *kirrel1* production seen in these myoblasts provided a valuable model to evaluate the importance of kirrel1 protein regulation for successful differentiation and fusion.

3.1.2 The CRISPR/Cas9-mediated disruption of the genomic kirrel1 sequence in C2C12s resulted in aberrant kirrel1B mRNA levels

Control C2C12s and the K1CRISPR cell line were differentiated over 4 days *in vitro* with total RNA collected on days 0, 2 and 4. Random hexamer primers were used to produce cDNA from these RNA samples with the resulting cDNA then being used in PCR reactions targeting the 3' end of the *kirrel1A* (primer set K1A) and *kirrel1B* (primer set K1B) mRNAs. These PCR products were run on acrylamide gels and the relative *kirrel1* band intensities calculated using *gapdh* as the loading control (figures 3.2). Unpaired student's t-tests were used to determine statistically significant differences in *kirrel1* mRNA levels between the K1CRISPR cell line and the control C2C12 line at each individual time point. One-way ANOVA was used to determine statistically significant changes in *kirrel1* mRNA levels over time for each individual cell line.

<u>*Kirrel1A* mRNA levels</u>: The *kirrel1A* expression profiles for both cell lines seemed to follow the same general pattern with slightly elevated mRNA levels on days 2 and 4 relative to day 0 (figure 3.2.A). The control C2C12s displayed statistically significant increases on both days 2 and 4 (**0.77**±0.01 (**mean**±SEM) and **0.77**±0.03) relative to day 0 (**0.59**±0.07; $p<0.05$), increases of 29% on both days 2 and 4. *Kirrel1A* levels in the K1CRISPR cell line were increased by 64% on day 2 (**0.94**±0.11) when compared to day 0 (**0.57**±0.03; $p<0.05$). These increases in *kirrel1A* mRNA production during myogenesis (in both cell lines) have also been Durcan *et al.*, (2013) (in regenerating mouse muscle). There were no statistically significant changes in *kirrel1A* mRNA levels in the control C2C12s vs the K1CRISPR line at any of the timepoints.

<u>*Kirrel1B* mRNA levels:</u> *Kirrel1B* mRNA levels in the differentiating control C2C12s remained relatively low when compared to *kirrel1A* levels (figure 3.2.B), a finding also demonstrated by Durcan *et al.*, (2013). Furthermore, these *kirrel1B* levels remained largely unchanged between days 0 and 4. However, *kirrel1B* levels in the K1CRISPR cell line were dramatically increased when compared to the control. The K1CRISPR *kirrel1B* mRNA values displayed statistically significant increases on days 0, 2 and 4

respectively (**1.19**±0.35, **1.65**±0.16 and **1.62**±0.24) when compared to the control C2C12 values at the same timepoints (**0.23**±0.06, **0.31**±0.045 and **0.21**±0.042; $p<0.05$), increases 409%, 425% and 682%.

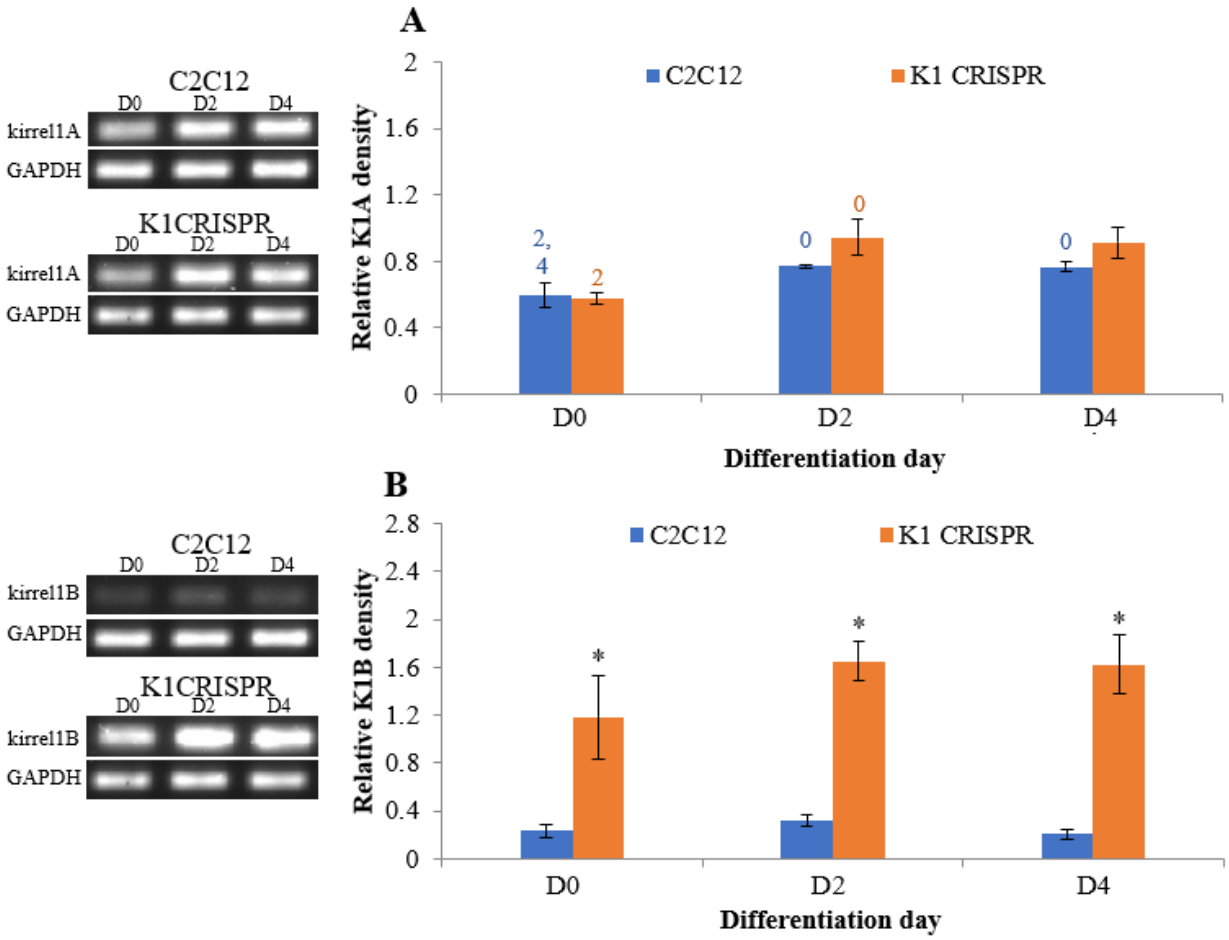


Figure 3.2 *Kirrel1* mRNA levels in differentiating C2C12s with CRISPR/Cas9-modified *kirrel1* alleles. *Kirrel1A* (A) and *kirrel1B* (B) mRNA levels in differentiating C2C12s were determined via semi-quantitative PCRs. Control C2C12s and the K1CRISPR cell line were differentiated *in vitro* over 4 days with RNA collected on days 0, 2 and 4. Densitometry analysis of agarose gels containing *kirrel1* PCR products was carried out using ImageJ and the relative *kirrel1* band intensities determined using *gapdh* as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 *kirrel1* mRNA levels at individual timepoints were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in *kirrel1* mRNA levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 2, $p<0.05$ when compared to D2; 4, $p<0.05$ when compared to D4.

The dramatic misregulation/upregulation of the *kirrel1B* variant was attributed to the CRISPR/Cas9-mediated modification of the *kirrel1* gene in the cell line. Modifications to the RSINE1 element within intron 2 (cut site 1) and/or to the proximal enhancer region downstream of exon 1 (cut site 2), may have altered gene expression; potentially via interrupted enhancer activity and/or by the misregulation of DNA folding and histone integration (nucleosome formation). Chromatin immunoprecipitation along with sequencing (ChIP-Seq) would need to be performed to assess any changes in the localization of regulatory, DNA-binding proteins (e.g., histones and transcription factors) at these modified sites. In addition to the

single modified *kirrel1* allele mentioned above, the deactivation of the other *kirrel1* allele (by removal of exons 1 and 2 by the CRISPR/Cas9) might have modified gene transcription and splicing.

3.1.3 The disruption of kirrel1 regulation by CRISRR/Cas9 reduced C2C12 myotube formation in vitro

Control C2C12s and the K1CRISPR cell line were differentiated over 4 days *in vitro* with phase-contrast images aquired on each day. Myotube analysis was carried out on these images using the imageJ analysis software. The control C2C12 line displayed typical myotube expansion with a steady increase in myotube size and total tube coverage seen after day 2 with no tubes seen on day 0 and 1 (figure 3.3). Conversely, the production of myotubes in the K1CRISPR line was severely inhibited with greatly reduced myotube coverage on days 2, 3 and 4 (**0.00**%±0.00%, **0.70**%±0.09% and **0.78**%±0.20%) when compared to the control (**5.20**%±0.55%, **10.4**%±0.72% and **26.2**%±0.91%; $p<0.05$) (figure 3.3.A).

Very few myotubes were produced by the modified C2C12s accounting for total tube coverage being the most reduced of the values featured in figure 3.3 (when compared to the control values). These sparse tubes were also much shorter and thinner than those produced by the control. Despite this severely inhibited myotube formation, the K1CRISPR myoblasts displayed an apparently healthy level of alignment and elongation.

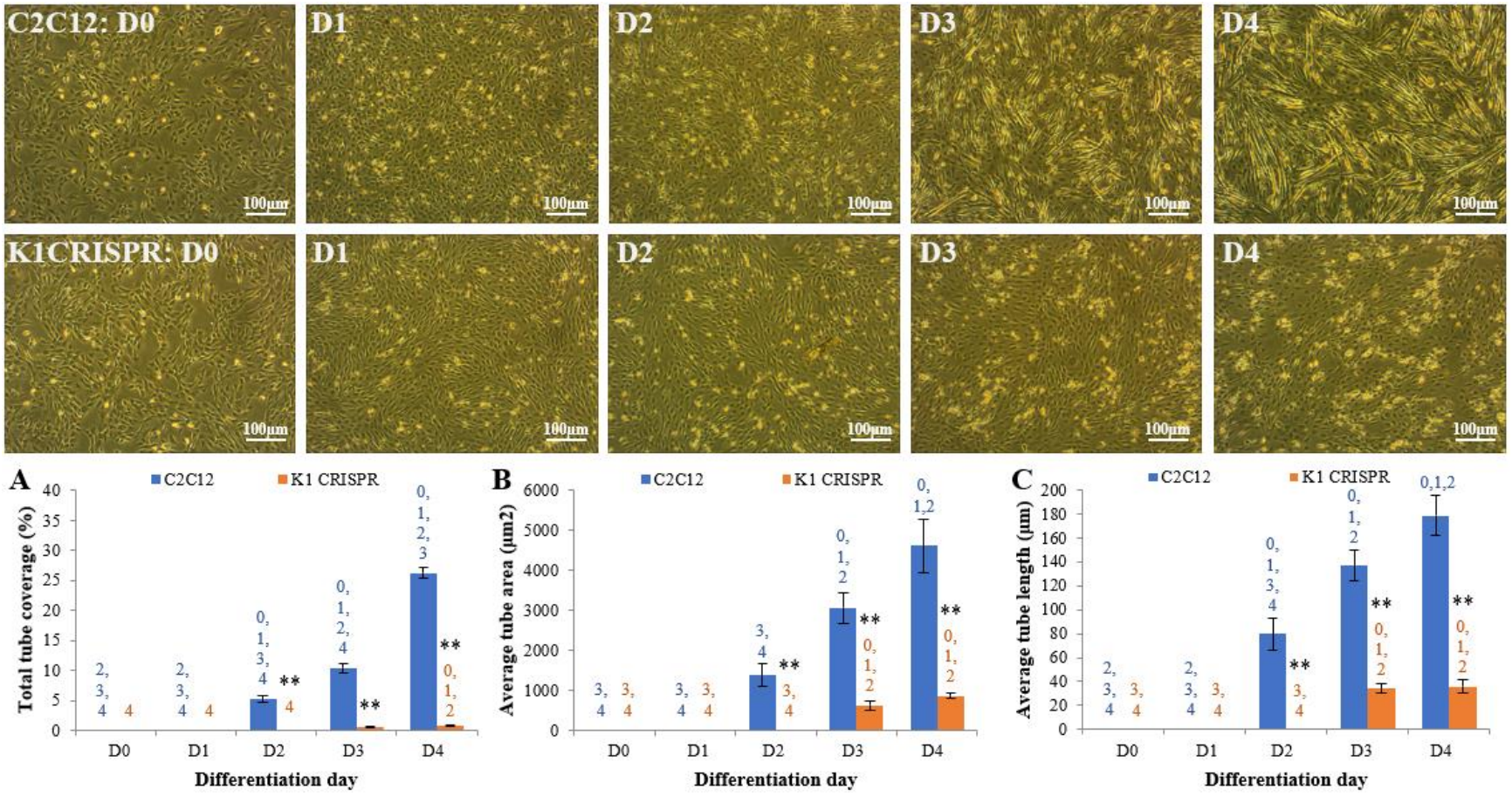


Figure 3.3 Myotube production by differentiating C2C12s with CRISPR/Cas9-modified *kirrel1* alleles. Unmodified C2C12s and the K1CRISPR cell line were differentiated *in vitro* over 4 days with images aquired daily at 40x magnification. A) Total myotube coverage was determined using an in-house ImageJ macro (see appendix X). B) Average individual tube areas determined by manually outlining the myotube structures using ImageJ and recording the area values. C) Average individual tube length determined by manually outlining the myotube structures using ImageJ and recording the Feret values. Data expressed as mean ±SEM; n=3. Statistically significant differences between the control C2C12 and K1CRISPR C2C12 myotube values at each timepoint were determined using an unpaired, 2-tailed Student's t-test. *p<0.05, **p<0.005. Statistically significant differences in myotube values between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, p<0.05 when compared to D0; 1, p<0.05 when compared to D1; 2, p<0.05 when compared to D2; 3, p<0.05 when compared to D3; 4, p<0.05 when compared to D4.

3.1.4 Immunohistochemical analysis of kirrel1 variant expression in differentiating C2C12s

The kirrel1A (NP_001164456.1) and kirrel1B (AAH23765.3) proteins consist of 789 and 634 amino acids respectively with predicted molecular weights of 87 kDa and 70 kDa (Durcan *et al.*, 2013). However, a number of publications have detected immunoreactive kirrel1 proteins between 80 and 125 kDa with the type/size of the kirrel1 isoforms seemingly dependent on both the organism type and the cell/tissue of origin (Liu *et al.*, 2003; Arif *et al.*, 2014; Durcan *et al.*, 2013). Mammalian brain, kidney and muscle tissue have been all shown to produce kirrel1 variants of different sizes (Liu *et al.*, 2003; Arif *et al.*, 2014; Durcan *et al.*, 2013; Solanki *et al.*, 2019). These size differences are in part due to tissue-specific alternative gene splicing, post-translational protein modifications (PTM) and/or the cleavage of extracellular Ig motifs on the kirrel1 molecules (Durcan *et al.*, 2013; Gerke *et al.*, 2015; Johnson *et al.*, 2017). Additionally, the members of the mammalian kirrel superfamily are able to form multimers with other kirrel proteins as well as complex with proteins such as nephrin, Grb2, PDZ, Myosin 1C (Myo1c), HGF and specific SH3-domain containing partners; therefore, the presence of these partners as well as potential differences in protein denaturing/disulphide reduction protocols during protein sample preparations may account for the different sizes of kirrel1 isoforms detected in the various publications (Arif *et al.*, 2014; Durcan *et al.*, 2013; Solanki *et al.*, 2019; Paul *et al.*, 2022).

Reducing western blots of mouse muscle tissue and C2C12 protein lysates have previously detected the kirrel1A isoform at 125 kDa while kirrel1B was seen at roughly 70 kDa (Durcan *et al.*, 2013). In the current study, the over expression of kirrel1A and kirrel1B in the K1A and K1B cell lines (also the K1CRISPR line) produced immunoreactive bands of increased intensities (relative to the controls) at 125 kDa and 84 kDa respectively. It was therefore decided to calculate kirrel1A and kirrel1B expression levels during differentiation using these band sizes. A single anti-kirrel1 antibody (Abcam ab156084) was used to detect a conserved region (aa505-534) that spans the transmembrane region and part of the cytoplasmic domain of both the kirrel1A and kirrel1B variants. One-way ANOVAs were used to determine statistically significant changes in kirrel1 protein levels between the different differentiation time points for each individual cell line. Unpaired student's t-tests were used to determine statistically significant differences in kirrel1 protein levels between control myoblasts and the modified C2C12 lines at each individual differentiation time point.

3.1.5 The CRISPR/Cas9-mediated disruption of kirrel1 expression in C2C12s moderately increased kirrel1A protein levels while kirrel1B expression increased dramatically. The CRISPR/Cas9 modifications appeared to inhibit the accumulation of a large kirrel1-containing multimer/complex that was seen to develope in differentiating myoblasts/myotubes

Control C2C12s and the K1CRISPR cell line were differentiated over 4 days *in vitro* (seen in figure 3.3) with protein lysates being collected on each day. Reducing western blots were carried out using these lysates to determine kirrel1A and kirrel1B protein levels. Non-reducing western blotting was used to determine the levels of a 350 kDa, kirrel1-containing complex that was seen to accumulate in differentiating myoblasts. Relative kirrel1 protein densities were calculated using GAPDH as a loading control.

Kirrel1A protein levels: Both the control and K1CRISPR myoblasts produced a suspected kirrel1A band of the same size (125 kDa), suggesting expression of the full-length, wild-type kirrel1A variant in the K1CRISPR myoblast line. The two cell lines expressed relatively stable levels of kirrel1A (125 kDa) between days 0 and 4 (figure 3.4). However, while not significant, kirrel1A levels in the control myoblasts appeared to increase slightly on days 2 and 3 relative to day 0 and gradually fell again to reach pre-fusion levels by day 4 (increases coinciding with the initiation of fusion); a similar pattern was also seen in the control myoblast lines in the shRNA and retroviral-overexpression sections (section 3.2 and 3.3). Durcan *et al.*, (2013) showed a similar pattern of kirrel1A expression in regenerating mouse muscle i.e., kirrel1A upregulation occurred during periods of increased myoblast fusion.

Statistically significant increases in kirrel1A protein levels were seen in the K1CRISPR line on days 0, 1, 3 and 4 (**0.94**±0.07, **0.95**±0.09, **1.05**±0.08 and **0.83**±0.09) when compared to the control values at the same time points (**0.46**±0.03, **0.67**±0.09, **0.49**±0.09 and **0.47**±0.08; $p<0.05$); corresponding to increases of 102%, 43%, 114% and 76% on days 0, 1, 3 and 4 respectively (figure 3.4).

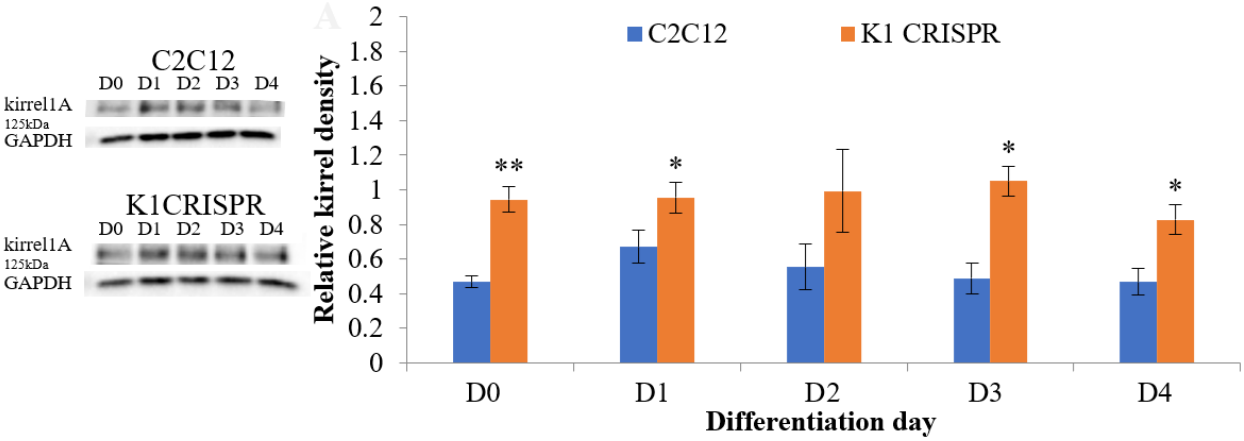


Figure 3.4 Kirrel1A (125 kDa) protein levels in differentiating C2C12s with CRISPR/Cas9-modified *kirrel1* alleles. Kirrel1 (125 kDa) levels in differentiating C2C12s were determined via western blotting under reducing conditions. Unmodified C2C12s and the K1CRISPR cell line were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative kirrel1A (125 kDa) band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 kirrel1A (125 kDa) levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in kirrel1A (125 kDa) levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

The increases in kirrel1A protein levels were not expected as there were no significant differences in *kirrel1A* mRNA levels in the K1CRISPR vs the control cell line (figure 3.2). Kirrel proteins localized to cell membranes are known to regulate the trafficking of additional kirrel molecules to the cell surface as well as regulating the rate at which kirrel molecules are moved out of the cell membrane (likely through vesicles) for recycling (Galletta *et al.*, 2004; Menon *et al.*, 2005) and since kirrel1B levels were seen to be dramatically increased in the K1CRISPR line (both mRNA and protein; figures 3.2 and 3.5), it is possible that elevated feedback from the overexpressed kirrel1B altered kirrel1A localization/trafficking and/or modified protein recycling. Additionally, since SINEs have been associated with the export of mRNAs from the nucleus and the processing of pre-mature mRNAs (Guo *et al.*, 2020), it is possible that modification of the RSINE1 element in intron 2 (present within the *kirrel1* gene and pre-mRNA) had an effect on the processing of the *kirrel1A* mRNA i.e., the rate at which pre-mature *kirrel1A* mRNA was converted to mature mRNA. The ratios of pre-mature to mature *kirrel1A* mRNA were not determined as the primers used to produce the amplicons targeted sequences common to both mRNA types i.e., the primers target regions contained within a single exon.

Kirrel1B protein levels: Kirrel1B (84 kDa) levels in the differentiating control C2C12s remained relatively low and constant during differentiation with the only statistically significant change in protein levels being seen on day 4 (figure 3.5). On day 4 (**0.17**±0.001), kirrel1B levels were seen to be decreased by 135% when compared to day 0 (**0.40**±0.01; $p<0.05$). Similarly, for the K1CRISPR line, kirrel1B protein levels between days 0 and 4 remained relatively stable (figure 3.5).

However, kirrel1B levels within the K1CRISPR myoblasts were dramatically increased on each day when compared to the control cell line. K1CRISPR kirrel1B protein levels were increased by 568%, 480%, 758%, 649% and 912% on days 0, 1, 2, 3 and 4 (**2.67**±0.27, **2.00**±0.05, **2.32**±0.29, **2.00**±0.22) respectively when compared to the control values at the same time points (**0.40**±0.01, **0.34**±0.05, **0.27**±0.05, **0.27**±0.04 and **0.17**±0.001; $p<0.05$). The increases in kirrel1B protein levels in the K1CRISPR myoblasts were congruent with the *kirrel1B* mRNA levels seen at the same timepoints (figure 3.2).

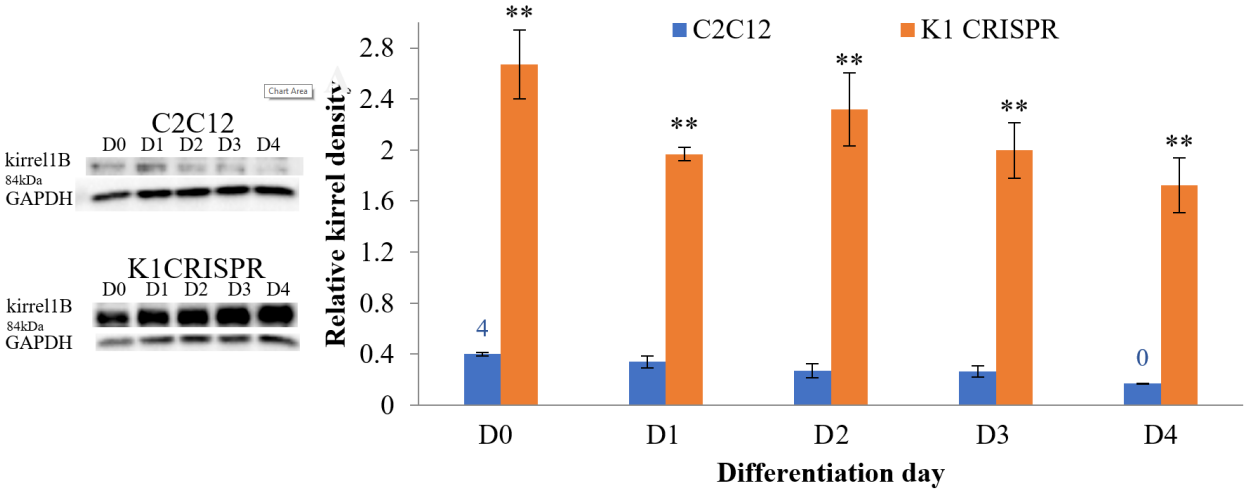


Figure 3.5 Kirrel1B (84 kDa) protein levels in differentiating C2C12s with CRISPR/Cas9-modified *kirrel1* alleles. Kirrel1B (84 kDa) levels in differentiating C2C12s were determined via western blotting under reducing conditions. Unmodified C2C12s and the K1CRISPR cell line were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative kirrel1B (84 kDa) band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 kirrel1B (84 kDa) levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in kirrel1B (84 kDa) levels between different the timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

The drastic increase in kirrel1B production in the K1CRISPR myoblasts presents a probable cause of the inhibited myotube formation seen in figure 3.3. The expression of kirrel1B has previously been shown to decrease as myoblast fusion events increase in injured mouse muscle tissue *in vivo* (Durcan *et al.*, 2013). Conversely, the *Drosophila* paralogs of kirrel1A, Duf and Rst, have been shown to be vital for myoblast fusion in the fly and although a direct role for kirrel1A in myoblast fusion has not been established in mammals, kirrel1A expression has been shown to peak in injured, recovering mouse muscle tissue (Durcan *et al.*, 2013). Based on previous findings in the fly, it would seem safe to assume that the evolutionarily conserved mammalian kirrel1A variant would be involved in myoblast fusion, the same cannot be said for the kirrel1B variant. The cytoplasmic domain of the kirrel1B protein is significantly different to that of kirrel1A with kirrel1B missing both a Grb2 and PDZ motif, these being replaced by a predicted phosphatidylinositol-3,4,5-triphosphate pleckstrin homology (PIP^3 PH) motif. The grossly overexpressed kirrel1B protein in the K1CRISPR cell line may have inhibited fusion by acting as a competitive inhibitor

of kirrel1A activity; both proteins have the same extracellular region and theoretical binding partners i.e., nephrin and other kirrel proteins. Kirrel proteins at the site of fusion focus migrate (in case of the FuRMAS) or are trafficked away from the site as fusion progresses; reducing steric hinderance and allowing for greater membrane-membrane intimacy, invasion and pore formation (Kim and Chen, 2019; Lee and Chen, 2019). Expression of the *Drosophila* paralogs of the kirrel proteins, Duf and Rst, has been shown to decrease during the final stages of the fusion process i.e., fusion pre formation and membrane integration (Bour *et al.*, 2000). If the drastically overexpressed kirrel1B protein (and potentially kirrel1A) in the K1CRISPR line could not be shuttled away from these sites (due to modification of protein function by CRISPR/Cas9 or due to excessive amounts), in turn preventing sufficient membrane-membrane closeness for fusion. Additionally, direct signalling by the cytoplasmic domain of the greatly overexpressed kirrel1B variant may have contributed to this inhibition; however, any direct signalling cascades initiated by this variant have not been identified in previous publications. Interestingly, and in contrast to the K1CRISPR line, the retroviral-overexpression of the *kirrel1B* isoform had little effect on the differentiation and fusion (section 3.3); although, it should be noted that the K1CRISPR line produced substantially more kirrel1B protein than the retroviral line (figure 3.5 and 3.19). Additionally, it is unknown whether the CRISPR/Cas9 modifications to the *kirrel1* gene affected post-translational processing of the protein, potentially affecting activity.

Kirrel1 complex levels: Following non-reducing western blots, a large immunoreactive protein band (350 kDa) was seen to accumulate in the control C2C12s as differentiation progressed (figure 3.6). This appears to be a novel finding. These apparent kirrel1 containing complexes/multimers were likely formed by homophilic, in-trans kirrel1 interactions and/or the binding of kirrel1 to other binding partners, e.g., nephrin, on the cell surface; although kirrel1 molecules have a higher binding affinity for nephrin when compared to other kirrel molecules (Abmayr and Pavlath, 2012). Additionally, kirrel1 does not appear to bind *in cis* with kirrel2 or kirrel3 *in vitro* (Gerke *et al.*, 2005). Nephrin is a 150 kDa protein, while the kirrel1 variants range is size between 80-125 kDA; these proposed kirrel1-kirrel1 or kirrel1-nephrin complexes could explain the 350 kDa bands seen on the gels. The absence/presence of these large kirrel1 complexes in reducing and non-reducing blots respectively has previously been observed by Gerke *et al.*, (2005) (in podocytes). Since the antibody used to identify the various kirrel1 variants recognized a conserved region common to both kirrel1A and kirrel1B, it was unclear which of the kirrel1 proteins (or whether multiple variants) were involved in the formation of these complexes. In the non-reducing gels in which the large complexes were seen, all kirrel1 bands smaller than these large multimers were of a weaker intensity when compared to the reduced gel, suggesting involvement of multiple variants in the make-up of the complexes. However, the lower intensities of the smaller kirrel1 bands (smaller than the 350 kDa band) seen in the non-

reduced gels may be a result of decreased binding affinity of the anti-kirrel1 antibodies to the native variants of the protein (due to non-linearization/conformational differences); rather than more kirrel1 proteins being locked up in the kirrel1 complexes. In the C2C12 line, this complex remained at low levels on day 0 and day 1 (**0.47**±0.11 and **0.40**±0.11) but was seen to accumulate as differentiation and fusion progressed with elevated levels seen after day 2; days 2, 3 and 4 having similar levels (**1.02**±0.10; **0.99**±0.16; **1.05**±0.21) (figure 3.6). No significant changes in the levels of this complex in the K1CRISPR line were seen from days 0 to 4 (figure 3.6). When compared to the control, the levels of this complex in the K1CRISPR myoblasts were decreased by 400%, 407 and 458% on days 2, 3 and 4 respectively ($p<0.05$).

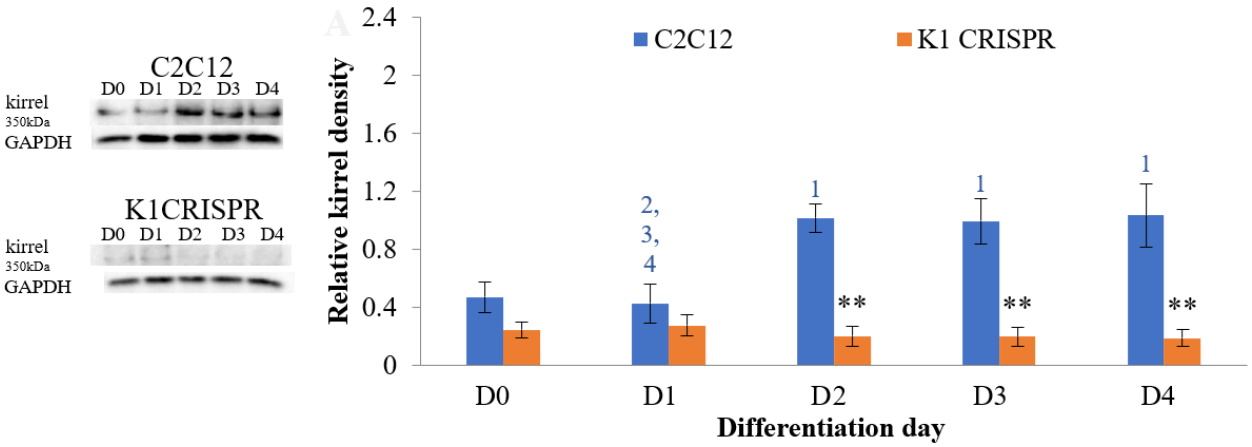


Figure 3.6 Kirrel1 multimer/complex (350 kDa) levels in differentiating C2C12s with CRISPR/Cas9-modified *kirrel1* alleles. Kirrel1 multimer/complex (350 kDa) levels in differentiating C2C12s were determined via western blotting under non-reducing conditions. Unmodified C2C12s and the K1CRISPR cell line were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative kirrel1 multimer/complex (350 kDa) band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 kirrel1 multimer/complex (350 kDa) levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in kirrel1 multimer/complex (350 kDa) levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

The lack of this complex in this cell line suggests a potential inhibition/inability of the overexpressed *kirrel1* variants to associate/bind to partner proteins as required for fusion. However, as will be covered later in the thesis (section 3.3), the retroviral overexpression of kirrel1A and kirrel1B had little effect on the formation of this complex as well as on the differentiation and fusion process. However, the levels of kirrel1B protein in the K1CRISPR myoblasts appeared to be substantially greater than those seen in the retroviral, overexpressing cells; especially on day 0. It is possible that the dramatic, seemingly uncontrolled overexpression of kirrel1B in the K1CRISPR cell line could have led to some manner of negative feedback (including trafficking of kirrel proteins), competitive inhibition of other kirrel variants/binding partners as well as creating physical interference at the site of fusion (kirrel molecules are trafficked away from the fusion site during the late stages of fusion). Furthermore, it is unclear whether modification of the *kirrel1*

gene in the K1CRISPR cell line had any effect on the post-translation modification of the protein; modifications that regulate kirrel1 activity and potentially the formation of these complexes.

3.1.6 The CRISPR/Cas9-mediated disruption of kirrel1 expression in C2C12s had little effect on the levels of Pax7, Myf5 and MyoD during differentiation while levels of myogenin and MyHC were significantly reduced in these myoblasts

Pax7, My5 and MyoD: The control and K1CRISPR lines displayed very similar patterns and levels of Pax7, My5 and MyoD during differentiation (figure 3.7). The only significant changes in the expression of any of these proteins in the K1CRISPR line relative to the control was seen for MyoD on day 1 with the K1CRISPR line producing roughly 72% less MyoD protein (**0.58**±0.04) relative to the control (**1.00**±0.04; $p<0.05$). Although both cell lines appeared to follow the same general Pax7 expression profile, only the K1CRISPR line displayed a statistically significant decrease in Pax7 after day 0. The decrease in Pax7 levels seen in the K1CRISPR cell line is typical in healthy, differentiating C2C12 myoblasts *in vitro* (Seale *et al.*, 2000; Chen *et al.*, 2010). With the relatively unchanged levels in these important myogenic regulatory proteins in the K1CRISPR line when compared to the control, it would seem that the inhibition of differentiation seen in this cell line was not as a result of the misregulation of any of these molecules.

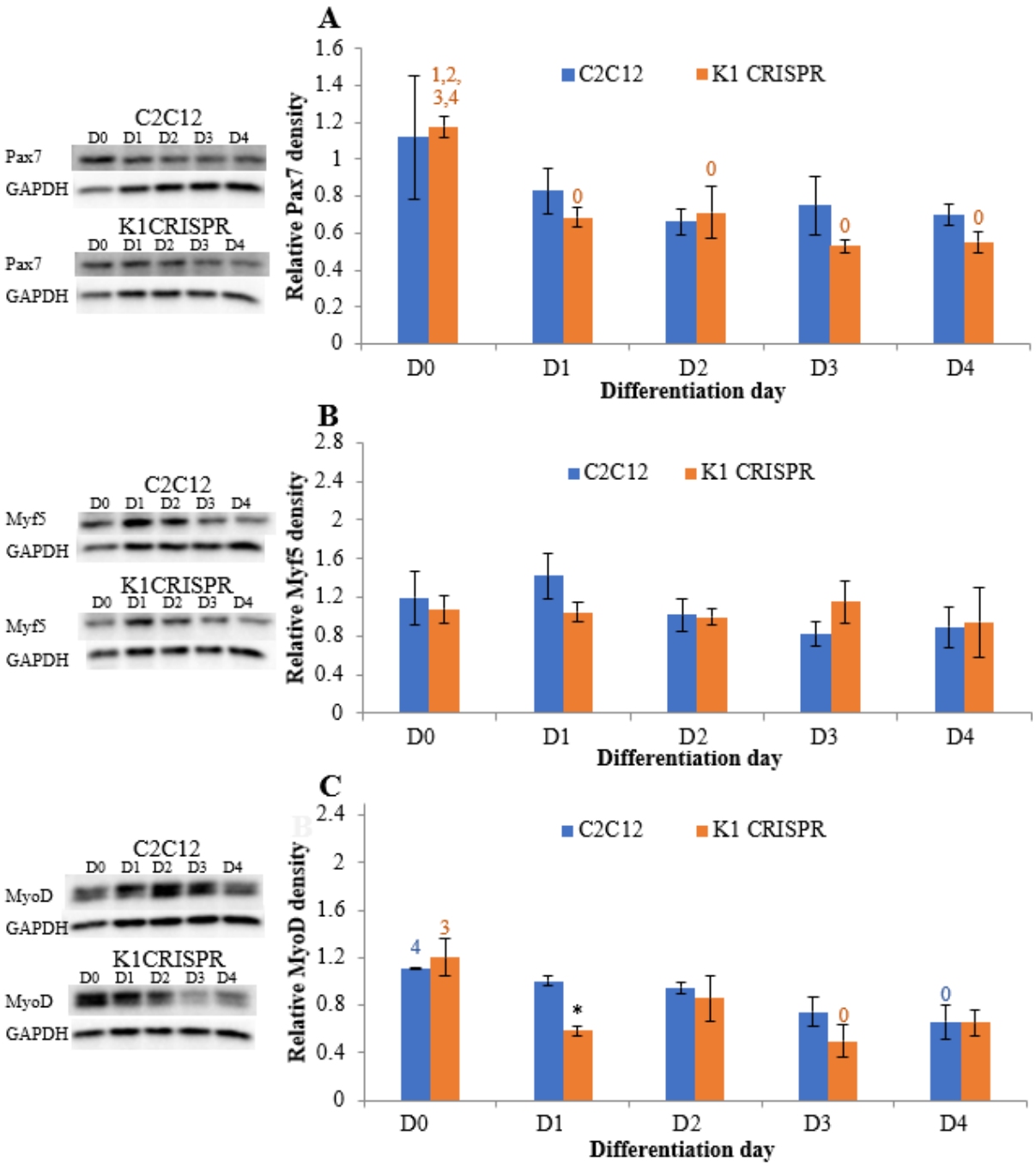


Figure 3.7 Pax7, Myf5 and MyoD levels in differentiating C2C12s with CRISPR/Cas9-modified *kirrel1* alleles. Pax7 (A), Myf5 (B) and MyoD (C) levels in differentiating C2C12s were determined via western blotting under non-reducing conditions. Unmodified C2C12s and the K1CRISPR cell line were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative Pax7/Myf5/MyoD band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 Pax7/Myf5/MyoD levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in Pax7/Myf5/MyoD levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

Myogenin protein levels: The myogenin expression profile in the control C2C12s followed a typical pattern for myoblasts differentiating *in vitro* (Panda *et al.*, 2014; Horibata *et al.*, 2020) with levels being significantly increased on days 2, 3 and 4 (**1.13**±0.07, **0.96**±0.04 and **0.95**±0.14) relative to day 0 (**0.09**±0.05; $p<0.05$) (figure 3.8.A). However, myogenin in the K1CRISPR myoblasts was significantly reduced on days 1, 2 and 3 (**0.05**±0.03, **0.07**±0.03 and **0.32**±0.19) when compared to the control values (**0.70**±0.28, **1.13**±0.07, and **0.96**±0.04; $p<0.05$), decreases of 1520%, 1480% and 200% respectively (figure 3.8.A). However, myogenin expression in the modified myoblasts appeared to recover by day 4, reaching similar levels to those in the control; **0.95**±0.14 and **0.82**±0.34 for the control and K1CRISPR lines respectively.

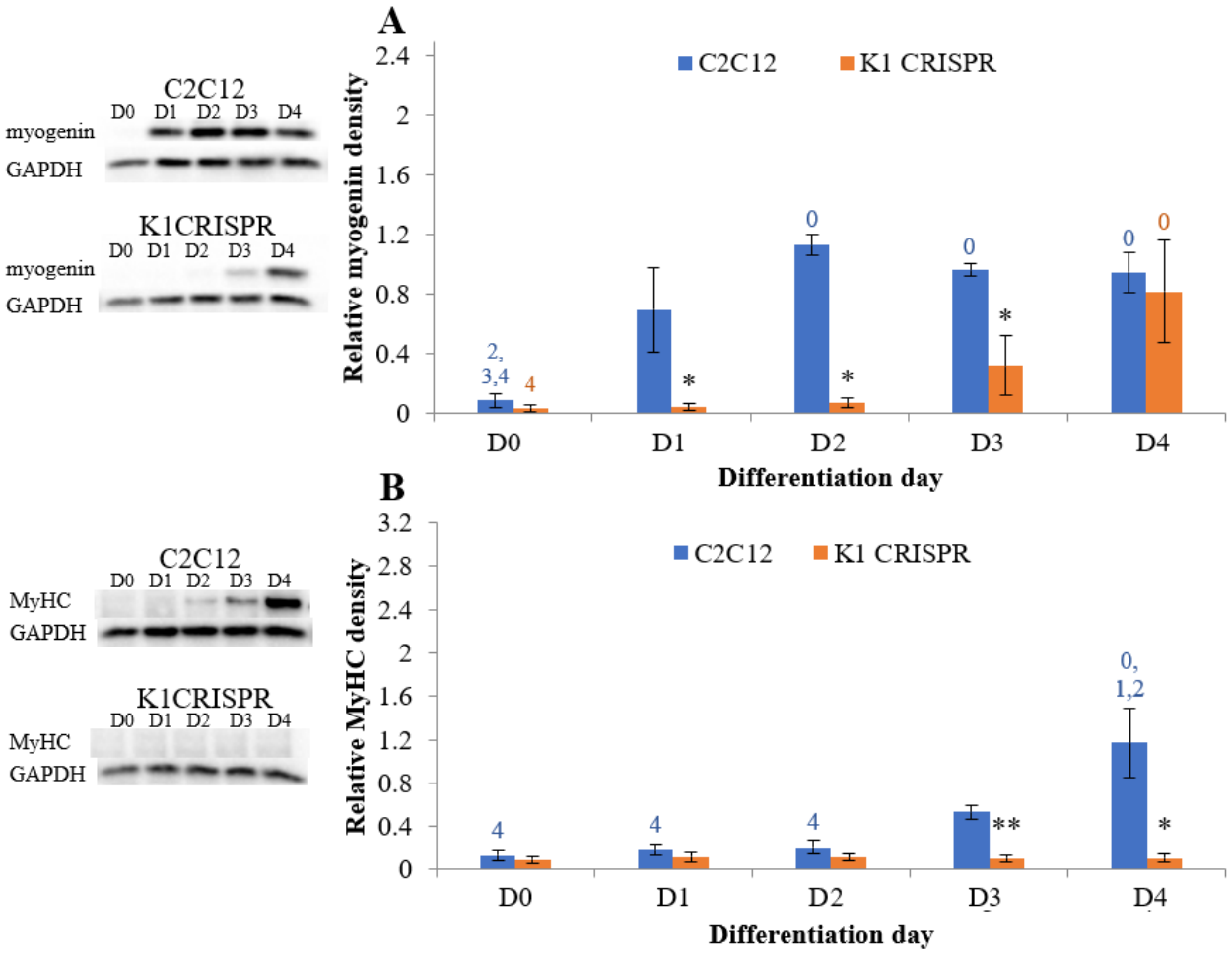


Figure 3.8 Myogenin and MyHC levels in differentiating C2C12s with CRISPR/Cas9-modified *kirrel1* alleles. Myogenin (A) and MyHC (B) levels in differentiating C2C12s were determined via western blotting under non-reducing conditions. Unmodified C2C12s and the K1CRISPR cell line were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative myogenin/MyHC band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 myogenin/MyHC levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in myogenin/MyHC levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

MyHC protein levels: MyHC levels in the control cell line followed a typical profile for healthy myoblasts differentiating and fusing *in vitro* (Kim *et al.*, 2015; Nie *et al.*, 2017) with levels on day 4 (**1.20**±0.29) being significantly increased relative to days 0, 1 and 2 (**0.13**±0.05, **0.19**±0.05 and **0.21**±0.07; $p<0.05$) (figure 3.8.B). MyHC levels in the K1CRISPR line remained low and unchanged from days 0 to 4 (figure 3.8.B) with significantly reduced levels on days 3 and 4 (**0.11**±0.03 and **0.11**±0.04) relative to the control (**0.53**±0.07 and **1.20**±0.29; $p<0.05$), decreases of 400% and 1039%. The drastic decrease in MyHC levels in these modified myoblasts further demonstrates the arrest of myotube production in the cell line (as seen in figure 3.3); this protein being an essential component of the contractile machinery within the myotubes with its expression being a hallmark of successful myoblast fusion.

Although not represented in the figures, the K1CRISPR cell line was differentiated to day 7 with severely inhibited myotube production still seen at this time point (even relative to the control line at day 4) (appendix III). This suggests that although myogenin levels appeared to recover in the K1CRISPR cell line by day 4, this recovery was not accompanied by a rescue in myotube formation. This result appears to be a novel finding with no past publications linking this kirrel variant to MRF expression being found, however, it has previously been shown that inhibition of kirrel3 expression by short interfering RNAs (siRNAs) in C2 myoblasts *in vitro* had no effect on myogenin levels during differentiation. Since cell surface interactions are major regulators of late-stage MRF expression e.g., myogenin, any modification of CAMs such as the kirrel molecules could have major effects on the production of these myogenic factors (Krauss, 2010; Knight and Kothary 2011). It has previously been shown by Laurin *et al.*, (2008), that the silencing of the Dock family of Rac GEFs in C2C12s (paralogues of the *Drosophila* Mbc protein) results in the delayed expression of myogenin and postpones exit from the cell cycle. The Dock proteins are in part regulated by the kirrel molecules which localize the Rac GEFs to the cell membrane, resulting in the activation of Rac and the progression of differentiation/fusion (Laurin *et al*, 2008; Haralalka *et al.*, 2011). It is possible that the reduced myogenin expression seen in the K1CRISPR line was a result of irregular Dock protein activity/localization caused by the misregulation of the kirrel1 variants. Alternatively, or in addition to the aforementioned direct signalling cascades, the inhibition of myogenin expression seen in the K1sh cells may be the result of less direct mechanisms e.g. sequestering of myogenin-promoting molecules/receptors by misregulated kirrel1 variants, crowding-out other myogenin-promoting molecules on the cell surface by kirrel1 molecules, alterations to kirrel1-regulated feedback loops (controlling the localization of cell-surface molecules) and/or the modification/interruption of spatiotemporal cell-cell interactions required for the progression of myogenesis (adhesion, alignment, degree of cell-cell proximity).

3.2 The shRNA-mediated knockdown of kirrel1A and kirrel1B

In light of the results covered in the previous section, it is clear that disruption of kirrel1 protein expression in the K1CRISPR cell line had a detrimental effect on myotube production. However, the precise roles of kirrel1A and kirrel1B during this process remain unclear. To establish the roles of these proteins during myogenesis, the activities of the *kirrel1A* and *kirrel1B* mRNA variants were separately knocked-down in C2C12s by use of short hairpin RNAs (shRNAs).

3.2.1 The treatment of C2C12s with shRNAs against kirrel1A or kirrel1B mRNAs reduced pre-fusion levels of kirrel1A mRNA but unchanged levels of the already lowly transcribed kirrel1B variant during differentiation in vitro

C2C12s were transfected with a plasmid cocktail consisting of 3 separate shRNA-producing vectors, all 3 targeting either *kirrel1A* or *kirrel1B*, resulting in two polyclonal cell lines that will be referred to as "K1Ash" (kirrel1A short hairpin) and "K1Bsh" (kirrel1B short hairpin) respectively. The sequences of these short hairpin RNAs produced by these plasmids can be seen in appendix IV. A control shRNA cell line was also produced and will be referred to as "Ctrlsh". These three cell lines were differentiated over 4 days *in vitro* with total RNA collected on days 0, 2 and 4. Using the same protocol as in section 3.1.2, cDNA was produced from these RNA samples and PCR reactions carried out to assess *kirrel1A* (primer set K1A) and *kirrel1B* (primer set K1B) levels.

Kirrel1A mRNA levels: Unlike with the C2C12 control and K1CRISPR cell lines in section 3.1.2, none of the three shRNA cell lines (Ctrlsh, K1Ash and K1Bsh) showed any statistically significant changes in *kirrel1A* mRNA levels between days 0 and 4 of differentiation. Although not statistically significant, *kirrel1A* levels in all three cell lines seemed to peak on day 2. The only statistically significant drop in *kirrel1A* levels (relative to the control cell line) was seen in the K1Ash myoblasts on day 0 (**0.28**±0.11), with a 130% decrease in *kirrel1A* mRNA at this timepoint relative to the Ctrlsh line (**0.65**±0.15; $p<0.05$) (figure 3.9.A). However, on days 2 and 4, *kirrel1A* mRNA levels in the K1Ash cell line appeared to recover, and reaching similar levels to those seen in the Ctrlsh line. The diminishing effectiveness of silencing over time is well documented when applying this strategy; shRNAs are increasingly degraded, sequestered and/or excreted from the cells after cessation of antibiotic selection (Applied Biosystems Inc., 2008). Significant reductions in shRNA effectiveness have been shown to typically manifest 5-7 days post transfection (Han, 2018). Differentiation of the three cell lines was initiated 4 days after transfection with the three days prior to this initiation allowing for the antibiotic selection of and expansion of the cell lines for plating. No statistically significant changes in *kirrel1A* levels were seen in the K1Bsh line relative to the control (figure 3.9.A).

Kirrel1B mRNA levels:As was seen in the C2C12 control line in the previous section (section 3.1.2), *Kirrel1B* mRNA levels in all three shRNA cell lines (Ctrsh, K1Ash and K1Bsh) were considerably lower than those of *kirrel1A* (figure 3.9.B). The K1Bsh cell line did not appear to have reduced *kirrel1B* levels when compared to the control. It is possible that due to the already low levels of *kirrel1B* mRNA, any interrupting effects by these shRNAs were not observed.

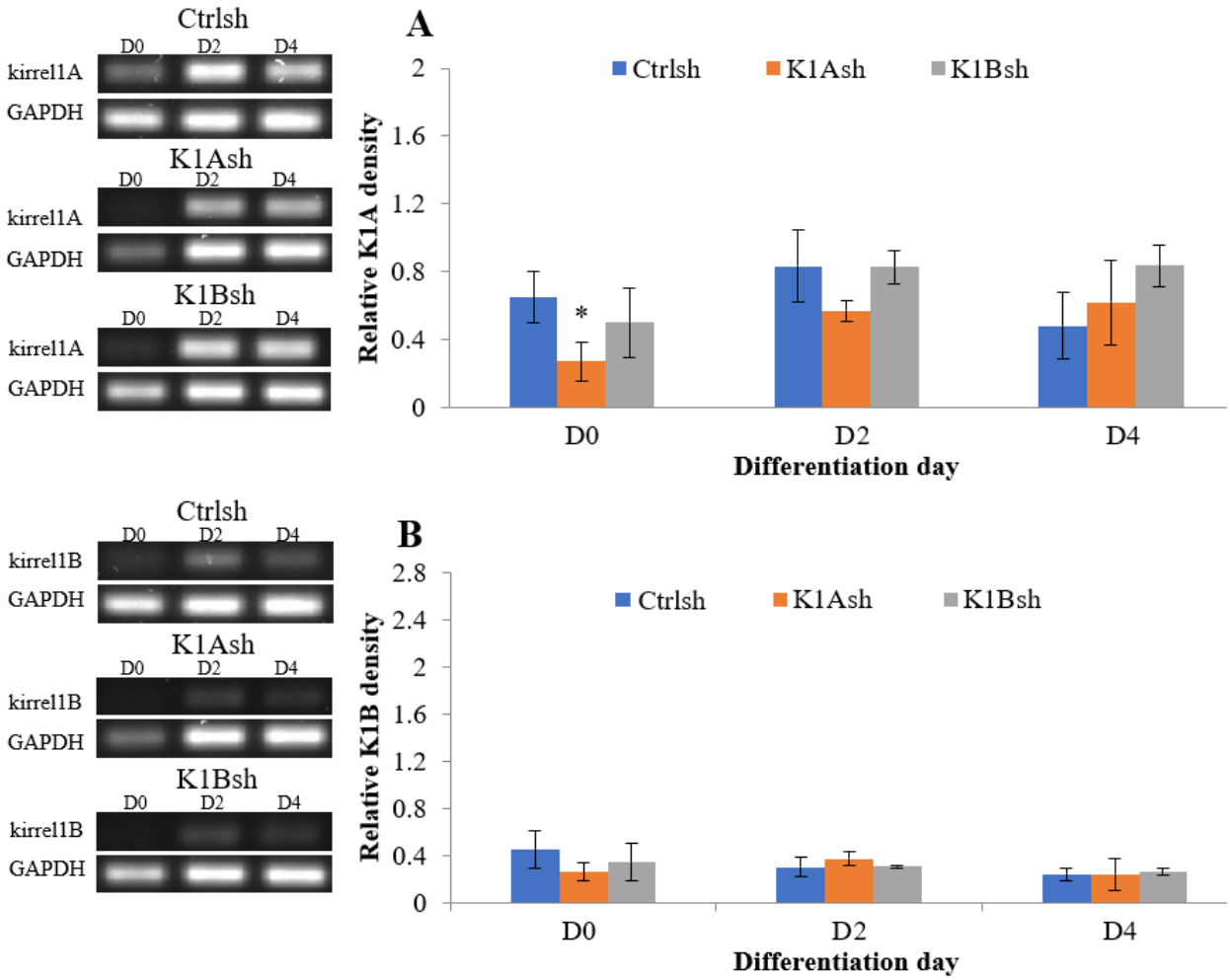


Figure 3.9 *Kirrel1* mRNA levels in differentiating C2C12s with *kirrel1A* or *kirrel1B* knockdown. *Kirrel1A* (A) and *kirrel1B* (B) mRNA levels in differentiating C2C12s were determined via semi-quantitative PCRs. The Ctrlsh, K1Ash and K1Bsh cell lines were differentiated *in vitro* over 4 days with RNA collected on days 0, 2 and 4. Densitometry analysis of agarose gels containing *kirrel1* PCR products was carried out using ImageJ and the relative *kirrel1* band intensities determined using *gapdh* as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 *kirrel1* mRNA levels at individual timepoints were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in *kirrel1* mRNA levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 2, $p<0.05$ when compared to D2; 4, $p<0.05$ when compared to D4.

3.2.2 The knockdown of kirrel1A protein expression severely inhibits the production of C2C12 myotubes in vitro

The Ctrlsh, k1Ash and K1Bsh cell lines were differentiated over 4 days *in vitro* with photographs taken on each day and myotube analysis carried out as described in section 3.1.3. Both the Ctrlsh and K1Bsh lines underwent seemingly successful differentiation and fusion with myotubes appearing from day 2 and increasing in number and size with each passing day, both cell lines reaching maximum coverage on day 4 (figure 3.10.A). When comparing these two lines to one another, the only statistically significant difference between the myotube values was seen when looking at the average individual tube area on day 4 where the K1Bsh myotubes (**4071 µm²**±360 µm²) were observed to be 35% larger than the control (**3021 µm²**±56 µm²; $p<0.05$) (figure 3.10.B). Conversely, the K1Ash cell line displayed severely inhibited myotube production with almost no myotubes seen after 4 days of differentiation (figure 3.10.C). K1Ash myotubes were only seen from day 3 onwards with these tubes being much smaller and fewer in number than those produced by the control and K1Bsh lines. On days 3 and 4, the total myotube coverage values for the K1Ash cell line (**1.00%**±0.31% and **2.11%**±1.04%) were seen to be significantly decreased when compared to the control values at the same time points (**6.30%**±0.34% and **16.7%**±2.10%; $p<0.05$). This novel finding is similar to that by Tamir-Livne *et al.,* (2017) who showed that the knockdown of *kirrel3* in mouse-derived C2 myoblasts resulted in severely inhibited myotube production.

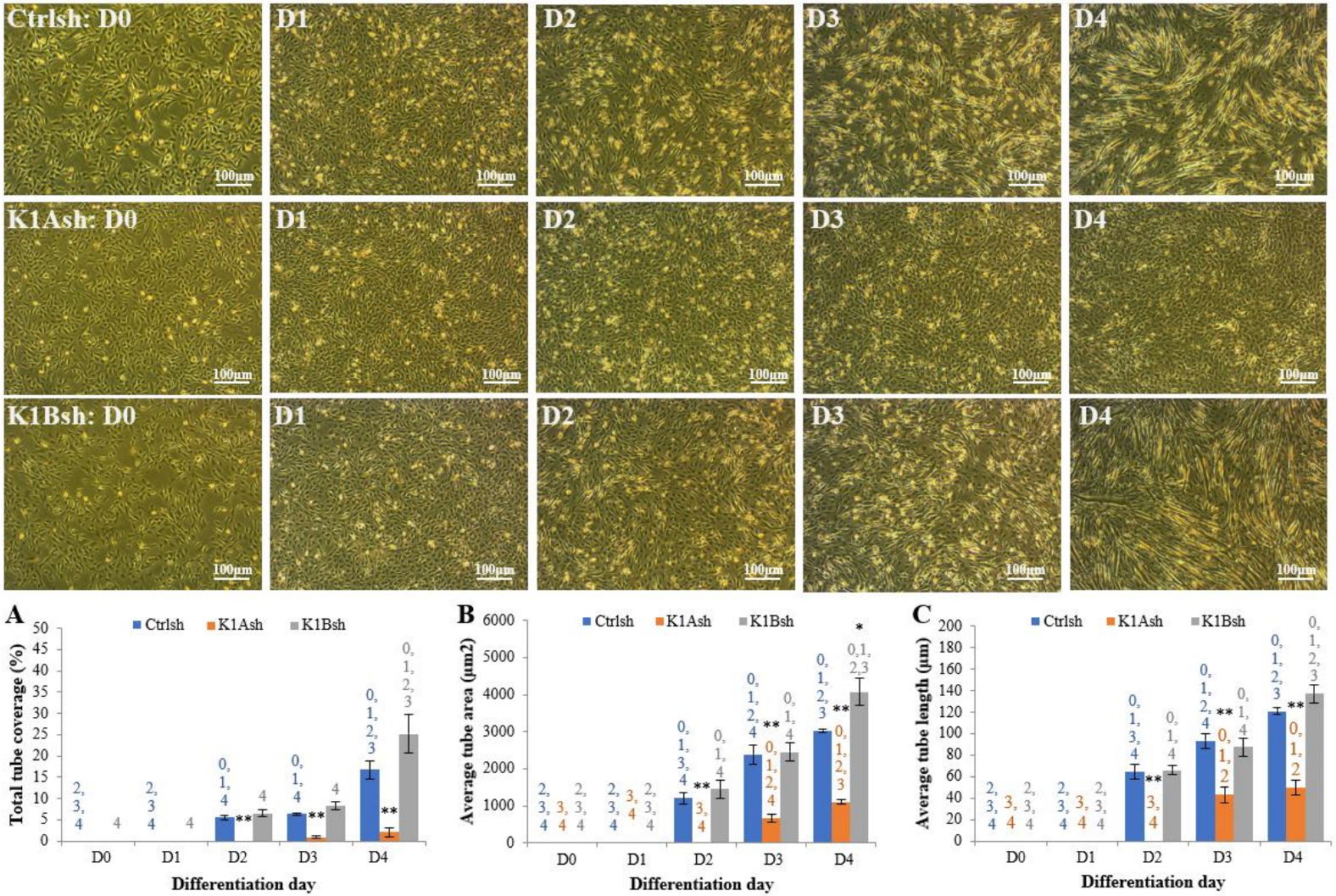


Figure 3.10 Myotube production by differentiating C2C12s with *kirrel1A* or *kirrel1B* knockdown. The Ctrlsh, K1Ash and K1Bsh cell lines were differentiated *in vitro* over 4 days with images aquired daily at 40x magnification. A) Total myotube coverage determined using an in-house ImageJ macro (appendix X). B) Average individual tube areas determined by manually outlining the myotube structures using ImageJ and recording the area values. C) Average individual tube length determined by manually outlining the myotube structures using ImageJ and recording the Feret values. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 myotube values at each timepoint were determined using an unpaired, 2-tailed Student's t-test. *p<0.05, **p<0.005. Statistically significant differences in myotube values between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, p<0.05 when compared to D0; 1, p<0.05 when compared to D1; 2, p<0.05 when compared to D2; 3, p<0.05 when compared to D3; 4, p<0.05 when compared to D4.

3.2.3 The treatment of C2C12s with shRNAs against kirrel1A or kirrel1B mRNAs resulted in reduced levels of the kirrel1A protein but relatively unchanged levels of the already lowly transcribed kirrel1B variant during differentiation in vitro. Kirrel1A antagonization prevented the formation of the 350 kDa kirrel1-containing multimers/complexes

Kirrel1A protein levels: Kirrel1A (125 kDa) levels in the Ctrlsh, K1Ash and K1Bsh cell lines remained relatively stable with no statistically significant changes in protein expression being observed between the differentiation timepoints in any of the three lines (figure 3.11). However, kirrel1A levels on days 0 and 2 in the K1Ash cell line (**0.18**±0.05 and **0.41**±0.03) were seen to be decreased relative to the Crtlsh control line values at the same timepoints (**0.43**±0.21 and **0.63**±0.03; $p<0.05$), drops of 133% and 54% on days 0 and 2 respectively.

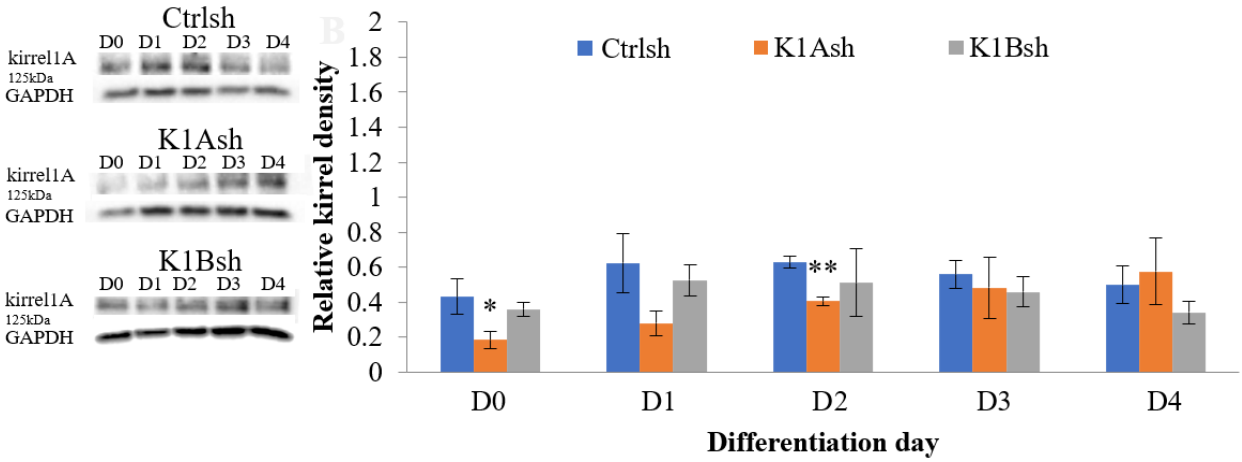


Figure 3.11 Kirrel1A (125 kDa) protein levels in differentiating C2C12s with *kirrel1A* or *kirrel1B* knockdown. Kirrel1A (125 kDa) levels in differentiating C2C12s were determined via western blotting under reducing conditions. The Ctrlsh, K1Ash and K1Bsh cell lines were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative kirrel1A (125 kDa) band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 kirrel1A (125 kDa) levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in kirrel1A (125 kDa) levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

It is apparent that this inhibition of kirrel1A production/regulation during the early stages of differentiation had a dramatic inhibitory effect on the production of myotubes in the K1Ash line as seen in figure 3.10. This result was not surprising as the need for nuanced kirrel3 expression during pre-fusion events in differentiating mouse myoblasts has been documented (Tamir-Livne *et al.*, 2017). Furthermore, the *Drosophila* paralogs of the kirrel proteins, Duf and Rst, are known to be vital for myoblast fusion (Kim *et al*., 2015; Rout *et al.*, 2022). Kirrel1A levels in the K1Ash myoblasts appear to recover by day 3, reaching similar levels to those in the Ctrlsh line on days 3 and 4, although this recovery was not accompanied by a rescue in myotube production. This gradual recovery of kirrel1A production in the K1Ash myoblasts was also demonstrated in the mRNA PCR gels (figure 3.9.A). No statistically significant changes in kirrel1A levels were observed in the K1Bsh cell line when compared to the Ctrlsh control myoblasts.

Kirrel1B protein levels: Kirrel1B (84 kDa) expression levels in the Ctrlsh, K1Ash and K1Bsh remained relatively stable from days 0 to 4 with no statistically significant changes in expression seen between any of the timepoints for any of the cell lines. All three cell lines expressed comparable amounts of kirrel1B on each day with the only significant change relative to the control being seen in the K1Bsh line on day 2. On day 2, kirrel1B levels in the K1Bsh line (**0.32**±0.05) were seen to be significantly reduced when compared to the Ctrlsh control line (**0.50**±0.15; $p<0.05$), a decrease of 60%.

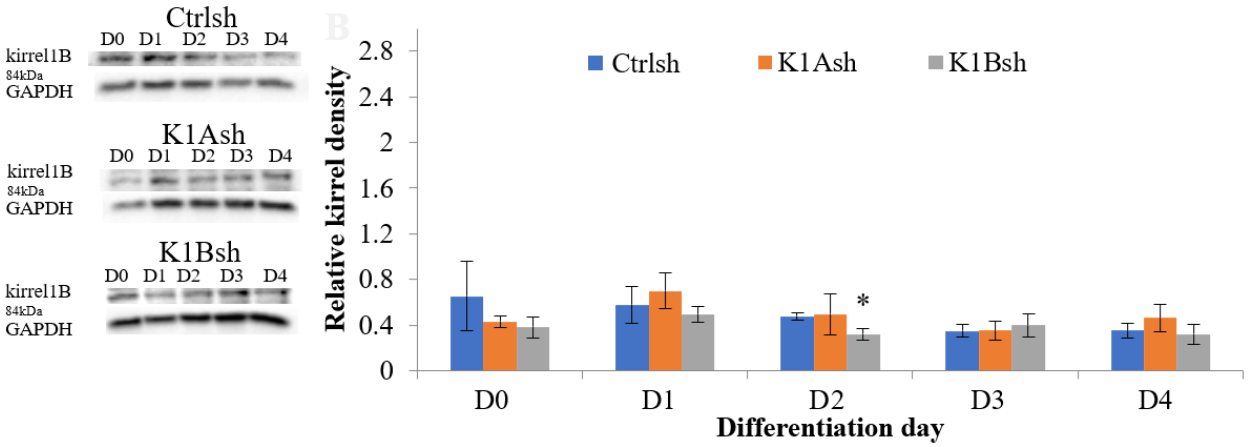


Figure 3.12 Kirrel1B (84 kDa) protein levels in differentiating C2C12s with *kirrel1A* or *kirrel1B* knockdown. Kirrel1B (84 kDa) levels in differentiating C2C12s were determined via western blotting under reducing conditions. The Ctrlsh, K1Ash and K1Bsh cell lines were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative kirrel1B (84 kDa) band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 kirrel1B (84 kDa) levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in kirrel1B (84 kDa) levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

The findings in these western blots appear to be backed up by the mRNA PCRs in which no statistically significant changes in *kirrel1B* levels were detected in the K1Ash or K1Bsh lines when compared to the control. In the previous mRNA PCR section (section 3.2.2), *kirrel1B* mRNA levels in the shRNA lines were significantly lower than those of *kirrel1A* and it would appear that the naturally low levels of *kirrel1B* were difficult to suppress with shRNAs. When looking at protein levels in the western blots (figures 3.11 and 3.12), the difference between the intensities of the kirrel1A (125 kDa) and kirrel1B (84 kDa) bands was not nearly as stark as the difference seen in the PCR gels (figure 3.9). It is possible that despite the relatively low levels of *kirrel1B* mRNA, mechanisms regulating the transcript's transcription allow for the production of protein levels similar to those of kirrel1A (125 kDa).

Kirrel1 complex levels: Following non-reducing western-blotting, the large, 350 kDa kirrel1-containing complexes were seen to accumulate in the successfully differentiated/fused Ctrlsh and K1Bsh lines (figure 3.13). In the Ctrlsh line, the level of this complex was seen to be significantly increased on day 4

(**1.52**±0.37) when compared to day 0 and 1 (**0.29**±0.10 and **0.24**±0.0; p<0.05), increases of 429% and 547%. The levels of this kirrel1-containning complex in the K1Bsh line were seen to be significantly increased on days 3 and 4 (**1.30**±0.28 and **1.34**±0.30) when compared to day 0 (**0.18**±0.03; p<0.05), increases of 647% and 669%. As was seen in the K1CRISPR cell line, the K1Ash did not form myotubes and did not appear to accumulate these large kirrel1 complexes as differentiation progressed. On day 4, the levels of this complex in the K1Ash cells (**0.26**±0.04) displayed a statistically significant decrease when compared to the Ctrlsh control (**1.52**±0.37: p<0.05), a decrease of 482%.

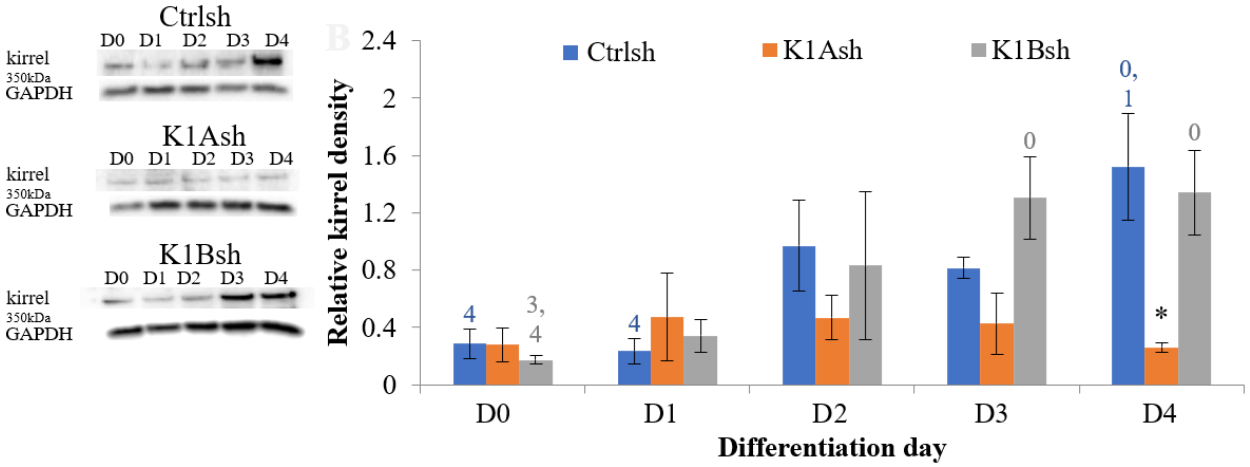


Figure 3.13 Kirrel1 multimer/complex (350 kDa) levels in differentiating C2C12s with *kirrel1A* or *kirrel1B* knockdown. Kirrel1 multimer/complex (350 kDa) levels in differentiating C2C12s were determined via western blotting under non-reducing conditions. The Ctrlsh, K1Ash and K1Bsh cell lines were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative kirrel1 multimer/complex (350 kDa) band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 kirrel1 multimer/complex (350 kDa) levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, p<0.05; **, p<0.005. Statistically significant differences in kirrel1 multimer/complex (350 kDa) levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, p<0.05 when compared to D0; 1, p<0.05 when compared to D1; 2, p<0.05 when compared to D2; 3, p<0.05 when compared to D3; 4, p<0.05 when compared to D4.

3.2.4 The treatment of C2C12s with shRNAs targeting kirrel1A had little effect on the expression of Pax7 and MyoD during differentiation in vitro whereas Myf5 expression was elevated. The expression of myogenin and MyHC was reduced in the kirrel1A knockdown myoblasts

Pax7 protein levels: Pax7 levels in the Ctrlsh, K1Ash and K1Bsh myoblasts appeared relatively stable between days 0 and 4 of differentiation (figure 3.14.A). Although not statistically significant, all three cell lines show a moderate trend of decreasing Pax7 levels with the progression of differentiation. Furthermore, Pax7 levels in the K1Ash and K1Bsh lines were unchanged when compared to the control values at each timepoint.

Myf5 protein levels: Myf5 protein expression in the K1Ash myoblasts appeared to be elevated relative to the control from day 1 onwards (figure 3.14.B). On days 1 and 4, Myf5 levels in the K1Ash cells (**2.40**±0.33 and **1.45**±0.25) were seen to be significantly increased when compared to the control values at the same time points (**1.28**±0.21 and **0.68**±0.21; $p<0.05$). Myf5 is typically one of the first of the MRFs to be upregulated following the activation of satellite cells *in vivo* and along with MyoD, is required for optimal myoblast expansion and differentiation (Singh and Dilworth, 2013; Zammit, 2017). The sequential expression of Myf5, followed closely by MyoD, ensures the initiation of the myogenic process (Singh and Dilworth, 2013; Zammit, 2017). The sustained/increased expression of Myf5 in the K1Ash line was likely not the cause of the inhibited differentiation and fusion but rather a result of this inhibition; perhaps being a compensatory action by the myoblasts trying to initiate myogenesis. The overexpression of Myf5 has not been shown to impede differentiation of C2C12 *in vitro* (Sapoznik *et al.*, 2018) while sustained Myf5 expression has been observed in spinal muscular atrophy (SMA) myoblasts that display inhibited fusion; this sustained expression not being the cause of the inhibition, rather a result (Verche *et al.*, 2017).

MyoD protein levels: No statistically significant changes in MyoD expression were seen between days 0 and 4 of differentiation of the Ctrlsh, K1Ash and K1Bsh cell lines (figure 3.14.C). Furthermore, MyoD levels in both the K1Ash and K1Bsh were unchanged when compared to those seen in the C2C12 control myoblasts at the same time points.

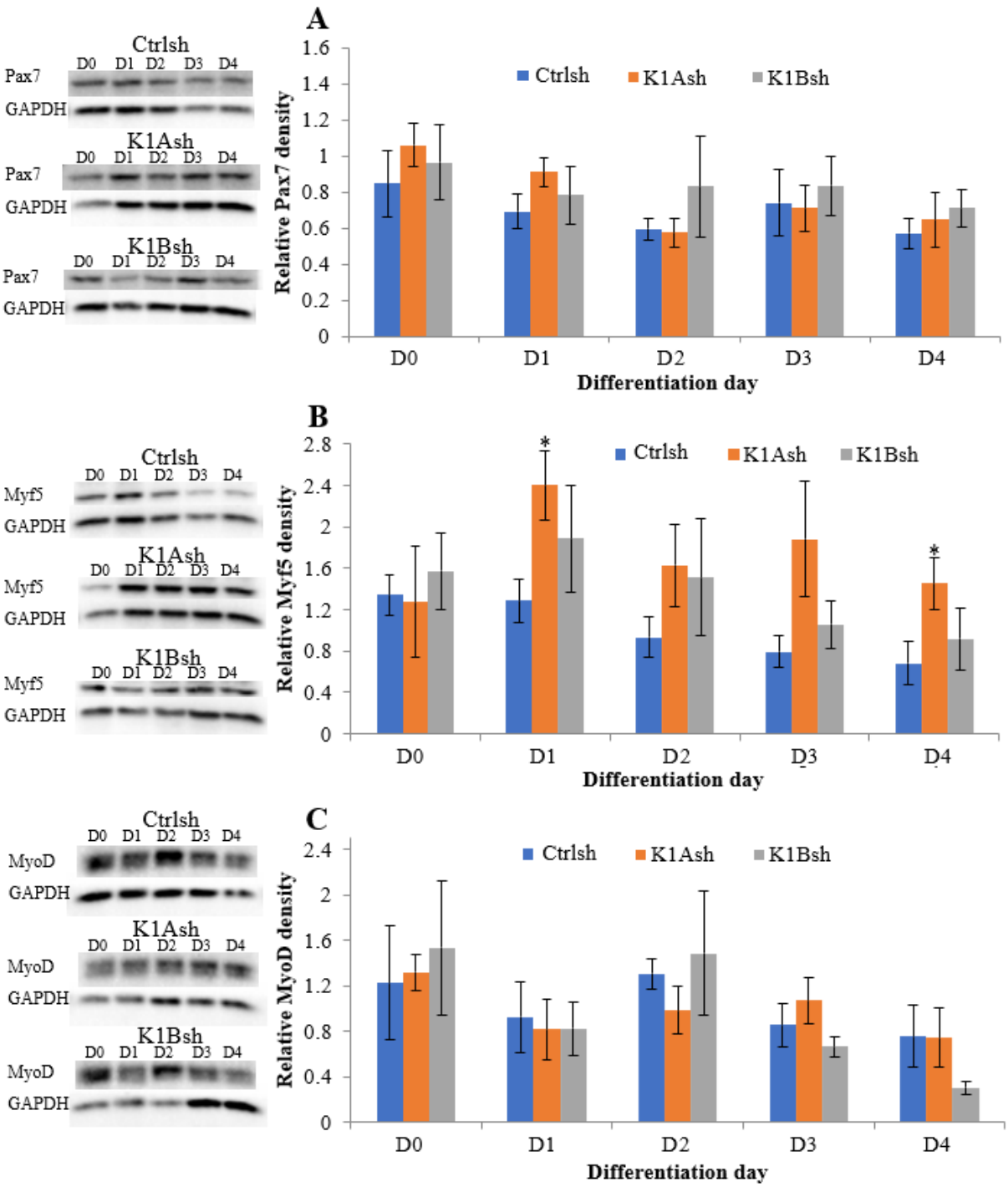


Figure 3.14 Pax7, Myf5 and MyoD levels in differentiating C2C12s with *kirrel1A* or *kirrel1B* knockdown. Pax7 (A), Myf5 (B) and MyoD (C) levels in differentiating C2C12s were determined via western blotting under non-reducing conditions. The Ctrlsh, K1Ash and K1Bsh cell lines were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative Pax7/Myf5/MyoD band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 Pax7/Myf5/MyoD levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in Pax7/Myf5/MyoD levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

Myogenin protein levels: Myogenin levels in the control Ctrlsh myoblasts were low on days 0 and 1 but increased dramatically on day 2 and remained elevated on days 3 and 4 (figure 3.15.A). Myogenin expression in the K1Ash myoblasts also appeared to peak by day 2 and remain elevated on days 3 and 4, however, these increases were not statistically significant relative to days 0 or 1, furthermore, the levels of myogenin in the K1Ash line were generally lower than those seen in the control with statistically significant decreases seen on days 3 and 4 (figure 3.15.A).

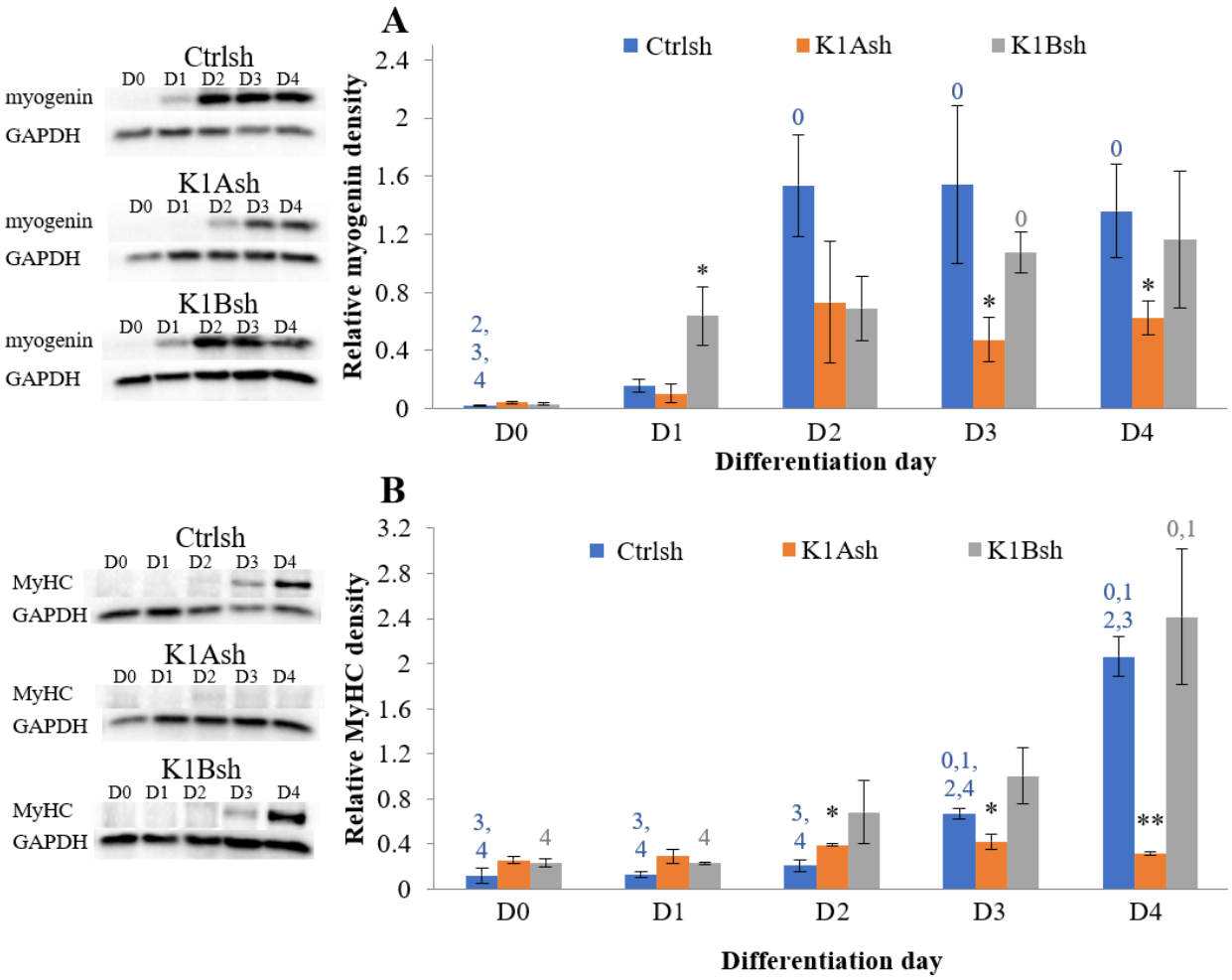


Figure 3.15 Myogenin and MyHC levels in differentiating C2C12s with *kirrel1A* or *kirrel1B* knockdown. Myogenin (A) and MyHC (B) levels in differentiating C2C12s were determined via western blotting under non-reducing conditions. The Ctrlsh, K1Ash and K1Bsh cell lines were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative myogenin/MyHC band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 myogenin/MyHC levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in myogenin/MyHC levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

MyHC protein levels: MyHC levels in the successfully differentiating and fusing Ctrlsh and K1Bsh myoblasts followed a typical pattern of expression (Kim *et al.,* 2015; Nie *et al.,* 2017) with maximums reached on day 4 (figure 3.15.B), the same day on which total myotube coverage values were highest. MyHC expression in the K1Ash cell line remained low on all days; as expected since no myotubes were seen to be formed by this cell line.

This result was somewhat unexpected as kirrel3 knockdown in C2 myoblasts had no effect on myogenin production (Tamir-Livne *et al.,* 2017). Similar to what was seen in the K1CRISPR line, it is possible that the knockdown of kirrel1A in the K1Ash cell line may have resulted in the misregulation of Dock protein activity/localization at the focal adhesions which in turn prevented the production of wild-type levels of myogenin. Furthermore, the silencing of Dock proteins has been shown to prevent the exiting of C2C12s from the cell cycle (Laurin *et al.,* 2008). It is therefore conceivable that the K1Ash myoblasts were unable to exit the cell cycle, potentially explaining the elevated/sustained expression of Myf5 seen in the previous figure (figure 3.14.B). The K1Bsh line displayed a seemingly healthy pattern of myogenin expression with elevated expression evident from day 1 onwards (figure 3.15.A). Interestingly, a statistically significant increase in myogenin expression was detected in the K1Bsh line (**0.55**±0.12) relative to the control (**0.16**±0.05; $p<0.05$) on day 1, an increase of 245%. This early, elevated myogenin expression on day 1 appeared to have little effect on the differentiation and fusion of the K1Bsh line as no major changes in MyHC (figure 3.15.B) or myotube production (figure 3.10) were observed relative to the control.

3.3 Retroviral overexpression of kirrel1A, kirrel1B and kirrel1A mutant

To compliment the CRISPR/Cas9 and shRNA knockdown results, kirrel1A, kirrel1B and a kirrel1A-mCherry mutant were overexpressed in differentiating C2C12s. Cell lines stably overexpressing these *kirrel1* variants were prepared via retroviral transduction and selection with puromycin. The *kirrel1A* and *kirrel1B* retroviral plasmids were created by RT-PCR cloning the mature mRNA sequences from proliferating C2C12 myoblasts which were then ligated into the pBABE-puro plasmid. The *kirrel1A* in the mCherry mutant plasmid (also using the pBABE-puro retroviral plasmid backbone) is of rat origin, having 95% homology with the mouse sequence.

3.3.1 Introduction of kirrel1A or kirrel1B coding sequences to the C2C12 genome via retroviral transduction increased kirrel1 mRNA levels during differentiation in vitro

Retroviral transduction was used to produce stable, polyclonal cell lines overexpressing either kirrel1A, kirrel1B or a mutant form of kirrel1A with an mCherry insertion between amino acids 575 and 576, a region located at the junction between the transmembrane and cytoplasmic domains of the CAM; these cell lines are referred to as "K1A", "K1B" and "mCherry" respectively. A control cell line was created using an empty version of the pBABE-puro retroviral plasmid and named "Pb". These four cell lines were differentiated over 4 days *in vitro* with total RNA collected on days 0, 2 and 4. Using the same protocol as in section 3.1.2, cDNA was produced from these RNA samples and PCR reactions carried out to assess *kirrel1A* (primer set K1A) and *kirrel1B* (primer set K1B) levels (figures 3.16).

<u>*Kirrel1A* mRNA levels</u>: Each of the four retroviral modified cell lines appeared to follow a similar pattern of *kirrel1A* production; these mRNAs being relatively low on day 0 and becoming elevated on days 2 and 4 (figure 3.16). However, only the Pb line displayed a statistically significant increase in *kirrel1A* levels on days 2 and 4 (**0.91**±0.06 and **0.91**±0.04) relative to day 0 (**0.34**±0.4; $p<0.05$). *Kirrel1A* mRNA levels in the K1A myoblasts on days 0 and 4 (**0.92**±0.15 and **1.30**±0.05) were significantly elevated when compared to the control values (**0.34**±0.4 and **0.91**±0.04; $p<0.05$), increases of 274% and 143% (figure 3.16.A). Although not significant, it also appeared as if *kirrel1A* levels were increased relative to the control on day 2 in the K1A cell line. No statistically significant differences in *kirrel1A* levels relative to the control were seen when looking at the K1B and mCherry cell lines.

<u>*Kirrel1B* mRNA levels:</u> *Kirrel1B* mRNA levels in the Pb control, K1A and mCherry cell lines all remained relatively low throughout differentiation with no statistically significant differences seen between the groups or between the timepoints when looking at the individual cell lines (figure 3.16.B). *Kirrel1B* mRNA levels in the K1B cell line however were significantly increased relative to the control on days 0, 2 and 4

with increases of 420%, 623% and 840% (figure 3.16.B). No statistically significant differences in *kirrel1B* levels were seen between the timepoints for the K1B cell line.

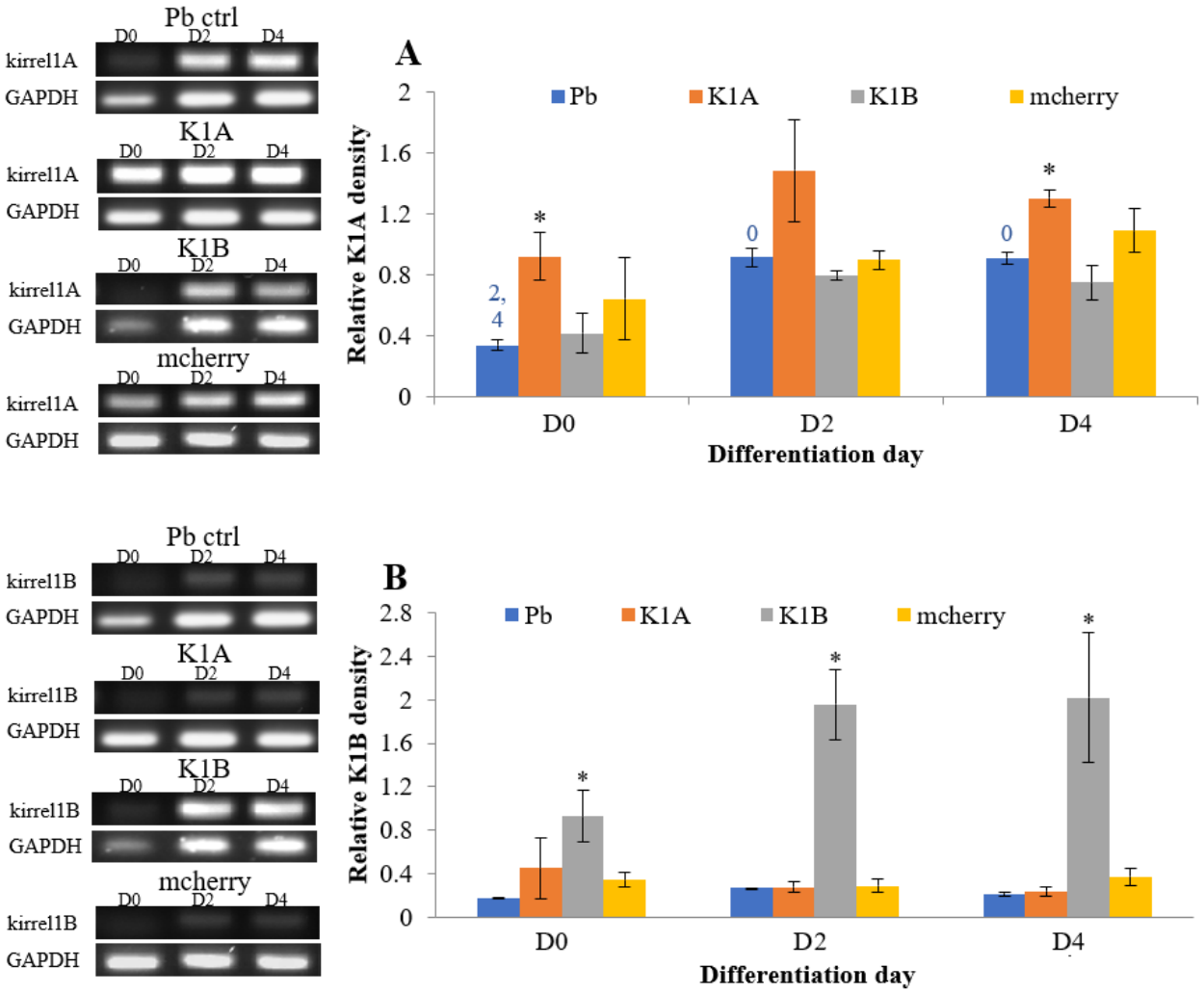


Figure 3.16 *Kirrel1* mRNA levels in differentiating C2C12s stably overexpressing the *kirrel1* variants. *Kirrel1A* (A) and *kirrel1B* (B) mRNA levels in differentiating C2C12s were determined via semi-quantitative PCRs. Control C2C12s and the K1A, K1B and mCherry cell lines were differentiated *in vitro* over 4 days with RNA collected on days 0, 2 and 4. Densitometry analysis of agarose gels containing *kirrel1* PCR products was carried out using ImageJ and the relative *kirrel1* band intensities determined using *gapdh* as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 *kirrel1* mRNA levels at individual timepoints were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in *kirrel1* mRNA levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 2, $p<0.05$ when compared to D2; 4, $p<0.05$ when compared to D4.

3.3.2 The overexpression of either kirrel1A or kirrel1B has little effect on C2C12 myotube formation in vitro while the expression of an mCherry-kirrel1A mutant completely inhibited myotube formation

The Pb control, K1A, K1B and mCherry cell lines were differentiated over 4 days *in vitro* with photographs taken on each day and myotube analysis carried out as described in section 3.1.5. No statistically significant changes in myotube coverage and size were seen in the K1A and K1B lines relative to the Pb control (figure 3.17); all 3 of these cell lines produced typical myotubes from day 2 onwards. Unexpectedly, the mCherry

mutant cell line displayed severely inhibited myotube formation with almost no myotubes being produced at all (figure 3.17). The mCherry myotubes that did form were much smaller than those of the control and only appeared from day 3 onwards. On day 4, the mCherry total tube coverage value was decreased by 2900% when compared to the control value.

It is possible that the modifications to the stably expressed *kirrel1A* variant in the mCherry cell line resulted in abnormal intracellular signalling/interactions and/or a loss of function during differentiation. The inability of the mCherry cell line to successfully differentiate and fuse was not predicted since the fluorescent mCherry insert (236 aa total, 26.7 kDa) is located between amino acids Val575 and Asn576 of kirrel1A (Arif *et al.*, 2014; Durcan *et al*., 2013); a location within the cytoplasmic tail (aa521-757) and upstream of any predicted tyrosine phosphorylation sites (31 predicted sites within the kirrel1 cytoplasmic tail according to the PTMcode 2 software, ptmcode.embl.de). Furthermore, this kirrel1A-mCherry mutant has been shown to localize correctly in human podocytes; the binding of Myoc1 to the cytoplasmic, C-terminal tail of kirrel1 (an interaction important for kirrel1 localization at the membrane) was unaffected by the addition of the mCherry insert (Arif *et al.*, 2014). However, any changes to other intracellular signalling/interactions caused by this modification have not been assessed in any previous publications with this study being the first time this plasmid has been used in SkM myoblasts. Although the mCherry insert was not predicted to have directly interrupted any sites that would normally undergo any post-translational modifications (PTMs), Solanki *et al.*, (2019) showed that a mutation to Ser573 in the human kirrel1 protein resulted in defective localization and function of the molecule in human podocytes *in vitro* and *in vivo* (Ser573 is also present in the mouse kirrel1 protein). However, the mechanism of this inhibited localization and functionality is unknown (Solanki *et al.*, 2019). Kirrel1 molecules with this mutation at Ser573 were seen to accumulate in early endosomes, the Golgi apparatus and lysosomes (instead of the cell membrane), resulting in a loss of intracellular junctions (Solanki *et al.*, 2019). It is possible that the mCherry insert at Val575 may have interrupted the functioning of the Ser573 in the kirrel1 molecule. Additionally, the mCherry insert is within 80 amino acids of two motifs that are predicted to undergo PTMs required for the regulation of protein activity. Tyr637 and Tyr638 are likely phosphorylated by the tyrosine-protein kinase Fyn, leading to the binding of SH2-SH3 domain containing proteins such as Grb2 (actin polymerization in podocytes), Nck1 and Nck2 (Dreadlock (Dck), the *Drosophila* paralog of the Nck proteins, is involved in actin reorganization in myoblasts during fusion) (Sellin *et al.*, 2003; Durcan *et al.*, 2013; Deng *et al.*, 2017; Paul *et al.*, 2022; uniprot.org database, query code Q80W68). A glycosylation site is also present at Asn503; such glycosylation sites are critical for the regulation of kirrel protein activity (Yesildag *et al.*, 2015; uniprot.org database, query code Q80W68). Although the distance between these motifs and the mCherry insert would suggest unhindered PTM, confirmational/folding changes caused by the addition of

the mCherry protein may have altered these PTMs as well as the binding of partner proteins close to this site e.g., the SH2-SH3 domain-containing proteins. Additionally, since the *kirrel1A* mRNA used in the production of the *kirrel1A-mCherry* plasmid was cloned from rat podocytes, it is possible that tissue or species-specific variations in *kirrel1A* mRNA sequence may have led to its inability to function in the mouse-muscle derived C2C12s. However, this hypothesis appears unlikely as a number of publications have demonstrated the high level of evolutionary conservation between *kirrel* paralog genes; this having been demonstrated with cross species, ectopic studies e.g., the ectopic expression of the mammalian kirrel proteins is able to compensate for the deletion of their paralogs in both *Drosophila* (Duf and Rst) as well as in *C. elegans* (SYG-1) (Neumann-Haefelin *et al*., 2010; Wanner *et al.*, 2011; Helmstädter *et al*., 2012).

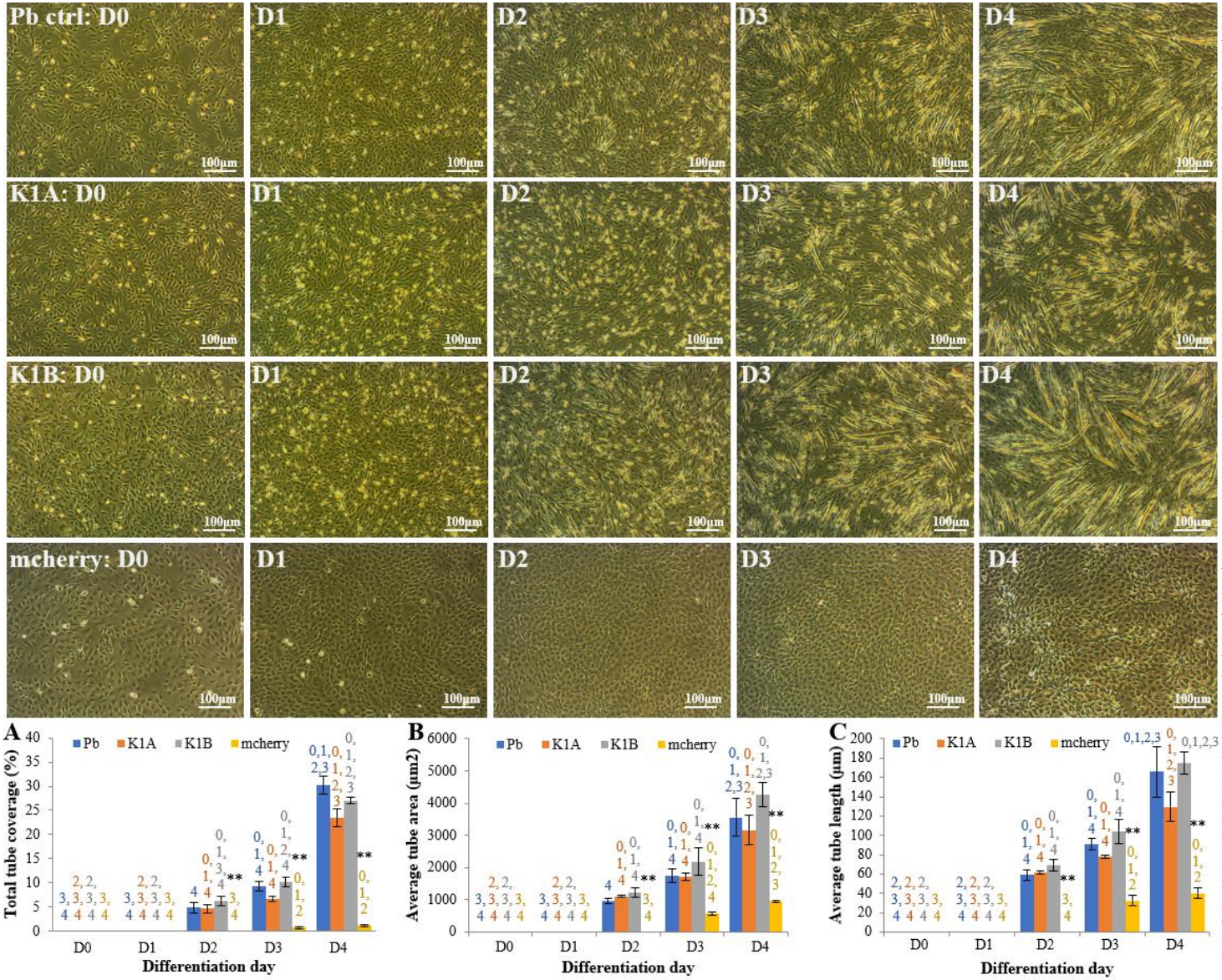


Figure 3.17 Myotube production by differentiating C2C12s stably overexpressing *kirrel1* variants. The Pb control, K1A, K1B and mCherry cell lines were differentiated *in vitro* over 4 days with images aquired daily at 40x magnification. A) Total myotube coverage determined using an in-house ImageJ macro (appendix X). B) Average individual tube areas were determined by manually outlining the myotube structures using ImageJ and recording the area values. C) Average individual tube length determined by manually outlining the myotube structures using ImageJ and recording the Feret values. Data expressed as mean ±SEM; n=3. Statistically significant differences between between control and modified C2C12 myotube values at each timepoint were determined using an unpaired, 2-tailed Student's t-test. *p<0.05, **p<0.005. Statistically significant differences in myotube values between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, p<0.05 when compared to D0; 1, p<0.05 when compared to D1; 2, p<0.05 when compared to D2; 3, p<0.05 when compared to D3; 4, p<0.05 when compared to D4.

3.3.3 The introduction of the kirrel1A or kirrel1B coding sequence to the C2C12 genome via retroviral transduction resulted in increased kirrel1A and kirrel1B protein levels. Kirrel1B overexpression also increased kirrel1A protein levels. Large kirrel1-containing complexes accumulated in differentiating control, kirrel1A overexpressing and kirrel11B overexpressing myoblasts but not in the kirrel1A-mCherry expressing cell line

No statistically significant fluctuations in kirrel1A (125 kDa) protein levels were seen in any of the four retroviral cell lines during the differentiation and fusion process (figure 3.18). However, while not significant, kirrel1A protein levels in both the Pb control and mCherry lines appeared to become elevated from day 2 onwards; increases that are congruent with the mRNA PCR results (figure 3.16).

Kirrel1A protein levels: Kirrel1A expression in the K1A cell line was significantly increased on days 0 and 1 (**1.01**±0.17 and **0.87**±0.12) when compared to the Pb line (**0.41**±0.11 and **0.52**±0.04, $p<0.05$), increases of 147% and 68%; the Pb kirrel1A values reached those of the K1A line by day 2. Unexpectedly, kirrel1A levels in the K1B line displayed statistically significant increases on each day of differentiation (D0-D4; **1.20**±0.24, **1.44**±0.33, **1.46**±0.28, **1.19**±0.20 and **0.84**±0.13) relative to the Pb control (**0.41**±0.11, **0.52**±0.04, **0.68**±0.07, **0.77**±0.16 and **0.59**±0.11; $p<0.05$), increases of 192%, 178%, 114%, 54% and 42%; these increases were not mirrored in the mRNA PCRs which showed no change in *kirrel1A* mRNA levels in the K1B line relative to the control (figure 3.16). Kirrel1A levels in the K1B line on day 2 (**1.46**±0.28) also displayed a statistically significant increase relative to the K1A line (**0.78**±0.07; $p<0.05$), an increase of 65%. Kirrel1A in the previously covered K1CRISPR cell line (did not form myotubes) were determined to be increased 102%, 43%, 114% and 76% on days 0, 1, 3 and 4 respectively when compared to its unmodified control. These increases are similar in scale to those seen in the K1A line (successfully produced myotubes), suggesting that this level of kirrel1A expression was unlikely the cause of the inhibited myogenesis displayed by the K1CRISPR line.

The kirrel1A (125 kDa) band seen in the mCherry blot likely represents the wild-type version of the protein and showed no increases in intensity vs. the Pb control. The fluorescent tag within the kirrel1A-mCherry protein expressed by the mCherry cell line was 26 kDa, however, no immunoreactive band was seen roughly 26 kDa above the wild-type 125 kDa band. It is possible that the fluorescent insert located between aa575-576 of the kirrel1 protein disrupted binding of the anti-kirrel1 antibody which recognizes a region between aa505-534. Furthermore, the levels of *kirrel1A* mRNA in the mCherry line were not elevated relative to the control (figure 3.16); it was therefore unclear to what degree the kirrel1A-mCherry mutant variant was expressed in the mCherry line. Fluorescent microscopy confirmed the expression of the fluorescent insert (appendix VII).

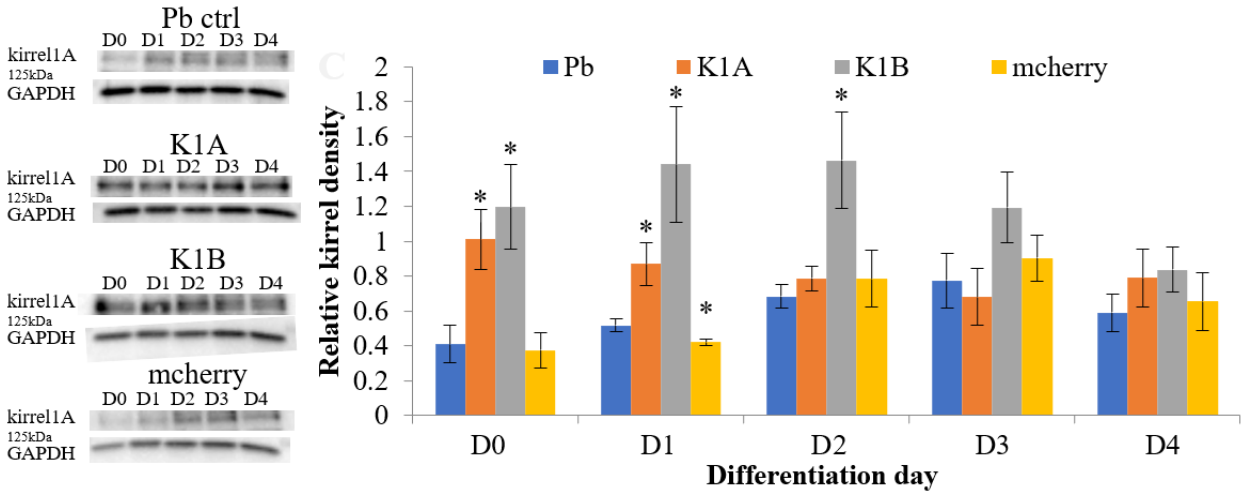


Figure 3.18 Kirrel1A (125 kDa) protein levels in differentiating C2C12s stably overexpressing *kirrel1* variants. Kirrel1A (125 kDa) levels in differentiating C2C12s were determined via western blotting under reducing conditions. The Pb, K1A, K1B and mCherry cell lines were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative kirrel1A (125 kDa) band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 kirrel1A (125 kDa) levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in kirrel1A (125 kDa) levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

Since there was no change in the levels of *kirrel1A* mRNA produced by the K1B line relative to control, it is possible that the upregulated kirrel1B protein promoted the build-up of kirrel1A e.g., via the alteration of kirrel1A trafficking and recycling (retention of kirrel1A at the cell membrane); the *Drosophila* paralogs of the kirrel molecules (Duf and Rst) have been shown to promote the transport of additional Duf and Rst proteins to the cell membrane via vesicle transport (Galletta *et al.*, 2004; Menon *et al.*, 2005).

As shown in figure 2.17, the increases in kirrel1A protein levels in the K1A and K1B lines appeared to have no effect the production of myotubes. A similar finding was made by Shilagardi *et al.*, (2013) who showed that the overexpression of Duf or Sns (Duf's binding partner) had no effect on the fusion of non-fusogenic *Drosophila* S2R+ cells (epithelial-derived). However, the expression of both Sns and the *C. elegans* fusion protein Eff-1 resulted in a 7-fold increase in fusion (relative to the expression of Eff-1 alone); the same was not found for Duf and Eff-1, although Duf expression did result in a significant increase in actin nucleation at the site of fusion (Shilagardi *et al.*, 2013).

Kirrel1B protein levels: All four of the retroviral cell lines produced relatively stable levels of kirrel1B (84 kDa) during differentiation with only the K1A line showing any statistically significant changes in expression between the timepoints; kirrel1B levels in the K1A line displayed a statistically significant drop on days 2 and 4 (**0.42**±0.03 and **0.46**±0.06) relative to day 0 (**0.88**±0.07). Kirrel1B levels in the K1B line (**1.69**±0.35, **1.42**±0.16, **1.38**±0.26, **1.54**±0.18 and **1.09**±0.02) showed statistically significant increases on each day of relative to the Pb control (**0.58**±0.08, **0.51**±0.05, **0.50**±0.13, **0.64**±0.13 and **0.58**±0.09; $p<0.05$); corresponding to increases of 189%, 179%, 179%, 139% and 87% on days 0, 1, 2, 3 and 4 respectively. However, unlike the K1CRISPR line which also expressed high levels of kirrel1B, the K1B line successfully produced myotubes (figure 3.17). However, the degree to which kirrel1B expression was elevated in the K1B myoblasts (relative to Pb) was not as drastic as that seen in the K1CRISPR line which displayed increases of 568%, 480%, 758%, 649% and 912% on days 0, 1, 2, 3 and 4 respectively when compared to its control cell line (figure 3.5). The stark differences in kirrel1B expression between the K1CRISPR and K1B lines could explain the contrasts in myotube formation.

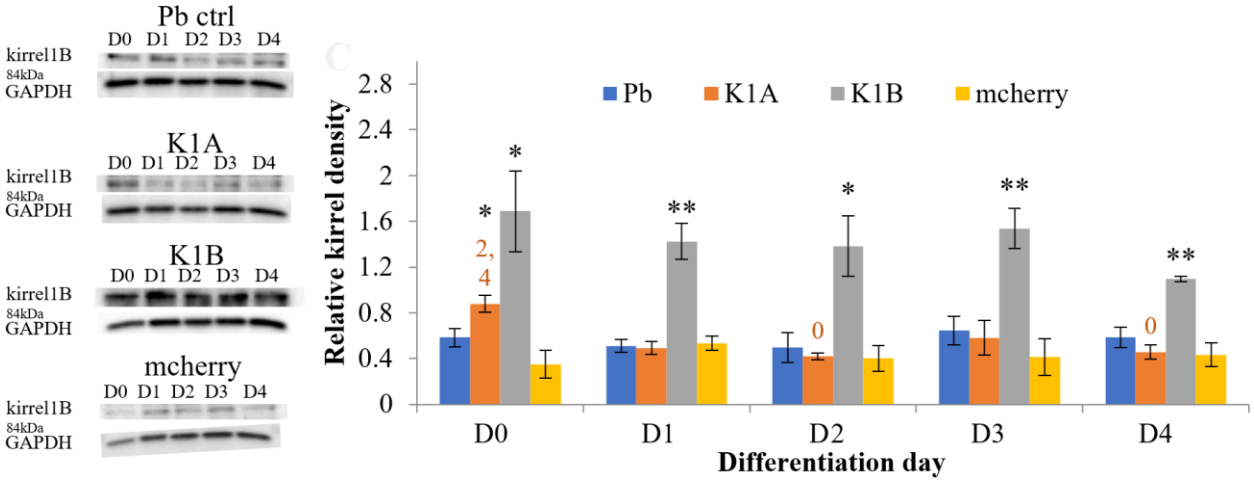


Figure 3.19 Kirrel1B (84 kDa) protein levels in differentiating C2C12s stably overexpressing *kirrel1* variants. Kirrel1B (84 kDa) levels in differentiating C2C12s were determined via western blotting under reducing conditions. The Pb, K1A, K1B and mCherry cell lines were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative kirrel1B (84 kDa) band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 kirrel1B (84 kDa) levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in kirrel1B (84 kDa) levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

Kirrel1 complex levels: Levels of the 350 kDa kirrel1-containing complex were seen to accumulate in the Pb control, K1A and K1B cell lines with peak levels being reached on day 4 (figure 3.20); all three cell lines successfully forming myotubes. As with the K1CRISPR and K1Ash lines (which also displayed inhibited myotube formation), the mCherry myoblasts did not produce increased levels of the complex as differentiation progressed with statistically significant decreases seen on days 3 and 4 (**0.59**±0.22 and **0.49**±0.18) relative to the Pb control (**1.73**±0.19 and **1.61**±0.19; $p<0.05$).

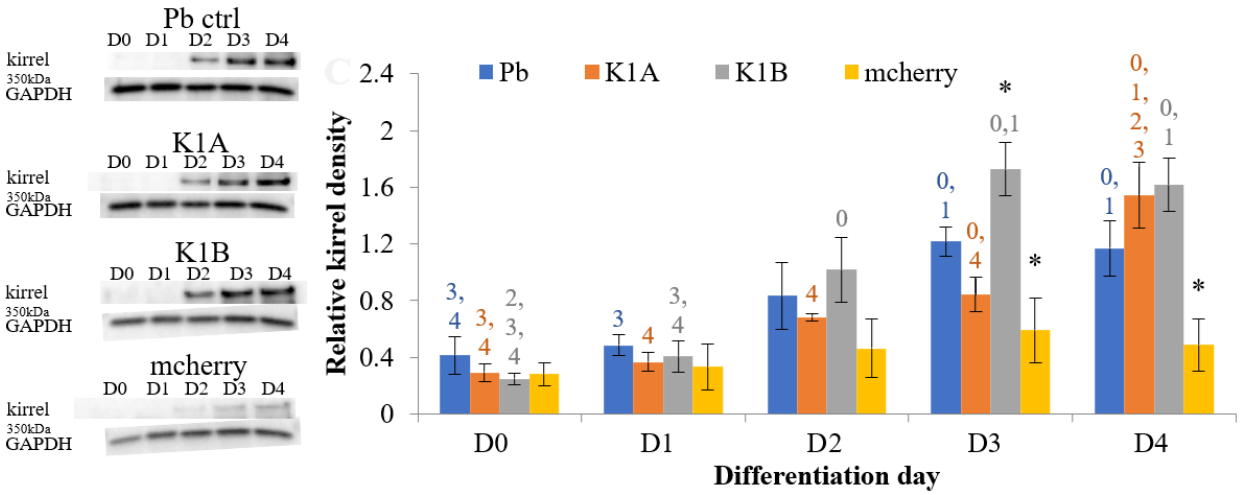


Figure 3.20 Kirrel1 multimer/complex (350 kDa) levels in differentiating C2C12s stably overexpressing *kirrel1* variants. Kirrel1 multimer/complex (350 kDa) levels in differentiating C2C12s were determined via western blotting under non-reducing conditions. The Pb, K1A, K1B and mCherry cell lines were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative kirrel1 multimer/complex (350 kDa) band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 kirrel1 multimer/complex (350 kDa) levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in kirrel1 multimer/complex (350 kDa) levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

Since overexpression of both kirrel1A and kirrel1B (both proteins in the K1B line) had little effect on the levels of this complex in any of the retroviral lines, it is possible that the limiting factor in its formation is a binding partner other than one of the kirrel1 variants e.g., the kirrel molecules preferred binding partner, nephrin. Therefore, in order to increase levels of this complex, it may be necessary to also overexpress nephrin. This being said, this dual-overexpression may still have little effect on cell fusion; as shown by Shilagardi *et al.*, (2013), overexpression of Sns, the *Drosophila* paralog of Nephrin, only enhanced fusion when expressed in conjunction with a fusion protein, however, this was carried out in non-fusogenic epithelial cells and may not be the case in myoblasts which are already prone to fuse.

3.3.4 The overexpression of kirrel1A and kirrel1B had little effect on the levels of Pax7 and Myf5 in differentiating myoblasts in vitro. Kirrel1B overexpression appears to promote earlier MyoD and myogenin upregulation. Expression of a kirrel1A-mCherry mutant resulted in lower levels of Pax7 and a delayed MyoD response and severely inhibited myogenin and MyHC production

Pax7 protein levels: Pax7 levels in the Pb control, K1A and K1B lines appeared to follow a typical pattern with peak expression being seen on day 0 and levels dropping from day 1 and reaching minimums by day 4 (figure 3.21.A). However, Pax7 levels in the mCherry line remained low throughout the differentiation process and most notably, on day 0 (**0.42**±0.16), the Pax7 level displayed a statistically significant drop relative to the control (**0.95**±0.14; $p<0.05$), a drop of 124%.

Myf5 protein levels: Myf5 expression in each of the retroviral cell lines remained relatively stable during differentiation (figure 3.21.B). A relatively minor but statistically significant increase in Myf5 expression was seen in the K1B line on day 1, however, overall, Myf5 expression appears unaffected in this cell line.

MyoD protein levels: Expression of MyoD in the Pb control, K1A and K1B lines followed the same general pattern with similar levels on each day of differentiation (Figure 3.21.C). However, a slight, statistically significant increase in MyoD was seen in the K1B line on day 1 (**1.35**±0.11) relative to the control (**0.93**±0.05; $p<0.05$). Perhaps as a result of this earlier MyoD expression, and as will be seen in the next figure (figure 3.22), the K1B line also appeared to upregulate myogenin earlier than the other cell lines. MyoD expression in the mCherry was drastically decreased on days 0, 1 and 2 (**0.29**±0.07, **0.16**±0.03 and **0.82**±0.43) relative to the control (**0.95**±0.09, **0.93**±0.05 and **1.74**±0.17), decreases of 222%, 471% and 113%. MyoD expression in the mCherry myoblasts starts to recover on D3 and approaches the control values by day 4.

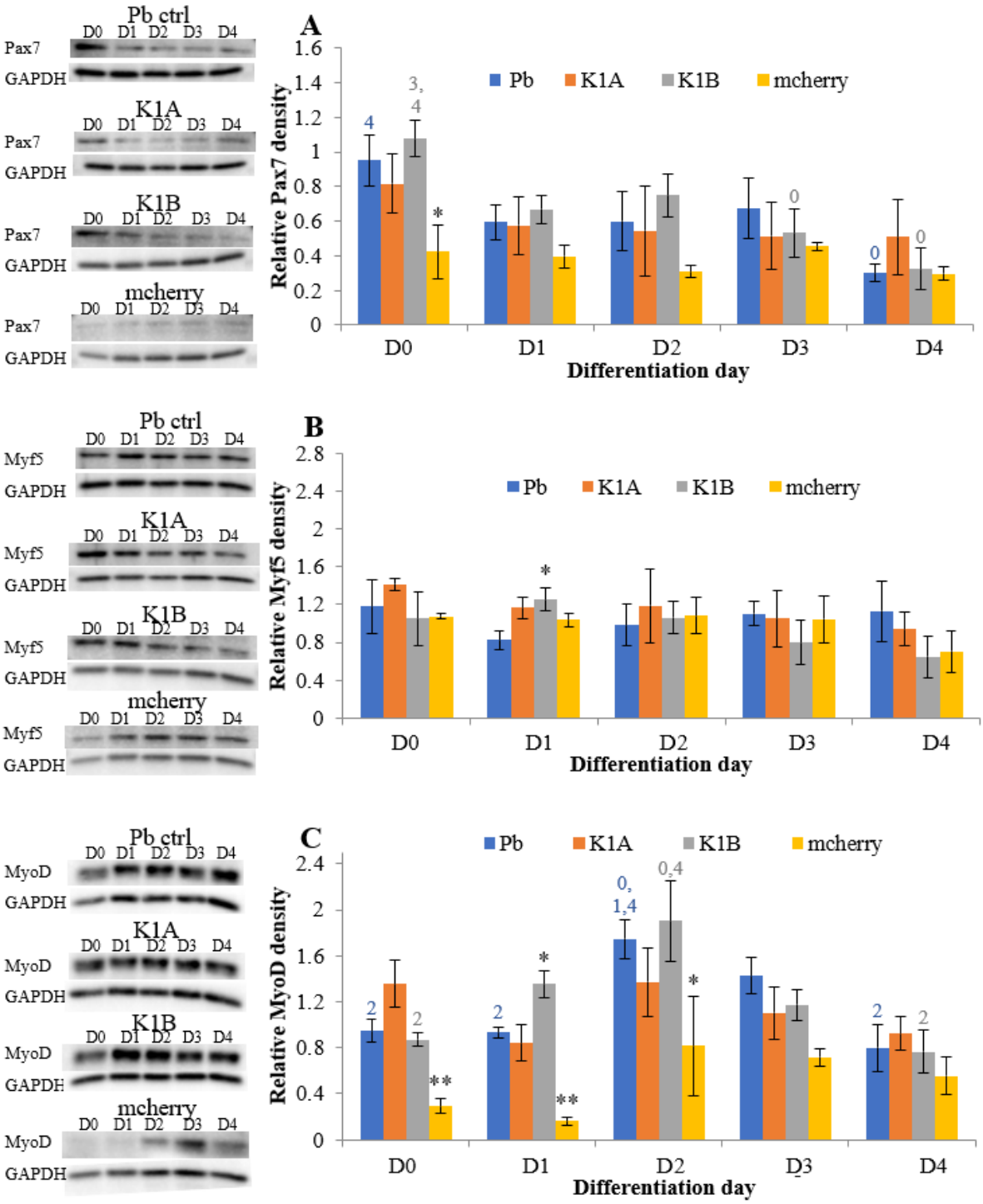


Figure 3.21 Pax7, Myf5 and MyoD levels in differentiating C2C12s stably overexpressing *kirrel1* variants. Pax7 (A), Myf5 (B) and MyoD (C) levels in differentiating C2C12s were determined via western blotting under non-reducing conditions. The Pb, K1A, K1B and mCherry cell lines were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative Pax7/Myf5/MyoD band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 Pax7/Myf5/MyoD levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in Pax7/Myf5/MyoD levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

Little research has been published regarding the kirrel proteins and MRF regulation (and vice versa) with no previous publications having looked at the kirrel1 variant in this regard. The most relevant finding so far has been made by Tamir-Livne *et al.,* (2017), who showed that kirrel3 expression is promoted by MyoD and that early MyoD expression appears to be low in kirrel3 knockdown myoblasts (the same was not seen in our *kirrel1* knockdown). Kirrel1 has also been shown to be a feedback regulator of hippo signalling in hek293 cells (kidney origin) (Paul *et al.,* 2022). Since the Hippo pathway directly modulates the transcription of *pax7*, *myf5* and *myod* in myoblasts (Yin *et al.,* 2013; Liu *et al.,* 2021; Paul *et al.,* 2022), it is possible that the modifications in the kirrel1A-mCherry mutant altered Pax7 and MyoD expression in a Hippo-dependent manner.

Myogenin protein levels: Myogenin levels in the Pb control, K1A and K1B lines were seen to peak between days 2 and 3 of differentiation (figure 3.22.A). Myogenin in the K1A line displayed a statistically significantly decrease on day 3 relative to the control but the reverse was true on day 4. This was unexpected as no effect on myotube production was seen in the K1A line. Myogenin in the K1B cell line appeared to become upregulated early with a statistically significant increase in its level on day 2 (**1.40**±0.24) when compared to control (**0.74**±0.10; $p<0.05$). This appeared to be the peak of myogenin expression in the K1B line; myogenin expression peaked on day 3 for the control. Perhaps as a result of the inhibited MyoD expression, myogenin levels in the mCherry line remained low throughout differentiation with statistically significant decreases seen on days 2, 3 and 4 (**0.05**±0.03, **0.29**±0.08 and **0.33**±0.06) relative to the control (**0.74**±0.10, **1.77**±0.11 and **0.71**±0.09; $p<0.05$).

MyHC protein levels: MyHC levels in the Pb control, K1A and K1B lines appeared healthy and peaked on day 4 in all three cell lines (figure 3.22). As expected, since no myotubes were produced, MyHC levels did not increase in the mCherry line.

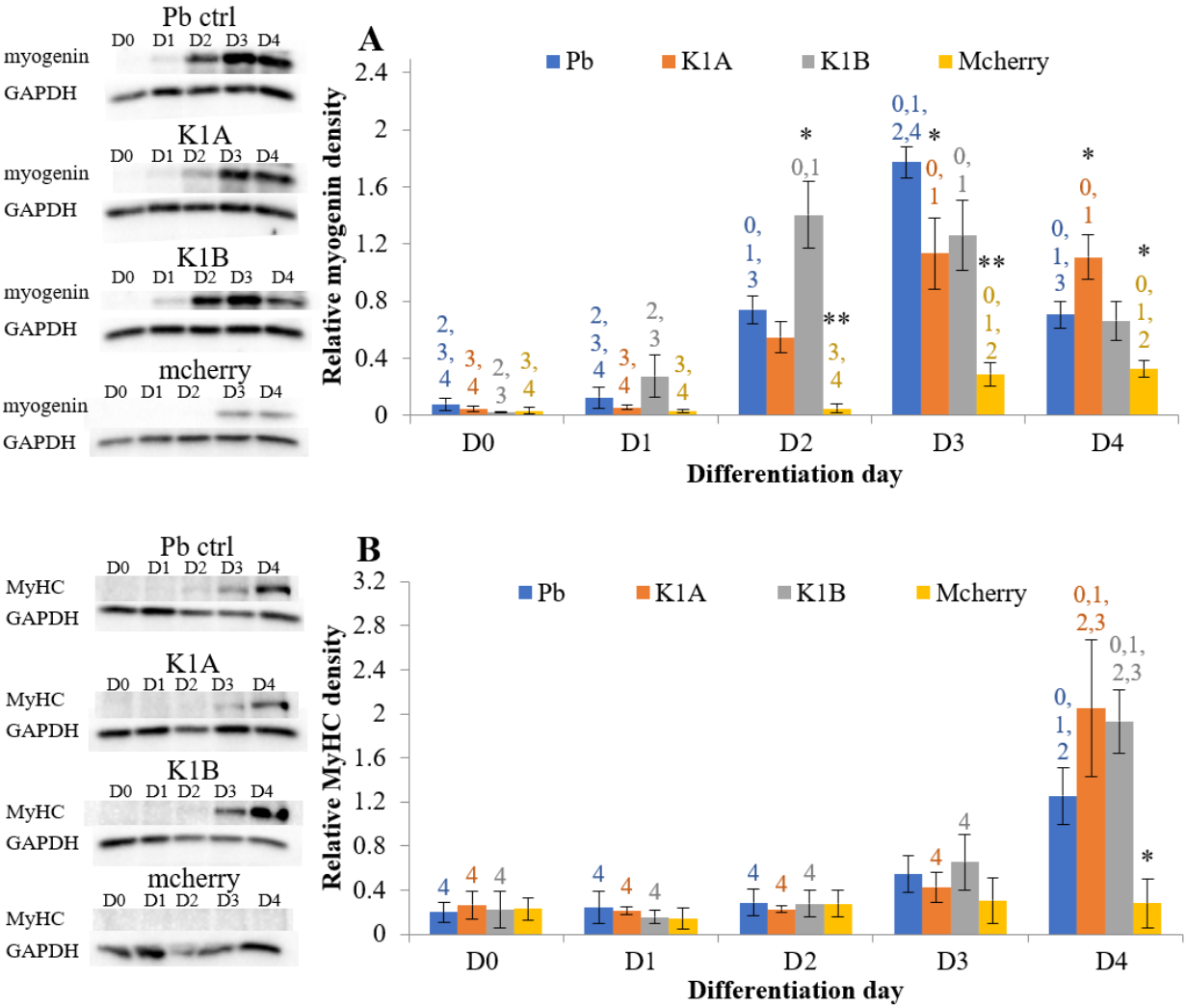


Figure 3.22 Myogenin and MyHC levels in differentiating C2C12s stably overexpressing *kirrel1* variants. Myogenin (A) and MyHC (B) levels in differentiating C2C12s were determined via western blotting under non-reducing conditions. The Pb, K1A, K1B and mCherry cell lines were differentiated *in vitro* over 4 days with protein lysates collected on each day. Densitometry analysis of the western blot images was carried out using ImageJ and the relative myogenin/MyHC band intensities determined using GAPDH as the loading control. Data expressed as mean ±SEM; n=3. Statistically significant differences between control and modified C2C12 myogenin/MyHC levels at each timepoint were determined using unpaired, 2-tailed Student's t-tests. *, $p<0.05$; **, $p<0.005$. Statistically significant differences in myogenin/MyHC levels between the different timepoints for individual cell lines were determined using one-way ANOVAs along with the Tukey's HSD post hoc test with differences indicated by coloured numbers. 0, $p<0.05$ when compared to D0; 1, $p<0.05$ when compared to D1; 2, $p<0.05$ when compared to D2; 3, $p<0.05$ when compared to D3; 4, $p<0.05$ when compared to D4.

CHAPTER FOUR

SYNTHESIS, CONCLUSIONS AND RECOMMENDATIONS

The novel results presented in this thesis clearly demonstrate the requirement for tightly-regulated kirrel1 activity during the production of mammalian SkM myotubes *in vitro*. Other than the work done by Tamir-Livne *et al.*, (2017) on kirrel3, this research appears to be the only other in which kirrel expression has been modified in mammalian myoblasts. Whilst the previous work knocked down kirrel3 using shRNA, the work reported here used 3 different approaches to better establish the individual roles of kirrel1A and 1B, to provide a more realistic understanding of the role of the kirrel1 variants.

***CRISPR/Cas9 modification*:** Looking at the CRISPR/Cas9 results in section 3.1, modifications to two predicted regulators of genomic *kirrel1* transcription (ENCODE-EM10E0734476 and an RSINE1 element) lead to aberrant levels of kirrel1A and kirrel1B protein in the modified C2C12s, resulting in an almost complete inhibition of myogenesis. This inhibition did not appear to be as a result of modified early MRF expression as the levels of Pax7, My5 and MyoD were unaffected by these modifications. However, myogenin expression was dramatically delayed in this cell line. The delay in myogenin expression may have been as a result of modified cell-cell adhesion/interactions, caused by aberrant kirrel1 levels on the cell surface; nuanced regulation of adhesion molecules is critical for successful terminal differentiation, alignment and fusion. Kirrel1A levels in the CRISPR modified myoblasts were higher than those in the relevant control but were similar to those seen in the kirrel1A overexpressing line (section 3.3). Notably, kirrel1B levels in the CRISPR-modified line were particularly high, even when compared to the kirrel1B overexpressing cell line (section 3.3). Since both the kirrel1A and kirrel1B retroviral-overexpressing myoblast lines still produced healthy myotubes, it would seem that the relatively extreme levels of kirrel1B protein in the CRISPR modified line prevented fusion; potentially due to competitive inhibition (with other kirrel variants or nephrin) or by increased steric hindrance at the cell surface. As the fusion process progresses, kirrel proteins are known to be shuttled away (in vesicles) from the direct site of fusion pore formation or moves outwards away from site as part of the expanding (the ring-shaped structure composing largely of the kirrel proteins that is essential for actin regulation at the fusion site). This process is required for membrane integration and pore formation (Abmayr and Pavlath, 2012; Kim *et al.*, 2015). In *C. elegans*, the in *trans* binding of membrane-embedded SYG-1 and SYG-2 molecules (paralogs of the kirrel proteins and nephrin respectively) has been shown to form a rigid L-shaped protein structure, holding the opposing cell membranes in place together at a distance of 45nm; this distance is too large to allow for fusion pore formation (1.2nm) (Abmayr and Pavlath, 2012; Kim *et al.*, 2015). It is possible that the greatly

overexpressed kirrel1B in the CRISPR line could not be transported away from the site at an appropriate rate and thus prevented closer membrane proximity and fusion.

The levels of *kirrel1B* mRNA in unmodified, wild-type C2C12s were seen to be much lower than those of *kirrel1A* on all days of differentiation. However, the same was not for protein expression as both kirrel1A and kirrel1B reached similar protein levels. This discrepancy may be a result of different mechanisms controlling the translation of the two *kirrel1* splice variants that were investigated here.

***Kirrel1A and kirrel1B* knockdown:** The knockdown of *kirrel1A* translation (by shRNAs) (section 3.2) resulted in the severe inhibition of myotube formation that also appeared to be unrelated to the expression levels of the early stage MRFs (Pax7, Myf5 and MyoD); similar to what was seen in the CRISPR modification experiments. Myogenin expression was also delayed in the knockdown line, potentially due to modified cell surface interactions during terminal differentiation and/or altered intracellular kirrel1A signalling/activity. Although kirrel1A expression in the knockdown cell line appeared to recover by day 3 of differentiation (reaching similar levels to the control), this recovery did not rescue myotube production. This suggests the critical role for kirrel1A during the stages leading up to fusion *in vitro* (beginning with cell adhesion). It is possible that by the time kirrel1A recovered in the cell line, the *in vitro* conditions had become unsuitable for fusion e.g., cell confluence becoming too high. Furthermore, cell-cell interactions/bonds that had become established by day 3 (not involving kirrel1A) may have inhibited any interactions that would ordinarily be initiated by the kirrel1A molecules during adhesion, differentiation and fusion. These findings, along with those by Tamir-Livne *et al.*, (2017) (kirrel3 knockdown inhibited myoblast fusion in MRF-independent manner), suggest that although all kirrel proteins (kirrel1, kirrel2 and kirrel3) are derived from the same ancestral gene, the molecules may not be able to compensate for the knockdown/knockout of fellow family members. This would point to unique roles for the different kirrel proteins during myogenesis.

***Retroviral overexpression of kirrel1A, kirrel1B and mutant variant*:** The retroviral overexpression of kirrel1A and kirrel1B (section 3.3) appeared to have little effect on the differentiation and fusion of C2C12 myoblasts *in vitro*. Our findings are consistent with those by Shilagardi *et al.*, (2013), who showed that neither the overexpression of the *Drosophila* kirrel paralog, Duf, nor the overexpression of Sns (nephrin paralog), were able to initiate cell fusion in non-fusogenic *Drosophila* epithelial cells. However, the overexpression of Sns and an ectopic fusion protein was shown to increase fusion by 7-fold when compared to the overexpression of the fusion protein alone (Shilagardi *et al.*, 2013). The same was not seen with the overexpression of Duf and the fusion protein; perhaps suggesting that in this case, wild-type Sns levels were a limiting factor (Shilagardi *et al.*, 2013). These results show that in non-fusogenic cell lines, the

overexpression of the kirrel paralogs alone is not enough to initiate fusion; however, the same has not been demonstrated in fusogenic mammalian cells. Considering both the findings in this thesis and those by Shilagardi *et al*, (2013), it is possible that, had we overexpressed the *kirrel1* variants along with nephrin and the fusion proteins Myomaker and Myomixer, fusion of the C2C12 myoblasts would have been significantly increased.

While the overexpression of the wild-type kirrel1 variants had little effect on myogenesis, the expression of a kirrel1A-mCherry mutant (section 3.3) dramatically reduced differentiation and myotube production. The early expression of Pax7 (day 0) and MyoD (days 0, 1 and 2) was significantly lower in the kirrel1A-mCherry mutant line when compared to the control. Perhaps as a result, the myogenin expression in this cell line was delayed and low on each day. It appeared that modification of the kirrel1A protein i.e., the addition of a fluorescent mCherry insert between Val-575 and Asn-576 of the kirrel1A molecule, had modified myogenesis in an MRF dependent manner. Although this site was not predicted to have any tyrosine phosphorylation sites or other active motifs/binding sites, conservation of function in this mutant has not been fully assessed with only its localization having been shown to be standard in human podocytes (Arif *et al.*, 2014). Furthermore, prior to our study, this kirrel1A-mCherry mutant had not been expressed in myoblasts and therefore any effects are unknown. Since it has been shown that wild-type kirrel1A is a feedback regulator of the partially defined Hippo pathway in Hek293 cells (Paul *et al.*, 2022) and that the pathway directly modulates the transcription of *pax7*, *myf5* and *myod* in myoblasts (Yin *et al.*, 2013; Liu *et al.*, 2021; Paul *et al.*, 2022); it is possible that expression of the kirrel1A-mCherry mutant protein modified Hippo signalling in the cell line, resulting in the abnormal expression of the MRFs.

***Formation of kirrel1 complexes during* myogenesis:** The current study has for the first time shown that large 350 kDa kirrel1-containing complexes accumulate in successfully differentiating and fusing myoblasts/myotubes; these complexes were not seen to accumulate in any of the cell lines that failed to form myotubes in this study i.e., the CRISPR modified, *kirrel1A* knockdown and mCherry mutant cell lines. It was unclear whether these complexes contained kirrel1A, kirrel1B or a combination of both; due to the antibody recognizing both variants. Furthermore, it is likely that these complexes would have also contained kirrel1's preferred binding partner nephrin, however, this was not demonstrated in this study.

***Conclusion and recommendations*:** Collectively, the results in this thesis demonstrate the need for kirrel1 activity during C2C12 myotube formation *in vitro*. Specifically, *kirrel1A* expression is required for successful myoblast fusion while the requirement for kirrel1B during this process remains unclear. Furthermore, the moderate overexpression of either kirrel1A or kirrel1B appeared to have little effect on fusion whereas the excessive expression of kirrel1B resulted in an almost complete inhibition of fusion; not

due to altered MRF expression but potentially due to increased steric hinderance at the cell surface (the same is not known for kirrel1A). Moreover, our results show that modifications to the cytoplasmic tail of the kirrel molecules (in the kirrel1A-mCherry mutant), even close to the membrane embedded region and away from any predicted active sites, risks disrupting kirrel activity and can result in significantly altered cell behaviour; in the current study, by modifying MRF expression.

The findings in this thesis highlight the need for further research into the kirrel family of proteins and cell fusion as the exact mechanisms of kirrel activity are still unclear. Since the process of cell fusion is poorly understood in all mammalian cell types, any findings regarding the kirrel proteins and myoblast fusion may also provide valuable insights into the mechanisms regulating this process in other types of fusion-capable cells e.g., cancer cells, epithelial cells liver cells, macrophages, trophoblasts. With regards to muscle tissue, a better understanding of fusion will allow for the development of interventions for diseases and conditions that result in impaired myogenesis e.g., cachexia, dystrophies and rhabdosarcomas.

With reference to the current study as well as to previous publications, a number of recommendations can be made regarding future work on the kirrel proteins and myogenesis. Firstly, in these experiments, *kirrel1* gene transcription was not fully knocked-out, but rather was modified with CRISPR and separately knocked down with shRNAs. A full gene knockout experiment, assuming similar results to those presented here, would further demonstrate the need for kirrel1 during myogenesis. Furthermore, in the experiments presented here, it proved difficult to knockdown *kirrel1B* levels and alternative silencing strategies may need to be employed to better define the role of this splice variant. More comprehensive investigations that look at all three kirrel isoforms i.e., kirrel1, kirrel2 and kirrel3 (as well as their splice variants), are required to determine the potentially unique roles of the proteins during fusion. These investigations would involve cell lines that have had multiple kirrel isoforms modified/knocked-out. Furthermore, it seems likely that to enhance fusion, multiple proteins involved in the process would need to be overexpressed/modified in the same cell line at specific times during fusion e.g., a combination of the kirrel molecules, nephrin and Myomaker/Myomixer; similar to the work by Shilagardi *et al.*, (2013). Determining how these actin nucleating and hemifusion proteins work together during myogenesis may help to develop expression strategies to enhance myotube production *in vitro* and *in vivo* for therapeutic interventions e.g., priming myoblasts and tissue for implantation *ex vivo* (cell therapies) as well as by direct *in vivo* genetic modification (gene therapies).

The localization of the kirrel proteins (and their partners such as nephrin) during fusion has not been extensively studied in mammalian cells. Further fluorescent microscopy analysis is required to evaluate intracellular kirrel transport (in vesicles), protein interactions/co-localization as well as to track kirrel-

mediated actin polymerization during fusion. Such studies would ideally confirm whether, like in the fly, if the kirrel molecules form a FuRMAS structure as well as whether the activity of the kirrel proteins is required for the formation of invasive, actin-driven protrusions during fusion. Additionally, fluorescent imaging may help determine the relationship between the kirrel molecules and Myomaker and Myomixer. A number of publications have highlighted the critical need for the *Drosophila* kirrel paralogs (Duf and Rst) for actin polymerization; however, very is little is known about whether these actin-regulating mechanisms are conserved in vertebrates.

Here, the mCherry mutant results highlight the importance of the integrity of the kirrel intracellular domain for myogenesis and that the full list of signalling cascades initiated by these molecules is not yet known; our study appears to be the first to directly link kirrel expression to MRF levels. A full range of signalling cascades initiated by the kirrel molecules and their splice variants is needed for a more holistic understanding of their functions.

Finally, determining the constituents and structure of the large kirrel1-containing complexes (accumulating in differentiating and fusing myoblasts) may reveal additional kirrel binding partners and give insights into membrane interactions on the cell surface. The makeup and structure of these complexes could be evaluated by use of x-ray crystallography, pull-down assays as well as a combination of affinity purification and mass spectrometry.

REFERENCES

Abmayr, S. M., & Pavlath, G. K. (2012). Myoblast fusion: lessons from flies and mice. *Development (Cambridge, England)*, *139*(4), 641–656.

Armand, A. S., Laziz, I., & Chanoine, C. (2006). FGF6 in myogenesis. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1763*(8), 773-778.

Asfour, H. A., Allouh, M. Z., & Said, R. S. (2018). Myogenic regulatory factors: The orchestrators of myogenesis after 30 years of discovery. *Experimental Biology and Medicine*, *243*(2), 118-128.

Beauregard, A., Curcio, M. J., & Belfort, M. (2008). The take and give between retrotransposable elements and their hosts. *Annual review of genetics*, *42*, 587–617.

Berger, S., Schäfer, G., Kesper, D. A., Holz, A., Eriksson, T., Palmer, R. H., Beck, L., Klämbt, C., Renkawitz-Pohl, R., & Onel, S. F. (2008). WASP and SCAR have distinct roles in activating the Arp2/3 complex during myoblast fusion. *Journal of cell science*, *121*(Pt 8), 1303–1313.

Bhat, N. R., Zhang, P., & Mohanty, S. B. (2007). p38 MAP kinase regulation of oligodendrocyte differentiation with CREB as a potential target. *Neurochemical research*, *32*(2), 293–302.

Blake, J. A., & Ziman, M. R. (2014). Pax genes: regulators of lineage specification and progenitor cell maintenance. *Development*, *141*(4), 737-751.

Bischoff, R. (1997). Chemotaxis of skeletal muscle satellite cells. *Developmental dynamics: an official publication of the American Association of Anatomists*, *208*(4), 505-515.

Bischoff, M. C., Lieb, S., Renkawitz-Pohl, R., & Bogdan, S. (2021). Filopodia-based contact stimulation of cell migration drives tissue morphogenesis. *Nature communications*, *12*(1), 791.

Bothe, I., Deng, S., & Baylies, M. (2014). PI(4,5)P2 regulates myoblast fusion through Arp2/3 regulator localization at the fusion site. *Development (Cambridge, England)*, *141*(11), 2289–2301.

Bour, B. A., Chakravarti, M., West, J. M., & Abmayr, S. M. (2000). Drosophila SNS, a member of the immunoglobulin superfamily that is essential for myoblast fusion. *Genes & development*, *14*(12), 1498–1511.

Brack, A. S., Conboy, I. M., Conboy, M. J., Shen, J., & Rando, T. A. (2008). A temporal switch from notch to Wnt signalling in muscle stem cells is necessary for normal adult myogenesis. *Cell stem cell*, *2*(1), 50-59.

Brunetti, T. M., Fremin, B. J., & Cripps, R. M. (2015). Identification of singles bar as a direct transcriptional target of Drosophila Myocyte enhancer factor-2 and a regulator of adult myoblast fusion. *Developmental biology*, *401*(2), 299–309.

Brzóska, E., Bello, V., Darribère, T., & Moraczewski, J. (2006). Integrin alpha3 subunit participates in myoblast adhesion and fusion in vitro. *Differentiation; research in biological diversity*, *74*(2-3), 105–118.

Cadot, B., Gache, V., Vasyutina, E., Falcone, S., Birchmeier, C., & Gomes, E. R. (2012). Nuclear movement during myotube formation is microtubule and dynein dependent and is regulated by Cdc42, Par6 and Par3. *EMBO reports*, *13*(8), 741–749.

Cardasis, C. A., & Cooper, G. W. (1975). An analysis of nuclear numbers in individual muscle fibers during differentiation and growth: A satellite cell- muscle fiber growth unit. *Journal of Experimental Zoology*, *191*(3), 347-357.

Cazzato, D., Assi, E., Moscheni, C., Brunelli, S., De Palma, C., Cervia, D., Perrotta, C., & Clementi, E. (2014). Nitric oxide drives embryonic myogenesis in chicken through the upregulation of myogenic differentiation factors. *Experimental cell research*, *320*(2), 269–280.

Chal, J., & Pourquié, O. (2017). Making muscle: skeletal myogenesis *in vivo* and *in vitro*. *Development (Cambridge, England)*, *144*(12), 2104–2122.

Chan, C. J., & Hiiragi, T. (2017). Keeping in Touch to Differentiate. *Developmental Cell*, *43*(2), 113-114.

Chanoine, C., Della Gaspera, B., & Charbonnier, F. (2004). Myogenic regulatory factors: redundant or specific functions? Lessons from Xenopus. *Developmental dynamics: an official publication of the American Association of Anatomists*, *231*(4), 662-670.

Chargé, S. B., & Rudnicki, M. A. (2004). Cellular and molecular regulation of muscle regeneration. *Physiological reviews*, *84*(1), 209-238.

Chen, A. E., Ginty, D. D., & Fan, C. M. (2005). Protein kinase A signalling via CREB controls myogenesis induced by Wnt proteins. *Nature*, *433*(7023), 317–322.

Chen, E. H., Pryce, B. A., Tzeng, J. A., Gonzalez, G. A., & Olson, E. N. (2003). Control of myoblast fusion by a guanine nucleotide exchange factor, loner, and its effector ARF6. *Cell*, *114*(6), 751–762.

Chen, J. F., Tao, Y., Li, J., Deng, Z., Yan, Z., Xiao, X., & Wang, D. Z. (2010). microRNA-1 and microRNA-206 regulate skeletal muscle satellite cell proliferation and differentiation by repressing Pax7. *The Journal of cell biology*, *190*(5), 867–879.

Chen, S. E., Jin, B., & Li, Y. P. (2007). TNF-α regulates myogenesis and muscle regeneration by activating p38 MAPK. *American Journal of Physiology-Cell Physiology*, *292*(5), C1660-C1671.

Choo, H. J., Canner, J. P., Vest, K. E., Thompson, Z., & Pavlath, G. K. (2017). A tale of two niches: differential functions for VCAM-1 in satellite cells under basal and injured conditions. *American journal of physiology. Cell physiology*, *313*(4), C392–C404.

Comai, G., & Tajbakhsh, S. (2014). Molecular and cellular regulation of skeletal myogenesis. *Current topics in developmental biology*, *110*, 1-73.

Cornelison, D. D. W., & Wold, B. J. (1997). Single-cell analysis of regulatory gene expression in quiescent and activated mouse skeletal muscle satellite cells. *Developmental biology*, *191*(2), 270-283.

Cretoiu, D., Pavelescu, L., Duica, F., Radu, M., Suciu, N., & Cretoiu, S. M. (2018). Myofibers. *Muscle Atrophy*, 23-46.

Dee, K., Freer, M., Mei, Y., & Weyman, C. M. (2002). Apoptosis coincident with the differentiation of skeletal myoblasts is delayed by caspase 3 inhibition and abrogated by MEK-independent constitutive Ras signalling. *Cell death and differentiation*, *9*(2), 209–218.

Deng, S., Azevedo, M., & Baylies, M. (2017). Acting on identity: Myoblast fusion and the formation of the syncytial muscle fiber. *Seminars in cell & developmental biology*, *72*, 45–55.

Deng, S., Bothe, I., & Baylies, M. K. (2015). The Formin Diaphanous Regulates Myoblast Fusion through Actin Polymerization and Arp2/3 Regulation. *PLoS genetics*, *11*(8), e1005381.

Desouza, M., Gunning, P. W., & Stehn, J. R. (2012). The actin cytoskeleton as a sensor and mediator of apoptosis. *Bioarchitecture*, *2*(3), 75–87.

Dong, L. Q., & Liu, F. (2005). PDK2: the missing piece in the receptor tyrosine kinase signalling pathway puzzle. *American journal of physiology. Endocrinology and metabolism*, *289*(2), E187–E196.

D'Souza-Schorey, C., & Chavrier, P. (2006). ARF proteins: roles in membrane traffic and beyond. *Nature reviews. Molecular cell biology*, *7*(5), 347–358.

Duan, R., Kim, J. H., Shilagardi, K., Schiffhauer, E. S., Lee, D. M., Son, S., Li, S., Thomas, C., Luo, T., Fletcher, D. A., Robinson, D. N., & Chen, E. H. (2018). Spectrin is a mechanoresponsive protein shaping fusogenic synapse architecture during myoblast fusion. *Nature cell biology*, *20*(6), 688–698.

Durcan, P. J., Al-Shanti, N., & Stewart, C. E. (2013). Identification and characterization of novel Kirrel isoform during myogenesis. *Physiological reports*, *1*(3), e00044.

Estrada, B., Maeland, A. D., Gisselbrecht, S. S., Bloor, J. W., Brown, N. H., & Michelson, A. M. (2007). The MARVEL domain protein, Singles Bar, is required for progression past the pre-fusion complex stage of myoblast fusion. *Developmental biology*, *307*(2), 328–339.

Faralli, H., & Dilworth, F. J. (2012). Turning on myogenin in muscle: a paradigm for understanding mechanisms of tissue-specific gene expression. *Comparative and functional genomics*, *2012*, 836374.

Fayard, E., Tintignac, L. A., Baudry, A., & Hemmings, B. A. (2005). Protein kinase B/Akt at a glance. *Journal of cell science*, *118*(Pt 24), 5675–5678.

Fischbach, K. F., Linneweber, G. A., Andlauer, T. F., Hertenstein, A., Bonengel, B., & Chaudhary, K. (2009). The irre cell recognition module (IRM) proteins. *Journal of neurogenetics*, *23*(1-2), 48–67.

Friedl, P., & Bröcker, E. B. (2000). The biology of cell locomotion within three-dimensional extracellular matrix. *Cellular and molecular life sciences CMLS*, *57*(1), 41-64.

Friedl, P., & Wolf, K. (2010). Plasticity of cell migration: a multiscale tuning model. *Journal of Cell Biology*, *188*(1), 11-19.

Gallant, N. D., Michael, K. E., & García, A. J. (2005). Cell adhesion strengthening: contributions of adhesive area, integrin binding, and focal adhesion assembly. *Molecular biology of the cell*, *16*(9), 4329-4340.

Galletta, B. J., Chakravarti, M., Banerjee, R., & Abmayr, S. M. (2004). SNS: Adhesive properties, localization requirements and ectodomain dependence in S2 cells and embryonic myoblasts. *Mechanisms of development*, *121*(12), 1455–1468.

Ganassi, M., Badodi, S., Ortuste Quiroga, H. P., Zammit, P. S., Hinits, Y., & Hughes, S. M. (2018). Myogenin promotes myocyte fusion to balance fibre number and size. *Nature communications*, *9*(1), 1-17.

Gayraud-Morel, B., Chrétien, F., Flamant, P., Gomès, D., Zammit, P. S., & Tajbakhsh, S. (2007). A role for the myogenic determination gene Myf5 in adult regenerative myogenesis. *Developmental biology*, *312*(1), 13-28.

Gerke, P., Sellin, L., Kretz, O., Petraschka, D., Zentgraf, H., Benzing, T., & Walz, G. (2005). NEPH2 is located at the glomerular slit diaphragm, interacts with nephrin and is cleaved from podocytes by metalloproteinases. *Journal of the American Society of Nephrology: JASN*, *16*(6), 1693–1702.

Gildor, B., Schejter, E. D., & Shilo, B. Z. (2012). Bidirectional Notch activation represses fusion competence in swarming adult Drosophila myoblasts. *Development (Cambridge, England)*, *139*(21), 4040–4050.

Goode, B. L., & Eck, M. J. (2007). Mechanism and function of formins in the control of actin assembly. *Annual review of biochemistry*, *76*, 593–627.

Griffin, C. A., Apponi, L. H., Long, K. K., & Pavlath, G. K. (2010). Chemokine expression and control of muscle cell migration during myogenesis. *Journal of cell science*, *123*(18), 3052-3060.

Guo, C. J., Xu, G., & Chen, L. L. (2020). Mechanisms of Long Noncoding RNA Nuclear Retention. *Trends in biochemical sciences*, *45*(11), 947–960.

Haralalka, S., Shelton, C., Cartwright, H. N., Katzfey, E., Janzen, E., & Abmayr, S. M. (2011). Asymmetric Mbc, active Rac1 and F-actin foci in the fusion-competent myoblasts during myoblast fusion in Drosophila. *Development (Cambridge, England)*, *138*(8), 1551–1562.

Harita, Y., Kurihara, H., Kosako, H., Tezuka, T., Sekine, T., Igarashi, T., & Hattori, S. (2008). Neph1, a component of the kidney slit diaphragm, is tyrosine-phosphorylated by the Src family tyrosine kinase and modulates intracellular signaling by binding to Grb2. *The Journal of biological chemistry*, *283*(14), 9177–9186.

Hawke, T. J., & Garry, D. J. (2001). Myogenic satellite cells: physiology to molecular biology. *Journal of applied physiology*, *91*(2), 534-551.

Hellmuth, A. E., & Allbrook, D. B. (1971). Muscle satellite cell numbers during the postnatal period. *Journal of anatomy*, *110*(Pt 3), 503.

Helmstädter, M., Lüthy, K., Gödel, M., Simons, M., Ashish, Nihalani, D., Rensing, S. A., Fischbach, K. F., & Huber, T. B. (2012). Functional study of mammalian Neph proteins in Drosophila melanogaster. *PloS one*, *7*(7), e40300.

Hernández-Hernández, J. M., García-González, E. G., Brun, C. E., & Rudnicki, M. A. (2017, December). The myogenic regulatory factors, determinants of muscle development, cell identity and regeneration. In *Seminars in cell & developmental biology* (Vol. 72, pp. 10-18). Academic Press.

Hindi, S. M., Tajrishi, M. M., & Kumar, A. (2013). Signaling mechanisms in mammalian myoblast fusion. *Science signaling*, *6*(272), re2.

Hirayama, E., & Kim, J. (2008). Identification and characterization of a novel neural cell adhesion molecule (NCAM)-associated protein from quail myoblasts: relationship to myotube formation and induction of neurite-like protrusions. *Differentiation; research in biological diversity*, *76*(3), 253–266.

Holterman, C. E., & Rudnicki, M. A. (2005, August). Molecular regulation of satellite cell function. In *Seminars in cell & developmental biology* (Vol. 16, No. 4-5, pp. 575-584). Academic Press.

Huard, J., Li, Y., & Fu, F. H. (2002). Muscle injuries and repair: current trends in research. *The Journal of bone and joint surgery. American volume*, *84*(5), 822–832.

Hudson, A. M., & Cooley, L. (2002). A subset of dynamic actin rearrangements in Drosophila requires the Arp2/3 complex. *The Journal of cell biology*, *156*(4), 677–687.

Humphries, J. D., Byron, A., & Humphries, M. J. (2006). Integrin ligands at a glance. *Journal of cell science*, *119*(19), 3901-3903.

Huttenlocher, A., & Horwitz, A. R. (2011). Integrins in cell migration. *Cold Spring Harbor perspectives in biology*, *3*(9), a005074.

Hynes, R. O. (2002). Integrins: bidirectional, allosteric signaling machines. *cell*, *110*(6), 673-687.

Jin, P., Duan, R., Luo, F., Zhang, G., Hong, S. N., & Chen, E. H. (2011). Competition between Blown fuse and WASP for WIP binding regulates the dynamics of WASP-dependent actin polymerization in vivo. *Developmental cell*, *20*(5), 623–638.

Jo, B. S., & Choi, S. S. (2015). Introns: The Functional Benefits of Introns in Genomes. *Genomics & informatics*, *13*(4), 112–118.

Johnson, N., Březinová, J., Stephens, E., Burbridge, E., Freeman, M., Adrain, C., & Strisovsky, K. (2017). Quantitative proteomics screen identifies a substrate repertoire of rhomboid protease RHBDL2 in human cells and implicates it in epithelial homeostasis. *Scientific reports*, *7*(1), 7283.

Jones, N. C., Tyner, K. J., Nibarger, L., Stanley, H. M., Cornelison, D. D., Fedorov, Y. V., & Olwin, B. B. (2005). The p38α/β MAPK functions as a molecular switch to activate the quiescent satellite cell. *The Journal of cell biology*, *169*(1), 105-116.

Joulia, D., Bernardi, H., Garandel, V., Rabenoelina, F., Vernus, B., & Cabello, G. (2003). Mechanisms involved in the inhibition of myoblast proliferation and differentiation by myostatin. *Experimental cell research*, *286*(2), 263-275.

Kaipa, B. R., Shao, H., Schäfer, G., Trinkewitz, T., Groth, V., Liu, J., Beck, L., Bogdan, S., Abmayr, S. M., & Önel, S. F. (2013). Dock mediates Scar- and WASp-dependent actin polymerization through interaction with cell adhesion molecules in founder cells and fusion-competent myoblasts. *Journal of cell science*, *126*(Pt 1), 360–372.

Kang, J. S., Yi, M. J., Zhang, W., Feinleib, J. L., Cole, F., & Krauss, R. S. (2004). Netrins and neogenin promote myotube formation. *The Journal of cell biology*, *167*(3), 493–504.

Karalaki, M., Fili, S., Philippou, A., & Koutsilieris, M. (2009). Muscle regeneration: cellular and molecular events. *In vivo*, *23*(5), 779-796.

Kassar-Duchossoy, L., Gayraud-Morel, B., Gomès, D., Rocancourt, D., Buckingham, M., Shinin, V., & Tajbakhsh, S. (2004). Mrf4 determines skeletal muscle identity in Myf5: Myod double-mutant mice. *Nature*, *431*(7007), 466-471.

Kawachi, H., & Fukusumi, Y. (2020). New insight into podocyte slit diaphragm, a therapeutic target of proteinuria. *Clinical and experimental nephrology*, *24*(3), 193–204.

Kim, J. H., & Chen, E. H. (2019). The fusogenic synapse at a glance. *Journal of cell science*, *132*(18)

Kim, J. H., Ren, Y., Ng, W. P., Li, S., Son, S., Kee, Y. S., Zhang, S., Zhang, G., Fletcher, D. A., Robinson, D. N., & Chen, E. H. (2015). Mechanical tension drives cell membrane fusion. *Developmental cell*, *32*(5), 561–573.

Kim, S., Shilagardi, K., Zhang, S., Hong, S. N., Sens, K. L., Bo, J., Gonzalez, G. A., & Chen, E. H. (2007). A critical function for the actin cytoskeleton in targeted exocytosis of prefusion vesicles during myoblast fusion. *Developmental cell*, *12*(4), 571–586.

Knight, J. D., & Kothary, R. (2011). The myogenic kinome: protein kinases critical to mammalian skeletal myogenesis. *Skeletal muscle*, *1*(1), 1-18.

Kook, S. H., Son, Y. O., Choi, K. C., Lee, H. J., Chung, W. T., Hwang, I. H., & Lee, J. C. (2008). Cyclic mechanical stress suppresses myogenic differentiation of adult bovine satellite cells through activation of extracellular signal-regulated kinase. *Molecular and cellular biochemistry*, *309*(1-2), 133–141.

Krauss, R. S. (2010). Regulation of promyogenic signal transduction by cell–cell contact and adhesion. *Experimental cell research*, *316*(18), 3042-3049.

Krauss, R. S., Joseph, G. A., & Goel, A. J. (2017). Keep your friends close: cell–cell contact and skeletal myogenesis. *Cold Spring Harbor Perspectives in Biology*, *9*(2), a029298.

Kuang, S., Kuroda, K., Le Grand, F., & Rudnicki, M. A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell*, *129*(5), 999-1010.

Lafreniere, J. F., Mills, P., Bouchentouf, M., & Tremblay, J. P. (2006). Interleukin-4 improves the migration of human myogenic precursor cells in vitro and in vivo. *Experimental cell research*, *312*(7), 1127-1141.

Lambrot, R., Chan, D., Shao, X., Aarabi, M., Kwan, T., Bourque, G., Moskovtsev, S., Librach, C., Trasler, J., Dumeaux, V., & Kimmins, S. (2021). Whole-genome sequencing of H3K4me3 and DNA methylation in human sperm reveals regions of overlap linked to fertility and development. *Cell reports*, *36*(3), 109418.

Laurin, M., Fradet, N., Blangy, A., Hall, A., Vuori, K., & Côté, J. F. (2008). The atypical Rac activator Dock180 (Dock1) regulates myoblast fusion in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(40), 15446–15451.

Le Grand, F., Jones, A. E., Seale, V., Scimè, A., & Rudnicki, M. A. (2009). Wnt7a activates the planar cell polarity pathway to drive the symmetric expansion of satellite stem cells. *Cell stem cell*, *4*(6), 535-547.

Le Grand, F., & Rudnicki, M. A. (2007). Skeletal muscle satellite cells and adult myogenesis. *Current opinion in cell biology*, *19*(6), 628-633.

Lee, D. M., & Chen, E. H. (2019). *Drosophila* Myoblast Fusion: Invasion and Resistance for the Ultimate Union. *Annual review of genetics*, *53*, 67–91.

Li, L., Heller-Harrison, R., Czech, M., & Olson, E. N. (1992). Cyclic AMP-dependent protein kinase inhibits the activity of myogenic helix-loop-helix proteins. *Molecular and cellular biology*, *12*(10), 4478–4485.

Liu, G., Kaw, B., Kurfis, J., Rahmanuddin, S., Kanwar, Y. S., & Chugh, S. S. (2003). Neph1 and nephrin interaction in the slit diaphragm is an important determinant of glomerular permeability. *The Journal of clinical investigation*, *112*(2), 209–221.

Liu, J., Huang, Z. P., Nie, M., Wang, G., Silva, W. J., Yang, Q., Freire, P. P., Hu, X., Chen, H., Deng, Z., Pu, W. T., & Wang, D. Z. (2020). Regulation of myonuclear positioning and muscle function by the skeletal muscle-specific CIP protein. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(32), 19254–19265.

Liu, Q., Pan, S., Liu, S., Zhang, S., Willerson, J. T., Martin, J. F., & Dixon, R. A. F. (2021). Suppressing Hippo signaling in the stem cell niche promotes skeletal muscle regeneration. *Stem cells (Dayton, Ohio)*, *39*(6), 737–749.

Luo, W., Li, E., Nie, Q., & Zhang, X. (2015). Myomaker, regulated by MYOD, MYOG and miR-140-3p, promotes chicken myoblast fusion. *International journal of molecular sciences*, *16*(11), 26186-26201.

Machado, M. C. R., Valer, F. B., Couto-Lima, C. A., & Ramos, R. G. P. (2018). Transcriptional cross-regulation of Irre Cell Recognition Module (IRM) members in the Drosophila pupal retina. *Mechanisms of development*, *154*, 193–202.

Martin, E. A., Muralidhar, S., Wang, Z., Cervantes, D. C., Basu, R., Taylor, M. R., Hunter, J., Cutforth, T., Wilke, S. A., Ghosh, A., & Williams, M. E. (2015). The intellectual disability gene Kirrel3 regulates target-specific mossy fiber synapse development in the hippocampus. *eLife*, *4*, e09395. f

Megeney, L. A., Kablar, B., Garrett, K., Anderson, J. E., & Rudnicki, M. A. (1996). MyoD is required for myogenic stem cell function in adult skeletal muscle. *Genes & development*, *10*(10), 1173-1183.

Menon, S. D., & Chia, W. (2001). Drosophila rolling pebbles: a multidomain protein required for myoblast fusion that recruits D-Titin in response to the myoblast attractant Dumbfounded. *Developmental cell*, *1*(5), 691–703.

Menon, S. D., Osman, Z., Chenchill, K., & Chia, W. (2005). A positive feedback loop between Dumbfounded and Rolling pebbles leads to myotube enlargement in Drosophila. *The Journal of cell biology*, *169*(6), 909–920.

Momiji, H., Hassall, K. L., Featherstone, K., McNamara, A. V., Patist, A. L., Spiller, D. G., Christian, H. C., White, M. R. H., Davis, J. R. E., Finkenstädt, B. F., & Rand, D. A. (2019). Disentangling juxtacrine from paracrine signalling in dynamic tissue. *PLoS computational biology*, *15*(6), e1007030.

Moo, E. K., Fortuna, R., Sibole, S. C., Abusara, Z., & Herzog, W. (2016). In vivo Sarcomere Lengths and Sarcomere Elongations Are Not Uniform across an Intact Muscle. *Frontiers in physiology*, *7*, 187.

Moore, C. A., Parkin, C. A., Bidet, Y., & Ingham, P. W. (2007). A role for the Myoblast city homologues Dock1 and Dock5 and the adaptor proteins Crk and Crk-like in zebrafish myoblast fusion. *Development (Cambridge, England)*, *134*(17), 3145–3153.

Moran, J. L., Li, Y., Hill, A. A., Mounts, W. M., & Miller, C. P. (2002). Gene expression changes during mouse skeletal myoblast differentiation revealed by transcriptional profiling. *Physiological genomics*, *10*(2), 103-111.

Mrozik, K. M., Blaschuk, O. W., Cheong, C. M., Zannettino, A. C. W., & Vandyke, K. (2018). N-cadherin in cancer metastasis, its emerging role in haematological malignancies and potential as a therapeutic target in cancer. *BMC cancer*, *18*(1), 939.

Neumann-Haefelin, E., Kramer-Zucker, A., Slanchev, K., Hartleben, B., Noutsou, F., Martin, K., Wanner, N., Ritter, A., Gödel, M., Pagel, P., Fu, X., Müller, A., Baumeister, R., Walz, G., & Huber, T. B. (2010). A model organism approach: defining the role of Neph proteins as regulators of neuron and kidney morphogenesis. *Human molecular genetics*, *19*(12), 2347–2359.

Nie, Y., Chen, H., Guo, C., Yuan, Z., Zhou, X., Zhang, Y., Zhang, X., Mo, D., & Chen, Y. (2017). Palmdelphin promotes myoblast differentiation and muscle regeneration. *Scientific reports*, *7*, 41608.

Nisar, S., Bhat, A. A., Singh, M., Karedath, T., Rizwan, A., Hashem, S., Bagga, P., Reddy, R., Jamal, F., Uddin, S., Chand, G., Bedognetti, D., El-Rifai, W., Frenneaux, M. P., Macha, M. A., Ahmed, I., & Haris, M. (2021). Insights Into the Role of CircRNAs: Biogenesis, Characterization, Functional, and Clinical Impact in Human Malignancies. *Frontiers in cell and developmental biology*, *9*, 617281.

Nowak, S. J., Nahirney, P. C., Hadjantonakis, A. K., & Baylies, M. K. (2009). Nap1-mediated actin remodeling is essential for mammalian myoblast fusion. *Journal of cell science*, *122*(Pt 18), 3282–3293.

Önel, S. F., Rust, M. B., Jacob, R., & Renkawitz-Pohl, R. (2014). Tethering membrane fusion: common and different players in myoblasts and at the synapse. *Journal of neurogenetics*, *28*(3-4), 302–315.

Pajcini, K. V., Pomerantz, J. H., Alkan, O., Doyonnas, R., & Blau, H. M. (2008). Myoblasts and macrophages share molecular components that contribute to cell-cell fusion. *The Journal of cell biology*, *180*(5), 1005–1019.

Paul, A., Annunziato, S., Lu, B., Sun, T., Evrova, O., Planas-Paz, L., Orsini, V., Terracciano, L. M., Charlat, O., Loureiro, Z. Y., Ji, L., Zamponi, R., Sigoillot, F., Lei, H., Lindeman, A., Russ, C., Reece-Hoyes, J. S., Nicholson, T. B., Tchorz, J. S., & Cong, F. (2022). Cell adhesion molecule KIRREL1 is a feedback regulator of Hippo signaling recruiting SAV1 to cell-cell contact sites. *Nature communications*, *13*(1), 930.

Perdiguero, E., Ruiz- Bonilla, V., Gresh, L., Hui, L., Ballestar, E., Sousa- Victor, P., Baeza-Raja, B., Jardí, M., Bosch-Comas, A., Esteller, M., Caelles, C., Serrano, A. L., Wagner, E. F., & Muñoz- Cánoves, P. (2007a). Genetic analysis of p38 MAP kinases in myogenesis: fundamental role of p38α in abrogating myoblast proliferation. *The EMBO journal*, *26*(5), 1245-1256.

Perdiguero, E., Ruiz-Bonilla, V., Serrano, A. L., & Muñoz-Cánoves, P. (2007b). Genetic deficiency of p38alpha reveals its critical role in myoblast cell cycle exit: the p38alpha-JNK connection. *Cell cycle (Georgetown, Tex.)*, *6*(11), 1298–1303.

Pizza, F. X., Martin, R. A., Springer, E. M., Leffler, M. S., Woelmer, B. R., Recker, I. J., & Leaman, D. W. (2017). Intercellular adhesion molecule-1 augments myoblast adhesion and fusion through homophilic trans-interactions. *Scientific reports*, *7*(1), 5094.

Plasschaert, R. N., Vigneau, S., Tempera, I., Gupta, R., Maksimoska, J., Everett, L., Davuluri, R., Mamorstein, R., Lieberman, P. M., Schultz, D., Hannenhalli, S., & Bartolomei, M. S. (2014). CTCF binding site sequence differences are associated with unique regulatory and functional trends during embryonic stem cell differentiation. *Nucleic acids research*, *42*(2), 774–789.

Przewoźniak, M., Czaplicka, I., Czerwińska, A. M., Markowska-Zagrajek, A., Moraczewski, J., Stremińska, W., Jańczyk-Ilach, K., Ciemerych, M. A., & Brzoska, E. (2013). Adhesion proteins--an impact on skeletal myoblast differentiation. *PloS one*, *8*(5), e61760.

Rajasekharan, S., & Kennedy, T. E. (2009). The netrin protein family. *Genome biology*, *10*(9), 239.

Randrianarison-Huetz, V., Papaefthymiou, A., Herledan, G., Noviello, C., Faradova, U., Collard, L., Pincini, A., Schol, E., Decaux, J. F., Maire, P., Vassilopoulos, S., & Sotiropoulos, A. (2018). Srf controls satellite cell fusion through the maintenance of actin architecture. *The Journal of cell biology*, *217*(2), 685–700.

Relaix, F., Rocancourt, D., Mansouri, A., & Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature*, *435*(7044), 948-953.

Richardson, B. E., Beckett, K., Nowak, S. J., & Baylies, M. K. (2007). SCAR/WAVE and Arp2/3 are crucial for cytoskeletal remodeling at the site of myoblast fusion. *Development (Cambridge, England)*, *134*(24), 4357–4367.

Rochlin, K., Yu, S., Roy, S., & Baylies, M. K. (2010). Myoblast fusion: when it takes more to make one. *Developmental biology*, *341*(1), 66–83.

Rosen, G. D., Sanes, J. R., LaChance, R., Cunningham, J. M., Roman, J., & Dean, D. C. (1992). Roles for the integrin VLA-4 and its counter receptor VCAM-1 in myogenesis. *Cell*, *69*(7), 1107–1119.

Rout, P., Preußner, M., & Önel, S. F. (2022). *Drosophila melanogaster*: A Model System to Study Distinct Genetic Programs in Myoblast Fusion. *Cells*, *11*(3), 321.

Rudnicki, M. A., Schnegelsberg, P. N., Stead, R. H., Braun, T., Arnold, H. H., & Jaenisch, R. (1993). MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell*, *75*(7), 1351-1359.

Ruiz-Bonilla, V., Perdiguero, E., Gresh, L., Serrano, A. L., Zamora, M., Sousa-Victor, P., Jardí, M., Wagner, E. F., & Muñoz-Cánoves, P. (2008). Efficient adult skeletal muscle regeneration in mice deficient in p38beta, p38gamma and p38delta MAP kinases. *Cell cycle (Georgetown, Tex.)*, *7*(14), 2208–2214.

Ruiz-Gómez, M., Coutts, N., Price, A., Taylor, M. V., & Bate, M. (2000). Drosophila dumbfounded: a myoblast attractant essential for fusion. *Cell*, *102*(2), 189–198.

Sampath, S. C., Sampath, S. C., & Millay, D. P. (2018). Myoblast fusion confusion: the resolution begins. *Skeletal muscle*, *8*(1), 1-10.

Sapoznik, E., Niu, G., Zhou, Y., Prim, P. M., Criswell, T. L., & Soker, S. (2018). A real-time monitoring platform of myogenesis regulators using double fluorescent labeling. *PloS one*, *13*(2), e0192654.

Sartorelli, V., & Caretti, G. (2005). Mechanisms underlying the transcriptional regulation of skeletal myogenesis. *Current opinion in genetics & development*, *15*(5), 528–535.

Schmalbruch, H., & Hellhammer, U. (1976). The number of satellite cells in normal human muscle. *The Anatomical Record*, *185*(3), 279-287.

Schöneich, C., Dremina, E., Galeva, N., & Sharov, V. (2014). Apoptosis in differentiating C2C12 muscle cells selectively targets Bcl-2-deficient myotubes. *Apoptosis : an international journal on programmed cell death*, *19*(1), 42–57.

Seale, P., Sabourin, L. A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., & Rudnicki, M. A. (2000). Pax7 is required for the specification of myogenic satellite cells. *Cell*, *102*(6), 777–786.

Segal, D., Dhanyasi, N., Schejter, E. D., & Shilo, B. Z. (2016). Adhesion and Fusion of Muscle Cells Are Promoted by Filopodia. *Developmental cell*, *38*(3), 291–304.

Sellin, L., Huber, T. B., Gerke, P., Quack, I., Pavenstädt, H., & Walz, G. (2003). NEPH1 defines a novel family of podocin interacting proteins. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, *17*(1), 115–117.

Sens, K. L., Zhang, S., Jin, P., Duan, R., Zhang, G., Luo, F., Parachini, L., & Chen, E. H. (2010). An invasive podosome-like structure promotes fusion pore formation during myoblast fusion. *The Journal of cell biology*, *191*(5), 1013–1027.

Shaul, O. (2017). How introns enhance gene expression. *The international journal of biochemistry & cell biology*, *91*(Pt B), 145–155.

Shelton, C., Kocherlakota, K. S., Zhuang, S., & Abmayr, S. M. (2009). The immunoglobulin superfamily member Hbs functions redundantly with Sns in interactions between founder and fusion-competent myoblasts. *Development (Cambridge, England)*, *136*(7), 1159–1168.

Shilagardi, K., Li, S., Luo, F., Marikar, F., Duan, R., Jin, P., Kim, J. H., Murnen, K., & Chen, E. H. (2013). Actin-propelled invasive membrane protrusions promote fusogenic protein engagement during cell-cell fusion. *Science (New York, N.Y.)*, *340*(6130), 359–363.

Singh, K., & Dilworth, F. J. (2013). Differential modulation of cell cycle progression distinguishes members of the myogenic regulatory factor family of transcription factors. *The FEBS journal*, *280*(17), 3991-4003.

Sohn, R. L., Huang, P., Kawahara, G., Mitchell, M., Guyon, J., Kalluri, R., Kunkel, L. M., & Gussoni, E. (2009). A role for nephrin, a renal protein, in vertebrate skeletal muscle cell fusion. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(23), 9274–9279.

Solanki, A. K., Widmeier, E., Arif, E., Sharma, S., Daga, A., Srivastava, P., Kwon, S. H., Hugo, H., Nakayama, M., Mann, N., Majmundar, A. J., Tan, W., Gee, H. Y., Sadowski, C. E., Rinat, C., Becker-Cohen, R., Bergmann, C., Rosen, S., Somers, M., Shril, S., … Nihalani, D. (2019). Mutations in KIRREL1, a slit diaphragm component, cause steroid-resistant nephrotic syndrome. *Kidney international*, *96*(4), 883–889.

Srinivas, B. P., Woo, J., Leong, W. Y., & Roy, S. (2007). A conserved molecular pathway mediates myoblast fusion in insects and vertebrates. *Nature genetics*, *39*(6), 781–786.

Stanton, L. A., Sabari, S., Sampaio, A. V., Underhill, T. M., & Beier, F. (2004). p38 MAP kinase signalling is required for hypertrophic chondrocyte differentiation. *The Biochemical journal*, *378*(Pt 1), 53–62.

Strünkelnberg, M., Bonengel, B., Moda, L. M., Hertenstein, A., de Couet, H. G., Ramos, R. G., & Fischbach, K. F. (2001). rst and its paralogue kirre act redundantly during embryonic muscle development in Drosophila. *Development (Cambridge, England)*, *128*(21), 4229–4239.

Strünkelnberg, M., de Couet, H. G., Hertenstein, A., & Fischbach, K. F. (2003). Interspecies comparison of a gene pair with partially redundant function: the rst and kirre genes in D. virilis and D. melanogaster. *Journal of molecular evolution*, *56*(2), 187–197.

Tajbakhsh, S. (2005). Skeletal muscle stem and progenitor cells: reconciling genetics and lineage. *Experimental cell research*, *306*(2), 364-372.

Tamir-Livne, Y., Mubariki, R., & Bengal, E. (2017). Adhesion molecule Kirrel3/Neph2 is required for the elongated shape of myocytes during skeletal muscle differentiation. *The International journal of developmental biology*, *61*(3-4-5), 337–345.

Tanaka, S., Terada, K., & Nohno, T. (2011). Canonical Wnt signaling is involved in switching from cell proliferation to myogenic differentiation of mouse myoblast cells. *Journal of molecular signaling*, *6*(1), 1-17.

Ustanina, S., Carvajal, J., Rigby, P., & Braun, T. (2007). The myogenic factor Myf5 supports efficient skeletal muscle regeneration by enabling transient myoblast amplification. *Stem cells*, *25*(8), 2006-2016.

van der Velden, J. L. J., Schols, A. M. W. J., Willems, J., Kelders, M. C. J. M., & Langen, R. C. J. (2008). Glycogen synthase kinase 3 suppresses myogenic differentiation through negative regulation of NFATc3. *The Journal of biological chemistry*, *283*(1), 358–366.

Vandamme, D., Lambert, E., Waterschoot, D., Cognard, C., Vandekerckhove, J., Ampe, C., Constantin, B., & Rommelaere, H. (2009). alpha-Skeletal muscle actin nemaline myopathy mutants cause cell death in cultured muscle cells. *Biochimica et biophysica acta*, *1793*(7), 1259–1271.

Vassetzky, N. S., & Kramerov, D. A. (2013). SINEBase: a database and tool for SINE analysis. *Nucleic acids research*, *41*(Database issue), D83–D89.

Vasyutina, E., Martarelli, B., Brakebusch, C., Wende, H., & Birchmeier, C. (2009). The small G-proteins Rac1 and Cdc42 are essential for myoblast fusion in the mouse. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(22), 8935–8940.

von Maltzahn, J., Chang, N. C., Bentzinger, C. F., & Rudnicki, M. A. (2012). Wnt signaling in myogenesis. *Trends in cell biology*, *22*(11), 602-609.

Wang, L. C., & Almazan, G. (2016). Cdon, a cell surface protein, mediates oligodendrocyte differentiation and myelination. *Glia*, *64*(6), 1021–1033.

Wang, L. H., & Baker, N. E. (2015). E proteins and ID proteins: helix-loop-helix partners in development and disease. *Developmental cell*, *35*(3), 269-280.

Wanner, N., Noutsou, F., Baumeister, R., Walz, G., Huber, T. B., & Neumann-Haefelin, E. (2011). Functional and spatial analysis of C. elegans SYG-1 and SYG-2, orthologs of the Neph/nephrin cell adhesion module directing selective synaptogenesis. *PloS one*, *6*(8), e23598.

Ye, M., Goudot, C., Hoyler, T., Lemoine, B., Amigorena, S., & Zueva, E. (2020). Specific subfamilies of transposable elements contribute to different domains of T lymphocyte enhancers. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(14), 7905–7916.

Yesildag, B., Bock, T., Herrmanns, K., Wollscheid, B., & Stoffel, M. (2015). Kin of IRRE-like Protein 2 Is a Phosphorylated Glycoprotein That Regulates Basal Insulin Secretion. *The Journal of biological chemistry*, *290*(43), 25891–25906.

Yin, H., Price, F., & Rudnicki, M. A. (2013). Satellite cells and the muscle stem cell niche. *Physiological reviews*, *93*(1), 23-67.

Zammit, P. S. (2017, December). Function of the myogenic regulatory factors Myf5, MyoD, Myogenin and MRF4 in skeletal muscle, satellite cells and regenerative myogenesis. In *Seminars in cell & developmental biology* (Vol. 72, pp. 19-32). Academic Press.

Zhang, Y., Featherstone, D., Davis, W., Rushton, E., & Broadie, K. (2000). Drosophila D-titin is required for myoblast fusion and skeletal muscle striation. *Journal of cell science*, *113 (Pt 17)*, 3103–3115.

Zhang, X. O., Pratt, H., & Weng, Z. (2021). Investigating the Potential Roles of SINEs in the Human Genome. *Annual review of genomics and human genetics*, *22*, 199–218.

Appendix I: DNA amplicon sequencing – C2C12 and K1CRISPR

Primer set 1: K1CRISPR amplicon
alignment to mouse chromosome 3
NCBI reference sequence: AC102115.11

K1 CRISPR amplicon: 481bp read

TCATCTCTTTGCCCCTAACTGGCCTCCTTATCAGMSCAACAMRRATCCTTGCTCCAACTTCT
CAACAGATCTAGTCCCCAACTACATACCAAATGCCTAATACTTAATATGTTTTGCATATGC
AGATGTATAGTATTGTGTGTGGTATTTGTGAAGTTTCTGTGTGGTAAGTATGTGGTGTATGT
ATGATGTGTGTGTTGTGTTTGAGGTGTGCATAGTGTATTTGCTGGGGGGGAGACTTGCTTA
TGTAGCCCAGGCTGGCCTTGAACTCACGCACACATGCAGACACCTGTCTCAGTAAACGTGT
AGGGTGTAATTTATCTATCCGGAGGGAAAACCAAAGAACTCGGTGGCTTTCTCTTGTGCCT
CCCCGAAGGCTTGCTACAGAAGGCAGCCACTGATCGATCTCTAAGCACTCAGGAACTTGA
GGTGCACTGGGRWKARGRYAMRKAGAATCRAAATAGAAGTGCATTCGGGTGGA

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Mus musculus chromosome 3, clone RP23-145A18, complete sequence	Mus musculus	375	643	99%	2e-99	88.44%	184436	AC102115.11

Mus musculus chromosome 3, clone RP23-145A18, complete sequence
Sequence ID: AC102115.11 Length: 184436 Number of Matches: 2

CDS starts at: 88951 of chr3

CRISPR cut site close to exon 1: 54715-5734 of chr3

CRISPR cut site in intron 2: 90354-90374 of chr3

Range 1: 90357 to 90660 GenBank Graphics ▼ Next Match ▲ P

Score	Expect	Identities	Gaps	Strand
375 bits(203)	2e-99	283/320(88%)	22/320(6%)	Plus/Minus

```
Query  2      CATCTCTTTGCCCCTAACTGGCCTCCTTATCAGMSCAACAMRRATCCTTGCTCCAACTTC  61
              ||||||  |||||||||||||||||||||||||  |||||    ||||||||||
Sbjct  90660  CATCTC-TTGCCCCTAACTGGCCTCCTTATCAGAGCAACAAGAATCCTTGCTC-------  90609

Query  62     TCAACAGATCTAGTCCCCAACTACATACCAAATGCCTAATACTTAATATGTTTTGCATAT  121
                ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  90608  --AACAGATCTAGTCCCCAATTACATACCAAATGCCTAATACTTAATATGTTTTGCATAT  90551

Query  122    GCAGatgtatagtattgtgtgtggtatttgtgaagtttctgtgtggtaagtatgtggtgt  181
              |||||||||||||   | | ||    ||  || | ||||||||||||||||||||||||||
Sbjct  90550  GCAGATGTATAGT--T-TCTG-AATAAATG-G-AGTTTCTGTGTGGTAAGTATGTGGTGT  90497

Query  182    atgtatgatgtgtgtgttgtgtttgaggtgtgcatagtgtatTTGCTgggggggg------  235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  90496  ATGTATGATGTGTGTGTTGTGTTTGAGGTGTGCATAGTGTATTTGCTGGGGGGGGGGGGG  90437

Query  236    AGACTTGCTTATGTAGCCCAGGCTGGCCTTGAACTCACGCACACATGCAGACACCTGTCT  295
              |||||||||||||||| |||||||||||||||||||  ||||||||||||||||||| ||
Sbjct  90436  AGACTTGCTTATGTATCCCAGGCTGGCCTTGAACTCACAAACACATGCAGACACCTGCCT  90377

Query  296    CAGTAAACGTGTAGGGTGTA  315
              ||||||||||||||||||||
Sbjct  90376  CAGTAAACGTGTAGGGTGTA  90357
```

Range 2: 54567 to 54731 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
267 bits(144)	1e-66	155/165(94%)	0/165(0%)	Plus/Minus

```
Query  316    ATTTATCTATCCGGAGGGAAAACCAAAGAACTCGGTGGCTTTCTCTTGTGCCTCCCCGAA  375
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54731  ATTTATCTATCCGGAGGGAAAACCAAAGAACTCGGTGGCTTTCTCTTGTGCCTCCCCGAA  54672

Query  376    GGCTTGCTACAGAAGGCAGCCACTGATCGATCTCTAAGCACTCAGGAACTTGAGGTGCAC  435
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54671  GGCTTGCTACAGAAGGCAGCCACTGATCGATCTCTAAGCACTCAGGAACTTGAGGTGCAC  54612

Query  436    TGGGrwkargryamrkaGAATCRAAATAGAAGTGCATTCGGGTGG  480
              ||||   | |  |  |||||| |||||||||||||||||||||
Sbjct  54611  TGGGGTGAAGACATGGAGAATCGAAATAGAAGTGCATTCGGGTGG  54567
```

<u>Primer set 2:</u> C2C12 amplicon
alignment to mouse chromosome 3
NCBI reference sequence: AC102115.11

<u>C2C12 amplicon:</u> 242bp read
GAAATCTTGACTGTCTTTGARMCCCTTCTGAGRRCYYYAGCCAGCCACATTGAGGCCCGCCT
CAGCGCTGGAGTCCGTAGCAGCCATGAAGAGGCGCAGTTAACTCTTCCTTTGTTCATCCCAG
ACATTTATCTATCCGGAGGGAAAACCAAAGAACTCGGTGGCTTTCTCTTGTGCCTCCCCGAA
GGCTTGCTACAGAAGGCAKCMMCTRATCGAYCTCTAAGCACTCAGGAACTTGAGGT

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Mus musculus chromosome 3, clone RP23-145A18, complete sequence	Mus musculus	398	398	100%	2e-106	94.63%	184436	AC102115.11

Mus musculus chromosome 3, clone RP23-145A18, complete sequence

Sequence ID: AC102115.11 **Length:** 184436 **Number of Matches:** 1

Range 1: 54616 to 54857 GenBank Graphics ▼ Next Match ▲ Pre

Score	Expect	Identities	Gaps	Strand
398 bits(215)	2e-106	229/242(95%)	0/242(0%)	Plus/Minus

```
Query  1      GAAATCTTGACTGTCTTTGARMCCCTTCTGAGRRCYYYAGCCAGCCACATTGAGGCCCGC  60
              ||||||||||||||||||||  ||||||||||  |  ||||||||||||||||||||||
Sbjct  54857  GAAATCTTGACTGTCTTTGAGCCCCTTCTGAGGGCTTCAGCCAGCCACATTGAGGCCCGC  54798

Query  61     CTCAGCGCTGGAGTCCGTAGCAGCCATGAAGAGGCGCAGTTAACTCTTCCTTTGTTCATC  120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Sbjct  54797  CTCAGCGCTGGAGTCCGTAGCAGCCATGAAGAGGCGCAGTTAACTCTTCCTTTGTTCACC  54738

Query  121    CCAGACATTTATCTATCCGGAGGGAAAACCAAAGAACTCGGTGGCTTTCTCTTGTGCCTC  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  54737  CCAGACATTTATCTATCCGGAGGGAAAACCAAAGAACTCGGTGGCTTTCTCTTGTGCCTC  54678

Query  181    CCCGAAGGCTTGCTACAGAAGGCAKCMMCTRATCGAYCTCTAAGCACTCAGGAACTTGAG  240
              |||||||||||||||||||||||  |  ||  |||||  ||||||||||||||||||||||
Sbjct  54677  CCCGAAGGCTTGCTACAGAAGGCAGCCACTGATCGATCTCTAAGCACTCAGGAACTTGAG  54618

Query  241    GT  242
              ||
Sbjct  54617  GT  54616
```

<u>CRISPR cut site close to exon 1:</u>
54715-5734 of chr3

<u>Primer set 3:</u> C2C12 amplicon – forward primer sequencing alignment to mouse chromosome 3
NCBI reference sequence: AC102115.11

<u>C2C12 amplicon</u> – <u>forward primer sequencing reaction:</u> 583bp read
CGCCAGGRACAGATCCTTGCTCCACTTCTCACAGATCTAGTCCCCAACTACATACCAAATGCCTAATA
CTTAWTWTGTTTTGCATATGCAGATGTATAGTATTGTGTGTGGTATTTGTGAAGTTTCTGTGTGGTAA
GTATGTGGTGTATGTATGATGTGTGTGTTGTGTTTGAGGTGTGCATAGTGTATTTGCTGGGGGGGAGA
CTTGCTTATGTAGCCCAGGCTGGCCTTGAACTCACGCACACATGCAGACACCTGTCTCAGTAAACGTG
TAGGGTGTAGGCAGGAGACAACTTATTCAAGAGTTTATTCTCTCCGTCCACCATGTAGGTGAGTTCTA
GGGAGGAAATTCGGGTTAACAGACAGTATTAAGTGCTTTTGCCTGCTGAGCCATCTCACCAGCCCATC
TCCTTCATAATTCACTGGGTATCYTCAAAAGACAAAGCCAGAGTCCTTCCCTCTATGTTTTKGTAGTM
GAWGAACAASGYACAATCWGTATTACAGTAAGTTCACAGAGTTCAGAAATCACACACCACACAGAC
TCTYCCTGTACAGGTAGCCCTCCTCACTACACMCCTGCAAR

	Description	Scientific Name	Common Name	Taxid	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☐	Mus musculus chromosome 3, clone RP23-14…	Mus mu…	house …	10090	848	848	95%	0.0	93.46%	184436	AC102115.11

Mus musculus chromosome 3, clone RP23-145A18, complete sequenc

Sequence ID: AC102115.11 Length: 184436 Number of Matches: 1

Range 1: 90056 to 90615 GenBank Graphics ▼ Next Match ▲ Previous

Score	Expect	Identities	Gaps	Strand
848 bits(459)	0.0	529/566(93%)	14/566(2%)	Plus/Minus

```
Query  25     CTT-CTC-ACAGATCTAGTCCCCAACTACATACCAAATGCCTAATACTTAWTWTGTTTTG
              ||| ||| |||||||||||||||||| |||||||||||||||||||||||| | |||||||
Sbjct  90615  CTTGCTCAACAGATCTAGTCCCCAATTACATACCAAATGCCTAATACTTAATATGTTTTG

Query  83     CATATGCAGatgtatagtattgtgtgtggtatttgtgaagtttctgtgtggtaagtatgt
              ||||||||||||||||||  | | ||   ||  || | ||||||||||||||||||||||
Sbjct  90555  CATATGCAGATGTATAGT--T-TCTG-AATAAATG-G-AGTTTCTGTGTGGTAAGTATGT

Query  143    ggtgtatgtatgatgtgtgtgttgtgtttgaggtgtgcatagtgtatTTGCTgggggggg-
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  90501  GGTGTATGTATGATGTGTGTGTTGTGTTTGAGGTGTGCATAGTGTATTTGCTGGGGGGGG

Query  202    -----AGACTTGCTTATGTAGCCCAGGCTGGCCTTGAACTCACGCACACATGCAGACACC
                   ||||||||||||||| |||||||||||||||||||||  ||||||||||||||||
Sbjct  90441  GGGGGAGACTTGCTTATGTATCCCAGGCTGGCCTTGAACTCACAAACACATGCAGACACC

Query  257    TGTCTCAGTAAACGTGTAGGGTGTAGGCAGGAGACAACTTATTCAAGAGTTTATTCTCTC
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  90381  TGCCTCAGTAAACGTGTAGGGTGTAGGCAGGAGACAACTTATTCAAGAGTTTATTCTCTC

Query  317    CGTCCACCATGTAGGTGAGTTCTAGGGAGGAAATTCGGGTTAACAGACAGTATTAAGTGC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  90321  CGTCCACCATGTAGGTGAGTTCTAGGGAGGAAATTCGGGTTAACAGACAGTATTAAGTGC

Query  377    TTTTGCCTGCTGAGCCATCTCACCAGCCCATCTCCTTCATAATTCACTGGGTATCYTCAA
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||
Sbjct  90261  TTTTGCCTGCTGAGCCATCTCACCAGCCCATCTCCTTCATAATTTACTGGGTATCCTCAA

Query  437    AAGACAAAGCCAGAGTCCTTCCCTCTATGTTTTKGTAGTMGAWGAACAASGYACAATCWG
              |||||||||||||||||||||||||||||||||||||| |||||  | |||||| | |||||| |
Sbjct  90201  AAGACAAAGCCAGAGTCCTTCCCTCTATGTTTTTGTAGTCAATGAACAAGGTACAATCAG

Query  497    TATTACAGTAAGTTCACAGAGTTCAGAAATCACACACCACACAGACTCTYCCTGTACAGG
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct  90141  TATTACAGTAAGTTCACAGAGTTCAGAAATCACACACCACACAGACTCTCCCTGTACAGG

Query  557    TAGCCCTCCTCACTACACMCCTGCAA  582
              |||||||||||||||||| ||||||||
Sbjct  90081  TAGCCCTCCTCACTACACACCTGCAA  90056
```

<u>CDS at</u>: 88951 of chr3

<u>Intron 2 cut at:</u> 90354-90374 of chr3

<u>Primer set 3:</u> C2C12 amplicon – reverse primer sequencing
alignment to mouse chromosome 3
NCBI reference sequence: AC102115.11

<u>C2C12 amplicon</u> – <u>reverse primer sequencing reaction (reverse compliment):</u> 758bp read
GTMTTCCCTTCTCCTTTYTMCACTAARGGAARRMATCMCGSCACMTGTTTCCCATCACAGCCTT
CAATTMCCTCCTCATTCAGRGGACCTGAARCACAGARCMGGCTCGCTGAKTGATTTAAGTTTAA
TTCATTTGTGTAATATCCCTGTTCAGAGCAATGTCTAATTAACACTTTTGTKGGGGGGAATARAA
CCTCCCATATTCATTAATGAATGGGGGAAGGGAGGAGAAGCCTTTGCAGGCAGCAACCCAGAG
TCTACAGGGCTGGSTCTATCTTGGAGATGTCTTAATGAGACATGTCTTYATAATGGGATTATTGC
TCCCTTTGGACTAGAAAAGCAAATGGGGCTGCGATCAGCTAGGGAATGGAGAGTCAGTAGTCC
ATGAGTCAGCAGGGCTGGGGACCTGAATTCTGGAAACTTCATTCAGGGAAGGGATACAGGGCT
AGCAAGGCTCTCTCGGAGCTGCTCTCTCCTGACTCTACTTCCAAGCCTTCCCAGGCCTGGCAGA
GAGAAACCACATCACAGATTTGTTTACGACATCACAATCTATAACATGATGCTTTATAAATGTT
AGTTTACACGTTTCTGGAAACGACAAGTGACCAACMCTGTTTAAGAGGCAAAGACTCTCAGAG
AAATGGAGGAACTCACCCAAAGTCACAAAGATAAGATAAGCAAMGRSTAAGTGTGGAGCCCA
CATCTTCTGACTCAGAAACCTGGGCTCTCAGGAGCAGGCTGATGCAGTCATGTTGSAT

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Mus musculus chromosome 3, clone RP23-145A18, complete sequence	Mus musculus	1245	1245	98%	0.0	95.74%	184436	AC102115.11

Mus musculus chromosome 3, clone RP23-145A18, complete sequence
Sequence ID: AC102115.11 Length: 184436 Number of Matches: 1

Range 1: 88982 to 89730 GenBank Graphics ▼ Next Match ▲ Pre

Score	Expect	Identities	Gaps	Strand
1245 bits(674)	0.0	719/751(96%)	8/751(1%)	Plus/Minus

```
Query  7      CCTTCTCCTTTYTMCACTAARGGAARRMATCMCGSCACMTGTTTCCCATCACAGCCTTCA  66
              ||||||||  ||  |  ||||  |  ||||    |||  ||  |||  ||||||||||||||||||||||
Sbjct  89730  CCTTCTCC-TTCTCCACT-AGGGAAGACATCCCGCCACCTGTTTCCCATCACAGCCTTCA  89673

Query  67     ATTMCCTCCTCATTCAGRGGACCTGAARCACAGARCMGGCTCGCTGAKTGATTTAAGTTT  126
              |||  |||||||||||||  |||||||||  ||||||  |  ||||||||||  ||||||||||||||
Sbjct  89672  ATTCCCTCCTCATTCAGGGGACCTGAAACACAGAGCCGGCTCGCTGATTGATTTAAGTTT  89613

Query  127    AATTCATTTGTGTAATATCCCTGTTCAGAGCAATGTCTAATTAACACTTTTGTKGGGGGG  186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Sbjct  89612  AATTCATTTGTGTAATATCCCTGTTCAGAGCAATGTCTAATTAACACTTTTGTTGGGGGG  89553

Query  187    AATARAACCTCCCATATTCATTAATGAATGGGGGAAGGGAGGAGAAGCCTTTGCAGGCAG  246
              ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89552  AATAAAACCTCCCATATTCATTAATGAATGGGGGAAGGGAGGAGAAGCCTTTGCAGGCAG  89493

Query  247    CAACCCAGAGTCTACAGGGCTGGSTCTATCTTGGAGATGTCTTAATGAGACATGTCTTYA  306
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||  |
Sbjct  89492  CAACCCAGAGTCTACAGGGCTGGCTCTATCTTGGAGATGTCTTAATGAGACATGTCTTCA  89433

Query  307    TAATGGGATTATTGCTCCCTTTGGACTAGAAAAGCAAATGGGGCTGCGATCAGCTAGGGA  366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89432  TAATGGGATTATTGCTCCCTTTGGACTAGAAAAGCAAATGGGGCTGCGATCAGCTAGGGA  89373

Query  367    ATGGAGAGTCAGTAGTCCATGAGTCAGCAGGGCTGGGGACCTGAATTCTGGAAACTTCAT  426
              ||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Sbjct  89372  ATGGAGAGTCAGTAGTCCATGAGTCAGCAGGGCTGGGGACCTGGATTCTGGAAACTTCAT  89313

Query  427    TCAGGGAAGGGATACAGGGCTAGCAAGGCTCTCTCGGAGCTGCTCTCTCCTGACTCTACT  486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89312  TCAGGGAAGGGATACAGGGCTAGCAAGGCTCTCTCGGAGCTGCTCTCTCCTGACTCTACT  89253

Query  487    TCCAAGCCTTCCCAGGCCTGGCAGAGAGAAACCACATCACAGATTTGTTTACGACATCAC  546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89252  TCCAAGCCTTCCCAGGCCTGGCAGAGAGAAACCACATCACAGATTTGTTTACGACATCAC  89193

Query  547    AATCTATAACATGATGCTTTATAAATGTTAGTTTACACGTTTCTGGAAACGACAAGTGAC  606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89192  AATCTATAACATGATGCTTTATAAATGTTAGTTTACACGTTTCTGGAAACGACAAGTGAC  89133

Query  607    CAACMCTGTTTAAGAGGCAAAGACTCTCAGAGAAATGGAGGAACTCACCCAAAGTCACAA  666
              ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89132  CAACACTGTTTAAGAGGCAAAGACTCTCAGAGAAATGGAGGAACTCACCCAAAGTCACAA  89073

Query  667    AGATAAGATAAGCAAMGRSTAAGTGTGGAGCCCACATCTTCTGACTCAGAAACCTGGGCT  726
              ||||||||||||||  |  ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89072  AGATAAGATAAGCAACGACTAAGTGTGGAGCCCACATCTTCTGACTCAGAAACCTGGGCT  89013

Query  727    -CTCAGG-AGCAGGCT-GAT-G-CAG-TCAT  751
               ||||||  ||||||||  |||  |  |||  ||||
Sbjct  89012  TCTCAGGGAGCAGGCTTGATTGGCAGGTCAT  88982
```

<u>CDS at</u>: 88951 of chr3

<u>Intron 2 cut at:</u>
90354-90374 of chr3

<u>Primer set 3:</u> K1CRISPR amplicon – forward primer sequencing alignment to mouse chromosome 3
NCBI reference sequence: AC102115.11

<u>K1CRISPR amplicon</u> – <u>forward primer sequencing reaction:</u> 398bp read
AAWCAAASAACAARCATCCTTGCTCCACTTCTMACAGATCTAGTCCCCAACTACATACCAAATGCCTA
ATACTTAWTATGTTTTGCATATGCARATGTATAGTATTGTGTGTGGTATTTGTGAAGTTTCTGTGTGGT
AAGTATGTGGTGTATGTATGATGTGTGTGTTGTGTTTGAGGTGTGCATAGTGTATTTGCTGGGGGGGA
GACTTGCTTATGTAGCCCAGGCTGGCCTTGAACTCACGCACACATGCAGACACCTGTCTCAGTAAACG
TGTAGGGTGCAAGACAGGCACCATTGTCTGATGAGAAAACTAGTGGCTCCCGAASATTCAATGGCAG
GGGAAAGTACACACTTTYCAATRCSWYCAGTGMCTCMKCCTTGAACTATTCCCCCTGA

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Mus musculus chromosome 3, clone RP23-145A18, complete sequence	Mus musculus	102	102	17%	3e-17	92.75%	184436	AC102115.11

Mus musculus chromosome 3, clone RP23-145A18, complete sequence

Sequence ID: AC102115.11 Length: 184436 Number of Matches: 1

Range 1: 89992 to 90060 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
102 bits(55)	3e-17	64/69(93%)	0/69(0%)	Plus/Minus

```
Query  281    TGCAAGACAGGCACCATTGTCTGATGAGAAAACTAGTGGCTCCCGAASATTCAATGGCAG  340
              ||||||||||||||||||||||||||||||||||||||||||| ||| | |||| |||||
Sbjct  90060  TGCAAGACAGGCACCATTGTCTGATGAGAAAACTAGTGGCTCCAGAAGACTCAACGGCAG  90001

Query  341    GGGAAAGTA  349
              ||||| |||
Sbjct  90000  GGGAAGGTA  89992
```

<u>CDS at</u>: 88951 of chr3

<u>Intron 2 cut at:</u>
90354-90374 of chr3

<u>Primer set 3:</u> K1CRISPR amplicon – reverse primer sequencing alignment to mouse chromosome 3
NCBI reference sequence: AC102115.11

<u>K1 CRISPR amplicon</u> – <u>reverse primer sequencing reaction:</u> 668bp read

TGAAAGCGCAGAGSCRSAGYCACTGATGATWTWAGWTTTATTCATTTGWGTAATATCCCTGTTCA
GARCAATGTCTAATTAACACTTTTGTTGGGGGGAATAAAACCTCCCATATTCATTAATGAATGRGG
GAAGGGAGGAGAAGCCTTTGCAGGCAGCAACCCAGAGTMTACAGGGGKGGSTCTATCTTGSMGAY
GTCTTAATGAGACATGTCTTCATAATGGGATTATTGCTCCCTTTGGACTAGAAAAGCAAATGGGGC
TGCGATCAGCTAGGGAATGGAGAGTCAGTAGTCCATGAGTCAGCAGGGCTGGGGACCTGAATTCT
GGAAACTTCATTCAGGGAAGGGATACAGGGCTAGCAAGGCTCTCTCGGAGCTGCTCTCTCCTGACT
CTACTTCCAAGCCTTCCCAGGCCTGGCAGAGAGAAACCACATCACAGATTTGTTTACGACATCACA
ATCTATAACATGATGCTTTATAAATGTTAGTTTACACGTTTCTGGAAACGACAAGTGACCAACACT
GTTTAAGAGGCAAAGACTCTCAGAGAAATGGAGGAACTCACCCAAAGTCACAAAGATAAGATAAG
CAAMGASTAAGTGTGGAGCCCACATCTTYTGACTYAGAAACCTGGGCTCTCAGGAGCAGGCTGAT
GCAGTCATRTSCA

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Mus musculus chromosome 3, clone RP23-145A18, complete sequence	Mus musculus	1075	1075	95%	0.0	95.99%	184436	AC102115.11

Mus musculus chromosome 3, clone RP23-145A18, complete sequence

Sequence ID: AC102115.11 Length: 184436 Number of Matches: 1

Range 1: 88982 to 89629 GenBank Graphics ▼ Next Match ▲ Pre

Score	Expect	Identities	Gaps	Strand
1075 bits(582)	0.0	622/648(96%)	7/648(1%)	Plus/Minus

```
Query  23     CTGA-TGATWTWAGWTTTATTCATTTGWGTAATATCCCTGTTCAGARCAATGTCTAATTA  81
              |||| |||| | || || |||||||| |||||||||||||||||| ||||||||||||
Sbjct  89629  CTGATTGATTTAAGTTTAATTCATTTGTGTAATATCCCTGTTCAGAGCAATGTCTAATTA  89570

Query  82     ACACTTTTGTTGGGGGGAATAAAACCTCCCATATTCATTAATGAATGRGGGAAGGGAGGA  141
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  89569  ACACTTTTGTTGGGGGGAATAAAACCTCCCATATTCATTAATGAATGGGGGAAGGGAGGA  89510

Query  142    GAAGCCTTTGCAGGCAGCAACCCAGAGTMTACAGGGGKGGSTCTATCTTGSMGAYGTCTT  201
              |||||||||||||||||||||||||||| ||||||| || ||||||||| || |||||
Sbjct  89509  GAAGCCTTTGCAGGCAGCAACCCAGAGTCTACAGGGCTGGCTCTATCTTGGAGATGTCTT  89450

Query  202    AATGAGACATGTCTTCATAATGGGATTATTGCTCCCTTTGGACTAGAAAAGCAAATGGGG  261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89449  AATGAGACATGTCTTCATAATGGGATTATTGCTCCCTTTGGACTAGAAAAGCAAATGGGG  89390

Query  262    CTGCGATCAGCTAGGGAATGGAGAGTCAGTAGTCCATGAGTCAGCAGGGCTGGGGACCTG  321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89389  CTGCGATCAGCTAGGGAATGGAGAGTCAGTAGTCCATGAGTCAGCAGGGCTGGGGACCTG  89330

Query  322    AATTCTGGAAACTTCATTCAGGGAAGGGATACAGGGCTAGCAAGGCTCTCTCGGAGCTGC  381
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89329  GATTCTGGAAACTTCATTCAGGGAAGGGATACAGGGCTAGCAAGGCTCTCTCGGAGCTGC  89270

Query  382    TCTCTCCTGACTCTACTTCCAAGCCTTCCCAGGCCTGGCAGAGAGAAACCACATCACAGA  441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89269  TCTCTCCTGACTCTACTTCCAAGCCTTCCCAGGCCTGGCAGAGAGAAACCACATCACAGA  89210

Query  442    TTTGTTTACGACATCACAATCTATAACATGATGCTTTATAAATGTTAGTTTACACGTTTC  501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89209  TTTGTTTACGACATCACAATCTATAACATGATGCTTTATAAATGTTAGTTTACACGTTTC  89150

Query  502    TGGAAACGACAAGTGACCAACACTGTTTAAGAGGCAAAGACTCTCAGAGAAATGGAGGAA  561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89149  TGGAAACGACAAGTGACCAACACTGTTTAAGAGGCAAAGACTCTCAGAGAAATGGAGGAA  89090

Query  562    CTCACCCAAAGTCACAAAGATAAGATAAGCAAMGASTAAGTGTGGAGCCCACATCTTYTG  621
              |||||||||||||||||||||||||||||| || |||||||||||||||||||||| ||
Sbjct  89089  CTCACCCAAAGTCACAAAGATAAGATAAGCAACGACTAAGTGTGGAGCCCACATCTTCTG  89030

Query  622    ACTYAGAAACCTGGGCT-CTCAGG-AGCAGGCT-GAT-G-CAG-TCAT  663
              ||| |||||||||||| |||||| |||||||| ||| | ||| ||||
Sbjct  89029  ACTCAGAAACCTGGGCTTCTCAGGGAGCAGGCTTGATTGGCAGGTCAT  88982
```

<u>CDS at</u>: 88951 of chr3

<u>Intron 2 cut at:</u> 90354-90374 of chr3

Appendix II: regulatory regions modified by Cas9 in the K1CRISPR line

Modified element 1: RSINE1 sequence in intron 2

Genomic position: chr3:87139676-87139828

GGGCTGGTGAGATGGCTCAGCAGGCAAAAGCACTTAATACTGTCTGTTAACCCGAATTTCCTCCCTAG
AACTCACCTACATGGTGGACGGAGAGAATAAACTCTTGAATAAGTTGTCTCCTGCCTACACCCTACAC
GTTTACTGAGGCAGGTG

RepeatMasker Information

Name: RSINE1 (link requires registration)
Family: B4
Class: SINE
SW Score: 513
Divergence: 31.5%
Deletions: 1.3%
Insertions: 6.5%
Begin in repeat: 3
End in repeat: 147
Left in repeat: 18
Position: chr3:87139676-87139828
Band: 3qF1
Genomic Size: 153
Strand: -

<u>Modified element 2:</u> EM10E0734476 enhancer sequence upstream of *kirrel1* exon 1 on the mouse chromosome 3

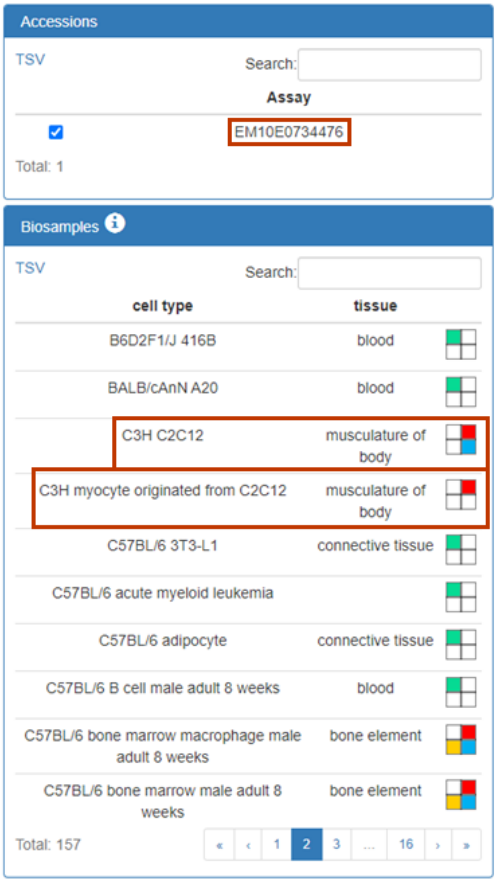
<u>Genomic position:</u> <u>chr3:87175201-87175507</u>

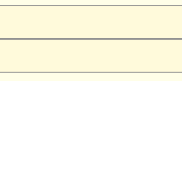
GGAGAAATCTTGACTGTCTTTGAGCCCCTTCTGAGGGCTTCAGCCAGCCACATTGAGGCCCGCCTCAGCGCTGGAG
TCCGTAGCAGCCATGAAGAGGCGCAGTTAACTCTTCCTTTGTTCACCCCAGACATTTATCTATCCGGAGGGAAAAC
CAAAGAACTCGGTGGCTTTCTCTTGTGCCTCCCCGAAGGCTTGCTACAGAAGGCAGCCACTGATCGATCTCTAAGC
ACTCAGGAACTTGAGGTGCACTGGGGTGAAGACATGGAGAATCGAAATAGAAGTGCATTCGGGTGGGTTCTCCAG
TCCA

ENCODE Candidate Cis-Regulatory Elements (cCREs) combined from all cell types (EM10E0734476)

cCRE details at ENCODE SCREEN: EM10E0734476
ENCODE Accession: EM10E0734476
Score: 204
Position: chr3:87175201-87175507
Band: 3qF1
Genomic Size: 307
View DNA for this feature (mm10/Mouse)

ENCODE classification	pELS,CTCF-bound
Max DNase Z-score	2.04098978295





Appendix III: day 7 differentiating K1CRISPR myoblasts

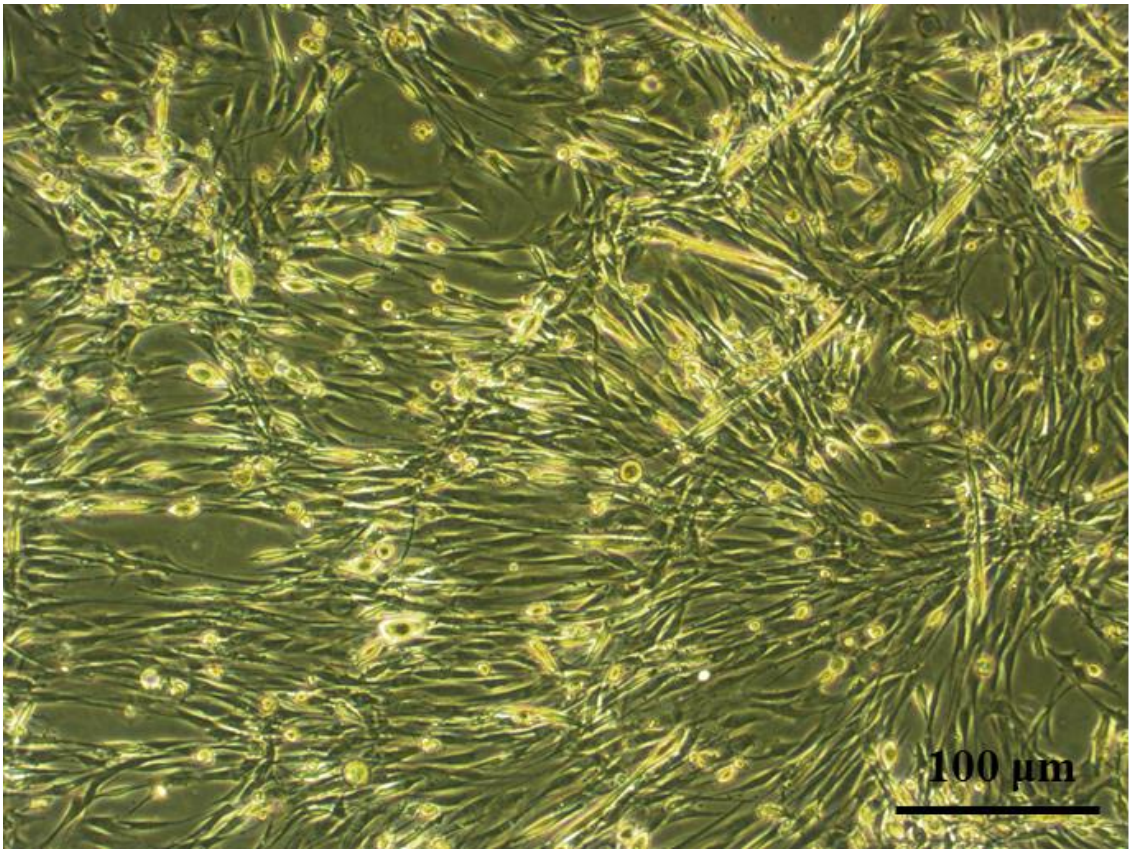


Figure I K1CRISPR myoblasts on day 7 of differentiation

Appendix IV: short hairpin RNA sequences

Kirrel1A-targeting short hairpins:

Plasmid 1 – forward oligo:
5’-GATCCCGCAGGAAAGAGAGAGGAAAGCTTCAAGAGAGCTTTCCTCTCTCTTTCCTGCTTTTTTGGAAA-3’
Plasmid 1 – reverse oligo:
5’-AGCTTTTCCAAAAAAGCAGGAAAGAGAGAGGAAAGCTCTCTTGAAGCTTTCCTCTCTCTTTCCTGCGG-3’

Plasmid 2 – forward oligo:
5’-GATCCCGAGCTCAACTAATGTGATTCTTTCAAGAGAAGAATCACATTAGTTGAGCTCTTTTTTGGAAA-3’
Plasmid 2 – reverse oligo:
5’-AGCTTTTCCAAAAAAGAGCTCAACTAATGTGATTCTTCTCTTGAAAGAATCACATTAGTTGAGCTCGG-3’

Plasmid 3 – forward oligo:
5’-GATCCCGCATAGCAATGTTAGGCAATTTTCAAGAGAAATTGCCTAACATTGCTATGCTTTTTTGGAAA-3’
Plasmid 3 – reverse oligo:
5’-AGCTTTTCCAAAAAAGCATAGCAATGTTAGGCAATTTCTCTTGAAAATTGCCTAACATTGCTATGCGG-3’

Kirrel1B-targeting short hairpins:

Plasmid 1 – forward oligo:
5’-GATCCCCGCTAATTCCTCTGTGCATTTAATTCAAGAGATTAAATGCACAGAGGAATTAGTTTTTTGGA-3’
Plasmid 1 – reverse oligo:
5’-AGCTTTTCCAAAAAACTAATTCCTCTGTGCATTTAATCTCTTGAATTAAATGCACAGAGGAATTAGGGC-3’

Plasmid 2 – forward oligo:
5’-GATCCCGCAGTTCCTATGTACCAACCTGCTCTTTCAAGAGAAGAGCAGGTTGGTACATAGGAACTGTTTTTTGGAAA-3’
Plasmid 2 – reverse oligo:
5’-AGCTTTTCCAAAAAACAGTTCCTATGTACCAACCTGCTCTTCTCTTGAAAGAGCAGGTTGGTACATAGGAACTGCGG-3’

Plasmid 3 – forward oligo:
5’-GATCCCACCAACCTGCTCTCAAATAAATTCAAGAGATTTATTTGAGAGCAGGTTGGTTTTTTTGGAAA-3’
Plasmid 3 – reverse oligo:
5’-AGCTTTTCCAAAAAAACCAACCTGCTCTCAAATAAATCTCTTGAATTTATTTGAGAGCAGGTTGGTGG-3’

Appendix V: day 7 differentiating K1Ash myoblasts

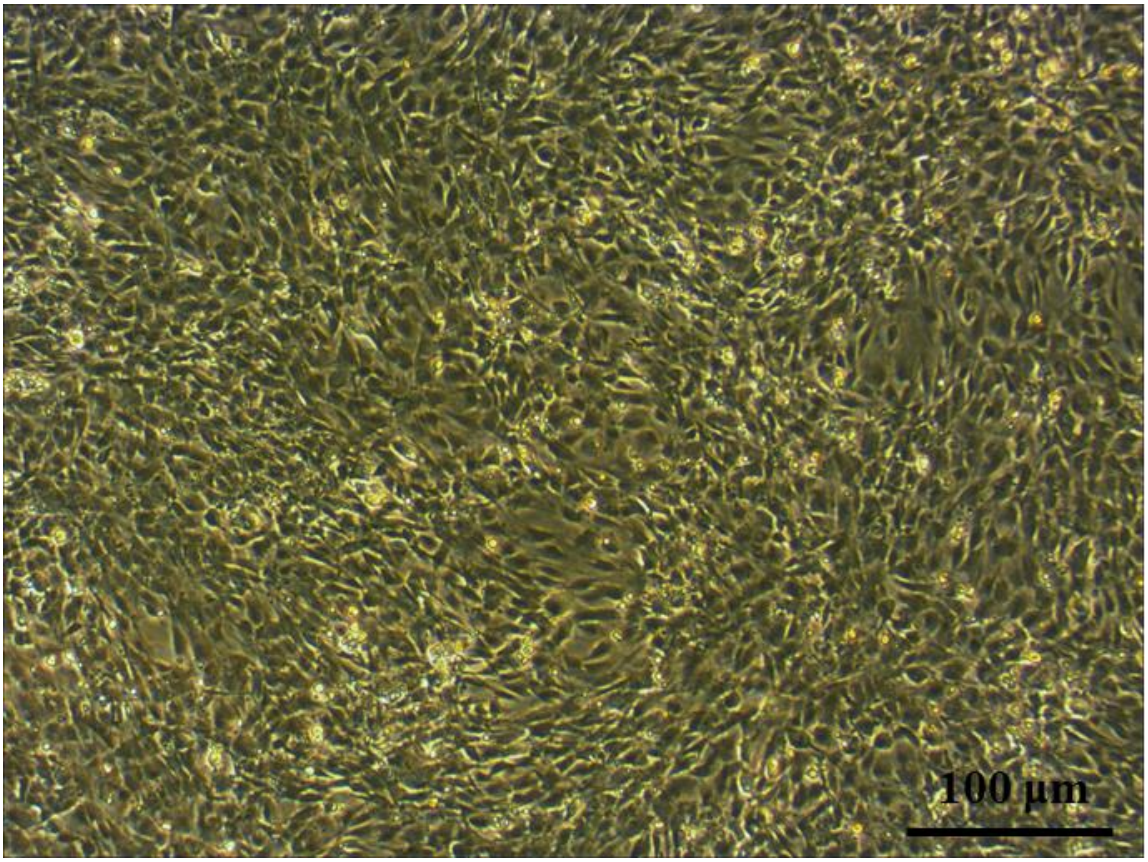


Figure II K1Ash myoblasts on day 7 of differentiation

Appendix VI: *kirrel1A* and *kirrel1B* retro-viral plasmid insert sequencing

Kirrel1A retroviral plasmid sequencing
read length: 2250 bp
alignment to NCBI reference mRNA sequence (mouse kirrel1A variant): NM_001170982.2

ACCCCGGATTCTAATGACCTGCCATCAAGCCTGCTCCCTGAGAAGCCCAGGTTTCTGAGTCAGAAGAT
GTGGGCTCCACACTTAGTCGTTGCTTATCTTATCTTTGTGACTTTGGCTTTGGCTTTGCCCGGGACTCAG
ACTCGCTTCAGCCAGGAGCCAGCTGATCAGACTGTGGTGGCCGGACAGCGGGCGGTGCTCCCGTGTGT
GCTCCTTAACTACTCTGGGATTGTACAGTGGACCAAGGACGGGCTGGCCCTGGGTATGGGCCAGGGCC
TCAAAGCCTGGCCACGGTACCGGGTCGTGGGCTCTGCGGATGCTGGGCAATACAACTTGGAGATCACA
GATGCCGAGCTGTCTGATGACGCTTCCTATGAGTGCCAGGCCACGGAGGCTGCCCTGCGCTCTCGGCG
GGCCAAACTCACCGTGCTCATTCCTCCAGAGGAAACAAGGATTGATGGGGGCCCGGTGATTCTGCTGC
AAGCAGGCACCCCCTACAACCTCACGTGCAGAGCATTTAATGCCAAACCTGCTGCCACCATCATTTGG
TTCCGAGATGGGACACAGCAGGAGGGGGCTGTGACTAGCACGGAGCTGCTGAAGGATGGGAAAAGGG
AGACCACAATCAGCCAACTGCTCATTGAGCCCACAGACCTAGACATTGGCCGCGTATTCACCTGTCGC
AGTATGAATGAGGCCATCCCCAATGGCAAGGAGACATCCATTGAGCTTGATGTGCACCACCCTCCCAC
AGTGACTCTGTCCATCGAGCCCCAGACAGTGCTGGAAGGCGAGCGTGTCATTTTTACATGCCAGGCCA
CAGCCAACCCAGAGATCTTGGGCTACAGGTGGGCCAAAGGGGGCTTCTTGATTGAAGACGCCCATGA
GAGTCGCTATGAGACAAACGTTGACTATTCCTTCTTCACGGAGCCTGTGTCTTGTGAGGTTTATAACAA
GTCGGGAGCACCAATGTCAGCACTTTAGTGAATGTCACTTCGCCCCCGGATTGTAGTTTACCCAAGCC
ACACACAGACATTGGATCTGATGTGACCCTCACTGTGTCTGGCTGGGAATCCTTCCTCACCTCACTGAC
AGAGACTCAACATGTCTGAACTTGGTTCTTGAGGTAACCAGCCATTCAACTGTGCTGAAGTCAGTGAC
AGCATGCTGCACTATACCTGCCGACATTCGTGCCTCGAATCGGAGTGCTGAGCGAGAGGTACCGCTTA
TGTAACGACTCTATCATTCCAGCGAGCGTACAGTTTGCTGTGAGAGTGATGGCGGTAAGTGGAGTGCT
TTTCGGAGTACCCCCACCTCCGGATCGAATTGCATGGCATGGAAGGAGAACTTCCTCGAGGTGGGGAC
CCTGGAACGCTACACCGTGGAGAGGACGAACTCAGGCAGCGGTGTGCTGTCCACGCTCACCATTAATA
ATGTCATGGAGGCGGACTTCCAGACCCACTACAACTGCACTGCCTGGAACAGCTTTGGACCAGGCACA
GCCATCATCCAGCTGGAAGAGCGAGAGGTGTTACCTGTGGGCATCATTGCCGGGGCCACCATCGGTGC
CGGCATCCTGGTCGTCTTCTCTTTTGCTGCCTTAGTGTTCTTCCTCTACCGACGTCGCAAAGGCAGTCGA
AAGGATGTGACGTTGAGGAAGCTGGACATCAAGGTGGAGACGGTGAATCGGGAGCCACTTACGATGC
ACTCTGACCGGGAGGATGATACTGCCAGCATTTCCACGGCAACGCGGGTCATGAAGGCCATCTACTCG
TCCTTTAAGGATGATGTGGATCTGAAGCAGGACCTGCGCTGTGACACCATTGACACCCGGGAAGAGTA
TGAGATGAAGGATCCCACCAATGGTTATTACAATGTGCGCGCCCACGAAGATCGCCCGTCCTCCAGGG
CGGTGCTGTATGCTGACTACCGTGCCCCTGGCCCTACTCGTTTTGATGGGCGCCCATCATCCCGCCTGT
CCCACTCCAGTGGTTATGCCCAGCTCAATACGTACAGCCGGGCCCCTGCCTCTGACTATGGCACAGAG
CCTACACCCTCTGGCCCTTCTGCTCCGGGTGGCACCGATACGACCAGCCAGCTGTCCTACGAGAACTAT
GAGAAGTTCAACTCCCACCCCTTTCCCGGGGCAGCTGGGTATCCTACATACCGTCTAGGCTACCCCCA
GGCCCCACCCTCTGGCCTGGAGAGGACCCCCTACGAAGCGTATGACCCTATTGGCAAGTATGCCACCG
CCACTCGGTTCTCCTACACCTCTCAGCCTCTAGCACCGCGTGCAGGACATATACGAACCACCA

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Mus musculus kirre like nephrin family adhesion molecule 1 (Kirrel), transcript variant 1, mRNA	Mus musculus	3875	3875	98%	0.0	97.31%	7076	NM_001170985.2

Mus musculus kirre like nephrin family adhesion molecule 1 (Kirrel), transcript variant 1, mRNA

Sequence ID: NM_001170985.2 Length: 7076 Number of Matches: 1

Range 1: 182 to 2479 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
3875 bits(2098)	0.0	2246/2308(97%)	48/2308(2%)	Plus/Plus

```
Query  10    TCTAATGACCTGCC-ATCAAGCCTGCTCCCTGAGAAGCCCAGGTTTCTGAGTCAGAAGAT  68
             ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  182   TCTAATGACCTGCCAATCAAGCCTGCTCCCTGAGAAGCCCAGGTTTCTGAGTCAGAAGAT  241

Query  69    GTGGGCTCCACACTTAGTCGTTGCTTATCTTATCTTTGTGACTTTGGCTTTGGCTTTGCC  128
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  242   GTGGGCTCCACACTTAGTCGTTGCTTATCTTATCTTTGTGACTTTGGCTTTGGCTTTGCC  301

Query  129   CGGGACTCAGACTCGCTTCAGCCAGGAGCCAGCTGATCAGACTGTGGTGGCCGGACAGCG  188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  302   CGGGACTCAGACTCGCTTCAGCCAGGAGCCAGCTGATCAGACTGTGGTGGCCGGACAGCG  361

Query  189   GGCGGTGCTCCCGTGTGTGCTCCTTAACTACTCTGGGATTGTACAGTGGACCAAGGACGG  248
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  362   GGCGGTGCTCCCGTGTGTGCTCCTTAACTACTCTGGGATTGTACAGTGGACCAAGGACGG  421

Query  249   GCTGGCCCTGGGTATGGGCCAGGGCCTCAAAGCCTGGCCACGGTACCGGGTCGTGGGCTC  308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  422   GCTGGCCCTGGGTATGGGCCAGGGCCTCAAAGCCTGGCCACGGTACCGGGTCGTGGGCTC  481

Query  309   TGCGGATGCTGGGCAATACAACTTGGAGATCACAGATGCCGAGCTGTCTGATGACGCTTC  368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  482   TGCGGATGCTGGGCAATACAACTTGGAGATCACAGATGCCGAGCTGTCTGATGACGCTTC  541

Query  369   CTATGAGTGCCAGGCCACGGAGGCTGCCCTGCGCTCTCGGCGGGCCAAACTCACCGTGCT  428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  542   CTATGAGTGCCAGGCCACGGAGGCTGCCCTGCGCTCTCGGCGGGCCAAACTCACCGTGCT  601

Query  429   CATTCCTCCAGAGGAAACAAGGATTGATGGGGGCCCGGTGATTCTGCTGCAAGCAGGCAC  488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  602   CATTCCTCCAGAGGAAACAAGGATTGATGGGGGCCCGGTGATTCTGCTGCAAGCAGGCAC  661

Query  489   CCCCTACAACCTCACGTGCAGAGCATTTAATGCCAAACCTGCTGCCACCATCATTTGGTT  548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  662   CCCCTACAACCTCACGTGCAGAGCATTTAATGCCAAACCTGCTGCCACCATCATTTGGTT  721

Query  549   CCGAGATGGGACACAGCAGGAGGGGGCTGTGACTAGCACGGAGCTGCTGAAGGATGGGAA  608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  722   CCGAGATGGGACACAGCAGGAGGGGGCTGTGACTAGCACGGAGCTGCTGAAGGATGGGAA  781

Query  609   AAGGGAGACCACAATCAGCCAACTGCTCATTGAGCCCACAGACCTAGACATTGGCCGCGT  668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  782   AAGGGAGACCACAATCAGCCAACTGCTCATTGAGCCCACAGACCTAGACATTGGCCGCGT  841

Query  669   ATTCACCTGTCGCAGTATGAATGAGGCCATCCCCAATGGCAAGGAGACATCCATTGAGCT  728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  842   ATTCACCTGTCGCAGTATGAATGAGGCCATCCCCAATGGCAAGGAGACATCCATTGAGCT  901

Query  729   TGATGTGCACCACCCTCCCACAGTGACTCTGTCCATCGAGCCCCAGACAGTGCTGGAAGG  788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  902   TGATGTGCACCACCCTCCCACAGTGACTCTGTCCATCGAGCCCCAGACAGTGCTGGAAGG  961

Query  789   CGAGCGTGTCATTTTTACATGCCAGGCCACAGCCAACCCAGAGATCTTGGGCTACAGGTG  848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  962   CGAGCGTGTCATTTTTACATGCCAGGCCACAGCCAACCCAGAGATCTTGGGCTACAGGTG  1021

Query  849   GGCCAAAGGGGGCTTCTTGATTGAAGACGCCCATGAGAGTCGCTATGAGACAAACGTTGA  908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1022  GGCCAAAGGGGGCTTCTTGATTGAAGACGCCCATGAGAGTCGCTATGAGACAAACGTTGA  1081
```

```
Query  909   CTATTCCTTCTTCACGGAGCCTGTGTCTTGTGAGGTTTATAAC-AAGTCGGGAGCACCAA  967
             ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Sbjct  1082  CTATTCCTTCTTCACGGAGCCTGTGTCTTGTGAGGTTTATAACAAAGTCGGGAGCACCAA  1141

Query  968   TGTCAGCACTTTAGTGAATG-TCACTTCG-CCCCCGGATTGTAGTTTACCCAA-GCC-AC  1023
             |||||||||||||||||||| |||||||| |||||||||||||||||||||| ||| ||
Sbjct  1142  TGTCAGCACTTTAGTGAATGTTCACTTCGCCCCCCGGATTGTAGTTTACCCAAAGCCCAC  1201

Query  1024  -A-CACAGACATTGGATCTGATGTGACCCTCAC-TGTGTCTGGC-TGGGAATCCTTCC-T  1078
              | |||||||||||||||||||||||||||||| |||||||||  |||||||||| || |
Sbjct  1202  CACCACAGACATTGGATCTGATGTGACCCTCACCTGTGTCTGGGTTGGGAATCCTCCCCT  1261

Query  1079  CACC-TCACTGACAGAGACTCAACATGT-CTGAA-CTTGGTTCTTGAGGTAACCAGCCAT  1135
             |||| ||||   | | ||| ||| | |  || || | |||| || ||| |||| ||| ||
Sbjct  1262  CACCCTCAC---CTG-GAC-CAAGAAGGACTCAAACATGGTCCT-GAG-TAAC-AGCAAT  1313

Query  1136  TCAACTG-TGCTGAAGTCAGTGAC--AG-CA--TGCTG-CA-CTATACCTGCCGA-CATT  1186
              |||||| ||||||||||||||||  || ||  ||||| || ||||||||||||  |  |
Sbjct  1314  -CAACTGTTGCTGAAGTCAGTGACCCAGGCAGATGCTGGCACCTATACCTGCCGGGCCAT  1372

Query  1187  CGTGCCTCGAATCGGAGT-GCTGAGCGAGAGGTACCGCTT-ATGTAA-CG-AC-TC-TAT  1240
             |||||||| |||||||| |||||||||||||||||||| |||||| || || || |||
Sbjct  1373  CGTGCCTCGGATCGGAGTGGCTGAGCGAGAGGTACCGCTTTATGTAAACGGACCTCCTAT  1432

Query  1241  CAT-TCCAGCGAG-C-GTACAGTTTGCTGTGAGA-GTGATGGCGGTAAG-TGGAGTGCTT  1295
             ||| ||||||||| | |||||||||||||||||| |||||||||||||| ||||||||||
Sbjct  1433  CATCTCCAGCGAGGCGGTACAGTTTGCTGTGAGAGGTGATGGCGGTAAGGTGGAGTGCTT  1492

Query  1296  T-TCGG-AGTACCCCCACCTCCGGATCGAATTGCAT-GGCATGGAAGGAGAACTTCCTCG  1352
             | |||| |||| |||||||||||||||||||||||| |||||||||||||||||||||||
Sbjct  1493  TATCGGGAGTA-CCCCACCTCCGGATCGAATTGCATGGGCATGGAAGGAGAACTTCCTCG  1551

Query  1353  AGGTGGGGACCCTGGAACGCTACACCGTGGAGAGGACGAACTCAGGCAGCGGTGTGCTGT  1412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1552  AGGTGGGGACCCTGGAACGCTACACCGTGGAGAGGACGAACTCAGGCAGCGGTGTGCTGT  1611

Query  1413  CCACGCTCACCATTAATAATGTCATGGAGGCGGACTTCCAGACCCACTACAACTGCACTG  1472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1612  CCACGCTCACCATTAATAATGTCATGGAGGCGGACTTCCAGACCCACTACAACTGCACTG  1671

Query  1473  CCTGGAACAGCTTTGGACCAGGCACAGCCATCATCCAGCTGGAAGAGCGAGAGGTGTTAC  1532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1672  CCTGGAACAGCTTTGGACCAGGCACAGCCATCATCCAGCTGGAAGAGCGAGAGGTGTTAC  1731

Query  1533  CTGTGGGCATCATTGCCGGGGCCACCATCGGTGCCGGCATCCTGGTCGTCTTCTCTTTTG  1592
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1732  CTGTGGGCATCATTGCCGGGGCCACCATCGGTGCCGGCATCCTGGTCGTCTTCTCTTTTG  1791

Query  1593  CTGCCTTAGTGTTCTTCCTCTACCGACGTCGCAAAGGCAGTCGAAAGGATGTGACGTTGA  1652
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1792  CTGCCTTAGTGTTCTTCCTCTACCGACGTCGCAAAGGCAGTCGAAAGGATGTGACGTTGA  1851

Query  1653  GGAAGCTGGACATCAAGGTGGAGACGGTGAATCGGGAGCCACTTACGATGCACTCTGACC  1712
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1852  GGAAGCTGGACATCAAGGTGGAGACGGTGAATCGGGAGCCACTTACGATGCACTCTGACC  1911

Query  1713  GGGAGGATGATACTGCCAGCATTTCCACGGCAACGCGGGTCATGAAGGCCATCTACTCGT  1772
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1912  GGGAGGATGATACTGCCAGCATTTCCACGGCAACGCGGGTCATGAAGGCCATCTACTCGT  1971

Query  1773  CCTTTAAGGATGATGTGGATCTGAAGCAGGACCTGCGCTGTGACACCATTGACACCCGGG  1832
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1972  CCTTTAAGGATGATGTGGATCTGAAGCAGGACCTGCGCTGTGACACCATTGACACCCGGG  2031

Query  1833  AAGAGTATGAGATGAAGGATCCCACCAATGGTTATTACAATGTGCGCGCCCACGAAGATC  1892
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2032  AAGAGTATGAGATGAAGGATCCCACCAATGGTTATTACAATGTGCGCGCCCACGAAGATC  2091
```

```
Query  1893  GCCCGTCCTCCAGGGCGGTGCTGTATGCTGACTACCGTGCCCCTGGCCCTACTCGTTTTG  1952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2092  GCCCGTCCTCCAGGGCGGTGCTGTATGCTGACTACCGTGCCCCTGGCCCTACTCGTTTTG  2151

Query  1953  ATGGGCGCCCATCATCCCGCCTGTCCCACTCCAGTGGTTATGCCCAGCTCAATACGTACA  2012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2152  ATGGGCGCCCATCATCCCGCCTGTCCCACTCCAGTGGTTATGCCCAGCTCAATACGTACA  2211

Query  2013  GCCGGGCCCCTGCCTCTGACTATGGCACAGAGCCTACACCCTCTGGCCCTTCTGCTCCGG  2072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2212  GCCGGGCCCCTGCCTCTGACTATGGCACAGAGCCTACACCCTCTGGCCCTTCTGCTCCGG  2271

Query  2073  GTGGCACCGATACGACCAGCCAGCTGTCCTACGAGAACTATGAGAAGTTCAACTCCCACC  2132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2272  GTGGCACCGATACGACCAGCCAGCTGTCCTACGAGAACTATGAGAAGTTCAACTCCCACC  2331

Query  2133  CCTTTCCCGGGGCAGCTGGGTATCCTACATACCGTCTAGGCTACCCCCAGGCCCCACCCT  2192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2332  CCTTTCCCGGGGCAGCTGGGTATCCTACATACCGTCTAGGCTACCCCCAGGCCCCACCCT  2391

Query  2193  CTGGCCTGGAGAGGACCCCCTACGAAGCGTATGACCCTATTGGCAAGTATGCCACCGCCA  2252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2392  CTGGCCTGGAGAGGACCCCCTACGAAGCGTATGACCCTATTGGCAAGTATGCCACCGCCA  2451

Query  2253  CTCGGTTCTCCTACACCTCTCAGC-CTC  2279
              |||||||||||||||||||||||| |||
Sbjct  2452  CTCGGTTCTCCTACACCTCTCAGCACTC  2479
```

Kirrel1B retroviral plasmid sequencing
read length: 1857 bp
alignment to NCBI reference mRNA sequence (mouse kirrel1B variant): BC023765.1

TACCCGCTCTATGACCTGCCATCAAGCCTGCTCCCTGAGAAGCCCAGGTTTCTGAGTCAGAAGATGTG
GGCTCCACACTTAGTCGTTGCTTATCTTATCTTTGTGACTTTGGCTTTGGCTTTGCCCGGGACTCAGACT
CGCTTCAGCCAGGAGCCAGCTGATCAGACTGTGGTGGCCGGACAGCGGGCGGTGCTCCCGTGTGTGCT
CCTTAACTACTCTGGGATTGTACAGTGGACCAAGGACGGGCTGGCCCTGGGTATGGGCCAGGGCCTCA
AAGCCTGGCCACGGTACCGGGTCGTGGGCTCTGCGGATGCTGGGCAATACAACTTGGAGATCACAGAT
GCCGAGCTGTCTGATGACGCTTCCTATGAGTGCCAGGCCACGGAGGCTGCCCTGCGCTCTCGGCGGGC
CAAACTCACCGTGCTCATTCCTCCAGAGGAAACAAGGATTGATGGGGGCCCGGTGATTCTGCTGCAAG
CAGGCACCCCCTACAACCTCACGTGCAGAGCATTTAATGCCAAACCTGCTGCCACCATCATTTGGTTCC
GAGATGGGACACAGCAGGAGGGGGCTGTGACTAGCACGGAGCTGCTGAAGGATGGGAAAAGGGAGA
CCACAATCAGCCAACTGCTCATTGAGCCCACAGACCTAGACATTGGCCGCGTATTCACCTGTCGCAGT
ATGAATGAGGCCGGGTATTACCTGTCCAGATTGAATGCATCGATGCAGAACAGCATGAGCTTGATTGC
CACCCTCCACCAGTGACTCTCATCGAGCCCAGACAGTGCTGTGCGAGGTCATTTTACATGCAGTCACA
GCACCAGAGATCTGGGCTACAGTGGCCAAAGGGGCTTCTGATGAAGACGCCCATGAGAGTCGCTATG
AGACAAACGTGACTATTCCTTCTTCACGGAGCCTGTGTCTTGTGAGGTTATAACAAAGTCGGGAGCAC
CAATGTCAGCACTTTAGTGAATGTTCACTTCGCCCCCCCGGATTGTAGTTTACCCAAAGCCCACCACCA
CAGACATTGGATCTGATGTGACCCTCACCTGTGTCTGGGTTGGGAATCCTCCCCTCACCCTCACCTGGA
CCAAGAAGGACTCAAACATGGTCCTGAGTAACAGCAATCAACTGTTGCTGAAGTCAGTGACCCAGGC
AGATGCTGGCACCTATACCTGCCGGGCCATCGTGCCTCGGATCGGAGTGGCTGAGCGAGAGGTACCGC
TTTATGTAAACGGACCTCCTATCATCTCCAGCGAGGCGGTACAGTTTGCTGTGAGAGGTGATGGCGGT
AAGGTGGAGTGCTTTATCGGGAGTACCCCACCTCCGGATCGAATTGCATGGGCATGGAAGGAGAACTT
CCTCGAGGTGGGGACCCTGGAACGCTACACCGTGGAGAGGACGAACTCAGGCAGCGGTGTGCTGTCC
ACGCTCACCATTAATAATGTCATGGAGGCGGACTTCCAGACCCACTACAACTGCACTGCCTGGAACAG
CTTTGGACCAGGCACAGCCATCATCCAGCTGGAAGAGCGAGAGGTGTTACCTGTGGGCATCATTGCCG
GGGCCACCATCGGTGCCGGCATCCTGGTCGTCTTCTCTTTTGCTGCCTTAGTGTTCTTCCTCTACCGACG
TCGCAAAGGCAGTCGAAAGGATGTGACGTTGAGGAAGCTGGACATCAAGGTGGAGACGGTGAATCGG
GAGCCACTTACGATGCACTCTGACCGGGAGGATGATACTGCCAGCATTTCCACGGCAACGCGGGTCAT
GAAGGCCATCTACTCGGTGAGGATAATGCTCCTTTCAACGTGGCCTAGGCTCTTCATTGTCCTACGGCC
AGCCCCACCAGGAAACGCA

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Mus musculus kin of IRRE like (Drosophila), mRNA (cDNA clone MGC:38329 IMAGE:5343720), comple...	Mus musculus	3090	3090	99%	0.0	96.90%	2228	BC023765.1

Mus musculus kin of IRRE like (Drosophila), mRNA (cDNA clone MGC:38329 IMAGE:5343720)

Sequence ID: BC023765.1 **Length:** 2228 **Number of Matches:** 1

Range 1: 205 to 2057 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
3090 bits(1673)	0.0	1815/1873(97%)	51/1873(2%)	Plus/Plus

```
Query  8     TCT-ATGACCTGCC-ATCAAGCCTGCTCCCTGAGAAGCCCAGGTTTCTGAGTCAGAAGAT  65
             ||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  205   TCTAATGACCTGCCAATCAAGCCTGCTCCCTGAGAAGCCCAGGTTTCTGAGTCAGAAGAT  264

Query  66    GTGGGCTCCACACTTAGTCGTTGCTTATCTTATCTTTGTGACTTTGGCTTTGGCTTTGCC  125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  265   GTGGGCTCCACACTTAGTCGTTGCTTATCTTATCTTTGTGACTTTGGCTTTGGCTTTGCC  324

Query  126   CGGGACTCAGACTCGCTTCAGCCAGGAGCCAGCTGATCAGACTGTGGTGGCCGGACAGCG  185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  325   CGGGACTCAGACTCGCTTCAGCCAGGAGCCAGCTGATCAGACTGTGGTGGCCGGACAGCG  384

Query  186   GGCGGTGCTCCCGTGTGTGCTCCTTAACTACTCTGGGATTGTACAGTGGACCAAGGACGG  245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  385   GGCGGTGCTCCCGTGTGTGCTCCTTAACTACTCTGGGATTGTACAGTGGACCAAGGACGG  444

Query  246   GCTGGCCCTGGGTATGGGCCAGGGCCTCAAAGCCTGGCCACGGTACCGGGTCGTGGGCTC  305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  445   GCTGGCCCTGGGTATGGGCCAGGGCCTCAAAGCCTGGCCACGGTACCGGGTCGTGGGCTC  504

Query  306   TGCGGATGCTGGGCAATACAACTTGGAGATCACAGATGCCGAGCTGTCTGATGACGCTTC  365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  505   TGCGGATGCTGGGCAATACAACTTGGAGATCACAGATGCCGAGCTGTCTGATGACGCTTC  564

Query  366   CTATGAGTGCCAGGCCACGGAGGCTGCCCTGCGCTCTCGGCGGGCCAAACTCACCGTGCT  425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  565   CTATGAGTGCCAGGCCACGGAGGCTGCCCTGCGCTCTCGGCGGGCCAAACTCACCGTGCT  624

Query  426   CATTCCTCCAGAGGAAACAAGGATTGATGGGGGCCCGGTGATTCTGCTGCAAGCAGGCAC  485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  625   CATTCCTCCAGAGGAAACAAGGATTGATGGGGGCCCGGTGATTCTGCTGCAAGCAGGCAC  684

Query  486   CCCCTACAACCTCACGTGCAGAGCATTTAATGCCAAACCTGCTGCCACCATCATTTGGTT  545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  685   CCCCTACAACCTCACGTGCAGAGCATTTAATGCCAAACCTGCTGCCACCATCATTTGGTT  744

Query  546   CCGAGATGGGACACAGCAGGAGGGGGCTGTGACTAGCACGGAGCTGCTGAAGGATGGGAA  605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  745   CCGAGATGGGACACAGCAGGAGGGGGCTGTGACTAGCACGGAGCTGCTGAAGGATGGGAA  804

Query  606   AAGGGAGACCACAATCAGCCAACTGCTCATTGAGCCCACAGACCTAGACATTGGCCGCGT  665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  805   AAGGGAGACCACAATCAGCCAACTGCTCATTGAGCCCACAGACCTAGACATTGGCCGCGT  864

Query  666   ATTCACCTGTCGCAGTATGAATGAGGCCGGGTATTACCTGTCCAGATTGAATGCATCGAT  725
             |||||||||||||||||||||||||||    | | ||   ||| | ||   |||   |
Sbjct  865   ATTCACCTGTCGCAGTATGAATGAGGCC----A-T-CC---CCA-A-TG---GCA---A-  906

Query  726   GCAGAACA-GCA-TGAGCTTGAT-TG--CCACCCTCCACCAGTGACTC--T-CATCGAG-  776
             | || |||  || |||||||||| ||  |||||||||  |||||||||  | |||||||
Sbjct  907   GGAG-ACATCCATTGAGCTTGATGTGCACCACCCTCCCACAGTGACTCTGTCCATCGAGC  965

Query  777   CCCAGACAGTGCTG---TGCGA--G-GTCA-TTTTACATG-CA-GTCACAG-C-A-CCAG  824
             |||||||||||||    ||||  | |||| ||||||||| || | ||||| | | ||||
Sbjct  966   CCCAGACAGTGCTGGAAGGCGAGCGTGTCATTTTTACATGCCAGGCCACAGCCAACCCAG  1025
```

```
Query  825   AGATC-TGGGCTACA-GT-GGCCAAA-GGGGCTTC-TGA-TGAAGACGCCCATGAGAGTC  878
             ||||| ||||||||| || |||||||| |||||||| ||| ||||||||||||||||||||
Sbjct  1026  AGATCTTGGGCTACAGGTGGGCCAAAGGGGGCTTCTTGATTGAAGACGCCCATGAGAGTC  1085

Query  879   GCTATGAGACAAACG-TGACTATTCCTTCTTCACGGAGCCTGTGTCTTGTGAGG-TTATA  936
             |||||||||||||| |||||||||||||||||||||||||||||||||||| |||||
Sbjct  1086  GCTATGAGACAAACGTTGACTATTCCTTCTTCACGGAGCCTGTGTCTTGTGAGGTTTATA  1145

Query  937   ACAAAGTCGGGAGCACCAATGTCAGCACTTTAGTGAATGTTCACTTCGcccccccGGATT  996
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  1146  ACAAAGTCGGGAGCACCAATGTCAGCACTTTAGTGAATGTTCACTTCG-CCCCCCGGATT  1204

Query  997   GTAGTTTACCCAAAGCCCACCACCACAGACATTGGATCTGATGTGACCCTCACCTGTGTC  1056
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Sbjct  1205  GTAGTTTACCCAAAGCCCACCACCACAGACATTGGATCTGATGTGACCCTCACTTGTGTC  1264

Query  1057  TGGGTTGGGAATCCTCCCCTCACCCTCACCTGGACCAAGAAGGACTCAAACATGGTCCTG  1116
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1265  TGGGTTGGGAATCCTCCCCTCACCCTCACCTGGACCAAGAAGGACTCAAACATGGTCCTG  1324

Query  1117  AGTAACAGCAATCAACTGTTGCTGAAGTCAGTGACCCAGGCAGATGCTGGCACCTATACC  1176
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1325  AGTAACAGCAATCAACTGTTGCTGAAGTCAGTGACCCAGGCAGATGCTGGCACCTATACC  1384

Query  1177  TGCCGGGCCATCGTGCCTCGGATCGGAGTGGCTGAGCGAGAGGTACCGCTTTATGTAAAC  1236
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1385  TGCCGGGCCATCGTGCCTCGGATCGGAGTGGCTGAGCGAGAGGTACCGCTTTATGTAAAC  1444

Query  1237  GGACCTCCTATCATCTCCAGCGAGGCGGTACAGTTTGCTGTGAGAGGTGATGGCGGTAAG  1296
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1445  GGACCTCCTATCATCTCCAGCGAGGCGGTACAGTTTGCTGTGAGAGGTGATGGCGGTAAG  1504

Query  1297  GTGGAGTGCTTTATCGGGAGTACCCCACCTCCGGATCGAATTGCATGGGCATGGAAGGAG  1356
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1505  GTGGAGTGCTTTATCGGGAGTACCCCACCTCCGGATCGAATTGCATGGGCATGGAAGGAG  1564

Query  1357  AACTTCCTCGAGGTGGGGACCCTGGAACGCTACACCGTGGAGAGGACGAACTCAGGCAGC  1416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1565  AACTTCCTCGAGGTGGGGACCCTGGAACGCTACACCGTGGAGAGGACGAACTCAGGCAGC  1624

Query  1417  GGTGTGCTGTCCACGCTCACCATTAATAATGTCATGGAGGCGGACTTCCAGACCCACTAC  1476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1625  GGTGTGCTGTCCACGCTCACCATTAATAATGTCATGGAGGCGGACTTCCAGACCCACTAC  1684

Query  1477  AACTGCACTGCCTGGAACAGCTTTGGACCAGGCACAGCCATCATCCAGCTGGAAGAGCGA  1536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1685  AACTGCACTGCCTGGAACAGCTTTGGACCAGGCACAGCCATCATCCAGCTGGAAGAGCGA  1744

Query  1537  GAGGTGTTACCTGTGGGCATCATTGCCGGGGCCACCATCGGTGCCGGCATCCTGGTCGTC  1596
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1745  GAGGTGTTACCTGTGGGCATCATTGCCGGGGCCACCATCGGTGCCGGCATCCTGGTCGTC  1804

Query  1597  TTCTCTTTTGCTGCCTTAGTGTTCTTCCTCTACCGACGTCGCAAAGGCAGTCGAAAGGAT  1656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1805  TTCTCTTTTGCTGCCTTAGTGTTCTTCCTCTACCGACGTCGCAAAGGCAGTCGAAAGGAT  1864

Query  1657  GTGACGTTGAGGAAGCTGGACATCAAGGTGGAGACGGTGAATCGGGAGCCACTTACGATG  1716
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1865  GTGACGTTGAGGAAGCTGGACATCAAGGTGGAGACGGTGAATCGGGAGCCACTTACGATG  1924

Query  1717  CACTCTGACCGGGAGGATGATACTGCCAGCATTTCCACGGCAACGCGGGTCATGAAGGCC  1776
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1925  CACTCTGACCGGGAGGATGATACTGCCAGCATTTCCACGGCAACGCGGGTCATGAAGGCC  1984

Query  1777  ATCTACTCGGTGAGGATAATGCTCCTTTCAACGTGGCCTAGGCTCTTCATTGTCCTACGG  1836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1985  ATCTACTCGGTGAGGATAATGCTCCTTTCAACGTGGCCTAGGCTCTTCATTGTCCTACGG  2044

Query  1837  CCAGCCCCACCAG  1849
             |||||||||||||
Sbjct  2045  CCAGCCCCACCAG  2057
```

APPENDIX VII: MICROSCOPY IMAGES OF KIRREL1A-MCHERRY MUTANT MYOBLASTS

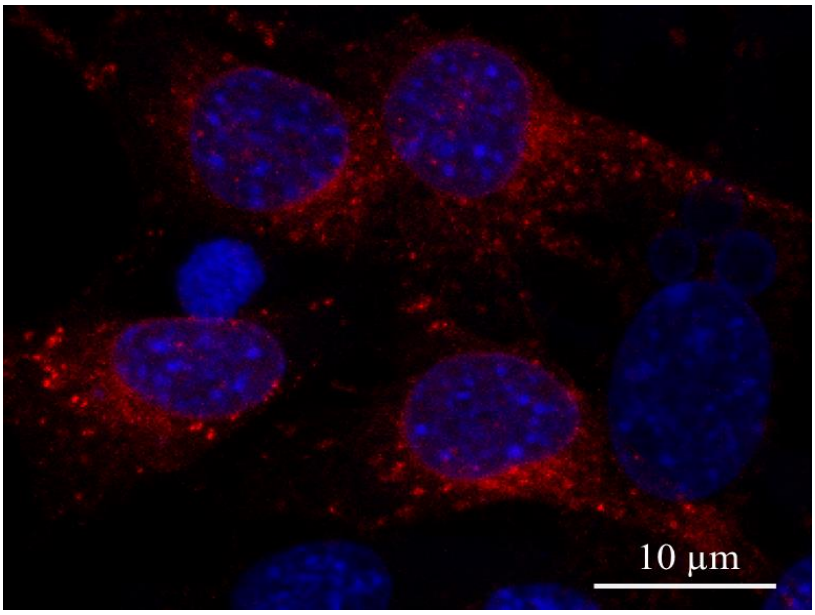


Figure III Confocal image of proliferating kirrel1A-mCherry mutant myoblasts.
Blue = Hoechst nuclear stain, Red = kirrel1A-mCherry

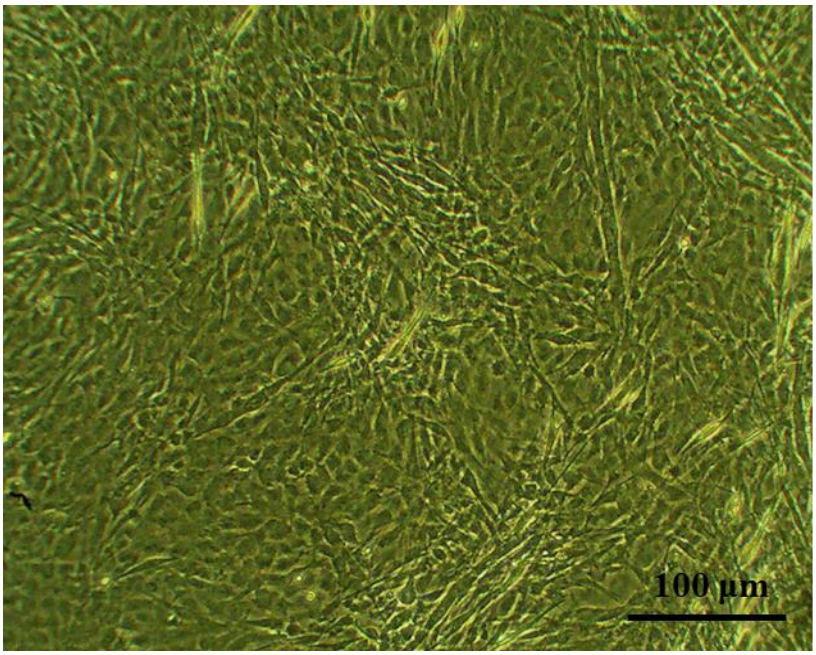


Figure IV Mcherry myoblasts on day 7 of differentiation

Appendix VIII: PCR thermal cycling conditions

Table I Genomic PCR cycling conditions

	Step	Duration
	1 cycle of 95°C	5 mins
25 cycles	95°C	30 secs
	57°C	30 secs
	72°C	1 min/kb
	Final extension 72 °C	8 mins

Table II Gene-cloning PCR cycling conditions

	Step	Duration
	1 cycle of 95°C	5 mins
30 cycles	95°C	30 secs
	57°C	30 secs
	72°C	1 min/kb
	Final extension 72 °C	8 mins

Table III Semi-quantitative PCR cycling conditions (*kirrel1A, kirrel1B* and *gapdh*)

	Step	Duration
	1 cycle of 95°C	5 mins
25 cycles	95°C	30 secs
	58°C	30 secs
	72°C	1 min/kb
	Final extension 72 °C	5 mins

Appendix IX: Western blot antibody dilutions

Table IV Western blotting antibody dilutions

target protein	primary antibody and dilution	secondary antibody dilution
1. kirrel1	rabbit polyclonal anti-kirrel IgG (Abcam, cat. 156084) - 1/800	HRP-linked goat anti-rabbit IgG (Cell Signalling Technology, cat. 7074S) - 1/10000
2. Pax7	mouse monoclonal anti-Pax7 IgG Developmental Studies hybridoma bank – no dilution required	HRP-linked horse anti-mouse IgG (Cell Signalling technology, cat. 7076S) – 1/10000
3. Myf5	mouse monoclonal anti-Myf5 IgG (Santa Cruz Biotechnology, cat. Sc-518039) – 1/800	HRP-linked horse anti-mouse IgG (Cell Signalling technology, cat. 7076S) – 1/10000
4. MyoD	mouse monoclonal anti-MyoD IgG (Dako, cat. M3512) – 1/800	HRP-linked horse anti-mouse IgG (Cell Signalling technology, cat. 7076S) – 1/10000
5. myogenin	rabbit polyclonal anti-myogenin IgG (Sigma-Aldrich, cat. SAB1305721) – 1/1000	HRP-linked goat anti-rabbit IgG (Cell Signalling Technology, cat. 7074S) - 1/10000
6. MyHC	rabbit polyclonal anti-fast MyHC IgG (Abcam, cat. Ab91506) – 1/800	HRP-linked goat anti-rabbit IgG (Cell Signalling Technology, cat. 7074S) - 1/10000
7. GAPDH	rabbit polyclonal anti-GAPDH IgG (Abcam, cat ab9485) – 1/2000	HRP-linked goat anti-rabbit IgG (Cell Signalling Technology, cat. 7074S) - 1/10000

Appendix X: ImageJ macros for myotube analysis and examples of processed phase contrast pictures

Differentiation day 0 and 1 ImageJ macro text commands

```
run("8-bit");
run("Despeckle");
run("Bandpass Filter...", "filter_large=30 filter_small=3 suppress=None tolerance=5 autoscale saturate");
run("Auto Local Threshold", "method=Sauvola radius=25 parameter_1=0 parameter_2=0 white");
run("Invert");
setOption("BlackBackground", false);
run("Close-");
run("Convert to Mask");
//run("Threshold...");
run("Analyze Particles...", "size=3500-4000 show=Overlay display clear summarize overlay");
```

Differentiation day 2 ImageJ macro text commands

```
run("8-bit");
run("Despeckle");
run("Bandpass Filter...", "filter_large=30 filter_small=3 suppress=None tolerance=5 autoscale saturate");
run("Auto Local Threshold", "method=Sauvola radius=25 parameter_1=0 parameter_2=0 white");
run("Invert");
setOption("BlackBackground", false);
run("Close-");
run("Convert to Mask");
//run("Threshold...");
run("Analyze Particles...", "size=2000-5000 show=Overlay display clear summarize overlay");
```

Differentiation day 3 and 4 ImageJ macro text commands

```
run("8-bit");
run("Despeckle");
run("Bandpass Filter...", "filter_large=30 filter_small=3 suppress=None tolerance=5 autoscale saturate");
run("Auto Local Threshold", "method=Sauvola radius=15 parameter_1=0 parameter_2=0 white");
run("Invert");
setOption("BlackBackground", false);
run("Close-");
run("Convert to Mask");
//run("Threshold...");
run("Analyze Particles...", "size=1500-Infinity show=Overlay display clear summarize overlay");
```

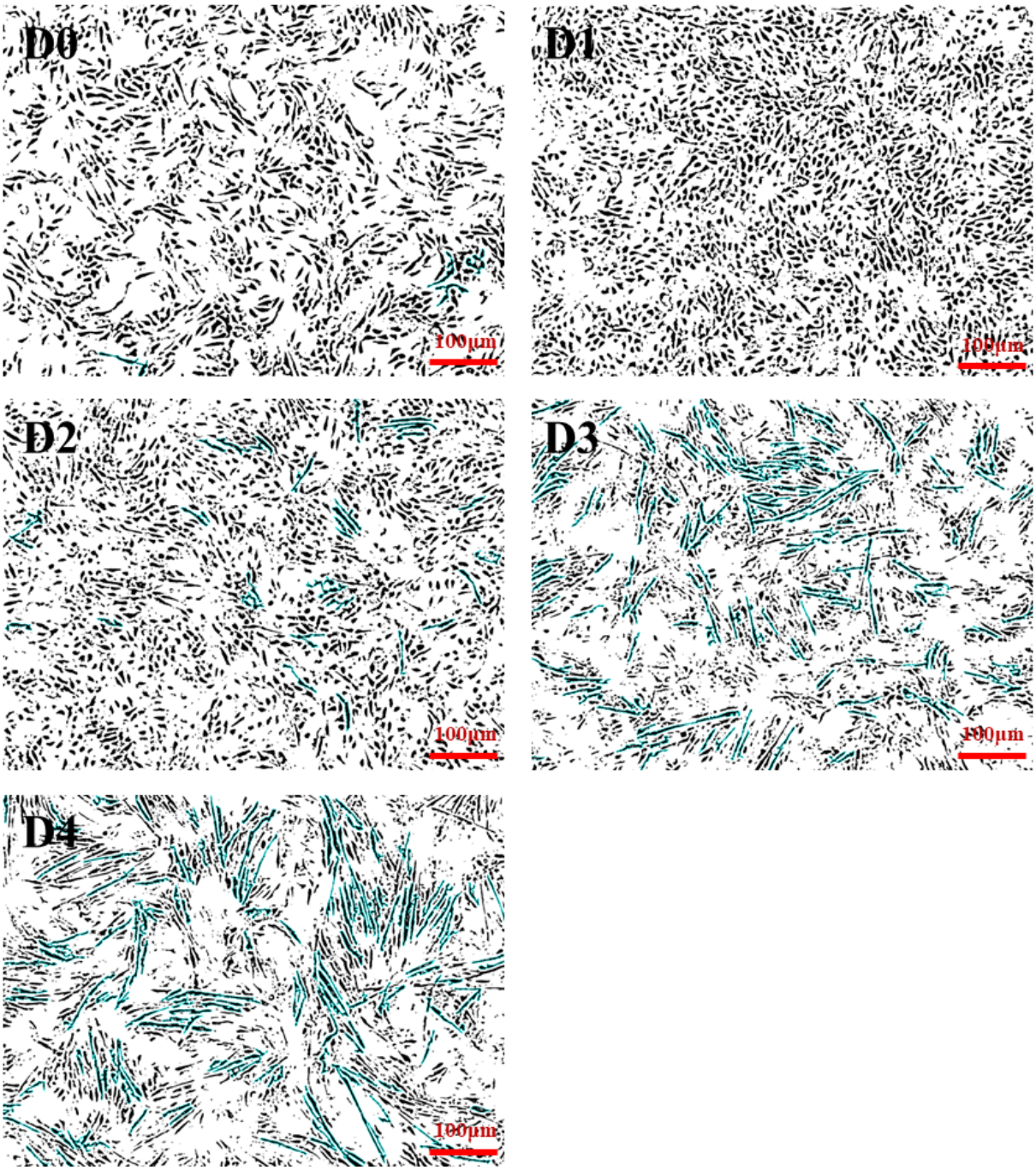


Figure V Examples of ImageJ processed pictures of differentiating myoblasts for myotube analysis