LncRNA-mediated Gene Regulation in Cardiopulmonary Development, Homeostasis and Disease

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich Biowissenschaften (FB15)

der Johann Wolfang Goethe-Universität

in Frankfurt am Main

von

Sandra Rogala

aus Hanau

Frankfurt 2023

Vom Fachbereich Biowissenschaften (FB15)

der Johann Wolfgang Goethe - Universität als Dissertation angenommen.

Dekan: Prof. Dr. Sven Klimpel

Gutachter: Prof. Dr. Michaela Müller-McNicoll

                  Prof. Dr. Jaya Krishnan

Externer Supervisor: Dr. Phillip Grote

Datum der Disputation:

Disclaimer:

While completing this work, two scientific publications, one about the Swhtr project and one about the Fendrr project, were handed in, one of which has been accepted for publication in Nucleic Acids Research and the other has been under revision:

*Ali T, **Rogala S**, Krause NM, Bains JK, Melissari MT, Währisch S, Schwalbe H, Herrmann BG, Grote P. *Fendrr* synergizes with Wnt signalling to regulate fibrosis related genes during lung development via its RNA:dsDNA triplex element. Nucleic Acids Res. 2023 Jul 7;51(12):6227-6237. doi: 10.1093/nar/gkad395. PMID: 37207329; PMCID: PMC10325902.
*equal contribution

Rogala S, Ali T, Melissari MT, Währisch S, Schuster P, Sarre A, Boettger T, Rogg EM, Kaur J, Krishnan K, Dimmeler S, Ounzain S, Pedrazzini T, Cordellini R, Dumbovic G, Herrmann BG, Grote P (2023) The lncRNA *Sweetheart* regulates compensatory cardiac hypertrophy after myocardial injury. (under revision in Nature Communications)

All data presented in these publications that has been generated without contribution of the author of this work are explained in the Introduction parts with added citations. A preliminary bioRxiv version (Rogala et al., 2022) was used for the Swhtr citation as the revision process was still ongoing at the time of completion of the thesis.
Hence, the author refrains from citation of her own publications within the Results parts, unless when referring to data from the Introduction. This is for reasons of readability and text esthetics.

Table of Contents

Fendrr synergizes with WNT signaling to regulate fibrosis related genes during lung development via its RNA:dsDNA Triplex Element

Initial characterization of the novel lncRNA *lncFsd2* in relation to its coding *cis*-located gene *Fsd2*

Table of Figures

Table of Tables

Abstract

Precise regulation of gene expression networks is required to develop and maintain a healthy organism before and after birth and throughout adulthood. Such networks are mostly comprised of regulatory proteins, but meanwhile many long non-coding transcripts (lncRNAs) are shown to participate in these regulatory processes. The functions and mechanisms of these lncRNAs vary greatly, however they are often associated with transcriptional regulation. Three lncRNAs, namely ***Sw**eet**h**eart **R**NA* (*Swhtr*), ***Fe**tal-lethal **n**oncoding **d**evelopmental **r**egulatory **R**NA / **F**oxf1 adjac**e**nt **n**on-Coding **d**evelopmental regulatory **R**NA* (*Fendrr*) and *lncFsd2*, were studied in this work to demonstrate the variety of cellular and biological processes that require lncRNA-mediated fine-tuning, in regard to the cardiopulmonary system.

Swhtr was found to be expressed exclusively in cardiomyocytes and became critical for regeneration after myocardial injury. Mice lacking *Swhtr* did not show issues under normal conditions, but failed to undergo compensatory hypertrophic remodeling after injury, leading to increased mortality. This effect was rescued by re-expressing *Swhtr*, demonstrating importance of the RNA. Genes dependent on *Swhtr* during cardiac stress were found to likely be regulated by NKX2-5 through physical interaction with *Swhtr*. *Fendrr* was found to be expressed in lung and interacted with target promoters through its RNA:dsDNA binding domain, the *FendrrBox*, which was partially required for *Fendrr* function. *Fendrr*, together with activated WNT signaling, regulated fibrosis related target genes via the *FendrrBox* in fibroblasts. *LncFsd2*, an ubiquitously expressed lncRNA, showed possible interaction with the striated muscle specific *Fsd2*, but its exact function and regulatory role remain unclear in muscle physiology. Immunoprecipitation and subcellular fractionation experiments suggest that *lncFsd2* might be involved in nuclear retention of *Fsd2* mRNA, thus fine-tuning FSD2 protein expression. These investigations have shed light on the roles of these lncRNAs in stress responses, fibrosis-related gene regulation, and localization processes, advancing our understanding of cardiovascular and pulmonary maintenance, reaction to injury, and diseases. The diverse and intricate roles of these three lncRNAs highlight how they influence various cellular processes and disease states, offering avenues for exploring lncRNA functions in different biological contexts.

Summary in German (Deutsche Zusammenfassung)

Lange nicht-kodierende RNAs (lncRNAs) sind eine heterogene Gruppe von RNA-Molekülen, die wie ihr Name bereits zeigt, nicht in Proteine übersetzt werden. LncRNA-Gene werden genauso transkribiert und verarbeitet wie protein-kodierende Gene, werden allerdings weniger effizient gespleißt und weisen generell eine niedrigere Transkriptionsrate auf. LncRNAs können verschiedene Funktionen in biologischen Prozessen wie Entwicklung, Organbildung, Homöostase und Stressreaktionen übernehmen. Da sie lediglich nach den Kriterien des niedrigen Kodierungspotentials und ihrer Länge, die 200 Nukleotide überschreitet, gruppiert werden, sind sie die vielfältigste Gruppe von nicht-kodierenden RNAs und können auf äußerst unterschiedliche Weisen wirken. Einige lncRNAs wirken als strukturelle Moleküle und organisieren Ribonukleoprotein-Komplexe oder ganze zelluläre Kompartimente. Andere lncRNAs dienen als Führungs- oder Verankerungsmoleküle für Proteine, während wieder andere als Schalter oder Regulatoren wirken, indem sie Konformationsänderungen in anderen RNA- oder DNA-Molekülen hervorrufen. Einige lncRNAs fungieren auch als Vorläufer für kleinere RNAs wie microRNAs (miRNAs).

Die Funktionen von lncRNAs sind oft gewebsspezifisch, da viele von ihnen von Enhancer-Elementen transkribiert werden, welche im Allgemeinen oft sehr spezifisch in bestimmten Geweben aktiv sind. Obwohl lncRNAs in der Regel wenig Sequenzkonservierung zwischen Spezies aufweisen, können einige von ihnen ähnliche Funktionen und strukturelle Merkmale in verschiedenen Organismen teilen. Ein bekanntes Beispiel für eine konservierte lncRNA ist *X-inactive specific transcript* (*XIST*), welches für die Inaktivierung des zweiten X-Chromosoms bei weiblichen Säugetieren verantwortlich ist. Es gibt jedoch auch lncRNAs, die artspezifisch sind und spezifische Funktionen in bestimmten Organismen haben. Die Konservierung von Funktion und Orthologen von lncRNAs zwischen Arten variiert daher und muss individuell bewertet werden. Insgesamt sind lncRNAs eine faszinierende und vielseitige Gruppe von RNA-Molekülen, die eine wichtige Rolle in der Regulation verschiedener biologischer Prozesse spielen können. Ihre genaue Funktion und ihr Mechanismus sind jedoch oft noch nicht vollständig verstanden und bieten ein reiches Forschungsgebiet.

Das kardiopulmonale System umfasst das Herz und die Lunge und ist entscheidend für die Sauerstoffversorgung des Körpers. Das Herz ist für die Pumpfunktion verantwortlich, die den Blutfluss durch den Körper aufrechterhält, während die Lunge den Gasaustausch ermöglicht, bei dem Sauerstoff aufgenommen und Kohlendioxid abgegeben wird. Das kardiopulmonale System arbeitet eng zusammen, um sicherzustellen, dass der Körper mit ausreichend Sauerstoff versorgt wird.

Störungen im kardiopulmonalen System können zu verschiedenen Erkrankungen führen, darunter Herzerkrankungen wie Herzinsuffizienz, koronare Herzkrankheit und angeborene Herzfehler sowie Lungenerkrankungen wie chronisch obstruktive Lungenerkrankung und Lungenkrebs. Die Behandlung von kardiopulmonalen Erkrankungen umfasst medizinische Interventionen wie Medikamente, chirurgische Eingriffe und Atemtherapie, um die Funktion des Herzens und der Lunge zu verbessern und die Symptome zu lindern. Ein gesundes kardiopulmonales System ist entscheidend für das allgemeine Wohlbefinden und die Lebensqualität einer Person. Faktoren wie zum Beispiel Mutationen in Genen, die eine entscheidende Rolle in der Entwicklung und Homöostase des kardiopulmonalen Systems haben, können verheerende Folgen haben. Mutationen im *NK2-Homeobox-Gen 5* (*NKX2-5*) beispielsweise sind mit verschiedenen Herzerkrankungen assoziiert, sowie die lncRNA *FENDRR*, die oft im Zusammenhang mit idiopathischer Lungenfibrose oder Krebs genannt wird.

In dieser Arbeit wurden drei verschiedene lncRNAs näher betrachtet und deren Funktion, insbesondere im kardiopulmonalen System, näher beleuchtet.

Das Projekt über die lncRNA *Sweetheart RNA* (*Swhtr*) konzentrierte sich darauf, die Rolle dieser lncRNA im Zusammenhang mit kardialem Stress und Krankheit am Model der Maus zu untersuchen. Die Ergebnisse zeigten, dass *Swhtr* in adulten Kardiomyozyten spezifisch exprimiert wird und sein Expressionsmuster dem seines benachbarten Transkriptionsfaktor-kodierenden Gens *Nkx2-5* folgt. Interessanterweise führte die genetische Inaktivierung von *Swhtr* während der embryonalen Entwicklung oder unter normalen Bedingungen bei adulten Herzen nicht zu offensichtlichen Defekten. Jedoch wurde deutlich, dass *Swhtr* unter Stressbedingungen eine wichtige Rolle spielt. Mäuse denen die *Swhtr* Expression aufgrund der genetischen Insertion eines starken

transkriptionellen Stopp-Signals fehlt, zeigten nach der Induktion eines akuten Myokardinfarkts (AMI) durch eine linke vordere absteigende (LAD) Koronararterienligatur eine reduzierte Überlebensrate und konnten keine kompensatorische kardiale Hypertrophie entwickeln. Diese Phänotypen konnten jedoch durch die Re-Expression von *Swhtr* durch zufällige Integration einem externen Ort im Genom wiederhergestellt werden, was die funktionale Bedeutung der RNA selbst betont.

In dieser Arbeit wurde auch die Interaktion zwischen *Swhtr* und NKX2-5 untersucht, was zeigte, dass *Swhtr* physisch an NKX2-5 bindet und diese Interaktion unter Stressbedingungen potenziell verstärkt wird. Interessanterweise war der *Swhtr*-spezifische Phänotyp in embryonalen und erwachsenen nicht-gestressten Herzen nicht erkennbar, was darauf hindeutet, dass eine Interaktion zwischen *Swhtr* und NKX2-5 nur unter stressbedingten Bedingungen erforderlich ist. Die Ergebnisse legen nahe, dass *Swhtr* eine Rolle bei der Initiierung einer hypertrophen Antwort im Herzen nach kardialem Stress spielt, was ein typischer kompensatorischer Mechanismus ist. Die Charakterisierung der *Swhtr*-abhängig regulierten Gene, enthüllte die Beteiligung des Phosphoinositid 3-Kinase (PI3-K)-Signalwegs sowie anderer Signalwege und biologischer Prozesse, die mit der Immunantwort, der Kalziumsignalisierung, der kardialen Muskelfunktion und der Ablagerung extrazellulärer Matrix in Verbindung stehen, welche für die hypertrophe Stressantwort unabdinglich sind. Einzelmolekulare fluoreszierende *In-situ*-Hybridisierung unter Hypoxie-induziertem Stress und normoxischen Kontrollkonditionen von primären Kardiomyozyten zeigten, dass *Swhtr* unter Standardbedingungen am Ort seiner Transkription verbleibt, erkennbar an zwei deutlichen fluoreszierenden Punkten, sich aber unter Stressbedingungen in mehrere Punkte zerstreut.

Darüber hinaus wurden zwei weitere zuvor beschriebene lncRNAs, *IRENE-SS* und *IRENE-div*, innerhalb des *Nkx2-5*-Lokus näher betrachtet. Die Funktion dieser lncRNAs sowie ihr Zusammenspiel mit *Swhtr* erfordern weitere Untersuchungen, da die zuvor beschriebenen Funktionen nur bedingt oder gar nicht bestätigt werden konnten. Allerdings hatte es den Anschein, dass *IRENE-div* eine unterrepräsentierte lange Isoform von *Swhtr* mit alternativem Transkriptionsstart und bislang unbekannter Funktion darstellt.

Bei der Untersuchung von *Fetal-lethal non-coding developmental regulatory RNA*/*Foxf1 adjacent non-coding developmental regulatory RNA* (*Fendrr*) lag der Fokus auf der Regulation von Genen im Zusammenhang mit Fibrose und der Rolle der RNA:dsDNA-Triplex-Bindungsdomäne von *Fendrr*, der sogenannten *FendrrBox*. Die Ergebnisse zeigten, dass *Fendrr* in Zusammenarbeit mit aktivierter WNT-Signalweg seine Ziel-Gene in Fibroblasten reguliert. Es wurde gezeigt, dass die *FendrrBox in vitro* Triplexe mit dem Promotor-DNA der Ziel Gene bildet und dass ihre Anwesenheit für die Regulation dieser Gene erforderlich ist.

Die Ergebnisse hoben die Bedeutung der WNT-Signalgebung bei der Regulation von Genen, die mit Fibrose in Verbindung stehen hervor, und wiesen darauf hin, dass *Fendrr* in dem verwendeten Zell-System, der NIH3T3 embryonalen Fibroblasten-Zelllinie, seine Funktion nur in Gegenwart eines aktivierten WNT-Signalweges entfaltet. Die WNT-Signalkaskade spielt eine entscheidende Rolle bei der Fibrose und ist an Prozessen wie Zellmigration, Ablagerung extrazellulärer Matrix und veränderten metabolischen Anforderungen beteiligt. Die Experimente zeigten, dass die Aktivierung von Fibrose assoziierten Genen durch *Fendrr* nur in Kombination mit aktiviertem WNT-Signalweg beobachtet wurde, was die synergistische und spezifische Rolle von *Fendrr* im Kontext der WNT-vermittelten Lungenentwicklung und fibrotischen Reizen betonte.

Die Konservierung von *Fendrr*/*FENDRR* beim Menschen wurde in verschiedenen Krankheitszusammenhängen beobachtet, einschließlich Krebs und idiopathischer Lungenfibrose (IPF). In Krebs wurde *FENDRR* mit Tumorinvasion, Aggressivität und schlechter Prognose in Verbindung gebracht. Bei IPF wurden erhöhte *FENDRR*-Spiegel in Lungengewebe, insbesondere in Endothelzellen und Fibroblasten, festgestellt. Signalwege, die an der Lungenentwicklung beteiligt sind, einschließlich der WNT-Signalgebung, spielen auch bei IPF eine Rolle. Darüber hinaus wurde in anderen Studien gezeigt, dass *FENDRR* in Fibroblasten β-Catenin-Spiegel reguliert und in einem Mausmodell für asbestinduzierte Fibrose fibrotische Läsionen reduzieren kann, was die Beobachtungen in dieser Arbeit untermauert und eine ähnliche Funktion von *FENDRR* im Menschen nahelegt. Die genauen Wirkungsmechanismen von *FENDRR* und seine subzelluläre Lokalisierung werden jedoch noch untersucht, und weitere Studien sind erforderlich, um seine Rolle in Krankheitszusammenhängen vollständig zu verstehen.

Vergleiche der Vor- und Nachteile der Plasmid-Überexpression und der dCas9-vermittelten Aktivierung (CRISPRa) Techniken für die Untersuchung von lncRNAs. Die Plasmid-Überexpression ermöglicht eine schnelle und zuverlässige Überexpression spezifischer Isoformen, spiegelt jedoch möglicherweise nicht genau das endogene Expressionsmuster und die Lokalisierung der lncRNA wider. Die CRISPRa-Technik ermöglicht dagegen die Überexpression der lncRNA von ihrem nativen Ort aus, wodurch ihr endogener Promotor und ihre Lokalisierung beibehalten werden, da das CRISPRa System als RNA-programmierbarer Transkriptionsaktivator angesehen werden kann. Die Funktion von RNA-basierten lncRNAs hängt stark von ihrer korrekten subzellulären Lokalisierung und ihrer Sekundärstruktur ab, aber auch Unterschiede in der Isoform-expression und strukturellen Integrität können die Funktion der lncRNA beeinflussen. Die teilweise erforderliche Rolle der *FendrrBox* für die Funktion von *Fendrr* und legte die Existenz von *Fendrr*-Isoformen mit *FendrrBox*-unabhängigen Funktionen nahe, wohingegen der WNT-abhängige *Fendrr* Effekt nur nach CRISPRa gezeigt werden konnte. Da dieser Effekt durch Entfernen der *FendrrBox* aufgehoben werden konnte und diese die DNA der Promotoren der Ziel-Gene binden kann, ist hier von einer RNA-abhängigen Wirkweise auszugehen. Weitere Untersuchungen sind erforderlich, um die vielfältigen Funktionen und Mechanismen von *Fendrr* und seinen Isoformen aufzudecken.

Abschließend, wurde eine unbeschriebene, neue lncRNA initial charakterisiert. Das Projekt der lncRNA *2900076A07Rik*, die hier als *lncFsd2* bezeichnet wird, konzentriere sich auf die Charakterisierung der Beziehung zwischen dem Protein-kodierenden Gen *Fsd2* dem *lncFsd2* benachbart ist und divergent exprimiert wird. Die Ergebnisse zeigten, dass beide Gene bei plazentaren Tieren hoch konserviert sind, wobei *lncFsd2* eine Sequenzähnlichkeit aufweist, die für eine nicht-kodierende RNA im Vergleich zu den meisten lncRNAs, als hoch anzusehen ist. *LncFsd2* und sein menschliches Ortholog weisen eine intronische miRNA auf, die bei initialer bioinformatischer Kalkulation keine Interaktion mit *Fsd2* zeigten. Konvergente lncRNAs zeigen oft eine negative Korrelation in der Expression mit ihren kodierenden Nachbarn, was auf eine potenzielle Interaktion

und Regulation hinweist, aufgrund dessen die Beziehung der divergenten Gene in den ersten Fokus rückte.

FSD2 weist im Herzen räumliche Nähe zu Ryanodin Rezeptor 2 (RYR2) und Cardiomyopathy Associated 5 (CMYA5), was auf seine Beteiligung an dem für die Organisation von Ryanodin-Rezeptor-Clustern verantwortlichen Myospryn-Komplex hinweist. Die Expressionsmuster von *Fsd2* und *lncFsd2* wurden untersucht, wobei *Fsd2* eine spezifische Anreicherung in Herz- und Skelettmuskulaturgeweben, insbesondere in Kardiomyozyten, zeigte. *LncFsd2* hingegen zeigte eine weit verbreitete Expression in verschiedenen Geweben, einschließlich des Herzens. Fraktionierung des Herzens in Kardiomyozyten und Nicht-Kardiomyozyten zeigte, dass *Fsd2* und *LncFsd2* im selben Zelltyp exprimiert werden können.

Die Interaktion zwischen *lncFsd2* und *Fsd2* wurde durch verschiedene experimentelle Ansätze untersucht, darunter dCas9-vermittelte Aktivierung oder Inaktivierung und Anti-sense Oligonukleotid (ASO)-vermittelte *lncFsd2*-Unterdrückung. Ferner wurden auch Herz- und Skelettmuskulaturgewebe von *lncFsd2* Mutanten Mäusen untersucht. Es wurde jedoch keine signifikante transkriptionelle Regulation von *Fsd2* durch *lncFsd2* beobachtet, wohingegen die FSD2 Proteinmenge im Herzen leicht erhöht ist. Da FSD2 im Zytoplasma sowie im Nukleus gefunden wurde, wurde die Proteinlokalisation in Abhängigkeit von *lncFsd2* betrachtet, unter der Annahme, dass die Lokalisation die Stabilität des Proteins und dahingehend die Menge leicht beeinflussen könnte. Hierbei wurde kein messbarer Unterschied deutlich. Weitere Ergebnisse zeigten darüber hinaus, dass etwa die Hälfte der *Fsd2*-mRNA im Zellkern verbleibt, was auf einen potenziellen Mechanismus hindeutet, bei dem *lncFsd2* an der Lokalisierung von *Fsd2* beteiligt sein könnte, was sich ebenfalls in subtilen Unterschieden in der Proteinmenge widerspiegeln könnte. Dies wird durch weitere Experimente unterstützt, die zeigten, dass lncFsd2 sich vorwiegend im Nukleus aufhält und an die *Fsd2* mRNA zu binden scheint, während keine physische Interaktion zwischen der lncRNA und dem Protein festgestellt werden konnte. Zukünftige Experimente wurden vorgeschlagen, um den genauen Mechanismus der Interaktion zwischen *lncFsd2* und *Fsd2* zu klären, einschließlich der subzellulären Fraktionierung der aus *lncFsd2* Mutanten Mäusen gewonnenen Kardiomyozyten. Darüber hinaus könnten die Rolle der miRNAs innerhalb von *lncFsd2* und die

Untersuchung von FSD2-unabhängigen Protein-Bindungspartner durch Massenspektrometrie weitere Erkenntnisse über die Funktionen von *IncFsd2* liefern.

Zusammenfassend tragen diese Studien dazu bei, unser Verständnis für die Rolle von lncRNAs in verschiedenen biologischen Prozessen zu erweitern. Sie zeigen die Bedeutung von Swhtr bei der Stressantwort des Herzens, die Rolle von Fendrr bei der Regulation von Genen im Zusammenhang mit Fibrose und die Beziehung zwischen Fsd2 und lncFsd2. LncRNA Biologie ist nicht nur wichtig um genetische Regulation besser zu verstehen, sondern auch Prozesse, die der Entwicklung und Homöstase von Organen, sowie deren Reaktion auf Verletzung und Stress zugrunde liegen. Weitere Untersuchungen der hier betrachteten lncRNAs sind jedoch erforderlich, um die genauen Mechanismen und Funktionen dieser lncRNAs vollständig zu verstehen und ihr Potenzial für therapeutische Interventionen zu nutzen, dennoch werden in dieser Arbeit die vielfältigen Funktionen von lncRNAs beleuchtet, sowie deren Implikationen in verschiedenste biologische Prozesse deutlich gemacht

Materials and Methods

Materials

General Laboratory equipment, chemicals and reagents

Table 1: Consumables

Product	Product Number	Manufacturer
Adhesion slides SuperFrost Plus™	631-9483	VWR (Darmstadt, Germany)
Blot Paper	1620118	Bio-Rad (Hercules, CA, USA)
Cell Lifter Corning®	3008	Corning (Corning, NY, USA)
Cell strainer Corning®	CLS431751	Corning (Corning, NY, USA)
Centrifuge Tubes Polypropylene CELLSTAR® (15 ml, 50 ml)	188261; 227261	Greiner (Kremsmünster, Austria)
Combitips® advanced (0.1 ml, 0.2 ml, 0.5 ml, 1 ml, 2.5 ml, 5 ml, 10 ml)	0030089766, 0030089774, 0030089782, 0030089790, 0030089804, 0030089812, 0030089820	Eppendorf (Hamburg, Germany)
Falcon™ Round-Bottom Polystyrene Test Tubes	352003	Thermo Fisher Scientific (Waltham, MA, USA)
Filter Tips TipOne®	S1121	Starlab (Hamburg, Germany)
gentleMACS™ C Tubes	130-096-334	Miltenyi Biomedicine GmbH (Bergisch Gladbach, Germany)
Homogenizer Tube CK14 (Soft Tissue, 2 ml)	P000912-LYSK0-A	Bertin Technologies (Montigny-le-Bretonneux, France)
Optical Adhesive Film Applied Biosystems™ MicroAmp™	4311971	Thermo Fisher Scientific (Waltham, MA, USA)
Pasteur pipettes	HXH7.1	Carl Roth (Karlsruhe, Germany)
PCR Tubes	I1402-3700	Starlab (Hamburg, Germany)
qPCR Plate 384-well	785290	Greiner (Kremsmünster, Austria)
qPCR Plate 96-well Applied Biosystems™ MicroAmp™	N8010560	Thermo Fisher Scientific (Waltham, MA, USA)
Reaction Tubes Protein LoBind®	0030108116	Eppendorf (Hamburg, Germany)
Reaction Tubes Safe-Lock (1.5 ml, 2.0 ml)	0030123328; 0030123344	Eppendorf (Hamburg, Germany)
Serological Pipettes Nunc™ (5 ml, 10 ml, 25 ml, 50 ml)	170366N; 170356N; 170357N; 170358N	Thermo Fisher Scientific (Waltham, MA, USA)
Tissue Culture Dishes (22.1 cm^2, 60.1 cm^2)	93060; 93100	TPP Techno Plastic Products (Trasadingen, Switzerland)
Tissue Culture Flask, 250 ml, 75 cm^2	658175	Greiner (Kremsmünster, Austria)
Tissue Culture Test Plates (6-well, 12-well, 24-well, 48-well, 96-well, 96-well U-bottom)	92006; 92012; 92024; 92048; 92096; 92097	TPP Techno Plastic Products (Trasadingen, Switzerland)
Transfer-membrane Immobilon®-P PVDF	T831.1	Carl Roth (Karlsruhe, Germany)

Table 2: Equipment

Instrument	Model	Manufacturer
Adjustable volume multi-channel pipette	ErgoOne®	Starlab (Hamburg, Germany)
Adjustable volume pipettes	Eppendorf Research® plus	Eppendorf (Hamburg, Germany)
Automated cell counter	Countess™ 3	Thermo Fisher Scientific (Waltham, MA, USA)
Cell Freezing Module	CoolCell® LX	Corning (Corning, NY, USA)
Centrifuge (falcons)	Megafuge 1.0R	Heraeus Instruments (Hanau, Germany)
Centrifuge (falcons)	Universal 320 R	Hettich (Tuttlingen, Germany)
Centrifuge (Mini)	3-1810	neoLab (Heidelberg, Germany)
Centrifuge (tabletop)	5430 R	Eppendorf (Hamburg, Germany)
CO_2 incubator	NU-5510	NuAire (Plymouth, MN, USA)
Confocal Microscope	CQ1	Yokogawa (Musashino, Japan)
Confocal Microscope	SP8	Leica Microsystems (Wetzlar, Germany)
Cryostat	CM3050 S	Leica Biosystems (Nussloch, Germany)
Dissociator	gentleMACS™	Miltenyi Biomedicine GmbH (Bergisch Gladbach, Germany)
Electroporation Transfection System	Neon™ MPK5000	Thermo Fisher Scientific (Waltham, MA, USA)
Flow Cytometer	BD FACSAria™ Fusion Flow	BD Biosciences (Franklin Lakes, NJ, USA)
Homogenizer	Minilys® (3D-Homogenizer)	Bertin Technologies (Montigny-le-Bretonneux, France)
Homogenizer (Dounce)	KIMBLE Dounce tissue grinder set	Thermo Fisher Scientific (Waltham, MA, USA)
Imaging system	C300	Azure Biosystems (Dublin, CA, USA)
Incubation chamber	Certomat® H	B. Braun Biotech (Melsungen, Germany)
Incubator (bacteria)	Heratherm™ Compact	Thermo Fisher Scientific (Waltham, MA, USA)
Individually ventilated cage	SealSafe PLUS AERO	Tecniplast (Hohenpeißenberg, Germany)
Inhalation anesthesia device	Trajan 808	Dräger Medical Deutschland GmbH (Lübeck, Germany)
Inverted Microscope (live cells)	EVOS® FL	Thermo Fisher Scientific (Waltham, MA, USA)
Inverted Microscope (live cells)	Primovert	Zeiss (Jena, Germany)
Laminar flow	CleanAir CA/Rev6	Rieger Industrievertretungen (Vienna, Austria)
Magnetic rack	DynaMag-2	Thermo Fisher (Waltham, MA, USA)
Microtome	Microm HM 430	Microm International (Walldorf, Germany)
Multipette	Stream	Eppendorf (Hamburg, Germany)
PCR system	Mastercycler X50i	Eppendorf (Hamburg, Germany)
pH meter	Φ 340	Beckman (Fullerton, CA, USA)
Physoxic and Hypoxic Atmosphere Workchamber and Hood	Baker Ruskinn InvivO$_2$® 400 Physoxia Workstation	Ruskinn Technology Ltd (Pencoed, UK)

Plate centrifuge	Centry 101	Gilson (Middleton, WI, USA)
Power supply	VWR® Power Source	VWR (Darmstadt, Germany)
Precision scales	AEJ-CM	Kern (Balingen, Germany)
Protein Electrophoresis System	Invitrogen™ XCell SureLock™ Mini-Cell	Thermo Fisher Scientific (Waltham, MA, USA)
Real time PCR system	StepOnePlus	Applied Biosystems (Foster City, CA, USA)
Real time PCR system	Viia 7	Applied Biosystems (Foster City, CA, USA)
Shaking plate	Certomat® R	Sartorius (Göttingen, Germany)
Spectrophotometer	NanoPhotometer® P330	Implen (Munich, Germany)
Spectropolarimeter	J-810	JASCO (Tokyo, Japan)
Thermocycler	Mastercycler® pro S	Eppendorf (Hamburg, Germany)
Tissue Embedding Center	TES 99	MEDITE Medical GmbH (Burgdorf, Germany)
Tissue Processing Center	TPC 15	MEDITE Medical GmbH (Burgdorf, Germany)
Ultrasound machine	Vevo 2100	VisualSonics (Toronto, Canada)
UV Trans illuminator	T2201	Sigma-Aldrich (St. Louis, MO, USA)
UV-crosslinker	Stratalinker 1800	Stratagene (La Jolla, CA, USA)
Vortexer	Vortex Genie 2	Scientific Industries (Bohemia, NY, USA)
Water bath	1008	GFL (Burgwedel, Germany)
Wet Electroblotting System	Mini Trans-Blot® Cell	Bio-Rad (Hercules, CA, USA)

Table 3: Cellculture media and supplements

Product	Product Number	Manufacturer
Calf Serum	C8056	Sigma-Aldrich (St. Louis, MO, USA)
CGP77675 (MEK-inhibitor)	SML0314	Sigma-Aldrich (St. Louis, MO, USA)
CHIR99021 (GSK3 inhibitor)	SML1046	Sigma-Aldrich (St. Louis, MO, USA)
Claycomb-Medium	51800C	Sigma-Aldrich (St. Louis, MO, USA)
EmbryoMax® Nucleosides (100X)	ES-008-D	Thermo Fisher Scientific (Waltham, MA, USA)
ESGRO® Recombinant Mouse LIF Protein	ESG1107	Sigma-Aldrich (St. Louis, MO, USA)
FBS Premium, South America origin, fetal bovine serum, 0.2 µm sterile filtered	P30-3301	PAN-Biotech (Aidenbach, Germany)
Gibco™ 2-Mercaptoethanol	21985023	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ B-27™ Supplement (50x)	17504044	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ DMEM/F-12	11320033	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ DPBS (1x)	14190144	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ KnockOut™ DMEM	10829018	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ N-2 Supplement (100X)	17502048	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ Neurobasal™-A Medium	10888022	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ Opti-MEM™ I Reduced Serum Medium, no phenol red	11520386	Thermo Fisher Scientific (Waltham, MA, USA)

Gibco™ Penicillin-Streptomycin (5.000 U/ml)	15070063	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ Sodium Pyruvate 100mM	11360070	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ DMEM, high glucose, GlutaMAX™ Supplement, Pyruvate	10569010	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ DMEM, high glucose, no glutamine	11500416	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ GlutaMAX™ Supplement	35050061	Thermo Fisher Scientific (Waltham, MA, USA)
Gibco™ MEM Non-Essential Amino Acids Solution (100X)	11140050	Thermo Fisher Scientific (Waltham, MA, USA)
L-Ascorbic acid	A92902	Sigma-Aldrich (St. Louis, MO, USA)
Pansera ES, South America origin, special designed bovine serum for embryonal stem cells, 0.2 μm sterile filtered	P30-2600	PAN-Biotech (Aidenbach, Germany)

Table 4: Chemicals and Reagents

Product	Product Number	Manufacturer
10X Orange Loading Dye	927-10100	LI-COR Biosciences (Lincoln, NE, USA)
2-Propanol	9781.1	Carl Roth (Karlsruhe, Germany)
Acrylamide/Bis-acrylamide, 30% solution	A3699	Merck (Darmstadt, Germany)
Ammonium Persulfate	17874	Thermo Fisher Scientific (Waltham, MA, USA)
ArciTect™ tracrRNA Kit	76018	STEMCELL Technologies (Vancouver, Canada)
autoMACS® Rinsing Solution	130-091-222	Miltenyi Biomedicine GmbH (Bergisch Gladbach, Germany)
Calcium Chloride	C1016	Sigma-Aldrich (St. Louis, MO, USA)
Carbenicillin disodium salt	6344.1	Carl Roth (Karlsruhe, Germany)
Chloroform	C2432	Sigma-Aldrich (St. Louis, MO, USA)
Citrate Buffer (pH 6.0), Concentrate	005000	Thermo Fisher Scientific (Waltham, MA, USA)
cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail	8820027712	Sigma-Aldrich (St. Louis, MO, USA)
DAPI	6335.1	Carl Roth (Karlsruhe, Germany)
Deionized formamide	F9037	Sigma-Aldrich (St. Louis, MO, USA)
DL-Dithiothreitol solution	646563	Sigma-Aldrich (St. Louis, MO, USA)
Eosin Y solution	X883.1	Carl Roth (Karlsruhe, Germany)
Ethanol	32221	Merck (Darmstadt, Germany)
Fibronectin, Bovine Plasma	341631	Merck (Darmstadt, Germany)
Fluorescence Mounting Medium	S3023	Agilent Technologies (Santa Clara, CA, USA)
Gelatin solution (2% in H_2O)	G1393	Sigma-Aldrich (St. Louis, MO, USA)
Gibco™ StemPro™ Accutase™	A1110501	Thermo Fisher Scientific (Waltham, MA, USA)
Glycerol	G5516	Sigma-Aldrich (St. Louis, MO, USA)
Glycine	G7126	Sigma-Aldrich (St. Louis, MO, USA)
Glycogen, Molecular Biology Grade	9005-79-2	Roche (Basel, Switzerland)
hBFGF	F0291	Sigma-Aldrich (St. Louis, MO, USA)
Hemalum solution acid acc. to Mayer	T865.1	Carl Roth (Karlsruhe, Germany)

HEPES solution, 1M	SRE0065	Sigma-Aldrich (St. Louis, MO, USA)
ImmEdge® Hydrophobic Barrier PAP Pen	H-4000	Vector Laboratories (Newark, CA, USA)
Invitrogen™ 1 Kb Plus DNA Ladder	10787026	Thermo Fisher Scientific (Waltham, MA, USA)
Invitrogen™ Dynabeads™ MyOne™ Streptavidin C1	65002	Thermo Fisher Scientific (Waltham, MA, USA)
Invitrogen™ Dynabeads™ MyOne™ Streptavidin C1	65001	Thermo Fisher Scientific (Waltham, MA, USA)
Invitrogen™ Lipofectamine™ 3000 Transfection Reagent	L3000008	Thermo Fisher Scientific (Waltham, MA, USA)
Invitrogen™ Lipofectamine™ RNAiMAX Transfection Reagent	13778150	Thermo Fisher Scientific (Waltham, MA, USA)
Invitrogen™ UltraPure™ 0,5 M EDTA, pH 8,0	15575020	Thermo Fisher Scientific (Waltham, MA, USA)
Invitrogen™ ytRNA	AM7119	Thermo Fisher Scientific (Waltham, MA, USA)
Lithium acetate	517992	Merck (Darmstadt, Germany)
Lithium chloride	L9650	Sigma-Aldrich (St. Louis, MO, USA)
MACS® BSA Stock Solution	130-091-376	Miltenyi Biomedicine GmbH (Bergisch Gladbach, Germany)
Magnesium Chloride	208337	Sigma-Aldrich (St. Louis, MO, USA)
Manganese(II) Chloride	244589	Sigma-Aldrich (St. Louis, MO, USA)
Methanol	322415	Merck (Darmstadt, Germany)
Midori Green Advance DNA Stain	52580	Nippongenetics (Tokyo, Japan)
Mitomycin C from *Streptomyces caespitosus*	M4287	Sigma-Aldrich (St. Louis, MO, USA)
NP-40 Surfact-Amps™ Detergent Solution	90982656	Thermo Fisher Scientific (Waltham, MA, USA)
NuPAGE™ 4 to 12%, Bis-Tris, 1.0–1.5 mm, Mini Protein Gels	NP0321	Thermo Fisher Scientific (Waltham, MA, USA)
NuPAGE™ LDS Sample Buffer (4X)	NP0007	Thermo Fisher Scientific (Waltham, MA, USA)
NuPAGE™ MOPS SDS Running Buffer (20X)	NP0001	Thermo Fisher Scientific (Waltham, MA, USA)
NuPAGE™ Sample Reducing Agent (10X)	NP0004	Thermo Fisher Scientific (Waltham, MA, USA)
NuPAGE™ Transfer Buffer (20X)	NP00061	Thermo Fisher Scientific (Waltham, MA, USA)
PageRuler™ Plus Prestained Protein Ladder, 10 bis 250 kDa	26619	Thermo Fisher Scientific (Waltham, MA, USA)
Pierce™ 16% Formaldehyde (w/v), Methanol-free	28906	Thermo Fisher Scientific (Waltham, MA, USA)
Pierce™ ECL Western Blotting Substrate	32109	Thermo Fisher Scientific (Waltham, MA, USA)
Pierce™ Protein A/G Magnetic Beads	91131725	Thermo Fisher Scientific (Waltham, MA, USA)
Potassium chloride	P9541	Sigma-Aldrich (St. Louis, MO, USA)
ProLong™ Gold Antifade Mountant	P36930	Life Technologies (Carlsbad, CA, USA)
Recombinant Human FGF-10 Protein	345-FG	R&D Systems (Hinnerup, Denmark)
Recombinant Mouse BMP-4 Protein	5020-BP-010	R&D Systems (Hinnerup, Denmark)
Red Blood Cell Lysis Solution (10x)	130-094-183	Miltenyi Biomedicine GmbH (Bergisch Gladbach, Germany)

RNasin® RNase Inhibition (10,000 units)	N2615	Promega (Madison, WI, USA)
ROTI®Histofix	P087.1	Carl Roth (Karlsruhe, Germany)
ROTI®Histol	6640.1	Carl Roth (Karlsruhe, Germany)
ROTI®Load 1, reducing	K929.1	Carl Roth (Karlsruhe, Germany)
Rotiphorese® 50x TAE buffer	CL86.1	Carl Roth (Karlsruhe, Germany)
Sakura Finetek™ Tissue-Tek™ O.C.T. Compound	12351753	Thermo Fisher Scientific (Waltham, MA, USA)
SDS Solution, 20% Sodium Dodecyl Sulfate Solution	10607633	Thermo Fisher Scientific (Waltham, MA, USA)
Skim Milk Powder	70166	Merck (Darmstadt, Germany)
Sodium chloride	S9888	Sigma-Aldrich (St. Louis, MO, USA)
Sodium deoxycholate	S1827	Sigma-Aldrich (St. Louis, MO, USA)
Stellaris® RNA FISH Hybridization Buffer	SMF-HB1-10	BioCat, Heidelberg, Germany
Stellaris® RNA FISH Wash Buffer A	SMF-WA1-60	BioCat, Heidelberg, Germany
Stellaris® RNA FISH Wash Buffer B	SMF-WB1-20	BioCat, Heidelberg, Germany
TEMED	2367.1	Carl Roth (Karlsruhe, Germany)
TRI Reagent®	T9424	Merck (Darmstadt, Germany)
TRIS base	4855.2	Carl Roth (Karlsruhe, Germany)
Triton™ X-100 solution	93443	Sigma-Aldrich (St. Louis, MO, USA)
Trypan Blue Stain (0.4%) for use with the Countess™ Automated Cell Counter	T10282	Thermo Fisher Scientific (Waltham, MA, USA)
Trypsin-EDTA (0,05 %), phenol red	25300054	Thermo Fisher Scientific (Waltham, MA, USA)
TWEEN® 20	P1379	Merck (Darmstadt, Germany)
Universal-Agarose, peqGOLD	732-2789	VWR (Darmstadt, Germany)
Urea	U5378	Sigma-Aldrich (St. Louis, MO, USA)

Table 5: Bacteria

Strain	Manufacturer
JM109 Competent *E. coli*	Promega (Madison, WI, USA)

Table 6: Bacterial media and supplements

Product	Product Number	Manufacturer
Agar-Agar	5210.1	Carl Roth (Karlsruhe, Germany)
Carbenicillin disodium salt	6344.1	Carl Roth (Karlsruhe, Germany)
Kanamycin sulfate from *Streptomyces kanamyceticus*	K4000	Sigma-Aldrich (St. Louis, MO, USA)
LB Broth (Luria/Miller)	X968.1	Carl Roth (Karlsruhe, Germany)
SOC Outgrowth Medium	B9020S	NEB (Ipswich, MA, USA)

Table 7: Plasmids

Plasmid Name	Plasmid Number	Supplier
B52 (empty plasmid backbone to express 2 sgRNAs)	100708	Addgene (Watertown, MA, USA)
CAG-Cas9-T2A-EGFP-ires-puro	78311	Addgene (Watertown, MA, USA)
dCas9-KRAB-MeCP2	110821	Addgene (Watertown, MA, USA)
ITPKA_3xmScarletI	112960	Addgene (Watertown, MA, USA)
pEGFP-N3	V012019	NovoPro Biosciences (Shanghai, China)

phbAct-Fendrr	-	Gift from Dr. Phillip Grote (Transgenic Core Facility; Georg-Speyer-Haus)
pRP[Exp]-Puro-CAG-dCAS9-VP64_T2A_MS2-p65-HSF1	-	Gift from Mohamed Nemir (Experimental Cardiology Unit Department of Medicine University; Lausanne Medical School)
sgRNA(MS2) cloning backbone	61424	Addgene (Watertown, MA, USA)

Table 8: Cell Lines and primary cells

Name	Species	Tissue	Supplier
C2C12	*Mus musculus*	Muscle	ATCC (Manassas, VA, USA)
HL-1	*Mus musculus*	Heart	Prof. W. Claycomb (New Orleans, LA, USA)
iPSC	*Homo sapiens*	PSC	EBiSC
mESC (F1G4)	*Mus musculus*	Blastocyst	(George et al., 2007)
MLg [Mlg 2908]	*Mus musculus*	Lung	ATCC (Manassas, VA, USA)
Neuro-2A	*Mus musculus*	Brain	ATCC (Manassas, VA, USA)
NIH3T3	*Mus musculus*	Embryo	ATCC (Manassas, VA, USA)
NMCMs	*Mus musculus*	Heart	Self-isolated

Table 9: Primary Antibodies

Antibody	Species	Product Number	Manufacturer	Application (Concentration)
Cardiac Troponin T Antibody, anti-human/mouse/rat, APC, REAfinity™	Human recombinant	130-120-403	Miltenyi Biomedicine GmbH (Bergisch Gladbach, Germany)	FACS sorting (1:10)
FSD2	Rabbit	25609-1-AP	Proteintech (Rosemont, IL, USA)	Western Blot (1:1000); RNA Immunoprecipitation (5 µg), Immunofluorescence (1:50)
NKX2-5	Goat	AF2444	R&D Systems (Hinnerup, Denmark)	RNA Immunoprecipitation (5 µg)
NKX2-5	Rabbit	5444	Cell Signaling Technology (Danvers, MA, USA)	Western Blot (1:1000)
Normal Goat IgG	Goat	AB-108-C	R&D Systems (Hinnerup, Denmark)	RNA Immunoprecipitation (5 µg)
Normal Rabbit IgG	Rabbit	2729	Cell Signaling Technology (Danvers, MA, USA)	RNA Immunoprecipitation (5 µg)
RYR2	Rabbit	PA5-87416	Thermo Fisher Scientific (Waltham, MA, USA)	Immunofluorescence (1:200)
Wheat Germ Agglutinin, Alexa Fluor™ 488 Conjugate	Wheat	W11261	Thermo Fisher Scientific (Waltham, MA, USA)	Immunofluorescence (1:100)

Table 10: Secondary Antibodies

Antibody	Species	Product Number	Manufacturer	Application
Anti-rabbit IgG, HRP-linked	Goat	7074	Cell Signaling Technology (Danvers, MA, USA)	Western Blot
Anti-Rabbit IgG (H+L), Alexa Fluor™ 488	Goat	A11008	Thermo Fisher Scientific (Waltham, MA, USA)	Immunofluorescence

Table 11: Enzymes

Enzyme	Product Number	Manufacturer	Application
Acc65I	R0599	NEB (Ipswich, MA, USA)	Restriction digest
BbsI	R0539	NEB (Ipswich, MA, USA)	Restriction digest
DraI	R0129	NEB (Ipswich, MA, USA)	Restriction digest
EcoRI-HF®	R3101	NEB (Ipswich, MA, USA)	Restriction digest
EcoRV	R0195	NEB (Ipswich, MA, USA)	Restriction digest
Esp3I	R0734	NEB (Ipswich, MA, USA)	Restriction digest
Expand™ Long Template PCR-System	11681834001	Roche (Basel, Switzerland)	RACE-PCR
Invitrogen™ RNAse H	18021071	Thermo Fisher Scientific (Waltham, MA, USA)	RNAse H protection assay
NdeI	R0111	NEB (Ipswich, MA, USA)	Restriction digest
Phusion® High-Fidelity DNA Polymerase	M0530	NEB (Ipswich, MA, USA)	PCR
Proteinase K	P8107	NEB (Ipswich, MA, USA)	Protein degradation (RIP)
Shrimp Alkaline Phosphatase (rSAP)	M0371	NEB (Ipswich, MA, USA)	Dephosphorylation
T4 DNA ligase	M0202	NEB (Ipswich, MA, USA)	Ligation

Table 12: Kits

Product	Product Number	Manufacturer
CD31 MicroBeads, mouse	130-097-418	Miltenyi Biomedicine GmbH (Bergisch Gladbach, Germany)
CD45 MicroBeads, mouse	130-109-682	Miltenyi Biomedicine GmbH (Bergisch Gladbach, Germany)
Invitrogen™ Neon™ Transfection System 10 µL Kit	MPK1096	Thermo Fisher Scientific (Waltham, MA, USA)
Neonatal Heart Dissociation Kit, mouse and rat	130-098-373	Miltenyi Biomedicine GmbH (Bergisch Gladbach, Germany)
QIAfilter Plasmid Maxi Kit	142919264	QIAGEN (Hilden, Germany)
QIAprep Spin Miniprep Kit	27104	QIAGEN (Hilden, Germany)
QIAquick Gel Extraction Kit	28706X4	QIAGEN (Hilden, Germany)
QIAquick PCR Purification Kit	28104	QIAGEN (Hilden, Germany)
QuantiTect Reverse Transcription Kit	205314	QIAGEN (Hilden, Germany)
REDExtract-N-Amp™ Tissue PCR Kit	XNAT-100RXN	Thermo Fisher Scientific (Waltham, MA, USA)
SMARTer® RACE 5'/3' Kit	634858	Takara (Kusatsu, Japan)

Oligonucleotides, gRNAs and Probes

Table 13: Murine Oligonucleotides for qRT-PCR

Gene	Forward	Reverse
Ager	TGGTCAGAACATCACAGCCC	CATTGGGGAGGATTCGAGCC
Akr1c14	TGGTCACTTCATCCCTGCAC	GCCTGGCCTACTTCCTCTTC
Axin2	GGCTCACCCATTTACCCAGG	GTTCCTGTCCCTCTGCTGAC
Cdh5	TTGGACCGAGAGAAACAGGC	CTTGCCCACTCGGATGTCTT
Clic5	ACTAATGGGGACGACAGGGA	ATTGAAGGTCAGGAAGGGCG
Ddr2	AGCGAGGTACAGGACTCCAT	TCCCCGCTCCTCCTCAATTA
Eaf1	AGAGCTGTGGTGTTGGCATT	AGAGAGCAAAACCTGGAGCC
Ef1a	TTCAAGTACGCCTGGGTCTT	ACACCAGCAGCAACAATCAG
Egr1	CAGCGCCTTCAATCCTCAAG	TAACTCGTCTCCACCATCGC
Emp2	GCTTCTCTGCTGACCTCTGG	CGAACCTCTCTCCCTGCTTG
Fendrr	CTGCCCGTGTGGTTATAATG	TGACTCTCAAGTGGGTGCTG
FendrrBox	TAGCCGTGAATGAGTCTGCC	GGGAGGGAGAAGAGAGGAG
Fn1	GAGTAGACCCCAGGCACCTA	GTGTGCTCTCCTGGTTCTCC
Foxf1	CAAAACAGTCACAACGGGCC	GCCTCACCTCACATCACACA
Fsd2	CTGTGAAGAGGCTGTGCTGA	GTTGAGGGCTCGCACATAGA
Hc	CTGCTTGAAAACACCCTGCC	AGCTGTCTGGACGTTTGAGG
Hprtf	GCTTTCCCTGGTTAAGCAGTACA	GAGAGGTCCTTTTCACCAGCAA
Id1	GAACCGCAAAGTGAGCAAGG	GATCGTCGGCTGGAACACAT
Id2	CACCCTGAACACGGACATCA	TCAACGTGTTCTCCTGGTGA
Id3	TCCTGCAGCGTGTCATAGAC	GGTCAGTGGCAAAAGCTCCT
Irene-DIV	TGGGAAAGGAGTGTGTGTCTG	GGAGGACCTTGTTACTGGCA
Irene-SS	TCTCTTTGGTCACTTGGGGC	GCCGCCGTTTATGCTGTAG
Ldha	TGTCTCCAGCAAAGACTACTGT	GACTGTACTTGACAATGTTGGGA
lncFsd2	GCCACACCCATGCCTTAGAT	CCTGGGTCCTGTCAGTTCAC
Malat1	GCAGGTGGGAGATGATGGTC	TAAACCCACATGCCTGACCC
Myl2	ATCGACAAGAATGACCTAAGGGA	ATTTTTCACGTTCACTCGTCCT
Myl7	GGCACAACGTGGCTCTTCTAA	TGCAGATGATCCCATCCCTGT
Nkx2-5	AAGTGCTCTCCTGCTTTCCCA	TTTGTCCAGCTCCACTGCCTT
Ptprc	GGCTTCAAGGAACCCAGGAA	GAGTGCCTTCCTCCATGCTT
Rpl10	GCTCCACCCTTTCCATGTCA	TGCAACTTGGTTCGGATGGA
Ryr1	CGGGATGACAACAAGAGGCT	ACGTCTTAGGCAGCTTCGTC
Ryr2	ACTGGCCGGGGATATTTCAG	CAGGCCACAACAGCTCTTTT
Serpinb6b	ATAAGCGTCTCCTCAGCCCT	CTTTTCCCCGAAGAGCCTGT
Sftpc	CCTTTTCCATCGGCTCCACT	GGTCTCTCCCGGAAGAATCG
Slc2a1	CAGTTCGGCTATAACACTGGTG	GCCCCCGACAGAGAAGATG
Swhtr	CCGCAAACCGAACAACTCAG	CTGCCTAAGTGGGGGAAGTG
Tnni1	CCATGGATCTGCGGGCCAACC	GCGGCCTTCCATGCCAGACA
Trim16	CCACACCAGGAGAACAGCAA	AGGTCCAACTGCATACACCG
Vnn1	ATGGGAGACAAGAAGCCGTG	TGACAGCGGGATCATGGAAG

Table 14: Human Oligonucleotides for qRT-PCR

Gene	Forward	Reverse
FSD2	AGCTGCTGATGTCTTCTGTGT	TGATGAAACCTCCTCCAAATGA
GAPDH	CCCACTCCTCCACCTTTGAC	GGAGATTCAGTGTGGTGGGG
IRENE-DIV	TCGCTTTCAGACTTCCGGTG	CCAGCGTCTCGATTCTCTCC
IRENE-SS	TCAGCTGGCACCAAATCCTT	CAGTCTCAGCTGTGCAGTCA
SNHG21	GAGTAGAGGTGAGGGGTCGT	AGCAAATTGAAGGCAGGGTT
SWHTR	CTCTCAAACGCGCTGAAACC	AAAGTCCAAGCTGCCCTCTC

Table 15: Oligonucleotides for Genotyping

Genotype	Forward	Reverse	Internal
FendrrBox	ATGCTTCCAAGGAAG GACGG	CTTGACGCCAAGCTCC TGTA	
lncFsd2⁻/⁻ (Promoter deletion)	GATCTGAGCGCTTTCT TATG	TCCACTGTCCTACATG TTTC	CAGTATTTGGCCACCTTG TG
lncFsd^{stop/stop} (Stop-signal insertion)	AACTTCCTCCCTACGG CTAA	AAAATCACCCACTGCG TGTG	
Swhtr^{3xpA/3xpA}	CCCGCATGAAGATTCT GGGA	GCCTGAAATGAGCCTT GGGA	CAGGTTGGAAGTGGAGC AG
Swhtr^{3xpA/3xpA;tg}	TGACCGAGTACAAGC CCAC	GCGCAGACAGGTCCC CAGAC	

Table 16: RNA pulldown probes (Biotinylated DNA oligos)

Probe	Type/Manufacturer	Sequence
lncFsd2_pulldown_AS2	Integrated DNA Technologies IDT (Coralville, IO, USA)	mAmCmCmUmCmUmGmAmUmA mAmAmUmCmUmUmAmCmUmG mUmC/iSp9//3deSBioTEG/
lncFsd2_pulldown_AS3	Integrated DNA Technologies IDT (Coralville, IO, USA)	mAmCmUmUmUmGmUmUmUmA mAmGmCmAmUmGmUmUmAmU mCmU/iSp9//3deSBioTEG/
Swhtr_Pulldown_Probe1	Integrated DNA Technologies IDT (Coralville, IO, USA)	mGmAmAmGmCmAmAmAmGmG mAmAmAmAmCmGmGmCmAmG mGmA/iSp9//3deSBioTEG/
Swhtr_Pulldown_Probe2	Integrated DNA Technologies IDT (Coralville, IO, USA)	mUmUmUmUmGmCmUmAmUm UmGmCmCmAmAmGmAmAmUm CmAmG/iSp9//3deSBioTEG/

Table 17: DNA oligonucleotides for CD-spectra and melting curve analysis

Probe	Sequence (GA-rich)	Sequence (CT-rich)
Ager	GGGAGTGAAGAGGGTGGTGGG	CCCACCACCCTCTTCACTCCC
Akr1c14	AGAAGAGAGAGTAAGGAAGGGGAAGA	TCTTCCCCTTCCTTACTCTCTCTTCT
Emp2	AGGAGAGAGAGGAGAGAGGGGAGAGAG GGG	CCCCTCTCTCCCCTCTCTCCTCTCTCT CCT
Fn1	GGAGGGGCGGGAGGAGTCGGAC	GTCCGACTCCTCCCGCCCCTCC
Foxf1	CCGAGCCGGGAGGAGGAGGAGGAGC AGGAGGGGAGGGAGGGGAGGGGGCT	AGCCCCCTCCCCTCCCTCCCCTCCT GCTCCTCCTCCTCCCGGCTCGG
Serpinb6 b	GGAAAGAAGAGAAGAGGAAGAGGGGG	CCCCCTCTTCCTCTTCTCTTCTTTCC
Trim16	GGAGTAGAGAGAAGGGGAGGGGGAGG	CCTCCCCCTCCCCTTCTCTCTACTCC

Table 18: RNA oligonucleotides for CD-spectra and melting curve analysis

Probe	Sequence
FendrrBox	UCUUCUCUCUCCUCUCUUCUCCCUCCCCUC

Table 19: CRISPR-Cas9 guide RNAs

sgRNA	Type/Manufacturer	Sequence
ctrl^{SAM}	DNA Oligonucleotide; Merck (Darmstadt, Germany)	GGGTCTTCGAGAAGACCT
FendrrBox deletion upstream	DNA Oligonucleotide; Merck (Darmstadt, Germany)	TCAGGCAACACTCACTGGAC
FendrrBox deletion downstream	DNA Oligonucleotide; Merck (Darmstadt, Germany)	GGGAAGACATGGGGGAGTAA
Fendrr^{SAM} sg1	DNA Oligonucleotide; Merck (Darmstadt, Germany)	GGCCTCCGACGCTGCGCGCC
Fendrr^{SAM} sg2	DNA Oligonucleotide; Merck (Darmstadt, Germany)	TCAACGTAAACACGTTCCGG
Fendrr^{SAM} sg3	DNA Oligonucleotide; Merck (Darmstadt, Germany)	AGTTGGCCTGATGCCCCTAT
Fsd2 sg1	DNA Oligonucleotide; Merck (Darmstadt, Germany)	CGTCCACCACAAATAGCAAC
Fsd2 sg2	DNA Oligonucleotide; Merck (Darmstadt, Germany)	CTTAACACTCAAGTGTCACC
Fsd2 sg3	DNA Oligonucleotide; Merck (Darmstadt, Germany)	GGGAAGTAGGGTGCCAAAGC
lncFsd2 pA insertion	ArciTect™ crRNA; STEMCELL Technologies (Vancouver, Canada)	GUUGCUAGCCCACGGAAUGG
lncFsd2 Promoter deletion (L) upstream 1	ArciTect™ crRNA; STEMCELL Technologies (Vancouver, Canada)	GCAACUUGAUGCCCUGUACC
lncFsd2 Promoter deletion (L) upstream 2	ArciTect™ crRNA; STEMCELL Technologies (Vancouver, Canada)	AUCUGGAUAGACUACUCGGU
lncFsd2 Promoter deletion (R) downstream 1	ArciTect™ crRNA; STEMCELL Technologies (Vancouver, Canada)	UGUAAGUUGCCUACAAGUCG
lncFsd2 Promoter deletion (R) downstream 2	ArciTect™ crRNA; STEMCELL Technologies (Vancouver, Canada)	GCUACACAGUACACCGGGAA
lncFsd2 sg4	DNA Oligonucleotide; Merck (Darmstadt, Germany)	ACTTCCGGCTTCTTAGCCGT
lncFsd2 sg5	DNA Oligonucleotide; Merck (Darmstadt, Germany)	GTTCCACGTCAGGCACCGCG

Table 20: RNAi

GapmeR	Type/Manufacturer	Sequence
GapmeR negative control A	LNA GapmeR; QIAGEN (Hilden, Germany)	AACACGTCTATACGC
lncFsd2 ASO1	LNA GapmeR; QIAGEN (Hilden, Germany)	CGCTAAAATTACTGCA
lncFsd2 ASO2	LNA GapmeR; QIAGEN (Hilden, Germany)	TAAGTGCACCTCACTA
Swhtr ASO1	LNA GapmeR; QIAGEN (Hilden, Germany)	TAGGATTGAGAGACGT
Swhtr ASO2	LNA GapmeR; QIAGEN (Hilden, Germany)	TTCGATGGATCATTAA

Table 21: Stellaris™ FISH probes Quasar® 670 (BioCat, Heidelberg, Germany)

Nkx2-5	*Swhtr*
CTGGAGTAGGGGGGATTCAG	ACCACTTACAACCTGATTCA
CAGGTGGGTAGCAGAGAGTG	ATGTGCACCTTGAAAGCTTG
GAGAAAGGCGTGGGTGTGAG	AAATGCTCCTTTTAAGGGCT
CAGGTTCAGGATGTCTTTGA	GTTCACACTAATTGGTGTGC
GAAAGCAGGAGAGCACTTGG	AAAGGCCCTATCGATTACTT
TAGGCTCCCGGGTAAAATGT	GAAGTGGAGCAGTTGAAGCA
GCGCACAGCTCTTTTTTATC	CAAAGAGTATCTGGGCCTAC
CTGCGAGAAGAGCACGCGTG	GTATTGCAGCCAAGAAGTGA
GGTACCGCTGTTGCTTGAAG	GGCAATTGGAAGGGAGCAAG
TGGACGTGAGCTTCAGCACG	CTCAGCAGTTTGAGTAATCC
TGGAACCAGATCTTGACCTG	CTACTTTGTTTCCGTACAGA
TCGCTTGCACTTGTAGCGAC	AACCTTTTGCTATTGCCAAG
AGAAGCTCCAGAGTCTGGTC	CCTGAAGGGAAGGACTAGTT
TTGTAGCCATAGGCATTGAG	CTGAACACTCTGCCTAAGTG
GACGCCAAAGTTCACGAAGT	TGTTAGAGGGAAAGACACCC
GACTCTGCACGGTGTTCAAG	TAGGACAGGAAGAGAGCCAT
TTCTCCTAAAGGTGGGAGTC	GAAAAGGTCTAAGTGTCCCT
TGGGTGTGAAATCTGAGGGA	TTACCAACCAGGCACTATAG
GGGATGGATCGGAGAAAGGT	GTCACAGAGATAGAGCAGGA
ACACCAGGCTACGTCAATAA	GTGTTCTTTCGTTCAGTTTG
GTGTGGAATCCGTCGAAAGT	CTAGTTTTCGGGTCAAGAGT
CGACGGCAAGACAACCAGGG	TCATGAACCCACATTGCATA
AACATAAATACGGGTGGGTG	GTCAAGAATGCTCAACCTGG
GAGGGTTCTCATTTCTTACA	GTTTCAGACTGTAGGCATTT
GTTAGCGCACTCACTTTAAT	CTTTCCATTTGATAGGCAGA
	GGGATACTTGCCTACATATT
	TATCCCTTTGGATTTTGAGT
	CCATCCAGAATGGATAGAGT
	AAATCCATAGCATTGGCTGG
	ATAGCTCGATTCTCTAGGTG
	CAACCTCAAGGCCAGAAGAA
	AGACTTTCCTTGGTTGAAGG
	AAATCGCTTAGACTGCGTGG
	GACCCAAAACAACAGCCATG
	CTAGTTCTTACCACCAAGTA
	GCCCAATAGTTAGAACATGC
	CTAGAAACCAAAGCCCTACT
	TTGTCCTGTGGACTTAACAC
	TCCTGGTGCTCAAGGTAAAA
	GATGTCCAGGATGGAAATGC
	AATACATTGGAACACCCTGC
	TTCCAATCTGTGCAGAAGTC
	AGGCCCAGGTAAAAGAACTC
	TTGTTATCTATTGTGCCACG
	ATGCATTACTATATAGCCCA
	ATTCACATTTATCCTGACCT
	TATCTTATTCCATGCTCACT
	TGCCAACTGAACCAGTTTTA

Computational analysis

Table 22: Software and online tools used for computational analysis

Application	Software/Version	Developer/Literature
Coding-Potential Assessment online tool	CPAT 3.0.0	(Wang et al., 2013)
Conservation analysis	WashU epigenome browser v54.0.4	(Li et al., 2019; Li et al., 2022)
CRISPR/Cas9 sgRNA design	CRISPOR v. 5.01	(Concordet and Haeussler, 2018)
DNA annotation, analysis	CLC Main Workbench 22.0	QIAGEN (Hilden, Germany)
Echocardiography analysis	Vevo Lab 1.7.0	FUJIFILM VisualSonics (Toronto, Canada)
FACS data analysis	FACSDiva	BD Biosciences (Erembodegen, Belgium)
Genome data, sequence acquisition	Ensembl Genome Browser	(Cunningham et al., 2022; Zerbino et al., 2018)
Genome data, sequence acquisition	UCSC Genome Browser	Santa Cruz (Dallas, TX, USA)
Graphs	RStudio 2023.03.1+446	(R Core Team, 2022)
Graphs/Statistics	GraphPad Prism 9.5.0	GraphPad Software (San Diego, CA, USA)
Image analysis, quantification	(Fiji Is Just) ImageJ 1.53f51	(Schindelin et al., 2012)
Mathematical analysis, summary of data, generation of papers	Microsoft® Office	Microsoft (Redmont, WA, USA)
Nuclear localization signal prediction	NLStradamus	(Nguyen Ba et al., 2009)
Plasmid database, analysis	SnapGene Viewer 6.1	GSL Biotech LLC (Boston, MA, USA)
Primer design	Primer3web 4.1.0	(Koressaar et al., 2018; Koressaar and Remm, 2007; Untergasser et al., 2012)
Probe design	Stellaris® Probe Designer	Biosearch Technologies (Novato, CA, USA)
qRT-PCR (384-well)	QuantStudio™ Real-Time PCR Software v1.3	Thermo Fisher Scientific (Waltham, MA, USA)
qRT-PCR (96-well)	StepOnePlus™ Real-Time PCR Software v2.3	Thermo Fisher Scientific (Waltham, MA, USA)
RNA-sequencing visualization, graphs	Integrated Genome Browser 9.1.10	(Freese et al., 2016)
RNA-sequencing visualization, nucleotide blast	Integrative Genomics Viewer 2.16.1	(Robinson et al., 2023; Robinson et al., 2011; Thorvaldsdottir et al., 2013)
Running RStudio	R 4.3.0	(R Core Team, 2022)
Thermal melting assay analysis	SigmaPlot 12.5	Systat Software GmbH (San Jose, CA, USA)

Methods

Cell culture

Growth and care of HL-1 cells

The murine cardiac muscle cell line HL-1 was established by William Claycomb (Claycomb et al., 1998) and received as a gift from Prof. Dr. Ralf Gilsbach's laboratory (Institute of Experimental cardiology, University Heidelberg). The cells were maintained in a humidified CO_2 incubator at 37 °C and 5% CO_2 and cultured in Claycomb medium (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS; PAN-Biotech), 1% Penicillin-Streptomycin (Thermo Fisher Scientific), 1% GlutaMAX™ Supplement (Thermo Fisher Scientific) and 0.1 mM Norepinephrine (Sigma-Aldrich; 10 mM stock in 30mM ascorbic acid (Sigma-Aldrich)). The fully supplemented cell culture medium will be referred to as Claycomb full medium. Culture dishes were coated with 0.1% Gelatin (Sigma-Aldrich) containing 5 µg/ml Fibronectin (Merck) for 30 minutes at 37 °C in advance to introducing cells.

For passaging, the culture medium was removed completely and the cells washed with DPBS (Thermo Fisher Scientific) once to remove the serum. The cells were incubated with Trypsin-EDTA (Thermo Fisher Scientific) for 5-10 minutes at 37 °C. Detachment of the cells was additionally encouraged by agitation of the flask. After verifying detachment of all cells, the reaction was stopped by adding double the amount of fresh medium containing FBS. The cells were collected in DPBS (Thermo Fisher Scientific) and centrifuged at 1000 rpm for 5 minutes. The cell pellet was resuspended in Claycomb full medium and seeded in the required density.

Transfection of HL-1 cells

HL-1 cells are hard to transfect, hence for transfection HL-1 cells were detached and cell culture plates (6-well) coated as described above. For LNA GapmeR (QIAGEN) mediated knockdown a reagent aided approach using Lipofectamine™ RNAiMAX transfection reagent (Thermo Fisher Scientific) was conducted following the manufacturer's guidelines with experimentally determined optimizations. For $3 * 10^5$ cells 5 µl of Lipofectamine™ RNAiMAX transfection reagent were diluted in 100 µl Opti-MEM (Thermo Fisher Scientific). Separately, the LNA GapmeRs were diluted in 100 µl Opti-MEM to achieve a

final concentration of 20 nM. The diluted reagents were mixed and incubated at room temperature for 15 minutes. After the incubation, the coating solution was aspirated from the prepared cell culture dishes and the transfection mix was added. Detached HL-1 cells were resuspended in transfection media (Opti-MEM containing 5% fetal bovine serum (FBS; PAN-Biotech) and 1% Penicillin-Streptomycin (Thermo Fisher Scientific)) and counted using the Countess™ 3 Automated Cell Counter. The cells were seeded on top of the transfection media in a density of $1.5 * 10^5$ cells/ml and left to attach. After 4 hours the medium was changed to Claycomb full medium. For knockdown under hypoxic conditions, the cells were moved into a Physoxic and Hypoxic Atmosphere Workchamber and Hood with adjustable O_2 levels (Ruskinn Technology Ltd) after changing the medium to Claycomb full medium.

Growth and care of NIH3T3 cells

The murine embryonic fibroblast cell line NIH3T3 was established by George Todaro and Howard Green (1962; Department of Pathology, New York University School of Medicine) and was received as a gift from Dr. Nuno Guimarães-Camboa's laboratory (Institute for Cardiovascular Regeneration, Goethe-University Frankfurt am Main). The cells were maintained in a humidified CO_2 incubator at 37 °C and 5% CO_2 and cultured in DMEM medium (Thermo Fisher Scientific) supplemented with 10% bovine calf serum (CS; Sigma-Aldrich), 1% Penicillin-Streptomycin (Thermo Fisher Scientific), 1% GlutaMAX™ Supplement (Thermo Fisher Scientific). The fully supplemented medium will be referred to as NIH3T3 medium.

For passaging, the culture medium was removed completely and the cells washed with DPBS (Thermo Fisher Scientific) once to remove the serum. The cells were incubated with Trypsin-EDTA (Thermo Fisher Scientific) for 3-4 minutes at 37 °C. After verifying detachment of all cells, the reaction was stopped by adding double the amount of fresh medium containing CS. The cells were collected in DPBS (Thermo Fisher Scientific) and centrifuged at 1200 rpm for 4 minutes. The cell pellet was resuspended in NIH3T3 medium and seeded in the required density.

Transfection and stimulation of NIH3T3 cells

For transfection, NIH3T3 cells were detached as described above. The cells were counted using the Countess™ 3 Automated Cell Counter and seeded in a density of 15.79 * 10^5 cells per cm^2. The cells were ready for transfection after being left to attach overnight. Transfection was carried out using Lipofectamine™ 3000 Transfection Reagent (Thermo Fisher Scientific) according to the manufacturer's guidelines. In brief, 1.5 µl Lipofectamine™ 3000 reagent were diluted in 100 µl Opti-MEM (Thermo Fisher Scientific). In a separate tube 1 µg total DNA was diluted in 100 µl and mixed with 2 µl p3000 reagent. For plasmid overexpression 1µg of the respective plasmid were used, for dCas9 mediated overactivation a 1:3 dCas9 to gRNA plasmid ratio was used. The diluted Lipofectamine™ 3000 Transfection Reagent was mixed with the DNA mixture and incubated at room temperature for 15 minutes. During the incubation the cells were washed with DPBS (Thermo Fisher Scientific) once and supplied with prewarmed Opti-MEM. Transfection mix was added to the cells and incubated for 4h at 37 °C. After the incubation, the media was changed with full media containing FGF (10 ng/ml bFGF, Sigma Aldrich; 25 ng/ml rhFGF, R&D Systems), BMP-4 (40 ng/ml, R&D Systems) or CHIR99021 (3 µM, Stemcell). The treatment was replenished by changing media after 24h, cells were harvested for RNA isolation after 48h.

Growth and care of MLg cells

MLg cells (Mlg 2908) are fibroblast cells derived from the lung of a neonatal mouse and were established by Yasuko Hirokawa (1964; Department of Oncology, Institute of Medical Science, University of Tokyo). For this work they were received as a kind gift from Prof. Dr. Saverino Bellusci and Stefano Rivetti (Institute for Lung Health, Department of Internal Medicine, Justus-Liebig-Universität Gießen).

The cells were maintained in a humidified CO_2 incubator at 37 °C and 5% CO_2 and cultured in DMEM medium (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS; PAN Biotech), 1% Penicillin-Streptomycin (Thermo Fisher Scientific), 1% GlutaMAX™ Supplement (Thermo Fisher Scientific). For passaging, the culture medium was removed completely and the cells washed with DPBS (Thermo Fisher Scientific) once to remove the serum. The cells were incubated with Trypsin-EDTA

(Thermo Fisher Scientific) for 3-4 minutes at 37 °C. After verifying detachment of all cells, the reaction was stopped by adding double the amount of fresh medium containing FBS. The cells were collected in DPBS (Thermo Fisher Scientific) and centrifuged at 1000 rpm for 5 minutes. The cell pellet was resuspended in supplemented DMEM medium and seeded in the required density.

Transfection was conducted similar to NIH3T3 cell transfection that was described above.

Growth and care of Neuro-2a cells

Neuro-2a (N2A) cells are a neuroblast cell line derived from a spontaneous brain tumor of a strain A albino mouse by R.J. Klebe and F.H. Ruddle (1969; Department of Biology, Yale University, New Haven). They are known to be a very suitable transfection host and were therefore mainly used for the testing of sgRNAs in this work. In the *lncFsd2* project they were additionally used as an *Fsd2* non-expressing cell line to study the effect of *Fsd2* on-set on *lncFsd2*.

The cells were maintained in a humidified CO_2 incubator at 37 °C and 5% CO_2 and cultured in a 1:1 mix of DMEM medium (Thermo Fisher Scientific) and Opti-MEM medium (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS; PAN Biotech) and 1% Penicillin-Streptomycin (Thermo Fisher Scientific). For passaging, the culture medium was removed completely, and the cells washed with DPBS (Thermo Fisher Scientific) once to remove the serum. The cells were incubated with Trypsin-EDTA (Thermo Fisher Scientific) for 3-4 minutes at 37 °C. After verifying detachment of all cells, the reaction was stopped by adding double the amount of fresh medium containing FBS. The cells were collected in DPBS (Thermo Fisher Scientific) and centrifuged at 1000 rpm for 3 minutes. The cell pellet was resuspended in supplemented DMEM medium and seeded in the required density.

Transfection and stimulation of Neuro-2a cells

For transfection, the cells were detached as described above, counted using a Countess™ 3 Automated Cell Counter and seeded in a density of $0.31 * 10^6$ cells/cm^2 to attach over-night. Transfection was carried out using Lipofectamine™ 3000 Transfection Reagent (Thermo Fisher Scientific) according to the manufacturer's guidelines. In brief,

3 μl Lipofectamine™ 3000 reagent were diluted in 100 μl Opti-MEM (Thermo Fisher Scientific). In a separate tube 1 μg total DNA was diluted in 100 μl and mixed with 2 μl p3000 reagent. The diluted Lipofectamine™ 3000 Transfection Reagent was mixed with the DNA mixture and incubated at room temperature for 15 minutes. During the incubation the cells were washed with DPBS (Thermo Fisher Scientific) once and supplied with prewarmed Transfection media consisting of Opti-MEM (Thermo Fisher Scientific) supplied with 5% fetal bovine serum (FBS; PAN Biotech) and 1% Penicillin-Streptomycin (Thermo Fisher Scientific). Transfection mix was added to the cells and incubated overnight. Media was changed to full culture media the next morning.

Growth and care of murine embryonic stem cells (mESCs)

F1G4 mESCs were originally derived from the blastocyst of a mouse of mixed C57BL/6Ncr x 129S6/SvEvTac genetic background. They were of pluripotent nature that could be remained in cell culture and thus could be used to generate blastocysts that can be re-transferred into pseudo-pregnant females, generating mutant embryos and offspring (described later).

The mESCs were maintained in a humidified incubator at 37 °C and 7% CO_2. They were always cultured on gelatin-coated dishes. For this purpose, an 0.1% sterile filtered gelatin solution was incubated on the respective dishes for 2 hours at 37 °C. After removing the gelatin solution and allowing the plates to dry for 2 additional hours at room temperature, the cells could be introduced. For general maintenance of the mESCs they were cultivated on a mitomycin C (Sigma-Aldrich) treated MEF feeder monolayer that was pre-seeded on coated plates. Feeders alone were cultivated in Feeder medium (DMEM (Thermo Fisher Scientific) containing 10% FBS, 1% GlutaMAX™ Supplement (Thermo Fisher Scientific) and 1% Penicillin-Streptomycin (Thermo Fisher Scientific)). Upon introducing mESCs to the feeders, they were cultivated in ES-medium (KnockOut™ DMEM (Thermo Fisher Scientific) containing 15% ESC-grade FBS (PAN Biotech), 1% GlutaMAX™ Supplement (Thermo Fisher Scientific), 1% Penicillin-Streptomycin (Thermo Fisher Scientific), 1% MEM-NEAA, 0.11 mM 2-Mercaptoethanol (Thermo Fisher Scientific), 1% EmbryoMax® Nucleosides and 10 ng/ml ESGRO® mLIF supplement). For passaging the cells were detached with Accutase™ (Thermo Fisher Scientific) at 37 °C

for 3-5 minutes. The reaction was stopped with equal amount of fresh ES-medium. The cells were washed in DPBS (Thermo Fisher Scientific), centrifuged at 1000 rpm for 5 minutes and resuspended in fresh ES-medium for seeding on a fresh plate with a pre-seeded feeder monolayer. The medium was replenished every day.

For experiments where the feeders needed to be removed, this was achieved by detaching the cells as described above and plating them in 2i medium (Consisting of an equal mixture of DMEM/F12 (Thermo Fisher Scientific) and Neurobasal®-A Medium (Thermo Fisher Scientific) containing 1% GlutaMAX™ Supplement (Thermo Fisher Scientific), 1% Penicillin-Streptomycin (Thermo Fisher Scientific), 1% MEM-NEAA, 0.11 mM 2-Mercaptoethanol (Thermo Fisher Scientific), 1% Sodium Pyruvate (Thermo Fisher Scientific), 1x N-2 Supplement (Thermo Fisher Scientific), 1x B-27™ Supplement (Thermo Fisher Scientific), 1 µM CGP77675 MEK-inhibitor (Sigma-Aldrich), 3 µM CHIR99021 GSK3-inhibitor (Sigma-Aldrich) and 10 ng/ml ESGRO® mLIF Supplement (Thermo Fisher Scientific) on 3 uncoated cell culture dishes for 20-60 minutes each at 37 °C before plating them on a gelatin coated cell culture dish. Detachment for sub-culturing was conducted in similar manner as for the mESCs on feeders and again, the medium was changed every day.

Generation of mutant mESCs using Neon™ Transfection

To obtain an 80 µM gRNA solution, 4 µl ArciTect™ crRNA was combined with 4 µl ArciTect™ tracrRNA, mixed with 2 µl 5x ArciTect™ Annealing Buffer and heated to 95 °C for 5 minutes. Subsequently, the temperature was reduced slowly at a rate of ~ 0.5 °C/minute until reaching room temperature before storing it on ice. The RNP complex was prepared by mixing 0.56 µl of the annealed gRNA with 0.9 µl ArciTect™ Cas9 Nuclease in 3.54 µl Resuspension Buffer R (STEMCELL) and subsequent incubation at room temperature for 20 minutes. For knockout the resulting RNP complex was diluted with 2.5 µl Resuspension Buffer R, for knock-in up to 2.5 µl of the required ssDNA template were added.

The feeders were removed from the mESC culture shortly before the transfection to minimize culture in 2i medium. For transfection the cells were detached as described above and counted. $0.3*10^6$ mESCs were resuspended in 7.5 µl Resuspension Buffer R

and combined with the RNP complex. With a Neon® pipette tip, 10 µl of the mixture were taken up carefully to avoid bubbles and placed in the electroporation chamber containing 3 ml Electrolytic Buffer E (STEMCELL). The mESCs were electroporated at 1400 V for 3 pulses with a pulse width of 10 ms. Afterwards the cells were resuspended in 40 ml fresh 2i medium and plated on 3 gelatin-coated 10cm cell culture plates. After 4-5 days of daily media changes clones were picked by flushing colonies off with a pipette and resuspending single colonies in 15 µl Accutase™ (Thermo Fisher Scientific) in a 96-well U-bottom plate (TPP). After 5 minutes of incubation at 37 °C the clones were resuspended into a gelatin coated 96-well flat bottom plate (TPP) containing fresh 2i medium. The cells were maintained in the same was as on dishes and expanded onto several plates, to use one as DNA plate for genotyping. To harvest DNA, an extraction buffer modified after Laird's (Laird et al., 1991) was used (100 mM Tris-HCl pH 8.0, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, 100 µg/ml Proteinase K). The medium was removed from the plate and the cells were washed once with DPBS. 100 µl of extraction buffer were added per well and the plate was incubated at 50 °C in a humid chamber over-night. Positive clones were expanded on feeder cells.

Isolation and handling of primary neonatal murine cardiomyocytes

Mice of the strain C57/Bl6J were timed mated. Neonatal mice were sacrificed by decapitation at p0 to collect the heart in HBSS (Thermo Fisher Scientific) on ice. Single cell solution was obtained using the Neonatal Heart Dissociation Kit (Miltenyi) following the manufacturer's guidelines. In brief, 2.3 ml of Buffer X were mixed with 62.5 µl Enzyme P and preheated at 37 °C for 5 minutes. In the meantime, 25 µl Buffer Y were mixed with 12.5 µl Enzyme A and 100 µl Enzyme D and added to the preheated enzyme mix. Washed hearts and enzyme mix were transferred to a gentleMACS™ C tube and incubated at 37 °C for 15 minutes. After attaching the gentleMACS™ C tube to the gentleMACS™ Dissociator a program termed 'mr_neoheart_01' was executed. This procedure was repeated 2 times before 7.5 ml DMEM (Thermo Fisher Scientific) containing 10 % fetal bovine serum (FBS, PAN Biotech), 1% Penicillin-Streptomycin (Thermo Fisher Scientific), 1% GlutaMAX™ Supplement (Thermo Fisher Scientific), 1% Non-Essential Amino Acids (Thermo Fisher Scientific) and 100 mg/l Sodium Pyruvate

(Thermo Fisher Scientific) were added to stop the enzymatic reaction. The described medium will be referred to as CM medium in the following. To obtain a single cell solution without debris, die suspension was filtered using a 70 µm cell-strainer (Corning). Cells were pelleted at 600xg for 5 minutes. PEB buffer (PBS pH7.4 containing 0.5% bovine serum albumin and 2 mM EDTA) was obtained by diluting MACS® BSA Stock Solution (Miltenyi) in autoMACS® Rinsing Solution (Miltenyi) 1:20. This buffer was used to resuspend the cell pellet before adding 10 ml of 1x Red Blood Cell Lysis Solution and incubating for 2 minutes at room temperature. The cells were pelleted by centrifugation at 600xg for 5 minutes and resuspended in DPBS containing 15 µl Enzyme A. The cell pellet resulting after centrifugation at 600xg for 5 minutes was resuspended in the appropriate amount of CM medium for further purpose.

Culture dishes were coated with 0.1% Gelatin (Sigma-Aldrich) containing 5 µg/ml Fibronectin (Merck) for 30 minutes at 37 °C in advance to introducing cells. The isolation was usually conducted in the evening to allow the cells to attach and recover overnight and immediately proceed with experiments the next morning.

Animals

Maintenance of mice

In this work mice of the strains C57/Bl6J and CD-1 were mainly used. For experiments, all mice that were used were of C57/Bl6J genetic background and male sex while CD-1 females were used as foster animals for embryo re-transfer or donors of two-cell-stage embryos for tetraploid or diploid morula aggregation (described later). Animals were maintained and handled according to Animal Welfare Act guidelines and with the principles of 3R in mind. The housing facility provided a controlled 12h/12h light-dark cycle and a temperature of 22-23 °C with relative humidity of 45-55 %. Fortified complete mouse feed (Ssniff #V1534-000) and tap-water were available *ad libitum*. All experiments were approved by the responsible authorities (License numbers: G0349/13, G0368/08, FU/1064).

Cardiac stress model – Left anterior descending artery (LAD) ligation

The cardiac stress model was a co-operation with Thierry Pedrazzini from the University of Lausanne (Department for Medicine).

In brief, the mouse was anesthetized using a combination of ketamine, xylazine, and acepromazine via an intraperitoneal injection. It was placed on a warming pad to maintain body temperature and underwent endotracheal intubation for artificial ventilation. Ocular gel was applied to hydrate the cornea during the procedure. The left thorax was shaved, disinfected, and a thoracotomy was performed. The left descending coronary artery was ligated with a silk suture to induce occlusion. After verification of occlusion, the chest and skin were sutured closed. The mouse was gradually taken off the ventilator, given intra-peritoneal glucose, and once spontaneous breathing resumed, the endotracheal tube was removed. The mouse was then placed back in its cage with access to standard food and water, and an analgesic drug was administered subcutaneously after the surgery (Rogala et al., 2022).

In vivo transthoracic ultrasound imaging

The transthoracic ultrasound imaging before and following LAD ligation was a co-operation with Thierry Pedrazzini from the University of Lausanne (Department for Medicine).

Briefly, transthoracic echocardiography was performed on mice using a 30 MHz probe and the Vevo 2100 Ultrasound machine. The mice were lightly anesthetized with 1-1.5% isoflurane to maintain a heart rate of 400-500 beats per minute. The procedure was conducted with the mice lying on a heated platform to maintain body temperature. Hair was removed using a depilatory agent, and ultrasound gel served as a coupling medium. The heart was imaged in 2D mode from the parasternal long-axis view. M-mode measurements were taken at specific locations in the left ventricle to assess dimensions such as IVS;d, IVS;s, LVPW;d, LVPW;s, LVID;d, and LVID;s. These measurements were averaged from three separate M-mode images. Additionally, left ventricular fractional shortening (%FS) and ejection fraction (%EF) were calculated using the M-mode data, with %EF derived from a specific formula based on left ventricular volumes at diastole and systole (Rogala et al., 2022).

Generation of embryos and strains from mESCs

For this work, considering the *Swhtr* project, only strains, no embryos, were generated. For this purpose, the genetically modified mESCs (generated by Sandra Währisch (*Swhtr³ˣᵖᴬ/³ˣᵖᴬ*) and Maria-Theodora Melissari (*Swhtr³ˣᵖᴬ/³ˣᵖᴬ;ᵗᵍ*)) were used for aggregation with diploid morula CD-1 embryos. In brief, the diploid morula was flushed from the oviducts of CD-1 donor mice and the zona pellucida was chemically removed. After bringing the embryo together with the genetically modified mESCs in a cell culture dish with indentation well, the resulting blastocysts can be surgically re-transferred into the uteri of pseudo-pregnant CD-1 foster animals (procedure done by Sandra Währisch). For the *lncFsd2* project, generation of embryos and a strain were attempted. Generation of the knock-out strain was conducted using diploid aggregation as described above. Generation of embryos was attempted by tetraploid and diploid aggregation. For tetraploid aggregation diploid two-cell-stage embryos were harvested from CD-1 donor mice and electro-fused to generate tetraploid four-cell-stage embryos. Two of these embryos were brought together with the genetically modified mESCs (generated by me) in a cell culture dish with indentation well. The resulting blastocyst was surgically re-transferred into the uteri of pseudo-pregnant CD-1 foster animals (procedure performed by Matthias Breiling).

Fractionation of the main cell types of the adult heart

Fractionation of the main cell types of the adult heart was conducted in co-operation with Eva-Maria Rogg (Institute of Cardiovascular Regeneration).

Adult mice were sacrificed by cervical dislocation and perfused with HBSS to remove blood. The hearts were subsequently collected in HBSS (gibco). Enzymatic digestion of the hearts into a single-cell suspension was conducted using the Multi Tissue Dissociation Kit 2 (Miltenyi #130-110-203). The cardiomyocyte fraction was obtained by pre-plating for 40 minutes whereas the non-adherent cells contained primarily cardiomyocytes. Endothelial cells (Miltenyi #130-097-418) and immune cells (Miltenyi #130-109-682) were obtained by magnetic separation using the MACS magnetic separation system. Fibroblasts were enriched from the unbound flow-through from the magnetic separation

by another pre-plating step in which predominantly the fibroblasts adhered after 40 minutes.

Preparation of heart slices and treatment

For preparation of the heart slices, 4 wild type and 4 $Swhtr^{3xpA/3xpA}$ mice were sacrificed by cervical dislocation when they reached 8 weeks of age. The heart was flushed and collected in HBSS before washing in BDM buffer (HBSS + 10 mM BDM) to remove as many blood cells as possible. Apex and base of the heart were manually removed before preparing the slices using a sharp scalpel. The slices were washed in BDM buffer once more before submerging them in culture medium consisting of DMEM (gibco) supplemented with 10 % Fetal Bovine Serum, 1% Non-Essential Amino Acids (gibco) and 1% Penicillin-Streptomycin (gibco). The heart slices were left to recover overnight and controlled for possible contamination. Remaining blood cells were removed by one wash with PBS before replenishing the culture medium. During the experiment, the media was refreshed every other day.

For the treatment, the slices were incubated in a humidified hypoxic chamber (3% O_2, 5% CO_2) at 37 °C for 7 days and moved to a humidified incubator with normoxic conditions (21% O_2, 5% CO_2) at 37 °C for 7 additional days. Slices that were incubated in a humidified incubator at normoxic conditions (21% O_2, 5% CO_2) at 37 °C for 14 days served as the control. Following the treatment, the slices were harvested in 1 ml TRI reagent (Sigma-Aldrich) and homogenized using Precellys® 2 mL Soft Tissue Homogenizing Ceramic Beads. RNA was extracted by Phenol/Chloroform extraction, which is described later.

FACS Sorting for Cardiomyocyte and Non-Cardiomyocyte cells of the embryonic heart

Timed mated female mice were sacrificed at day E12.5 after positive plug check and gain of weight. The hearts of the embryos were collected in HBSS on ice and washed with fresh ice-cold HBSS twice to remove the blood. All centrifugation steps were carried out at room temperature for 5 minutes at 1000xg. A single cell suspension was obtained by using the dissociation enzymes of the Neonatal Heart Dissociation Kit (Miltenyi). After

passing the cells through an 80 µm cell strainer (Corning) and washing the cells, the pellet was resuspended in 4% ROTI®Histofix at room temperature for 5 minutes. For permeabilization, the cells were washed in PBS and resuspended in PBS containing 1% PFA (Thermo Fisher Scientific) and 0.2% TWEEN® 20 for 5 minutes at room temperature. The cells were washed in PBS and stained with an APC conjugated cardiac troponin antibody (cTNT, Miltenyi) diluted 1:10 in PBS for 20 minutes at room temperature in the dark. After washing the cells with PBS three times, the pellet was resuspended in 500 µl PBS and used for FACS sorting. The gates were adjusted for single cells to avoid debris and cell clumps (Appendix Fig. 1) and subsequently the cells were sorted for APC positive and negative cells that represented the cardiomyocyte (APC positive) and non-cardiomyocyte (APC negative) fraction.

Magnetic separation of Cardiomyocyte and Non-Cardiomyocyte cells of the neonatal heart

The hearts of neonatal p0 mice were collected in HBSS on ice and washed twice with fresh ice-cold HBSS to remove remaining blood. A single cells suspension was obtained using the Neonatal Heart Dissociation Kit (Miltenyi) as described above for the isolation of murine primary neonatal cardiomyocytes. The cells were counted and resuspended in 90 µl of buffer and mixed with 10 µl Neonatal Cardiomyocyte Isolation Cocktail per 5×10^6 cells. After 15 minutes incubation at 4 °C in the dark, the volume was adjusted to 500 µl and the cells used for magnetic separation after equilibration of the MS Columns with PEB buffer on a MACS separator with magnetic field. The unlabeled cells of the first to fourth flow-through represented the population enriched for cardiomyocytes. The labeled non-cardiomyocyte fraction was obtained by removing the column from the magnetic field and flushing out the cells by forcing 1 ml PEB buffer through the column using the plunger. The cells were resuspended in 1 ml TRI reagent and used for Phenol/Chloroform extraction, which is described later.

Histology

Tissue preparation and embedding for cryo-sections

After cervical dislocation, the chest of the mice was open to briefly perfuse the heart with DPBS followed by 50 mM KCl (Sigma-Aldrich). The heart was carefully removed using scissors and fixed in 4% PFA (Thermo Fisher Scientific) in PBS on ice. The skeletal muscle was obtained from the right hindleg. After removal of the skin, the quadriceps tendon was released with an incision and the quadriceps gently removed from the femur. To remove any remaining fur or blood cells, the skeletal muscle was washed in PBS before collection in 4% PFA. The fixation was continued at 4 °C over-night. The heart was washed in PBS 3 times and incubated in 5% sucrose (Sigma-Aldrich) in PBS at 4 °C for 2 hours followed by 2 hours incubation in 12% sucrose in PBS at 4 °C. Subsequently, the heart was moved to 20% sucrose in PBS and incubated at 4 °C over-night. Initially, the tissue floated and indicated by sinking to the bottom of the vessel successful impregnation of the tissue with sucrose. After washing the tissue in a 1:1 O.C.T.-20% sucrose mixture, the tissue was placed in a mold, submerged in 1:1 O.C.T.-20% sucrose mixture. Using a metal block cooled down by dry ice, the tissue was frozen slowly, allowing adjustments of orientation if necessary. After freezing, sections of 10 µm were obtained using a cryostat (Leica Biosystems). The sections were placed on SuperFrost® Plus adhesion slides (VWR) and kept at -20 °C until further use.

Immunofluorescence of frozen sections

The frozen sections were warmed at room temperature for 2-5 minutes. To wash off the O.C.T. compound, the sections were submerged in PBS for 5 minutes. Permeabilization was conducted by incubation of the sections with PBS containing 0.5% Triton-X for 30 minutes. After administration of a hydrophobic barrier using ImmEDGE™ Hydrophobic Barrier Pen (Vector Laboratories), the tissue was covered with blocking solution (PBS, 10% normal goat serum, 5% skim milk, 0.1% Triton-X) for 60 minutes. The blocking solution was gently flicked away and replaced by blocking solution containing the primary antibody of interest in an experimentally determined dilution. The sections were incubated at 4 °C over-night in a humidity chamber. The slides were washed three times in PBS containing 0.1% Triton-X for 5, 10 and 15 minutes respectively. The secondary antibody

was diluted appropriately in blocking solution together with 1µg/ml DAPI (Carl Roth) and incubated on the sections for 60 minutes at room temperature in a humidity chamber. The slides were washed three times in PBS containing 0.1% Triton-X for 5, 10 and 15 minutes respectively. After removing the hydrophobic barrier, die slides were mounted with fluorescence mounting medium (Agilent Technologies).

Tissue preparation and embedding for paraffin-sections

After cervical dislocation, the chest of the mice was open to briefly perfuse the heart with DPBS followed by 50 mM KCl (Sigma-Aldrich). The heart was carefully removed using scissors and fixed in 4% PFA (Thermo Fisher Scientific) in PBS at 4 °C over-night. After fixation the tissue was dehydrated using an automated tissue processing center (MEDITE) and embedded in liquid paraffin using a tissue embedding center (MEDITE). The blocks were left to harden at room temperature and could be kept at room temperature until further use. Sections of 4 nm were obtained using a microtome (Microm International) and stored at room temperature until staining.

H&E Staining of Paraffin Sections

To remove the paraffin, the slides were immersed in ROTI®Histol (Carl Roth) two times for 5 minutes each. To rehydrate the sections, they were immersed in a descending ethanol gradient (100% twice, 96%, 70% and 50% EtOH) each step incubating for 2 minutes. After immersing the sections in distilled water for 2 minutes, they were immersed in Hemalum solution acid according to Mayer (Carl Roth), rinsed in distilled water for 2 minutes and then incubated with tap water for 1-15 minutes. The sections were washed in distilled water for 2 minutes and incubated in 0.5% Eosin Y solution (Carl Roth) for 3 minutes. After washing the sections with tap water for 1 minute and distilled water for 2 minutes, they were dehydrated in an ascending ethanol gradient (50%, 70%, 96% and twice 100% EtOH) for 2 minutes each step. After incubation in ROTI®Histol (Carl Roth) twice for 5 minutes, the sections were mounted and left to dry for imaging.

WGA Staining of Paraffin Sections

To remove the paraffin and rehydrate the sections, the slides were immersed in ROTI®Histol (Carl Roth) two times for 5 minutes each and in a descending ethanol gradient (100% twice, 80% and 50% EtOH) and distilled water each step incubating for 2 minutes. The slides were immersed in PBS containing 0.1% TritonX-100 for 5 minutes and incubated in prewarmed citrate buffer (Thermo Fisher Scientific) for 30 minutes in a steam cooker. After the sections were cooled down at room temperature for additional 30 minutes, they were washed 3 times with PBS containing 0.1% TritonX-100 for 5 minutes each. The sections were covered WGA-Alexa488 primary antibody (Thermo Fisher Scientific) diluted 1:100 in PBLEC buffer (PBS containing 1mM $MgCl_2$, 1mM $CaCl_2$, 0.1mM $MnCl_2$, 1% TritonX-100) for 1 hour at room temperature and subsequently washed with PBS containing 0.1% TritonX-100 3 times for 5 minutes. Subsequently the slides were covered with PBS containing 0.1% TritonX-100 and 1µg/ml DAPI (Carl Roth) for 10 minutes at room temperature, washed with PBS containing 0.1% TritonX-100 3 times for 5 minutes and were mounted using fluorescence mounting medium (Agilent Technologies).

Molecular biology

Phenol/Chloroform extraction of RNA

RNA was purified using Phenol/Chloroform extraction from samples harvested in TRI reagent (Sigma-Aldrich). 100 µl Chloroform (Sigma-Aldrich) were added per 500 µl of TRI reagent and mixed vigorously. To separate the phases, mixed samples were centrifuged at 13500 rpm at 4 °C for 15 minutes. The aqueous phase was moved to a fresh reaction tube and 0.7 volumes of 2-Propanol (Carl Roth) were added. After mixing by inverting, the samples were incubated at -20 °C for at least 1 hour, up to over-night. Subsequently, the RNA was pelleted at 13500 rpm at 4 °C for 20 minutes. The pellet was washed with 70% EtOH at room temperature and centrifuged for 10 minutes at 13500 rpm at 4 °C. The RNA pellet was resuspended in the appropriate amount of nuclease free water. The RNA amount was quantified using a spectrophotometer and adjusted to the required concentration.

Reverse transcription of RNA and analysis by quantitative Realtime PCR

500 - 1000 ng of RNA were typically used for reverse transcription of RNA into cDNA. For that purpose, the QuantiTect Reverse Transcription Kit (QIAGEN) was used following the manufacturer's guidelines. In brief, the RNA was incubated with gDNA wipeout buffer for 2-5 minutes at 42 °C before adding 1µl RT primer mix, 1 µl Quantiscript Reverse Transcriptase and 4 µl Quantiscript RT Buffer per sample. The mixture was incubated at 42 °C for 30 minutes followed by heat inactivation at 95 °C for 3 minutes.

For quantitative Realtime PCR (qRT-PCR) Fast SYBR™ Green Master Mix was used for analysis on either the StepOnePlus™ Real-Time 96-well PCR System (15 µl reactions) or the ViiA™ 7 Real-Time 384-well PCR System (10 µl reactions). For analysis of murine samples, the genes *Rpl10* and *Eaf1* were used as housekeeping genes, while for analysis of human samples *GAPDH* was used as housekeeping gene.

Transformation of plasmids into competent bacteria and colony picking

Ligated samples were combined with 25 µl competent JM109 *E.coli* cells (Promega) and incubated on ice for 30 minutes. To transform the plasmid into the bacteria, a heat shock was performed at 42 °C in a water bath for 75 seconds. The bacteria were left to recover on ice before adding 900 µl of SOC Outgrowth Medium (NEB) and incubating the bacteria at 800 rpm at 37 °C for 1 hour. Subsequently, the bacteria were spread onto LB-Agar plates containing 50 µg/ml Carbenicillin (Carl Roth) and incubated at 37 °C over-night. The bacterial colonies were picked from the plate by dipping a sterile pipette tip into the colony and resuspending it in LB-medium containing 100 µg/ml Carbenicillin for inoculation at 500 rpm at 37 °C over-night. For miniprep of bacterial plasmid DNA, the QIAprep® Spin Miniprep Kit (QIAGEN) following the manufacturer's guidelines.

Cloning of *FendrrBox* overexpression plasmid

The regular *Fendrr* overexpression plasmid (phbAct-Fendrr) was generated by and received from Dr. Phillip Grote (Transgenic Core Facility, Georg-Speyer-Haus Frankfurt am Main). A map of the plasmid can be found attached to the Appendix (Appendix Fig. 3). To remove the *FendrrBox*, Acc65I and EcoRV restriction sites were made use of. 5 µg of the phbAct-Fendrr plasmid were digested using 10 U Acc65I and 10 U EcoRV at 37 °C

over-night followed by dephosphorylation using 10 U of rSAP (NEB). The fragments were separated on an agarose gel and the required backbone was purified by QIAquick Gel Extraction Kit (QIAGEN).

The insert was amplified by Phusion® (NEB) PCR to amplify the regions upstream and downstream of the *FendrrBox*, using primers with complementary overlaps for subsequent fusion PCR to merge the fragment. To amplify the part upstream of the *FendrrBox*, the following primers were used: Boxdel_left_fw: TGCCCACAGCACCCACTTGA; Boxdel_left_rv: TGTTCAGGCAACACTCACTGGTAAGGGGTGTGGCGTAAAATATAATC. To amplify the part downstream of the *FendrrBox*, the following primers were used: Boxdel_right_fw: GATTATATTTTACGCCACACCCCTTACCAGTGAGTGTTGCCTGAAC; Boxdel_right_rv: ATCAAGCTTATCGATACCGT, whereas the reverse primer contained a 3' phosphorylation. The fragments were retrieved from the agarose gel using the QIAquick Gel Extraction Kit (QIAGEN). For the fusion 5 ng of each fragment were pooled as template and the primers Boxdel_right_fw and Boxdel_left_rv were used. Absence of the *FendrrBox* was verified by Sanger sequencing (Microsynth Seqlab) using the amplification primers before digesting the insert with 10 U Acc65I at 37 °C over-night. After purification of the insert with the QIAquick PCR purification Kit (QIAGEN). Ligation of 700 ng of the insert with 100 ng of the backbone was carried out using T4 ligase (NEB) for 2 hours at room temperature. Transformation and colony picking were carried out as described above.

To identify positive clones, the purified plasmids were digested using 10 U DraI at 37 °C for 1 hour as it cut within the *FendrrBox*. Positive clones were investigated for point mutations that the PCR amplification might have introduced by Sanger sequencing and subsequently used for further experiments. For cryo-conservation as bacterial glycerol stocks, 500 µl of the bacterial suspension were mixed with 500µl 50% glycerol (Sigma Aldrich) and stored at -80 °C. For inoculation of frozen glycerol stock, some of the frozen bacteria were scratched into fresh LB-medium containing 100 µg/ml Carbenicillin and incubated at 37 °C at 500 rpm over-night.

Cloning of dCas9 sgRNA plasmids

The respective dCas9 sgRNAs were ordered as forward and reverse complimentary oligonucleotides (MERCK) with the respective 3' and 5' overhangs complimentary to the desired insertion site of the plasmid. As backbone the sgRNA(MS2) cloning backbone (addgene #61424) was used. Linearization was carried out at 37 °C over-night using 10 U BbsI restriction enzyme (NEB) per 5 µg plasmid, followed by purification of the DNA using the QIAquick PCR purification Kit (QIAGEN) following the manufacturer's guidelines.

To anneal the oligonucleotides, they were mixed in equal ratio in annealing buffer (10 mM Tris-HCl pH7.5, 50 mM NaCl, 1 mM EDTA) and heated to 95 °C for 5 minutes. Subsequently, the temperature was reduced slowly at a rate of ~ 0.5 °C/minute until reaching room temperature. 100 ng of the linearized plasmid were ligated with 5 µl of the annealed oligonucleotides for 2 hours at room temperature using T4 ligase (NEB). Transformation and picking of clones was carried out as described above. To identify positive clones, 100 ng of the purified plasmid DNA was sent for Sanger Sequencing (Microsynth Seqlab) using the following sequencing primer: ACTATCATATGCTTACCGTAAC. Conservation of positive clones was carried out as described above.

Cloning of *FendrrBox* deletion sgRNA plasmid with selection fluorescence

The respective Cas9 sgRNAs were ordered as forward and reverse complimentary oligonucleotides (MERCK) with the respective 3' and 5' overhangs complimentary to the desired insertion site of the plasmid. As backbone the B52 empty plasmid backbone to express 2 sgRNAs (addgene #100708) was used for cloning one sgRNA into the BbsI restriction site and one sgRNA into the Esp3I restriction site. As the Esp3I (NEB) digestion to add the second sgRNA removed one guanine, it was added to the oligonucleotide in order to keep the gRNA scaffold intact. Furthermore, as the sticky ends of the two restriction digests were too similar and hence, cloning of both sgRNAs did not yield positive colonies when conducted simultaneously, the sgRNAs were cloned into the plasmid one after another. This was achieved by following the same steps as described above for cloning dCas9 sgRNA plasmids. For identification of positive clones 100 ng of

the purified plasmid DNA was sent for Sanger Sequencing (Microsynth Seqlab) using the following sequencing primer: ACTATCATATGCTTACCGTAAC.

For insertion of the mScarlet red fluorescence marker, the EcoRI-HF (NEB) restriction site of the plasmid was used, once both *FendrrBox* sgRNAs were successfully cloned into the plasmid. Linearization was carried out using 10 U of restriction enzyme per 5 µg plasmid at 37 °C over-night followed by dephosphorylation using 10 U of rSAP (NEB) to prevent self-ligation of the plasmid. The rSAP reaction was heat inactivated at 65 °C for 5 minutes before purification of the DNA by using the QIAquick PCR purification Kit (QIAGEN) following the manufacturer's guidelines.

The insert was generated by Phusion® (NEB) PCR amplification of the CMV promoter and the *mScarlet* with an SV40 poly(A) signal using 10 ng per reaction of the ITPKA_3xmScarletI plasmid (addgene #112960) as template. Complementary overhangs for following fusion PCR were added to the oligonucleotides. The CMV promoter was amplified using the following primers: CMV_insert_fw: ATATGAATTCTGGATAACCGTATTACCGCC; CMV_insert_rv: GCCTCGCCCTTGCTCACCATGTTCACTAAACCAGCTCTGC whereas the forward primer contained an EcoRI restriction site. The *mScarlet* sequence was amplified using the following primers: mScarlet_insert_fw: GCAGAGCTGGTTTAGTGAACATGGTGAGCAAGGGCGAGGC; mScarlet_insert_rv: ATATGAATTCACGCCTTAAGATACATTGAT whereas the reverse primer contained an EcoRI restriction site. The resulting PCR products were excised from the agarose gel and purified using the QIAquick Gel Extraction Kit (QIAGEN). 5 ng of each PCR product were pooled and used as template for a fusion Phusion® (NEB) PCR using the CMV_insert_fw and the mScarlet_insert_rv primer. The resulting band was excised from the agarose gel and purified using the QIAquick Gel Extraction Kit (QIAGEN) and sent for Sanger Sequencing (Microsynth Seqlab) using the amplification primers. The insert was digested using 10 U EcoRI-HF (NEB) at 37 °C over-night and purified using the QIAquick PCR purification Kit (QIAGEN). 100 ng linearized dephosphorylated backbone were ligated with 700 ng digested insert for 2h at room temperature using T4 ligase (NEB). The plasmid was transformed into competent JM109 *E. coli* bacteria (Promega) and DNA was isolated by miniprep as described above. Positive clones were identified by control

restriction digest with 10 U NdeI (NEB) at 37 °C for 2 hours, expecting bands of 2077 bp, 359 bp and 201 bp in the plasmid without insert, 2077 bp, 1477 bp, 359 bp and 289 bp in the plasmid with the flipped insert and 2077 bp, 1322 bp, 444 bp and 359 bp in the plasmid with the non-flipped insert. The DNA from the positive clones was sent for Sanger Sequencing (Microsynth Seqlab) to control for point mutations that might have been inserted during PCR amplification of the insert and test-transfected into Neuro-2A cells to verify presence of mScarlet red fluorescence as described above (Appendix Fig. 2).

Western Blot (Bio-Rad System)

Gels for Western Blot with the Bio-Rad system were self-made. For two 10% resolving gels, 6 ml clean water was mixed with 5 ml Acrylamide/Bis mix (30/0.8%; Merck), 3.75 ml 1.5 M Tris-HCl pH 8.8, 150 µl 10% SDS (Thermo Fisher Scientific), 70 µl 40% APS (Thermo Fisher Scientific) on ice. 30 µl of TEMED (Carl Roth) were added and the mixture poured into a gel mold, adding 1 ml 2-Propanol on top. After successful polymerization of the gel, the 2-Propanol was removed and the gel washed with clean water. For two 5% stacking gels 5.7 ml water were mixed with 1.7 ml Acrylamide/Bis mix (30/0.8%; MERCK), 2.5 ml 0.5 M Tris-HCl pH 6.8, 100 µl 10% SDS (Thermo Fisher Scientific), 50 µl 40% APS (Thermo Fisher Scientific) on ice. 20 µl of TEMED (Carl Roth) were added and the mixture was poured on top of the resolving gel, carefully inserting a comb to create pockets.

The protein lysate was boiled at 95 °C in 1x ROTI®Load 1 (Carl Roth) for 10 minutes and centrifuged at 4 °C for 10 minutes at full speed before applying it to the gel. The gel electrophoresis was carried out in Tris-Glycine-SDS running buffer (25 mM Tris (Carl Roth), 192 mM Glycine (Sigma Aldrich), 0.1% SDS (Thermo Fisher Scientific), pH 8.3) at 80 V for 10 minutes followed by 120 V until the samples were separated satisfactorily. The PVDF membrane (Carl Roth) was activated in Methanol (Merck) for 5 minutes before layering it with blotting paper (Bio-Rad), foam pads and the membrane into the cassette and inserting it next to a cooling unit (Bio-Rad) into the wet blotting chamber (Bio-Rad) in pre-cooled transfer buffer (25 mM Tris (Carl Roth), 192 mM Glycine (Sigma Aldrich); pH 8.3). Blotting was carried out at 100 V for 45-60 minutes. Success of the transfer was validated by transfer of the pre-stained protein ladder (Thermo Fisher Scientific) before blocking the membrane in 5% milk (Merck) in TBS-T (20 mM Tris (Carl Roth), 150 mM

NaCl (Sigma Aldrich), pH 7.6; containing 0.1% TWEEN® 20 (Merck)). The experimentally pre-determined concentration for the primary antibody was added to 3% milk (Merck) in TBS-T and incubated on the membrane at 4 °C over-night. The membrane was washed with TBS-T at room temperature 3 times for 15 minutes each before adding the secondary antibody in a concentration of 1:500 to 0.5 % milk (Merck) in TBS-T and incubating it on the membrane for 1 hour at room temperature. After washing the membrane with TBS-T at room temperature 3 times for 15 minutes each, the membrane was developed using the Pierce™ ECL Western Blotting Substrate (Thermo Fisher Scientific) by mixing the peroxide solution and the luminol enhancer to equal parts and incubating it on the membrane for 30-60 seconds before imaging using the Azure C300 Chemiluminescent Western Blot Imager (Azure Biosystems).

Western Blot (NuPage™)

Gels for the Thermo Fisher Scientific XCell SureLock™ system were 4-12% Bis-Tris pre-cast protein gels (Thermo Fisher Scientific). The protein lysate was heated to 65 °C for 10 minutes in 1x NuPAGE™ LDS Sample Buffer (Thermo Fisher Scientific) containing 1x NuPAGE™ Sample Reducing Agent (Thermo Fisher Scientific) and centrifuged at 4 °C at full speed to remove debris before administration to the gel pockets. The electrophoresis was carried out in 1x MOPS SDS Running buffer (Thermo Fisher Scientific) at 120 V for 1-2 hours until the separation of the proteins was satisfactory. The blotting of the proteins and subsequent antibody-coupling as well as developing was carried out as described above, with the only change being the usage of 1x NuPAGE™ Transfer Buffer (Thermo Fisher Scientific) instead of the home-made buffer.

RNA immunoprecipitation (RIP)

For RIP from native cardiac tissue, 300 mg of the heart were lysed in 250 µl lysis buffer (10 mM Tris-HCl, 150 mM NaCl (Sigma Aldrich), 1% NP-40 (Thermo Fisher Scientific), 0.5% Triton-X100 (Sigma-Aldrich), 1 mM EDTA (Thermo Fisher Scientific), 5% Glycerol (Sigma Aldrich)) by Dounce homogenization. The lysate was cleared by centrifugation at full speed for 10 minutes at 4 °C.

For RIP from HL-1 cardiomyocytes $1*10^6$ cells were washed with DPBS and UV-crosslinked (Stratagene UV Stratalinker 1800) at 100 mJ. After scraping the cells into ice cold DPBS the pellet was resuspended in 250 µl lysis buffer (described above) and prepared similarly to cardiac tissue lysate.

For pulldown of the respective proteins 50 µl magnetic Pierce™ Protein A/G Magnetic Beads (Thermo Fisher Scientific) were used per reaction. To prepare the beads they were washed 3 times with binding buffer (10 mM Tris-HCl, 150 mM NaCl (Sigma Aldrich), 1% NP-40 (Thermo Fisher Scientific), 0.5% Triton-X100 (Sigma Aldrich), 1 mM EDTA (Thermo Fisher Scientific), 1x cOmplete™ Protease Inhibitors (Sigma Aldrich)) and incubated with 2.5 µg of the antibody diluted in 1 ml of the binding buffer at room temperature for 1 hour on a rotating wheel. As negative control a normal IgG of the respective species of the specific primary antibody was used in the same concentration. After washing the beads 3 times with binding buffer, the beads were resuspended in 900 µl binding buffer and added to 100 µl of the clear lysate. The pull-down was conducted at 4°C over-night on a rotating wheel. The beads were washed two times with binding buffer, one time with a high salt buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl (Sigma Aldrich), 10 mM EDTA (Thermo Fisher Scientific)) and LiCl buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl (Sigma Aldrich), 0.5% NP-40 (Thermo Fisher Scientific), 1% Sodium deoxycholate (Sigma Aldrich)) and two subsequent washes with binding buffer. For analysis of protein, the beads were resuspended in the respective loading buffer and used for Western Blot (described above). For analysis of RNA, the beads were resuspended in Protein degradation buffer (200 mM Tris-HCl pH 8.0, 300 mM NaCl (Sigma Aldrich), 25 mM EDTA (Thermo Fisher Scientific), 2% SDS (Thermo Fisher Scientific), 6.4 U Proteinase K (NEB)) and incubated at 50 °C for 30 min at 450 rpm. 1 ml of TRI reagent (Merck) were added to use the samples for Phenol/Chloroform extraction as described above.

RNAse H protection assay

10 Million cells were harvested in lysis buffer (50mM Tris-HCl pH 8, 150 mM NaCl (Sigma-Aldrich), 0.5% NP-40 (Thermo Fisher Scientific), 1x cOmplete™ Protease Inhibitors (Sigma-Aldrich)) and cleared by centrifugation at full speed for 10 minutes at 4 °C. The

lysate was adjusted to a total volume of 1 ml (final concentration: 60 mM NaCl (Sigma-Aldrich), 50 mM Tris-HCl pH 8, 75 mM KCl (Sigma-Aldrich), 3 mM MgCl$_2$ (Sigma-Aldrich), 10 mM DDT (Merck)). 100 µl of the adjusted lysate were mixed with 1 µl of the anti-sense oligonucleotide (100 µM) to be tested. As control 100 µl of the adjusted lysate were mixed with 1 µl of the corresponding sense oligonucleotide (100 µM). After 2 hours of incubation at 4 °C on a turning wheel, 2 U of RNAse H (Thermo Fisher Scientific) were added and incubated at 37 °C for 20 minutes. 1 ml TRI reagent (Merck) were added per sample and Phenol/Chloroform extraction and qRT-PCR analysis were conducted as described above, using qRT-PCR primers designed to flank the binding region of the oligonucleotide to be tested. Upon normalization to the sense control, working probes/oligonucleotides showed a normalized expression level of <1. The probes with values closest to 0 were ordered with methylation to avoid degradation and a destiobiotin label separated from the oligonucleotide by an iSP9 spacer to allow better detection (IDT).

RNA pulldown

For RNA pulldown, 75 µl magnetic Dynabeads™ MyOne™ Streptavidin C1 beads were used per sample. To prepare the beads they were washed in wash buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl (Sigma-Aldrich), 1 mM EDTA (Thermo Fisher Scientific), 0.05% NP-40 (Thermo Fisher Scientific)) 3 times and blocked with 0.2 mg/ml ytRNA (Thermo Fisher Scientific) and 0.2 mg/ml Glycogen (Roche) in wash buffer at 4 °C for 2 hours on a turning wheel. The beads were washed 3 times in wash buffer and kept on a turning wheel at 4 °C until needed.

The lysate was prepared by scraping the HL-1 cardiomyocytes into ice-cold DPBS after UV-crosslinking (Stratagene UV Stratalinker 1800) at 100 mJ. The cells were washed in DPBS and resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl (Sigma-Aldrich), 1 mM EDTA (Thermo Fisher Scientific), 1% NP-40 (Thermo Fisher Scientific), 1x cOmplete™ Protease Inhibitors (Sigma-Aldrich)). After 30-60 minutes of incubation on ice the lysate was homogenized by Dounce and debris was removed by centrifugation at full speed for 10 minutes at 4 °C. The supernatant was adjusted to a volume of 1 ml (final concentration: RNAse H buffer 1x, 150 mM NaCl (Sigma-Aldrich), 160 U RNAse inhibitors (Thermo Fisher Scientific), 1x protease inhibitors (Sigma-Aldrich)) and incubated with 1/3

of the pre-blocked beads for 2 hours at 4 °C on a turning wheel. After removing the beads, the supernatant was mixed with 200 pmol of the respective pulldown probes and incubated at 4 °C over-night on a turning wheel. Subsequently, the probes were coupled to the beads by adding the pre-blocked beads and incubating the mixture for 1 hour at 37 °C on a turning wheel. The beads were washed 6 times in wash buffer before resuspending them in Protein degradation buffer (200 mM Tris-HCl pH 8.0, 300 mM NaCl (Sigma Aldrich), 25 mM EDTA (Thermo Fisher Scientific), 2% SDS (Thermo Fisher Scientific), 6.4 U Proteinase K (NEB)). Following incubation at 50 °C for 30 min at 450 rpm, 1 ml of TRI reagent (Merck) was added to use the samples for Phenol/Chloroform extraction as described above.

Subcellular Fractionation

$2-5*10^6$ cells were used per replicate of fractionation. The cells were detached and washed in ice-cold DPBS once. The pellet was resuspended in cytoplasmic lysis buffer (50 mM Tris-HCl pH 8.0, 15 mM NaCl (Sigma-Aldrich), 1.5 mM $MgCl_2$ (Sigma-Aldrich), 0.5% NP-40 (Thermo Fisher Scientific)). After 20 minutes incubation on ice the samples were centrifuged at 1000xg for 4 minutes at 4 °C. The supernatant represented the cytoplasmic fraction and was transferred to a fresh tube. 1 ml of TRI reagent (Merck) were added per 200 µl lysate. The pellet was resuspended in fresh cytoplasmic lysis buffer and incubated on ice for 20 minutes. For separation of only cytoplasm and nucleus, the sample was centrifuged at 1000xg for 4 minutes at 4 °C, resuspended in fresh cytoplasmic lysis buffer and transferred to a fresh reaction tube. After centrifugation the nuclear pellet was resuspended in 1 ml TRI reagent (Merck).

For separation of nucleoplasm and chromatin, the nuclear pellet was resuspended in glycerol buffer (20 mM Tris-HCl pH 8.0, 75 mM NaCl (Sigma-Aldrich), 0.5 mM EDTA (Thermo Fisher Scientific), 0.85 mM DTT (Thermo Fisher Scientific), 50% Glycerol (Sigma-Aldrich)) and mixed with equal parts nuclear lysis buffer (20 mM HEPES (Sigma-Aldrich), 300 mM NaCl (Sigma-Aldrich), 0.5 mM EDTA (Thermo Fisher Scientific), 0.85 mM DTT (Thermo Fisher Scientific), 7.5 mM $MgCl_2$ (Sigma-Aldrich), 1 M Urea (Sigma-Aldrich), 2% NP-40 (Thermo Fisher Scientific), 2% Triton-X100 (Sigma-Aldrich), 0.1% SDS (Thermo Fisher Scientific)). The samples were incubated on ice for 20 minutes with

pulsed vortexing in between and were then centrifuged at 13500 rpm for 2 minutes at 4 °C. The supernatant, that represented the nucleoplasmic fraction, was transferred to a fresh tube and was mixed with 1 ml Trizol per 200 µl. The remaining pellet was washed with DPBS 6 times without resuspending it and was then transferred to a homogenizer tube with 1 ml TRI reagent (Sigma-Aldrich) for homogenization for 30 seconds on the medium intensity (Bertin Technologies).

Single molecular fluorescent *in situ* hybridization (smFISH)

Probes were designed using the Stellaris probe designer online tool (biosearchtech.com/support/tools/design-software/stellaris-probe-designer) and dissolved in TE buffer to a concentration of 12.5 µM. The neonatal cardiomyocytes were isolated as described above and seeded onto gelatin-fibronectin coated coverslips for incubation using the conditions specified for the required experiment. Cells incubated under hypoxic conditions were fixed in the hypoxic chamber to avoid oxygen shock. For fixation the cells were washed with PBS (Thermo Fisher Scientific) and submerged in fixation buffer (PBS containing 3.7% formaldehyde (Thermo Fisher Scientific) at room temperature for 10 minutes. The coverslips were washed with PBS twice and permeabilized in 70% EtOH for at least one hour at 4 °C. After 5 minutes washing with Wash Buffer A (1x Stellaris® RNA FISH Wash Buffer A (BioCat) in nuclease-free water containing 10% deionized formamide (Sigma-Aldrich)) the coverslips were hybridized with the probes (125 nM) in hybridization buffer (Stellaris® RNA FISH Hybridization Buffer (BioCat) containing 10% deionized formamide) at 37 °C for at least 4 hours in the dark. After washing the coverslips with Wash Buffer A at 37 °C for 30 minutes in the dark, 5 ng/ml DAPI nuclear stain (Carl Roth) was added to Wash Buffer A and subsequently incubated on the coverslips for 30 minutes at 37 °C in the dark. The coverslips were equilibrated in glucose buffer (2x SSC, 0.4% glucose, 20 mM Tris-HCl pH 8.0) and subsequently incubated in glucose buffer supplemented with 1:100 glucose oxidase and catalase. The coverslips were washed with Wash Buffer B (1x Stellaris® RNA FISH Wash Buffer B in nuclease-free water) at room temperature for 5 minutes. For mounting a drop of ProLong™ Gold Antifade Mountant (Life Technologies) was added to a slide and the coverslip carefully added on top with the cells facing downwards. Images were acquired

using a Nikon ECLIPSE Ti2 widefield microscope with a Nikon Plan Apo λ 100x/1.45-numerical aperture oil objective lens and a Nikon DS-Qi2 camera. Z-stacks of 200 nm step size were acquired, deconvoluted (Type Richardson-Lucy) and used to create maximum intensity projections by ImageJ/Fiji software. Subsequently, the smFISH signals were quantified manually using ImageJ/Fiji.

Rapid amplification of cDNA ends (RACE)

For determination of RNA 3'-ends, the SMARTer® RACE 5'/3' Kit (Takara) was used following the manufacturer's guidelines. In brief, total RNA was isolated from E13.5 hearts and the concentration was measured. 1 µg of the RNA was mixed with 1 µl 3'-CDS Primer A in water to a total volume of 12 µl and incubated at 72 °C for 3 minutes followed by 42 °C for 2 minutes. The denatured RNA was mixed with 4 µl 5X First-Strand Buffer, 0.5 µl -100 mM DTT, 1 µl 20 mM dNTPs, 0.5 µl 40U/µl RNase Inhibitors and 2 µl SMARTScribe Reverse Transcriptase. After 90 minutes of incubation at 42 °C followed by 10 minutes of 72°C and subsequent dilution in Tricine-EDTA buffer. Obtained RACE samples were used for PCR amplification using the universal primer contained in the kit and the gene specific 3' primers with either the Expand™ Long Template PCR-System (Roche) or Phusion® Polymerase (NEB) with the following primers. Fsd2_RACE_pA1_GSP: AGGTTCCTACTGCTAGCCGA, Fsd2_RACE_pA2_GSP: GCCCAGGGAGACACTCTTCT, Fsd2_RACE_pA3_GSP: CTGTCTGACCTCCAAGCACC

Subsequently, samples were analyzed by Sanger-sequencing.

Circular dichroism (CD) spectra and melting curve analysis

In vitro triplex formation assay using CD spectra and melting curve analysis was a co-operation with Jasleen Kaur and Nina M. Krause (Center for Biomolecular Magnetic Resonance, Institute for Organic Chemistry and Chemical Biology, Goethe University).

The CD spectra were acquired using a Jasco J-810 spectropolarimeter, recording from 210 to 320 nm at 25 °C using a 1 cm path length quartz cuvette. The recordings were conducted on 8 µM samples of either DNA duplex or RNA:dsDNA triplex in a buffer containing 20 mM LiOAc and 10 mM $MgCl_2$ at pH 5.5 using an excess of 5 ng RNA to

hybridize RNA:dsDNA triplex. Hybridization was conducted using complimentary DNA strands and incubating them at 95 °C for 5 minutes followed by slowly cooling them down to room temperature. For additional formation of the RNA:dsDNA triplex, the RNA was added to the duplex followed by 1h incubation at 60 °C and subsequent cooling to room temperature. 8 scans were acquired and smoothed using Savitzky-Golay filters. Conversion of observed ellipticities recorded in millidegree (mdeg) to molar ellipticity was done using the following formula:

$$[\theta] = deg*cm^2*dmol^{-1}$$

To generate melting curves, samples were exposed to gradually increasing temperatures (1 °C/min) in a range of 5-95 °C at constant wavelength. Conversion of melting temperature data to normalized ellipticity was conducted and evaluated using SigmaPlot12.5 software applying the following formula:

$$f \frac{a}{\left(1 + \exp\left(-a\frac{(x-x0)}{b}\right)\right)} + \frac{c}{\left(1 + \exp\left(-a\frac{(x-x2)}{d}\right)\right)}$$

Bioinformatic analysis

Statistical analysis

Statistical analysis was conducted using integrated features of GraphPad Prism 9.5.0 version for Windows, GraphPad Software, San Diego, California USA, (www.graphpad.com). Used tests and parameters were given in figure descriptions and were conducted using default parameters unless stated differently.

Analysis of *Swhtr* Heart Slice RNA-sequencing

Bioinformatic analysis of raw RNA-sequencing data obtained from Novogene (25 million paired-end reads of 150 bp) was conducted by Dr. Tamer Ali as published in Rogala et al., 2022 and has been deposited to GEO (GSE200380). In brief, STAR aligner was used for read mapping, applying –outSAMtype BAM SortedByCoordinate (Dobin et al., 2013). Quantification of reads and alignment was conducted with GenomicAlignments Bioconductor package (Lawrence et al., 2013). After filtering processes, DESeq2 (cutoffs: Padj < 0.05; log2FC = 0.58) was used for analysis of gene deregulation (Love et al.,

2014). Using clusterProfiler and enrichplot Bioconductor packages, GO-terms of commonly deregulation biological processes could be obtained (Wu et al., 2021).

Analysis of NKX2-5 occupation on *Swhtr* dependent genes

NKX2-5 occupation on *Swhtr* dependent genes was analyzed by Dr. Tamer Ali as published in Rogala et al., 2022. In brief, a publicly available NKX2-5 ChIP-seq dataset from adult heart apex tissue (GSM3518668) was downloaded, aligned to mm10 by Bowtie2 (Langmead and Salzberg, 2012), converted to sorted bam files using Samtools (Danecek et al., 2021) and filtered and organized using bedtools (Quinlan and Hall, 2010). Peaks that were called using MACS3 peak caller (Zhang et al., 2008) were overlapped with *Swhtr* dependent genes using extended gene co-ordinates (+/- 1000 bp) using the IRanges Bioconductor package (Zhu, 2013).

General Introduction

The non-coding genome

Diversity and features of non-coding RNAs (ncRNAs)

The genome encodes for genes that serve as the template for transcription into RNAs. For a long time, it was believed that the purpose of these RNAs is to be translated into proteins that execute all functions required for survival of an organism. Later it was discovered that beside protein-coding genes, also non-protein-coding genes exist, that exert a function without being translated into proteins. These non-coding genes can be divided into different subgroups. The most abundant ones are a group of non-coding RNAs (ncRNAs) that are referred to as housekeeping ncRNAs. These contain ribosomal RNAs and transfer RNAs which both are indispensable for the translational process of proteins, demonstrating the importance of the non-coding genome. The remaining ncRNAs are categorized by their length into small ncRNAs of less than 200 nucleotides (nts) and long ncRNAs (lncRNAs) that exceed 200 nt. The most intensively studied group of small ncRNAs are microRNAs (miRNAs). They are single-stranded RNA molecules of around 22 nts and are frequently discussed in the context of RNA interference (RNAi), post-translational regulation and thus translation efficiency of proteins (Filipowicz et al., 2008). Small nuclear RNAs (snRNAs) or small nucleolar RNAs (snoRNA) are other examples of small ncRNAs of 60-150 nts that can be found predominantly within the nucleus and serve as components of the spliceosome or nuclear organelles. While both major classes of ncRNAs, small and long ncRNAs, are very divers in terms of discussed functions, due to categorization by an arbitrary length of 200 nts or more, lncRNAs are the most incoherent and versatile group of ncRNAs. They have been implicated in many different biological processes like development, organogenesis but also homeostasis of organ function and stress reaction and disease.

LncRNAs and their mechanisms

Within the human transcriptome, lncRNA genes are almost as abundant as protein-coding genes and yield half the number of transcripts due to alternative splicing as protein-coding genes (Lagarde et al., 2017; Mele et al., 2017). Commonly, lncRNAs are transcribed at

lower levels compared to mRNAs. The transcription of lncRNAs and mRNAs is mostly carried out by the RNA polymerase 2 multiprotein complex and they undergo the same transcriptional and post-transcriptional processing with only minor to no difference in length, whereas lncRNAs have on average fewer and longer exons with less efficient splicing (Lagarde et al., 2017; Mele et al., 2017). Lower level of RNA expression in lncRNAs is associated with an increased presence of repressive H3K9me3 marks at active lncRNA promoters compared to those of mRNAs, and lower frequency of transcription factor binding sites (Mele et al., 2017). Together with many lncRNAs being transcribed from enhancer elements that can regulate transcription levels of RNAs in a tissue-specific manner, that results in lncRNAs being the most inherently tissue-specific group of RNAs (Mattioli et al., 2019).
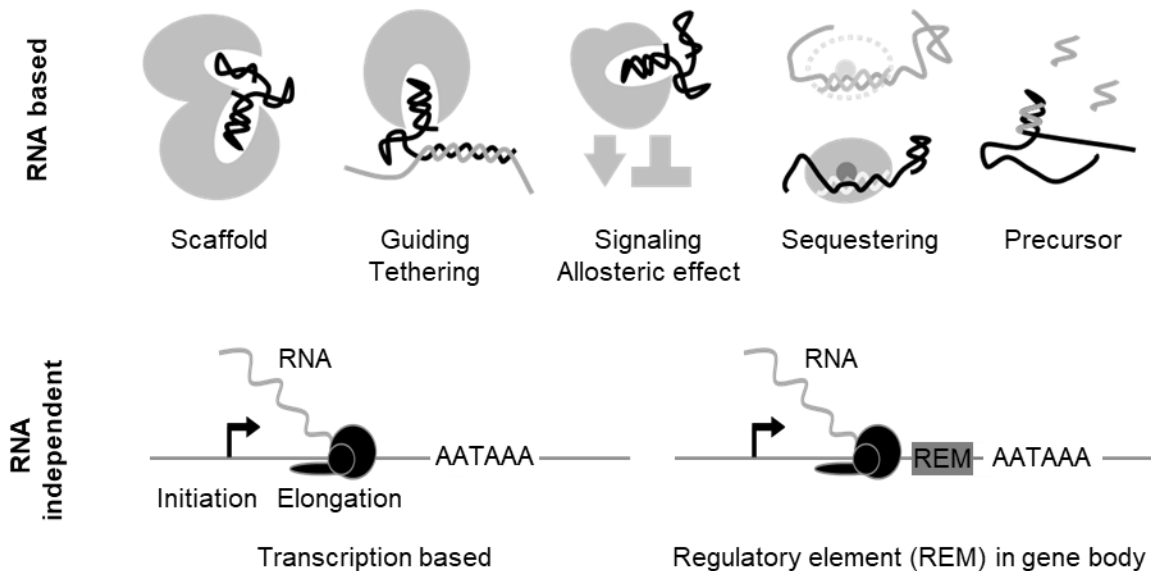
Fig. 1: Diversity of lncRNA mechanisms. LncRNAs can act as an RNA molecule themselves (top) or via their transcription or gene bodies (bottom). AATAAA=transcriptional termination signal. This figure was generated by modifying a figure taken from Grote and Ali; 2019 and a figure that was a kind gift from Prof. Dr. Petr Svoboda (Czech Academy of Sciences, Institute of Molecular Genetics).

The mechanisms of action of lncRNAs are as diverse as their length. The most important differentiation is, whether the transcribed RNA serves a function or whether the act of transcription itself exerts a regulatory effect with the produced RNA being a mere byproduct (Fig. 1).

The RNA itself can for example scaffold RNP complexes and by such even entire cellular compartments, such as *NEAT1*, which is essential for the formation of nuclear

paraspeckles and by such influencing processing of miRNAs (Clemson et al., 2009; Jiang et al., 2017). Another possible mechanism is the tethering or guiding of proteins, like the essential lncRNA *Fendrr* can tether histone modification complexes to the promoters of its target genes via an RNA:dsDNA triplex formation domain (Grote and Herrmann, 2013). Allosteric effects of RNAs have been most commonly discussed in bacteria, but also in plants and fungi. Riboswitches for example can interact with various co-factors such as co-enzymes, nucleotide derivatives, amino acids or signaling molecules thus inducing conformational changes that lead to alteration of transcription or translation of coding RNAs (McCown et al., 2017). Although specific lncRNA switches have not been found in mammals yet, the possibility of their existence is not excluded (Venkata Subbaiah et al., 2019). In the context of sequestering functions, the lncRNA *NORAD* is a well-studied example. Upon induction of DNA damage *NORAD* becomes activated, sequestering PUMILIO proteins and thus promoting DNA stability by inhibition of PUMILIO-associated repression of mitosis-associated genes and DNA repair and replication factors (Lee et al., 2016; Tichon et al., 2016). Some lncRNAs, however, serve as precursors, for example for miRNAs. One such example is lncRNA *H19*. Despite being one of the most abundant and conserved lncRNA transcripts in embryonic development, no function could be determined for a long time until it was recognized that the lncRNA serves as precursor for miRNA *miR-675* that is embedded in the first exon and becomes processed dynamically to regulate growth of the placenta before birth (Cai and Cullen, 2007; Keniry et al., 2012).

In contrast, the produced RNA can also be irrelevant for the function of the lncRNA gene itself. The lncRNA *Lockd*, which is located 4 kb downstream of the protein-coding gene *Cdkn1b*, has no effect on its neighboring gene, as evident by no detectable deregulation after insertion of a transcriptional stop signal close to the TSS. Deletion of the entire genomic locus encoding for *Lockd* and its regulatory elements that physically associate with the *Cdkn1b* promoter in chromosome conformation capture data leads to strong reduction of *Cdkn1b* (Paralkar et al., 2016). Hence, *Lockd* lncRNA and its transcription are dispensable for its regulatory effect on *Cdkn1b*. *XIST* is one of the most extensively studied lncRNAs and is responsible for inactivation of the second X-chromosome in female individuals via an RNA-based mechanism (Brockdorff et al., 1992; Brown et al.,

1992). Its regulation, on the other hand, is at least in parts controlled by the 140 kb upstream located *Ftx* lncRNA locus. Loss of the RNA does not show an influence on *XIST*, while loss of transcriptional activity by deletion of the promoter or CRISPRi leads to comparable loss of *XIST* expression, demonstrating that *Ftx* relies on the act of transcription for regulation of its target gene with a dispensable RNA by-product (Furlan et al., 2018).

An example for a lncRNA locus with dual function, is the *Haunt* locus, downstream of the *HOXA1* locus of the *HOXA* gene cluster. While the *Haunt* RNA has an inhibitory effect on the *HOXA* gene cluster, the DNA locus seems to contain regulatory elements to facilitate *HOXA* induction in embryonic stem cells (ESCs) during retinoic acid-induced differentiation. These effects are only observed during ESC differentiation and are absent during ESC maintenance and self-renewal but also during LIF-withdrawal induced differentiation (Yin et al., 2015). This demonstrates that lncRNA loci can exert multiple functions and use different mechanisms to perform several and sometimes even opposing functions. Furthermore, it elucidates a possible requirement of secondary and very specific co-factors or signaling pathways.

Conservation of LncRNAs

LncRNAs are typically regarded to be poorly conserved between species at least in regard to their sequence. This low sequence conservation does not always result in separate functions. Some lncRNAs, despite poor sequence similarity to their orthologous counterparts share secondary structure and function (Bryzghalov et al., 2021; Johnsson et al., 2014). Furthermore, protein-coding genes that are associated with lncRNAs in *cis* in humans often have orthologous protein-coding genes within the mouse genome that are also associated with lncRNAs in *cis*. Hence, conservation of organization of protein-coding genes with anti-sense or divergent non-coding genes is higher than the low sequence similarity leads to believe. Often these relationships are not obvious as the genes have been discovered independently of each other and follow varying nomenclature.

However, organizational similarity does not always imply shared functionality. Some lncRNAs exert the same function in many different organisms, such as the most well-

studied lncRNA *XIST*, which is responsible for inactivation of the second X-chromosome across species of female mammals (Chaumeil et al., 2011; Penny et al., 1996). Other lncRNAs do not have orthologues, such as human specific, obesity-associated *LINC01018* that interacts with the RNA-binding protein ELAV-like protein 1 (HuR) in order to regulate the transcriptomic network of fatty acid oxidation in the liver (Ruan et al., 2020). With no validation, opposing or at least varying functions cannot be excluded, either. In conclusion, whether or not a lncRNA has conserved function and orthologues between species has to be evaluated individually.

Aim of this work

Gene regulatory networks are the foundation of a healthy, functioning organism and thus life itself. All processes rely on transcriptional regulation of genes that influence the fate or behavior of cells. Highly plastic events such as embryogenesis are dependent on gene regulatory networks to control cell fate determination and lineage commitment but also homeostasis of an organism and reaction to external stimuli rely on the transcriptomic environment and adaptational changes thereof. LncRNAs are now recognized for taking part in fine-tuning such gene programs and have been an emerging target of interest in basic science but are now also recognized for their disease relevance.

This work aims to demonstrate the impact lncRNAs can have on life and maintenance of the organism particularly in regard to function of the cardiopulmonary system. Three lncRNAs will be studied in more detail in order to elucidate their function. In particularly their implication in the cardiopulmonary system in terms of homeostasis and disease stimuli will be studied. More specific introductions and experimental aims regarding the specific context of the respective lncRNA will be given within each chapter.

Sweetheart lncRNA regulates compensatory cardiac hypertrophy after myocardial injury

Introduction

The essential cardiac transcription factor coding gene *NK2 Homeobox 5 (Nkx2-5)*

Homeobox proteins are transcription factor proteins that contain a homeobox domain that is responsible for the binding of DNA. Typically, these transcription factors can regulate gene programs by direct binding to gene promoters and their activation or repression. The *NK2 Homeobox 5* (*Nkx2-5*) encodes for a well-studied cardiac core transcription factor that is crucial for embryonic viability and development of the heart (Tanaka et al., 1999). Evolutionary the gene is well conserved between mammals, even sharing homology with the Drosophila melanogaster cardiac homeobox-containing gene *tinman* that is essential for heart development (Bodmer, 1993). Complete ablation of *Nkx2-5* in mice leads to early embryonic lethality but also shows that *Nkx2-5* is not required for initial looping of the heart, whereas it does regulate its own locus but also expression of other cardiac transcription factors and by such, cardiomyocyte lineage commitment (Tanaka et al., 1999). For this purpose, it acts in concert with co-factors such as GATA-4 or TBX5 to regulate cardiac maturation associated gene programs (Akazawa and Komuro, 2005). As *Nkx2-5* deficient mice die before extra-cardiac organs can be developed, *Nkx2-5* has only been characterized scarcely in an adult context despite its abundance in adult heart tissue. It is, however, known that *NKX2-5* mutations are frequently found in patients suffering from cardiac diseases, such as dilated cardiomyopathy (Sveinbjornsson et al., 2018).

Nkx2-5's divergently expressed lncRNA *Sweetheart RNA* (*Swhtr*)

While *Nkx2-5* has been studied well, at least in an embryonic context, the lncRNAs contained within its genomic vicinity have been studied to a lesser extent. *Sweetheart RNA* (*Swhtr*) was via 5' and 3' RACE PCR determined to be a 9809 bp long lncRNA divergently expressed from *Nkx2-5* that partially overlaps with *IRENE-div* while not

48

sharing a transcriptional start site with the eRNA (Rogala et al., 2022; Salamon et al., 2020). The *Swhtr* transcriptional start contains a GATA-4 bound previously described element (Lien et al., 1999) that most likely serves as the lncRNA promoter that *Swhtr* shares with an anti-sense transcript that was termed *IRENE-SS* (Fig. 2). The eRNAs *IRENE-SS* and *IRENE-div* have been described to have opposing functions on *Nkx2-5* expression, where *IRENE-SS* knockdown decreased *Nkx2-5* levels and *IRENE-div* knockdown increased *Nkx2-5* levels in HL-1 cardiomyocytes *in vitro* (Salamon et al., 2020).
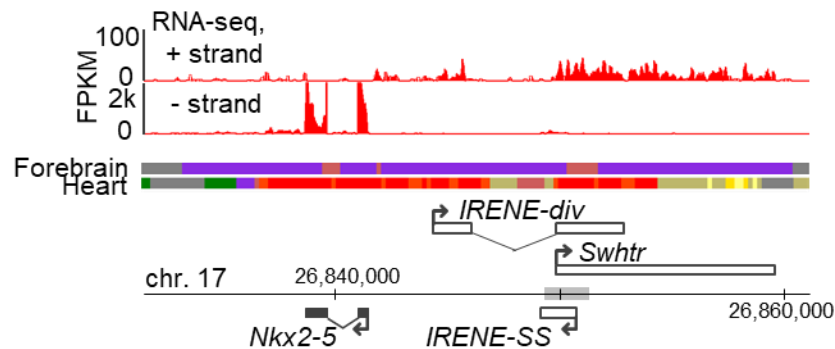

Fig. 2: *Swhtr* is divergently expressed from *Nkx2-5* in the heart. Strand specific RNA-sequencing from murine E9.5 hearts with chromatin state tracks from forebrain where the locus is off and heart where the locus is on (ENCODE in ChromHMM colors (red = TSS, yellow/orange = enhancers, light grey = poised enhancer, purple = heterochromatin, dark grey = quiescent)). mRNAs (dark grey) and lncRNAs (no fill) present in the locus are depicted with corresponding *mm10* co-ordinates (adapted from Rogala et al., 2022)

Expression of *Swhtr* started from early embryonic development (E8.25) and depicted expression in the early inflow tract but later also in the atria and ventricles of the heart while not being present in the epicardium. The lncRNA was found to be chromatin associated and most likely resides at its own locus of production in primary murine cardiomyocytes. Genetic ablation of *Swhtr* in mice by insertion of a premature transcriptional stop signal led to complete loss of *Swhtr* lncRNA following the stop signal with no effect on *Nkx2-5* expression and no phenotype in the mutant embryos (Fig. 3). Whether *Swhtr* is also present abundantly in adult murine hearts and whether it might serve a function in that context has not been studied. Hence, this work focused on identifying a role of *Swhtr* in heart maintenance and disease settings.
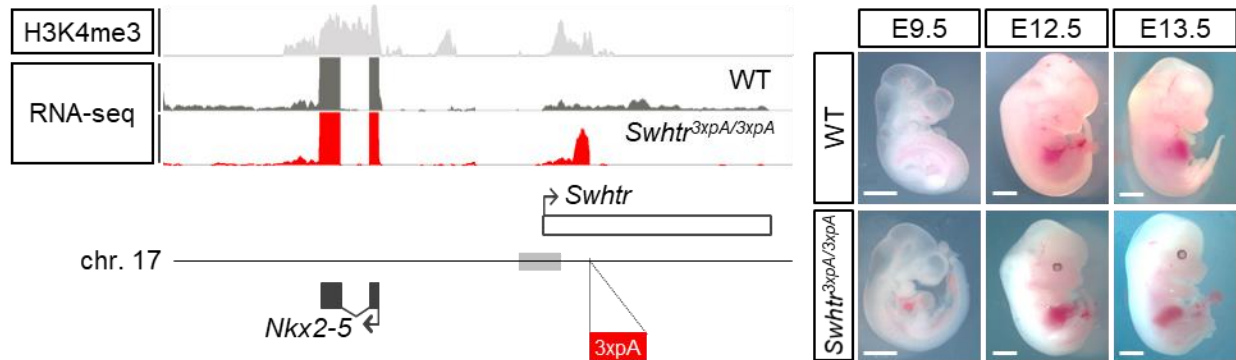
Fig. 3: The transcriptional stop mutant of Swhtr shows no overt phenotype. ChIP-sequencing (H3K4me3) and RNA-sequencing from WT and *Swhtr³ˣᵖᴬ/³ˣᵖᴬ* mice with schematic of the mutation (left) and pictures of WT and mutant embryos of the indicated embryonic stage (right). The white line represents 1 mm. (adapted from Rogala et al., 2022)

Cardiovascular disease (CVD)

Cardiovascular disease (CVD) is an umbrella term describing disorders of the heart and vessels. This work focuses on diseases directly affecting the heart. These can be distinguished into CVDs of the blood vessels that supply the heart itself, such as coronary heart disease, but also defects of the heart muscle and malformations of heart morphology often from birth, such as congenital heart disease. As previously mentioned, mutations in *NKX2-5* are often associated with heart malfunctions such as congenital heart disease or dilated cardiomyopathy (Costa et al., 2013; Sveinbjornsson et al., 2018). CVD is the leading cause of mortality worldwide, according to the World Health Organization (WHO). The major behavioral risk factors increasing the chance of developing cardiac disease are poor diet and obesity, physical inactivity, use of tobacco and exceeding use of alcohol, making it even more common issue in the developed world. Especially, acute manifestations of CVD such as acute myocardial infarctions (AMI) and stroke made up for 85% CVD related deaths globally in 2019. AMI is defined by restricted blood-supply of the cardiac muscle by a physical blockage affecting one or more coronary arteries, leading to damage of the tissue and subsequent scar formation. With appropriate medical care 28-day survival rates after AMI range between 93 – 96% showing no major differences in long term survival between males and females with respect to age at the time of the incident (Gardarsdottir et al., 2022; Rosengren et al., 2001). This high survival rate, however, correlates with a high subsequent occurrence of heart failure and substantial left ventricular damage leading to increased risk of secondary CVD incidents

with decreased prognosis for long-term survival (Guidry et al., 1999; Spencer et al., 2002; Steg et al., 2004). Hence, understanding of the compensatory mechanisms following AMI is crucial.

Aim of the study

Many lncRNA genes are not essential for life and do not show overt phenotype when being depleted (George et al., 2019; Goudarzi et al., 2019; Sauvageau et al., 2013). It is possible, that while the lncRNA is not required for embryonic development, it might serve a function in the adult. An example for such a lncRNA can be found with the conserved and well-studied lncRNA *Malat1* that is dispensable for embryonic development and viability, but regulates its *cis*-located genes in adult mouse tissue (Zhang et al., 2012). Furthermore, *Malat1* has been implicated in various disease settings, especially cancer (Arun et al., 2020). Often it is inferred that lncRNAs that are dispensable for development and life might be involved in regulation of stress response or might become important in disease context, as for example the lncRNA *Mhrt* that regulates BRG1 chromatin binding after TAC induced pathological stress and as such has an effect on stress-induced heart failure (Han et al., 2014).

This study aims to identify a function of *Swhtr* in maintenance of heart function and cardiac stress settings by *in vitro* and *in vivo* methods using the mouse as a model organism. As lncRNAs are often implied in *cis*-regulatory processes, an involvement of *Nkx2-5* in the process will be investigated in greater detail.

Results

Nkx2-5 and *Swhtr* are abundant in postnatal murine hearts

Sweetheart RNA (*Swhtr*) is a lncRNA divergently expressed from the cardiac core transcription factor coding gene *Nkx2-5*. Characterization showed expression specifically in the heart of developing embryos from day E8.25 onwards. Upon dissection of the hearts of embryos of day 10.5 and 14.5 after fertilization as well as those of neonatal mice, 1-week-old mice and 8-week-old adult mice, RNA levels were analyzed by qRT-PCR. This demonstrated that *Nkx2-5* as well as *Swhtr* are abundantly expressed in adult heart tissue. Additionally, it was found that the expression pattern of the lncRNA changes comparably to the mRNA (Fig. 4) suggesting a synergistic role of the two genes.

In embryos, *Nkx2-5* expression has been detected in extracardiac tissue (Kasahara et al., 1998), however, depletion of *Nkx2-5* in mice led to embryonic lethality before extracardiac tissues can be developed, thus only *Nkx2-5* function in the embryonic heart has been characterized in greater detail (Tanaka et al., 1999). To validate the heart specificity of *Swhtr* that was found in embryos (Rogala et al., 2022) and examine the RNA level of *Nkx2-5*, several tissues of adult wild type (WT) mice were collected for qRT-PCR analysis. While *Nkx2-5* RNA was found in the heart and brain, *Swhtr* was found exclusively in the heart, pointing towards the possibility of a heart specific interaction of the two genes (Fig. 4).
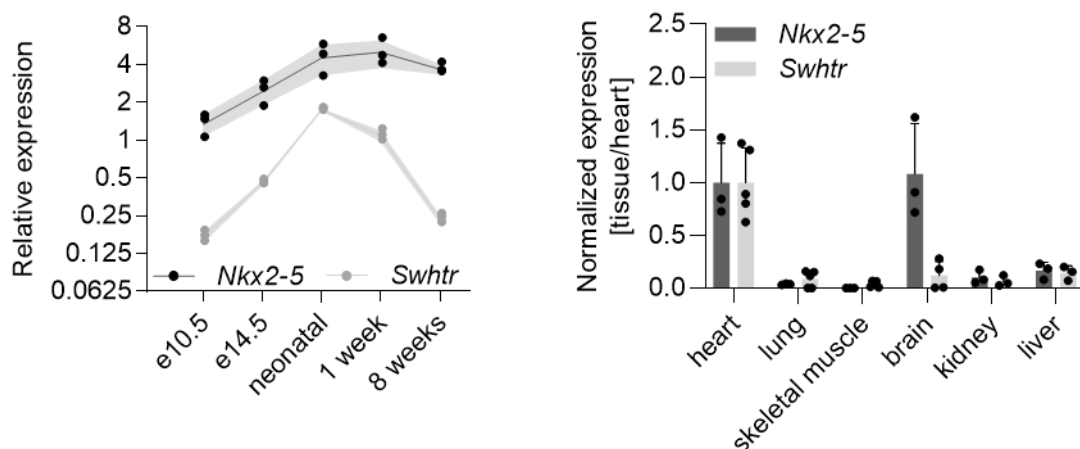

Fig. 4: *Swhtr* and *Nkx2-5* expression pattern and tissue specificity in adult murine hearts. qRT-PCR expression pattern analysis of *Swhtr* and *Nkx2-5* in the hearts of E10.5 embryos to adult mice (left; n=3) and *Swhtr* RNA levels in tissues dissected from adult mice (right; n=3-5; dots represent n number).

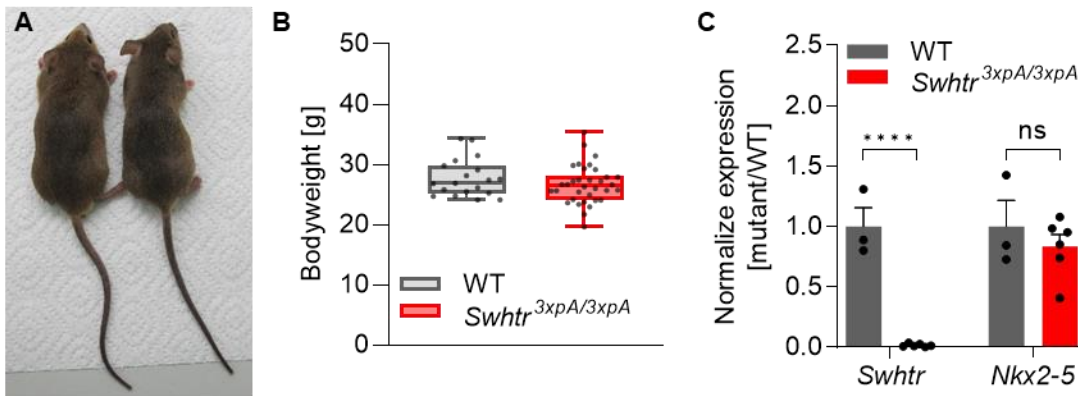*Swhtr* is dispensable for life and heart maintenance under standard conditions

Fig. 5: Adult *Swhtr^{3xpA/3xpA}* mice display no overt phenotype or deregulation of *Nkx2-5* in the heart. (**A**) 8-week-old adult *Swhtr^{3xpA/3xpA}* founder mice. (**B**) Bodyweight of *Swhtr^{3xpA/3xpA}* backcrossed eight times (C57Bl6J) mice (WT n=19; *Swhtr^{3xpA/3xpA}* n=34). (**C**) qRT-PCR analysis of *Swhtr* and *Nkx2-5* normalized RNA level in hearts isolated from 10-week-old adult mice (WT n=3; *Swhtr^{3xpA/3xpA}* n=5) normalized to the WT. Statistical significance was assessed by Two-way ANOVA. ns = not significant, **** < 0.0001

To generate a null mutant of *Swhtr* a strong transcriptional stop signal was inserted downstream of the *Swhtr* TSS to terminate transcription avoiding interference with the activity of the GATA4 bound enhancer (Lien et al., 1999) contained in the *Swhtr* TSS. The investigation of loss of *Swhtr* in a developmental context did not reveal any defect dependent on its presence and was not sufficient to identify the function of the lncRNA (Rogala et al., 2022). In the following, the role of *Swhtr* in the adult animals was characterized. The *Swhtr^{3xpA/3xpA}* transcriptional stop mutants (*Swhtr Null*) were born following standard mendelian ratios and without any overt differences to their WT littermates in terms of appearance or bodyweight (Fig. 5). In embryo hearts it could be shown previously that *Nkx2-5* RNA level remains unaffected by loss of *Swhtr* (Rogala et al., 2022). To examine whether this holds true in an adult context, hearts from adult WT and *Swhtr Null* animals were isolated and used for qRT-PCR analysis. Hereby, it could be confirmed that genetic ablation of *Swhtr* does not have an effect on *Nkx2-5* RNA level under standard conditions in adult hearts (Fig. 5).

The adult heart was inspected in more detail by using echocardiography measurements as the method would allow to detect minor defects within the heart without manifestation of an extrinsically visible phenotype in a non-invasive manner. The function of the hearts of 8-week-old adult WT and *Swhtr Null* mice was assessed by measuring standard

parameters such as heart rate and ejection fraction. Furthermore, the phenotype of the heart was evaluated by measurement of LV volume, mass and thickness of the IVS. Neither of the parameters varied significantly between the two genotypes, indicating normal function of the heart independent of *Swhtr* presence (Fig. 6, more shown in Appendix Fig. 4). In conclusion, it could be shown that *Swhtr* expression is not required for viability and heart maintenance under standard conditions.
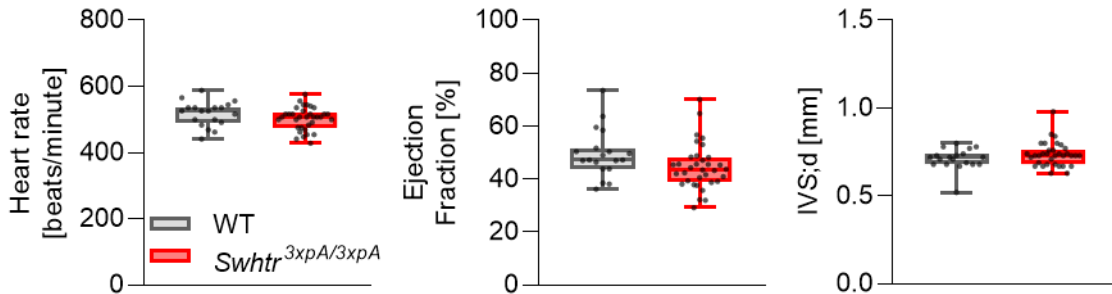

Fig. 6: Selected heart parameters from adult WT and *Swhtr*$^{3xpA/3xpA}$ mice. Depicted are heart rate, ejection fraction and IVS thickness assessed by echocardiography of 8-week-old adult mice (WT n=19; *Swhtr*$^{3xpA/3xpA}$ n=34). The figure legend applies to all graphs. Statistical significance was tested by Two-way ANOVA. No statistically significant differences could be detected.

Compensatory response after AMI requires *Swhtr*

Many lncRNAs are not required for life and their depletion depicts mild phenotypes or none at all (Sauvageau et al., 2013). Often those lncRNAs are theorized to serve a purpose under stress conditions, however, experiments proving those hypotheses are not always undertaken. In the murine cardiac system, stress can be induced by various methods ranging from mild stress models such as exercise to more severe stress models like induction of pressure overload by aortic constriction. The most drastic stress model is the induction of acute myocardial infarction (AMI) that is achieved by ligation of the left anterior descending artery (LAD ligation), which restricts blood supply to the left ventricle (Kolk et al., 2009).

In order to investigate whether *Swhtr* could be involved in a cardiac stress reaction, LAD ligation as the most severe stress model was chosen, due to the absence of even a mild phenotype in the heart under standard laboratory conditions. For this purpose, 8-week-old mice underwent LAD ligation surgery. Heart parameters were assessed by echocardiography before the surgery, as well as 7 and 14 days following the surgery (Fig.

7). Although AMI poses a threat to life, survival and recovery is possible and even common, which was also reflected by survival for at least 2 weeks after LAD ligation of the majority of WT mice. In contrast, the ratio of AMI induced mortality was significantly increased in the *Swhtr Null* mice with comparable size of AMI between the groups (Fig. 7, Appendix Fig. 5).
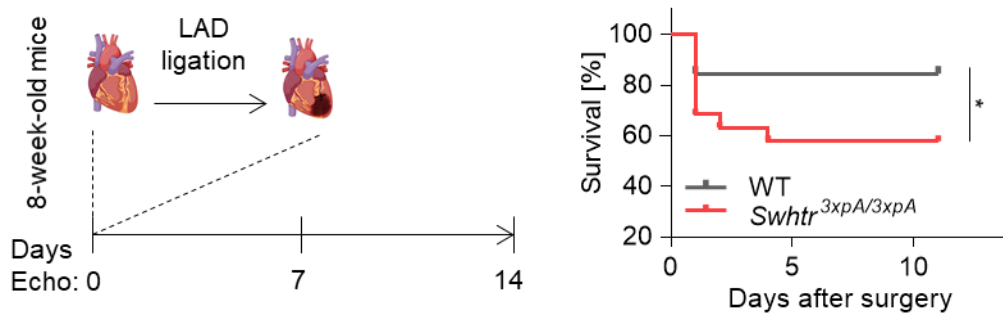

Fig. 7: LAD ligation induced AMI in adult WT and *Swhtr$^{3xpA/3xpA}$* mice. Schematic of the analysis setup for echocardiography and LAD ligation in adult mice (left). Reduced survival of *Swhtr$^{3xpA/3xpA}$* mice compared to WT mice after LAD ligation (right; WT n=13; *Swhtr$^{3xpA/3xpA}$* n=19). Statistical significance was tested by Kaplan-Meier Simple Survival Analysis. * < 0.05

Analysis of heart specific parameters revealed no differences between the genotypes in most parameters (Fig. 8; more shown in Appendix Fig. 5), except for a slight decrease in ejection fraction, that did not show any significance. One outstanding difference that could be observed was the size of the IVS. AMI leads to formation of scar tissue in the left ventricle resulting in a loss of flexibility. To preserve the pumping function of the heart for maintained supply of blood to the body, the remaining tissue is known to grow in size (Rubin et al., 1983). Here, one week after AMI, the size of the IVS decreased comparably in both genotypes, however, whereas it recovered in the WT, the hearts lacking *Swhtr* fail to do so (Fig. 8).
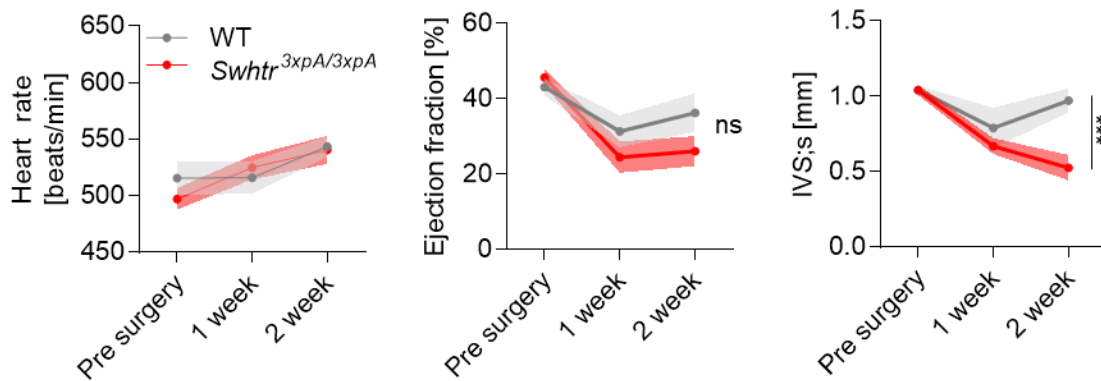
Fig. 8: Heart function after AMI. Selected heart specific parameters in mice (n=9 animals per genotype) before and after (1 and 2 weeks) LAD ligation obtained by echocardiography analysis. The figure legend applies to all three graphs. Statistical significance was tested by Two-way ANOVA. ns = not significant, *** < 0.001

Swhtr RNA is required for compensatory response after AMI

In order to determine whether the observed function is tied to the transcription of the *Swhtr* RNA or a function of the RNA itself, a rescue mouse line was generated. For that purpose, a bacterial artificial chromosome containing the 190 kb genomic region surrounding the *Nkx2-5* locus was modified to contain a yellow fluorescent *H2Bvenus* construct instead of *Nkx2-5* ($P_{Nkx2-5}H2Bvenus$) to avoid adding a third copy of the essential gene. This construct was randomly inserted into the genome of mice that were selected for insertion of only one copy of the BAC and later cross-bred into the *Swhtr Null* (*Swhtr³ˣᵖᴬ/³ˣᵖᴬ*) background, resulting in a mouse line with blocked *Swhtr* transcription from the endogenous locus and re-expression of *Swhtr* from an exogenous locus. The expression pattern of the transgene mimicked the WT situation and could be verified by monitoring yellow fluorescence that was present as a consequence of *Nkx2-5* promoter activity. These *Swhtr³ˣᵖᴬ/³ˣᵖᴬ;ᵗᵍ* mice are referred to as *Swhtr Rescue* mice in the following.
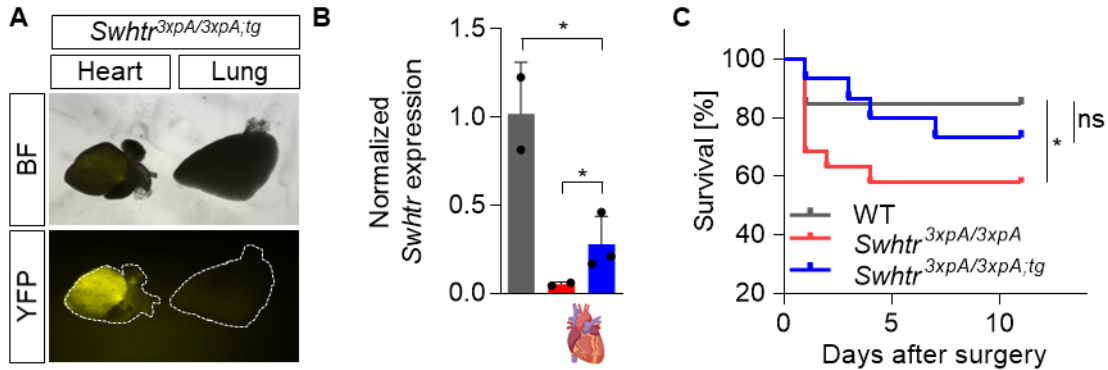
Fig. 9: Rescue of reduced survival after AMI by re-expression of *Swhtr* from an exogenous locus.
(**A**) Verification of presence and activity of the rescue transgene by heart specific expression of H2BVENUS. BF = Brightfield, YFP = Yellow fluorescent protein (VENUS) (**B**) qRT-PCR analysis of *Swhtr* expression in hearts isolated from 9-week-old adult WT (grey), *Swhtr³ˣᵖᴬ/³ˣᵖᴬ* (red) and *Swhtr³ˣᵖᴬ/³ˣᵖᴬ;tg* (blue) mice normalized to the WT expression. The figure legend is the same as in **C**. Statistical significance tested by One-way ANOVA. * < 0.05. (**C**) Survival of *Swhtr³ˣᵖᴬ/³ˣᵖᴬ;tg* mice compared to WT and *Swhtr³ˣᵖᴬ/³ˣᵖᴬ* mice after LAD ligation (WT n=13; *Swhtr³ˣᵖᴬ/³ˣᵖᴬ* n=19; *Swhtr³ˣᵖᴬ/³ˣᵖᴬ;tg* n=15). Statistical significance was tested by Kaplan-Meier Simple Survival Analysis. * < 0.05.

Activity of the transgene could be detected by presence of yellow fluorescence in the heart (Fig. 9A), but also by qRT-PCR analysis of adult hearts (Fig. 9B). In order to determine a dependance on *Swhtr* for stress reaction after injury *Swhtr Rescue* mice underwent the same LAD ligation procedure as previously described for WT and *Swhtr Null* mice. Compared to WT mice, the *Swhtr Rescue* mice do not display significantly increased lethality after AMI (Fig. 9C). The heart parameters determined by echocardiography were comparable to WT mice before AMI, as well as one or two weeks after AMI, also in respect to IVS size (Fig. 10 more shown in Appendix Fig. 5). The ejection fraction of *Swhtr Rescue* mice closely resembles that of WT hearts, suggesting that the observed decrease, despite not significant, might be a relevant manifestation of *Swhtr* dependent heart function after AMI. As *Swhtr Null* animals depicted increased mortality after AMI, especially during the first days following the operations suggests that *Swhtr* is already required for early response to AMI. However, this furthermore puts a compositional bias onto the analysis of heart parameters, as it can be hypothesized that those animals that do not survive, also would display the strongest manifestation of the phenotype but are excluded from echocardiographic measurements. In summary, it can be concluded that the RNA of the *Swhtr* locus is important compensatory response of the cardiac tissue to stress.
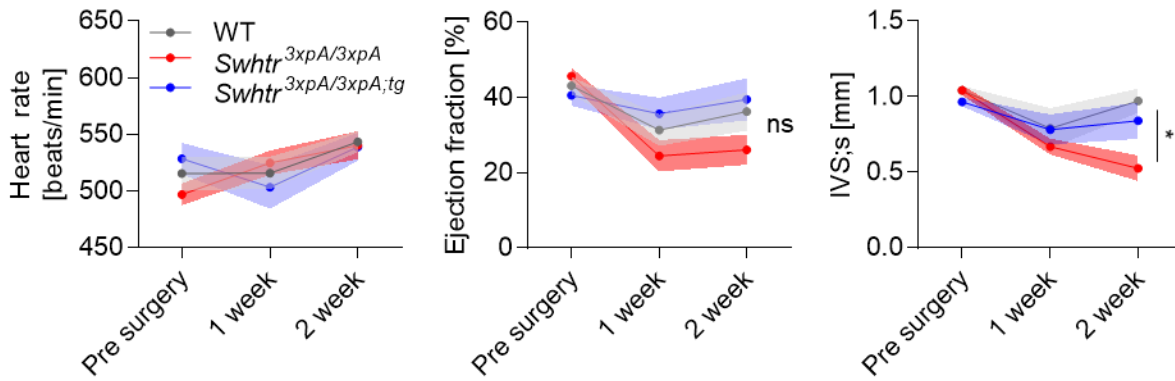
Fig. 10: Rescue of *Swhtr Null* phenotype after AMI. Selected heart specific parameters in mice (n=9 animals per genotype) before and after (1 and 2 weeks) LAD ligation obtained by echocardiography analysis. The figure legend applies for all graphs Statistical significance was tested by Two-way ANOVA. ns = not significant, *** < 0.001

Hypertrophic re-modeling after AMI is dependent on *Swhtr*

Thickening of tissue can happen either by hyperplasia, which is an increase of cell number through mitosis, or hypertrophy which is increase of cell size. As terminally differentiated cardiomyocytes typically do not divide anymore, hyperplasia would more likely be displayed by other cell-types present in the heart, while hypertrophy is a mechanism commonly facilitated by cardiomyocytes. Furthermore, hypertrophy is commonly found in cardiac tissue after myocardial infarction (Rubin et al., 1983). In order to identify the underlying mechanism behind the *Swhtr Null* phenotype, first the cell-type expressing *Swhtr* in the adult heart was determined.
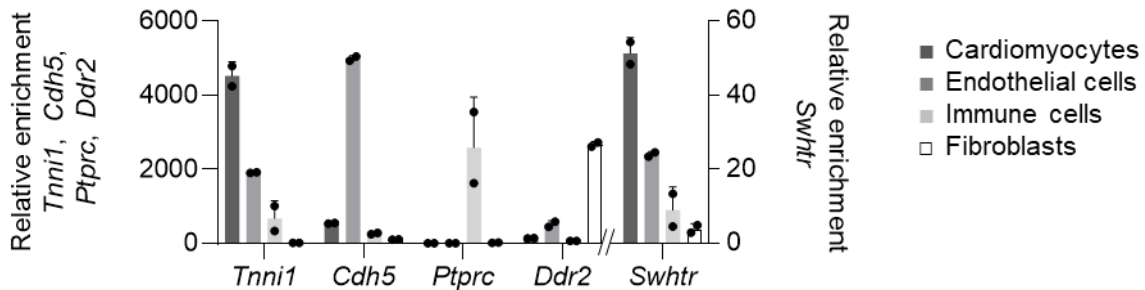

Fig. 11: Cardiomyocyte specificity of *Swhtr*. qRT-PCR analysis of *Swhtr* compared to known marker genes in the 4 major cell types obtained from adult mouse hearts (n=2).

For this purpose, the heart of adult mice was fractionated into the four cell-types that comprise the majority of the heart tissue. Namely, those are cardiomyocytes, endothelial

cells, immune cells and fibroblasts. These cell types express known marker genes that were used to verify successful enrichment for specific cell types. qRT-PCR revealed a *Swhtr* expression pattern resembling the cardiomyocyte marker *Tnni1*, demonstrating cardiomyocyte specificity of *Swhtr* (Fig. 11). This led towards the conclusion that the *Swhtr Null* animals do not display cardiac hyperplasia after AMI, hence the hearts from mice two weeks post-AMI were embedded and sectioned for histological analysis. To further exclude hyperplasia as the underlying mechanism of the increased thickness of the IVS, the number of nuclei was counted in H&E-stained sections of the heart assuming the number of nuclei would increase should there be a greater number of cells present.
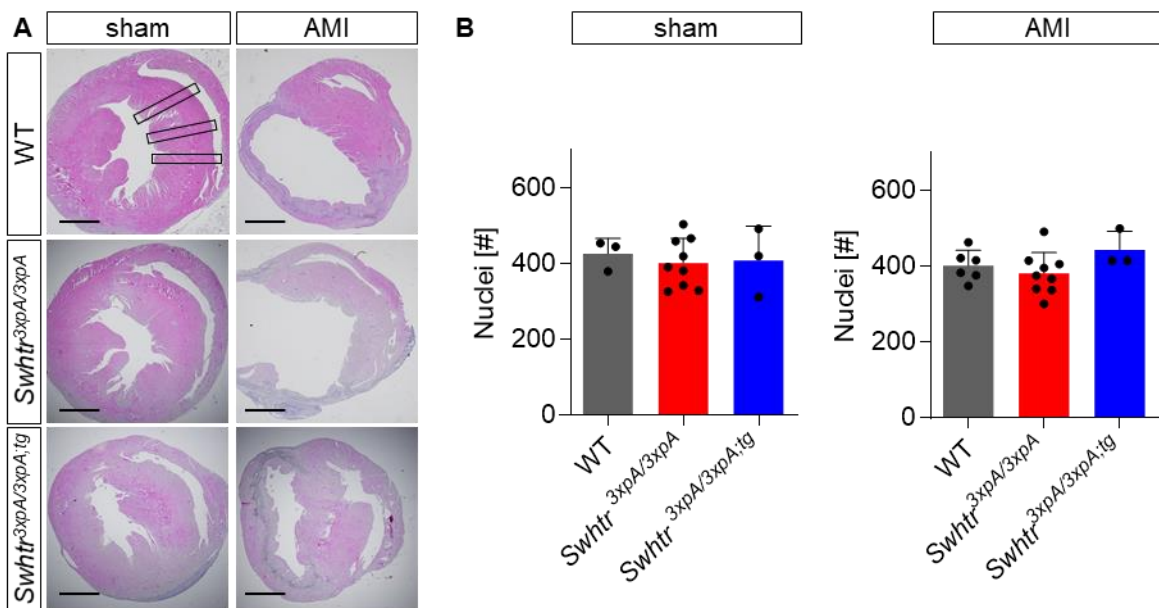
Fig. 12: The cell number does not increase in *Swhtr* dependent thickening of the IVS after AMI. (A) Representative H&E-stained sections of sham and AMI hearts of the indicated genotype. The boxes in the WT sham indicate the fixed areas used the attain average number of nuclei per IVS. The black line represents 2 mm. (**B**) Counted nuclei in the IVS of sham and AMI hearts of the indicated genotype. The number of dots indicated n number. No statistically significant differences could be detected. Dots indicate number of n.

The number of nuclei was counted in three areas of a fixed size in each IVS and the average number of nuclei from those three areas was used as the final count (Fig. 12A). As expected from the *Swhtr* cardiomyocyte specificity, the number of nuclei remained unchanged between genotypes after sham operations and after AMI (Fig. 12B). Consequently, the size of the cells contained in the IVS was investigated. The cell borders can be visualized using a Wheat Germ Agglutinin (WGA) antibody on fixed sections from

LAD operated and sham animals. Hereby, the presence of enlarged cells in AMI compared to sham heart IVS, evident of hypertrophy, could be verified in WT and *Swhtr Rescue* but not *Swhtr Null* heart sections. Quantification of the cell size showed no difference between the cell size in the IVS of sham operated mice, while the cell size in the IVS of *Swhtr Null* AMI hearts was significantly reduced compared to WT and *Swhtr Rescue* (Fig. 13A,B). In summary, it can be concluded that *Swhtr* RNA is required for compensatory hypertrophic re-modeling after myocardial infarction.
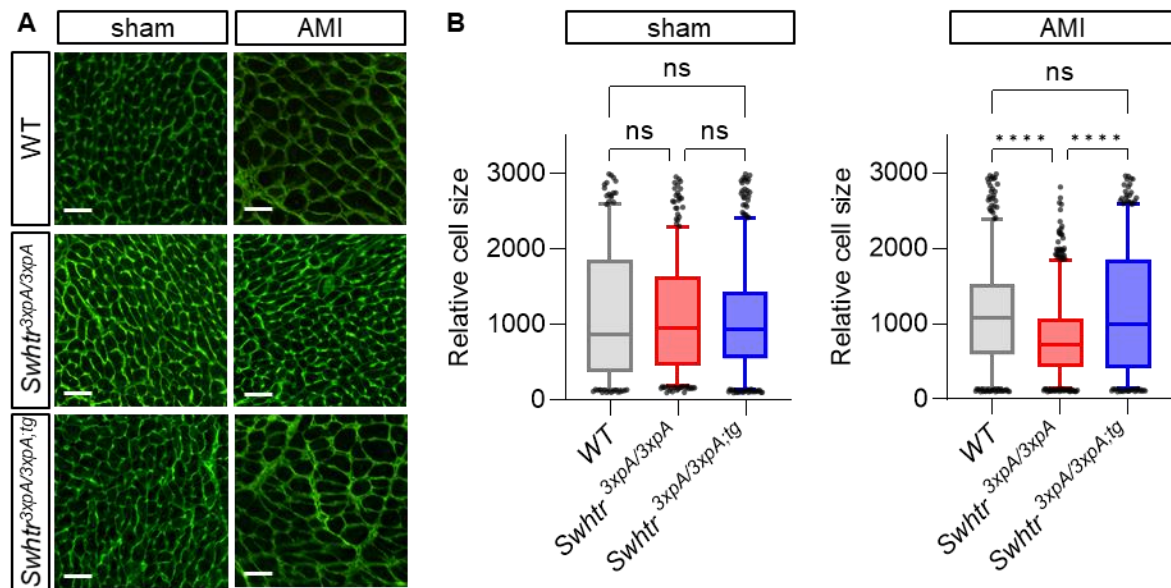
Fig: 13: *Swhtr^3xpA/3xpA* mice do not display compensatory hypertrophy after AMI. (**A**) Representative sections of Wheat Germ Agglutinin-stained cell borders in the IVS of mice of the indicated genotype 2 weeks after sham or AMI. The white line represents 50 µm. (**B**) Automated quantification of relative cell size in the IVS of mice of the indicated genotype 2 weeks post sham or AMI of three representative animals per genotype. Statistical significance was tested by One-way ANOVA. ns = not significant, **** < 0.0001

Swhtr does not regulate *Nkx2-5* RNA levels

Despite not seeing an effect of loss of *Swhtr* on *Nkx2-5* under normal conditions, an involvement of *Nkx2-5* in the *Swhtr* mediated stress response could not be ruled out. Furthermore, two additional enhancer RNAs contained in the Nkx2-5 locus have been described to regulate *Nkx2-5*, *IRENE-div* and *IRENE-SS*. These two lncRNAs are located in close proximity to *Swhtr*, while *IRENE-div* even partially overlaps with the *Swhtr* 5'-end (Fig. 2). Due to this overlap, the expression of the *IRENE* RNAs was investigated in the *Swhtr Null* mouse heart to exclude an involvement of *IRENE* deregulation as the cause for the observed phenotype. For qRT-PCR, primers were carefully designed to measure

expression of only *IRENE-div* or *Swhtr* as both are located on the same strand divergent from *Nkx2-5*. For this purpose, the intron-containing structure that clearly distinguishes *IRENE-div* from *Swhtr* was used to design specific exon-spanning primers for *IRENE-div* and specific primers for *Swhtr* downstream of the 3'-end of *IRENE-div*. Using these primers for qRT-PCR on tissue derived from WT and *Swhtr Null* mouse hearts, neither *IRENE-SS*, nor *IRENE-div* show any significant changes in response to insertion of the transcriptional stop within the *Swhtr* sequence (Fig. 14).
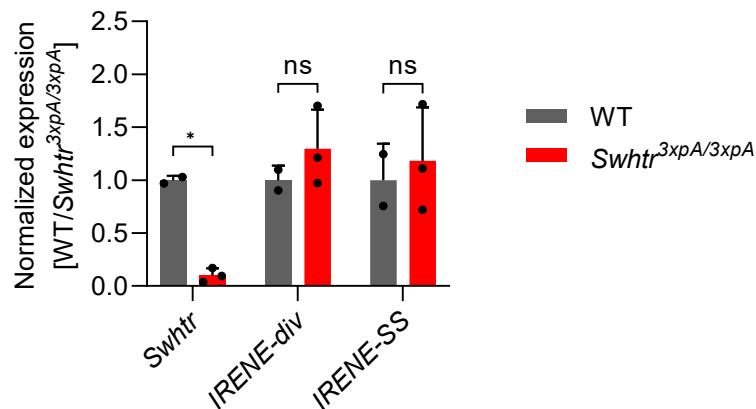
Fig. 14: *IRENE* RNAs are unaffected by loss of *Swhtr*. *Swhtr*, *IRENE-div* and *IRENE-SS* RNA levels assessed by qRT-PCR in hearts of WT and *Swhtr Null* mice. Statistical significance assessed by Two-way ANOVA. ns = not significant, * < 0.05 (n: WT=2, *Swhtr$^{3xpA/3xpA}$*=3)

To further investigate the possibility of *Swhtr* dependent regulation of *Nkx2-5* RNA levels, an *ex vivo* cardiomyocyte stress model was used. One of the common challenges after AMI are low blood oxygen levels (hypoxemia) and hence hypoxic stress (Valencia and Burgess, 1969). Thus, neonatal cardiomyocytes were isolated from WT and *Swhtr Null* mice and exposed to hypoxic stress using a chemical inducer of hypoxic stress, namely deferoxamine (DFO) diluted in DMSO. Hence, DMSO alone was used as normoxic control. qRT-PCR analysis of two known hypoxia response genes, namely *Slc2a1* and *Ldha*, that are targets of HIF1A was performed as well as *Swhtr* and *Nkx2-5*. As the *IRENE* RNAs were not altered in the *Swhtr Null* mouse hearts, they were not investigated in this experiment. Both control genes were upregulated after stimulation with DFO, revealing a successful induction of the hypoxic stress response by the treatment (Fig. 15).

While *Swhtr* seems to be highly responsive to hypoxic stress, *Nkx2-5* does not react considerably in this setting. Except from *Swhtr* itself, no overt differences in expression patterns of the examined genes could be observed (Fig. 15).
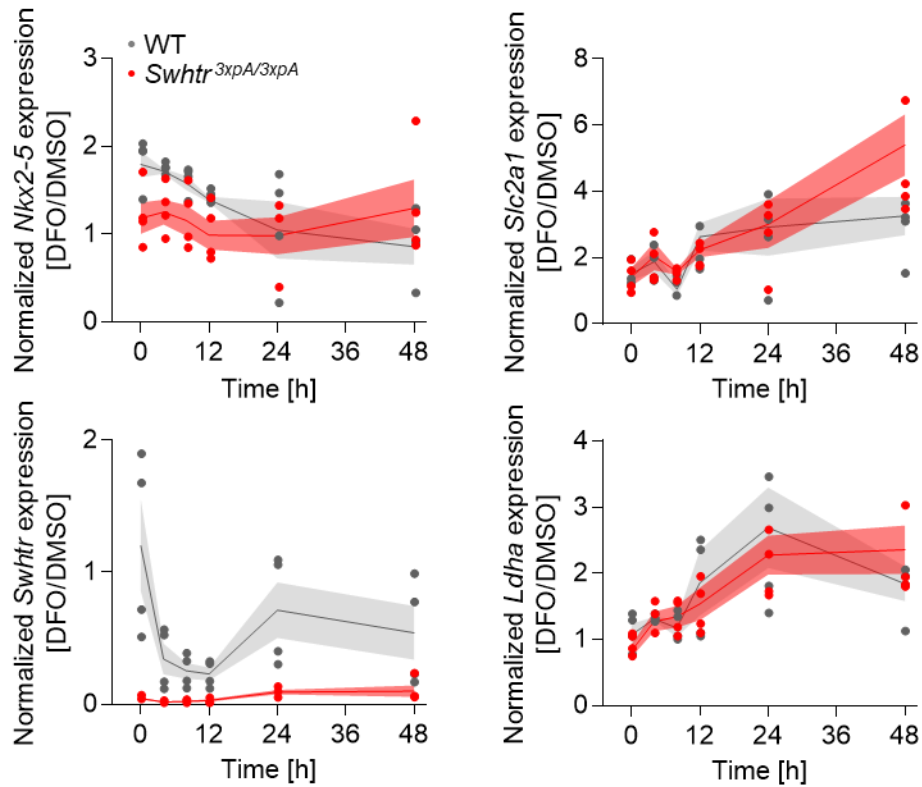
Fig. 15: DFO-induced hypoxic stress reaction of neonatal cardiomyocytes. qRT-PCR timeline analysis of indicated genes in neonatal cardiomyocytes isolated from WT or *Swhtr$^{3xpA/3xpA}$* mice normalized to the WT DMSO control. 100 µM DFO were added to the ca (n=4)

Furthermore, the isolation of neonatal cardiomyocytes led to variation in terms of cardiomyocyte content that might lead to the observed high spread of expression patterns between replicates. To circumvent this effect in a more homogenous cell population and examine whether acute loss of *Swhtr* might influence *Nkx2-5*, HL-1 cardiomyocytes were used as a model system for antisense oligo (ASO) mediated knockdown of *Swhtr* RNA. Two different ASOs located towards the *Swhtr* 3'-end were transfected individually, both achieving considerable loss of *Swhtr* after only 6 hours with maximum loss between 12 and 36 hours after transfection (Fig. 16).
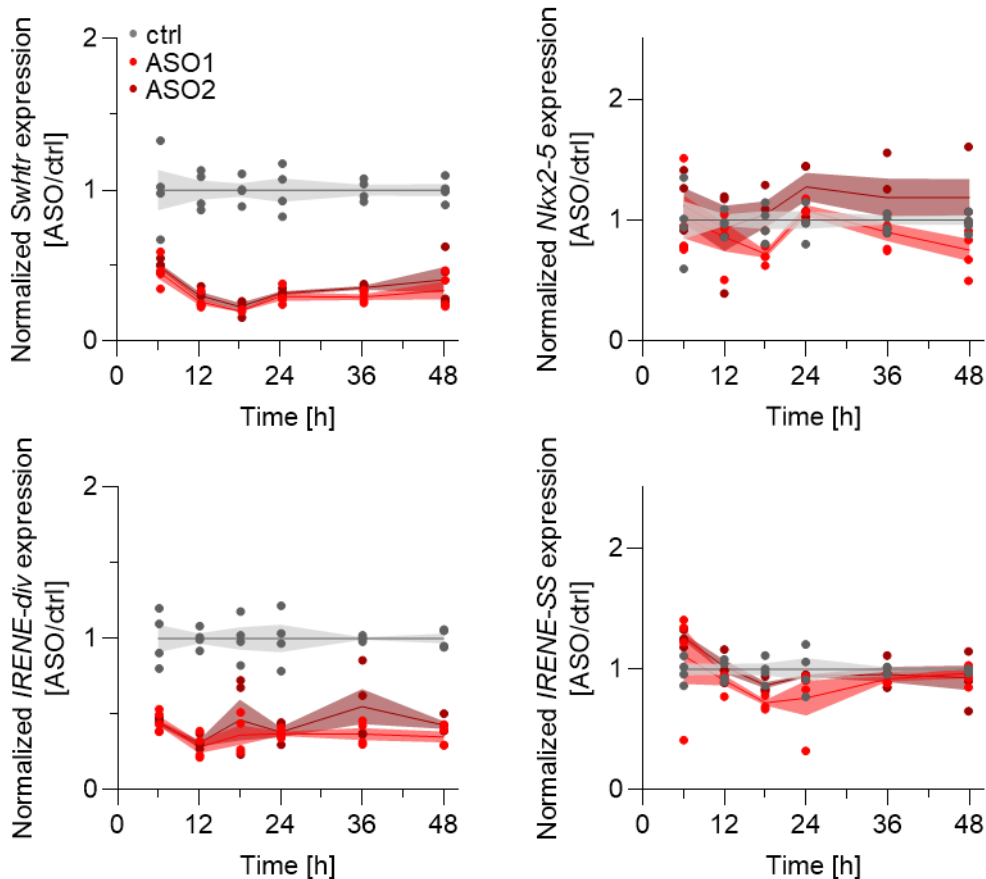
Fig. 16: qRT-PCR timeline of selected target genes in HL-1 cardiomyocytes after ASO mediated *Swhtr* knockdown. No significant changes could be detected in *Nkx2-5* RNA levels at any given timepoint, despite corresponding significant knockdown of *Swhtr* ($p < 0.0001$). *IRENE-div* was also downregulated in knockdown samples compared to ctrl at any measured timepoint ($p < 0.0001$). Statistical significance assessed by Two-way ANOVA and given in figure description when significant. (n=4)

Despite high knockdown efficiency of the ASOs, no effect on *Nkx2-5* RNA levels could be observed. Counterintuitively, however, *IRENE-div* RNA levels were strongly affected by the *Swhtr* ASOs, despite location of the ASOs downstream of the *IRENE-div* 3'-end leading to the conclusion of *IRENE-div* being an alternative transcriptional version of *Swhtr* with a different annotation and stop signal usage than proposed in the paper (Salamon et al., 2020). Contrary to what Salomon et al show, despite usage of the same cell line, reduction of *IRENE-div* did not lead to an effect on *Nkx2-5*.
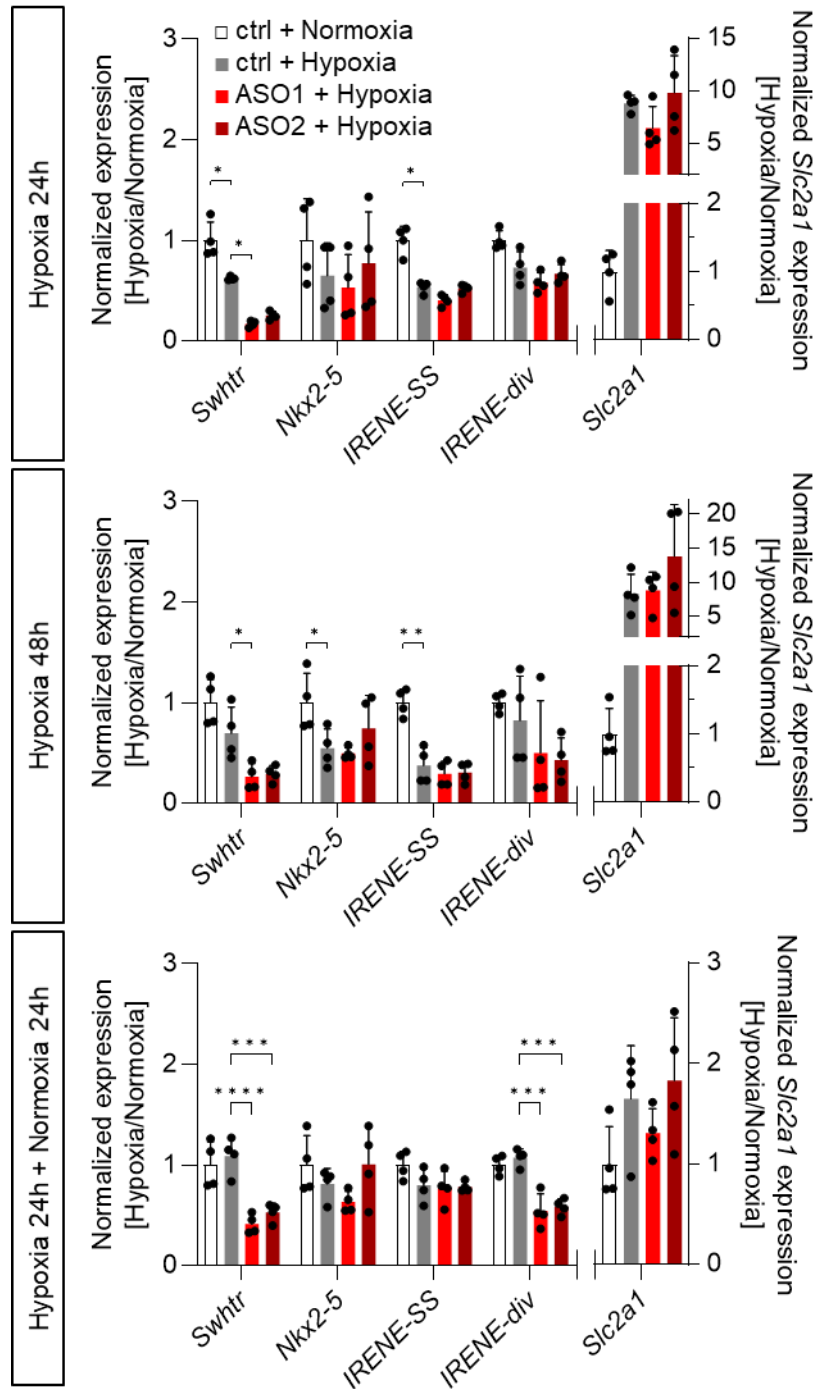
Fig. 17: _Swhtr_ knockdown in combination with hypoxic stress and subsequent regeneration in HL-1 cardiomyocytes. qRT-PCR analysis of selected genes of interest and the hypoxia response gene _Slc2a1_. Cells were incubated for the indicated amount of time under Normoxia (21% O_2) or Hypoxia (1% O_2). Statistical significance of difference to hypoxic incubated control (grey) was tested by Two-way ANOVA and is given when significant. * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001 (n=4)

To further investigate whether short term loss of *Swhtr* might influence *Nkx2-5* RNA under hypoxic stress, ASO mediated knockdown was applied to HL-1 cardiomyocytes in combination with hypoxic stress. For that purpose, the cells were transfected with the knockdown ASOs and left to attach for 4 hours. Subsequently the cells were supplied with fully supplemented HL-1 culture medium and moved to a chamber with controllable atmosphere to induce hypoxic stress. Strongly upregulated *Slc2a1* compared to the normoxic control verified successful induction of hypoxic stress as well as return to baseline after 24h of return to normoxic conditions demonstrated recovery (Fig. 17).

All genes of interest were downregulated in response to hypoxic stress, including *IRENE-SS* and *IRENE-div*. The two *IRENE* RNAs have been described to have opposing functions on *Nkx2-5* RNA levels, which is contradicted by these experiments. In neither of the conditions, *Nkx2-5* RNA levels differ from the control GapmeR transfected cells, incubated in the same way as the *Swhtr* knockdown samples, despite downregulation of *IRENE-div* being caused by *Swhtr* ASOs. In conclusion, *Swhtr* does not regulate *Nkx2-5* RNA levels while the results of Salomon at al could also not be verified.

Swhtr interacts with NKX2-5 to mediate cardiac stress response

To unravel the mechanism behind *Swhtr* dependent stress response an *ex vivo* model of cardiac stress mimicking the *in vivo* AMI situation was established. This was utilized to avoid compositional bias of potential results. AMI leads to loss of flexibility of the LV due to the formation of scar tissue after the injury. Each of these scars varies in size and is highly individual and heterogenous potentially leading to a compositional bias of biopsies taken from such tissues. To circumvent this issue, cardiac slices from adult WT and *Swhtr Null* mice were obtained and maintained in culture. Stress was induced by incubation under hypoxic conditions for one week followed by a subsequent regeneration period of one week under normoxic conditions. After two weeks in cell culture, no necrosis could be observed in neither of the heart slices independent of genotype (Fig. 18). Cardiac slices that were kept in culture under non-stressed normoxic conditions for 14 days served as control. Total RNA was extracted from these samples and used to obtain transcriptomic profiles by bulk RNA sequencing.
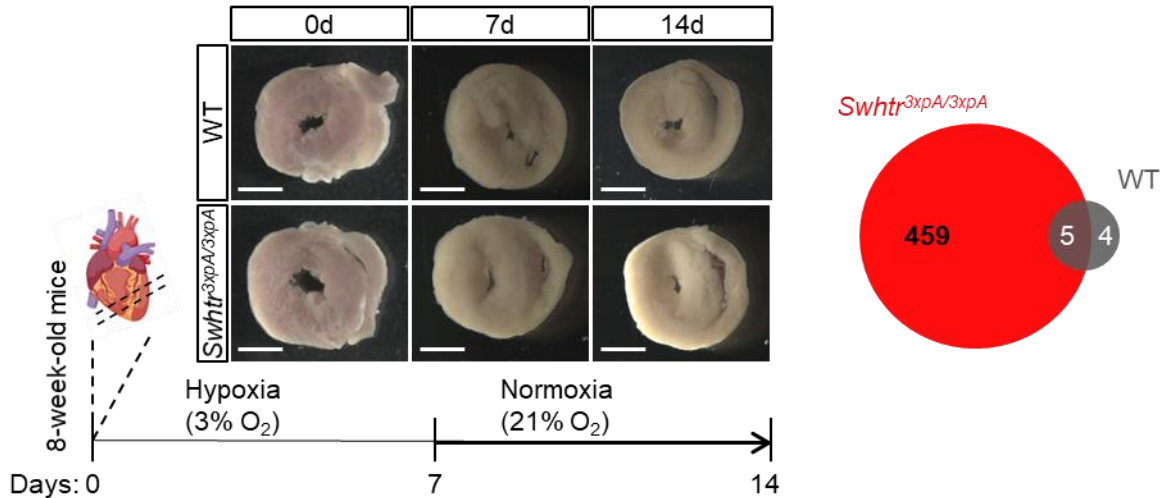
Fig. 18: Swhtr dependent deregulation of genes after mimicking AMI stress experiments without formation of scar tissue *ex vivo*. Depicted is a schematic of the experimental procedure and representative pictures of cardiac slices after indicated times of indicated genotype (left). The white line represents 500 mm. VENN diagram of the number of deregulated genes (right) after 7d of hypoxic stress (3% O_2) followed by 7d of regeneration (21% O_2) as compared to 14d normoxic control (21% O_2) of 4 slices obtained from independent animals of the indicated genotype [stress/ctrl].

From all the samples sufficient amount of RNA of good quality could be obtained. Strikingly, transcriptomic analysis of these samples revealed almost no deregulated genes in the WT cardiac slices as compared to normoxic control, indicating capability to compensate for hypoxic stress and ability to return to baseline expression patterns after stress. In contrast, in *Swhtr Null* cardiac slices a high number of genes was persistently deregulated, demonstrating the inability of these slices to adapt to stress (Fig. 18).

Furthermore, GO-term analysis of these genes revealed enrichment of the *Swhtr* dependent genes in biological processes that are associated with cardiac stress response, most of which are involved in hypertrophic remodeling of cardiac tissue (Fig. 19). The processes marked in green (Fig. 19), termed 'Gliogenesis body fluid levels' by the algorithm encompasses genes involved in immune response and inflammatory response, both of which are known key players in initial response to acute myocardial infarction (Nian et al., 2004). Calcium signaling (Fig. 19, brown) influences muscle contraction (Fig. 19, blue) which is responsible for the maintenance of cardiac function and supply of the organs with oxygen; a process that must function or be compensated for in order to preserve heart function (Schaub et al., 1998). Cardiac heart morphogenesis (Fig. 19, turquoise) undergoes changes after cardiac injury. Deregulation of the genes

involved in this process supports the previous findings of absence of hypertrophic changes. Another hallmark of hypertrophy is excessive deposition of extracellular matrix (Silva et al., 2020), another pathway impaired by loss of *Swhtr* (Fig. 19, pink).
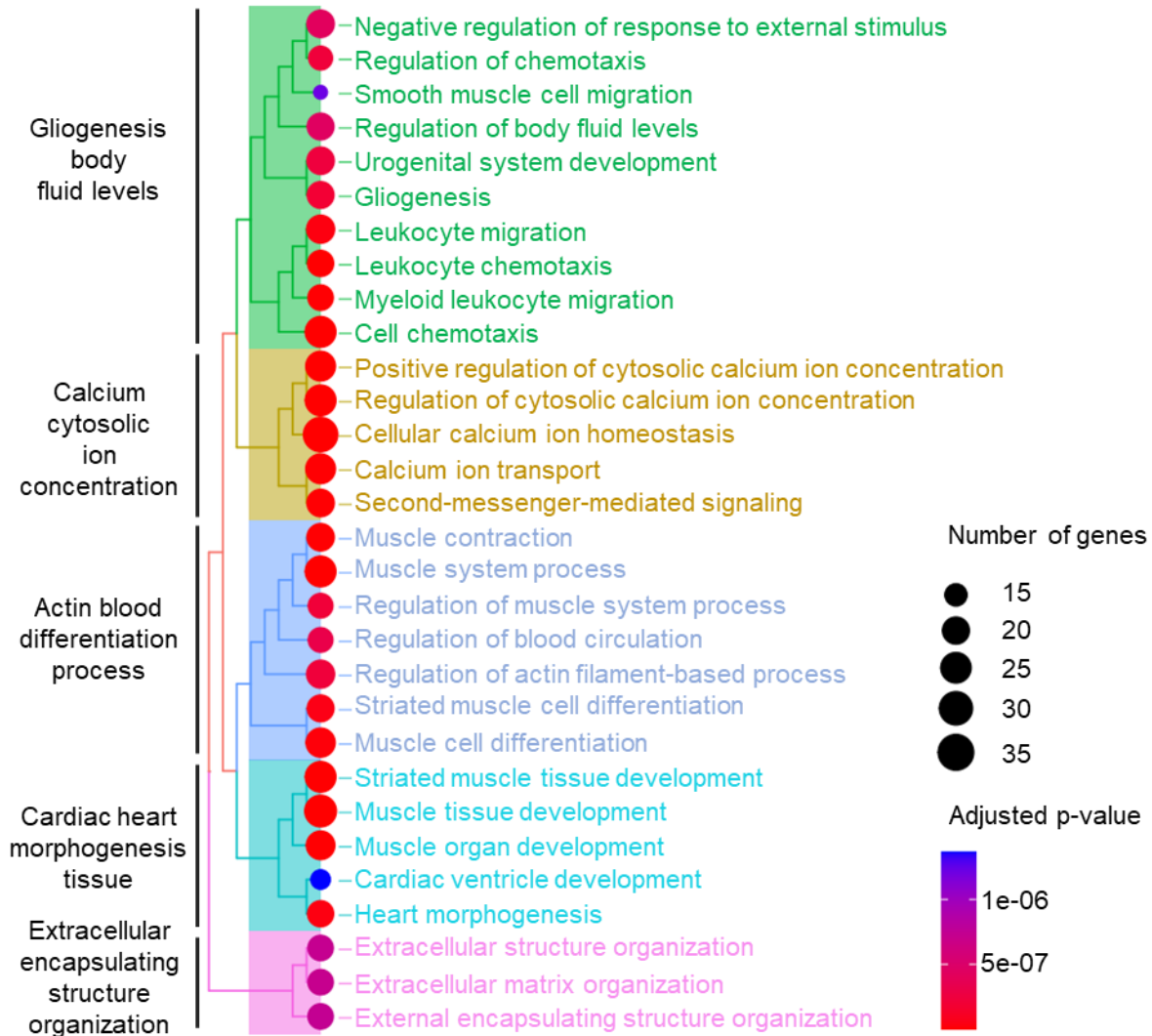
Fig. 19: GO-term enrichment analysis of *Swhtr* dependent genes.

Malfunction of hypertrophic response is furthermore supported by KEGG-pathway analysis of the deregulated genes. In this analysis, signaling pathways involved in inflammatory response, cardiac muscle contraction and cardiomyopathy, glucose metabolism, response to oxygen levels and most notably, hypertrophic cardiomyopathy as one of the top hits (Fig. 20).
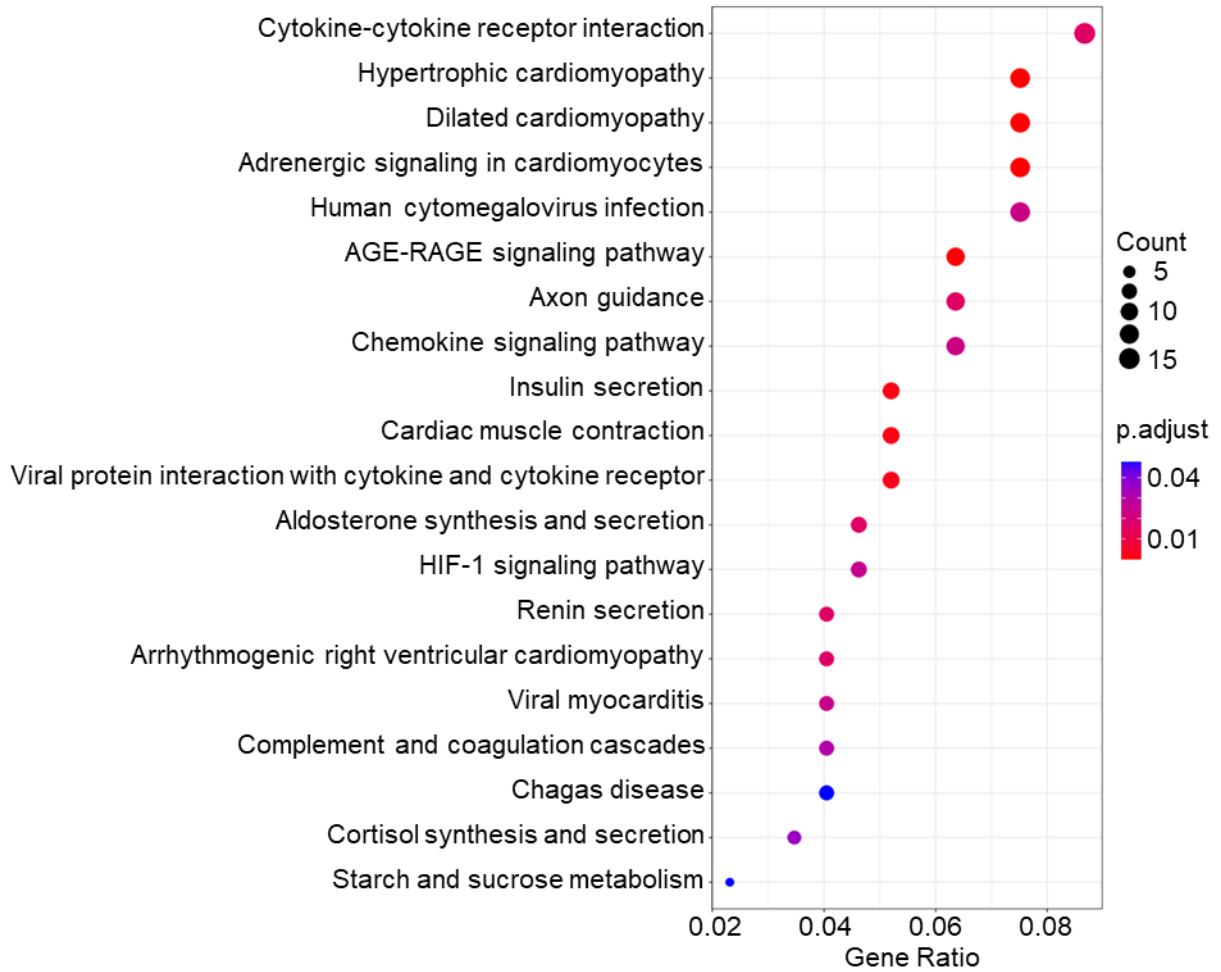
Fig. 20: KEGG-pathway analysis of Swhtr dependent genes.

The rescuable phenotype of loss of *Swhtr* under cardiac stress demonstrates the ability of *Swhtr* to act in *trans*, opposing an effect of *Swhtr* on *Nkx2-5* as excluded within further experiments as well. To investigate posttranslational interaction of *Swhtr* with NKX2-5 protein, an available NKX2-5 ChIP-seq dataset performed on adult cardiomyocytes (Akerberg et al., 2019) was analyzed and revealed a significantly increased occupation of NKX2-5 on *Swhtr* dependent genes (Fig. 21A,B). Almost all *Swhtr* dependent genes (98.71%) showed occupation of NKX2-5 within their gene bodies while only 69.08% of all genes expressed in the heart slices that are not deregulated are occupied by NKX2-5 (Fig. 21B). This indicates that the lack of *Swhtr* impairs NKX2-5 mediated response to cardiac stress, suggesting that *Swhtr* might act together with NKX2-5 to regulate cardiac stress response. Hence, a physical interaction of *Swhtr* with its neighboring gene was not ruled out and investigated further.
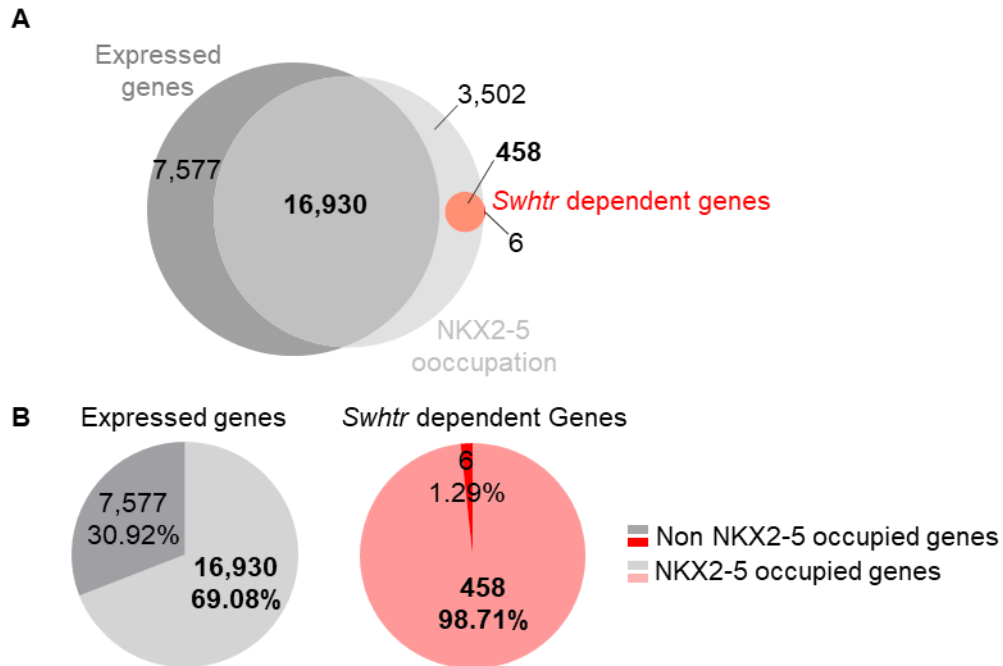
Fig. 21: *Swhtr* dependent hypoxic stress genes are significantly occupied by NKX2-5. (**A**) VENN diagram of all expressed genes that are not dysregulated (dark grey), NKX2-5 occupied gene bodies (light grey) and the 464 Swhtr dependent genes (red) analyzed from a public available NKX2-5 ChIP-seq dataset from adult cardiomyocytes. (**B**) Pie charts comparing NKX2-5 occupation on expressed genes that are not deregulated (grey) to NKX2-5 occupation on *Swhtr* dependent genes (red). Percental distribution was tested for significant differences between *Swhtr* dependent genes and expressed genes by binomial test (p < 0.0001).

The result of NKX2-5 RIP performed by Salomon et al pointed against an interaction of *Swhtr* with NKX2-5 under standard conditions (Salamon et al., 2020). In their paper it was shown that *IRENE-div* does not bind to NKX2-5 but a different RNA binding protein. As there is a partial overlap of the *IRENE-div* annotation with *Swhtr*, the primers the group used to detect their lncRNA were extracted from the paper and aligned to *Swhtr* revealing that the group simultaneously measured RNA levels of *IRENE-div* and also *Swhtr*. Their experiments, however, were performed under standard culture conditions without stress. As loss of *Swhtr* does not show any effect under standard conditions, there might be the possibility of *Swhtr* binding to NKX2-5 only under stressed conditions. This was tested by performing NKX2-5 RIP in HL-1 cardiomyocytes under normoxic and hypoxic conditions.
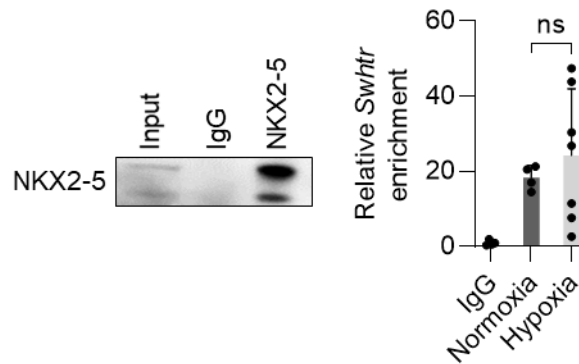
Fig. 22: *Swhtr* binds to NKX2-5. Representative Western Blot analysis showing successful NKX2-5 enrichment compared to input with respective qRT-PCR analysis of Swhtr enrichment in the pulldown samples performed from HL-1 cardiomyocytes incubated for 24h at normoxic (21% O_2) or hypoxic (1% O_2) conditions. (Number of n indicated by dots; n=4-7).

In contrast to what Salomon et al found, *Swhtr* RNA was enriched on NKX2-5 RIP even under standard conditions. When repeated under stressed conditions, this interaction seemed to increase, although high variation between samples impedes a definitive statement. What became clear, however, was the presence of physical interaction between *Swhtr* and NKX2-5 under stressed and standard conditions (Fig. 22). In accordance with high occupation of NKX2-5 on *Swhtr* dependent genes (Fig. 21) it could be concluded that *Swhtr* has an auxiliary effect on NKX2-5 mediated cardiac stress response.

Stress-induced dispersion of *Swhtr*

By subcellular fractionation, it was shown previously that in embryonic cardiomyocytes *Swhtr* localizes mainly to the chromatin fraction of the cell (Rogala et al., 2022). A similar distribution could also be shown for HL-1 cardiomyocytes with the majority of *Swhtr* bound to chromatin and a small portion of *Swhtr* in the nucleoplasm (Fig. 23).
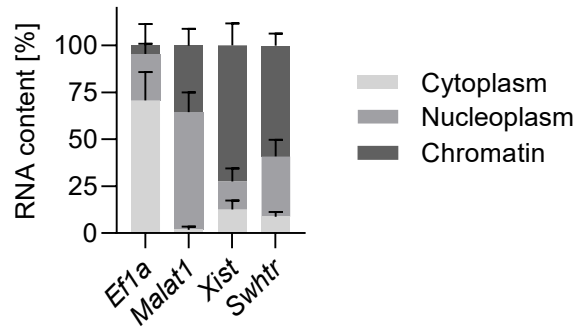
Fig. 23: Subcellular fractionation of HL-1 cardiomyocytes. Note that *Swhtr* localizes mainly to the chromatin fraction (n=2).

To unravel how *Swhtr* could interact with protein binding partners despite being bound most likely to its transcriptional locus, the behavior of *Swhtr* under standard and stressed conditions was investigated in more detail. Single-molecular fluorescent *in situ* hybridization (smFISH) stainings were conducted using HL-1 cardiomyocytes to visualize localization of the RNA. The cells were seeded on coated glass coverslips and subsequently incubated at 1% O_2 hypoxic conditions or normoxic control for 24 and 48h. The stressed cells were fixed within the hypoxic atmosphere to avoid oxygen shock before the staining.
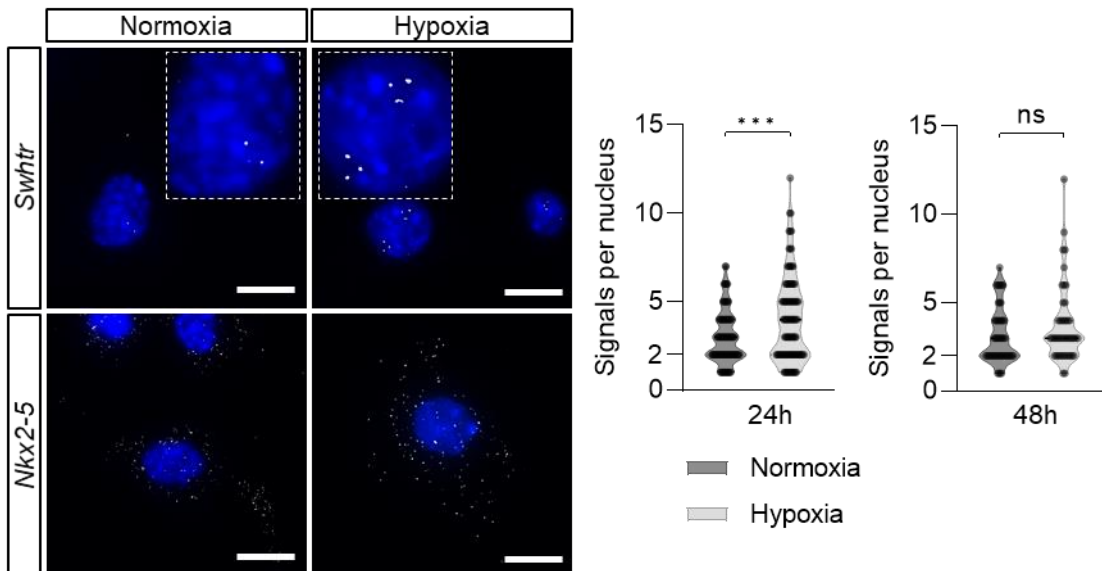
Fig. 24: *Swhtr* disperses under hypoxic stress in HL-1 cardiomyocytes. Representative pictures of smFISH staining of *Swhtr* and *Nkx2-5* in HL-1 cardiomyocytes with corresponding quantification (n > 63) of *Swhtr* signals per nucleus after 24h or 48h of Normoxia (21% O_2) or Hypoxia (1% O_2). The white line represents 10 µm. Statistical significance tested by t-test analysis. ns = not significant; *** < 0.001.

In the normoxic control cells, two distinct signals can be found on average reminiscent of the two transcription loci present in wild type cells. Under standard conditions some cells showed more than two dots, which is in accordance with a minor fraction of *Swhtr* being present in the nucleoplasm (Fig. 24), however, after applying stress to the cells, *Swhtr* disperses from its locus as evident by significantly increased multiple signals in the nucleus after 24 hours of stress. The number of signals in the nucleus was still slightly increased after 48 hours of stress, however, no statistically significant difference could be detected anymore. This suggests a role of *Swhtr* in fast response to stress. As it was shown previously that *Swhtr* decreases under hypoxic stress (Fig. 15,17) this behavior was unexpected. Hence, the cells positive and negative for *Swhtr* expression were quantified as well.
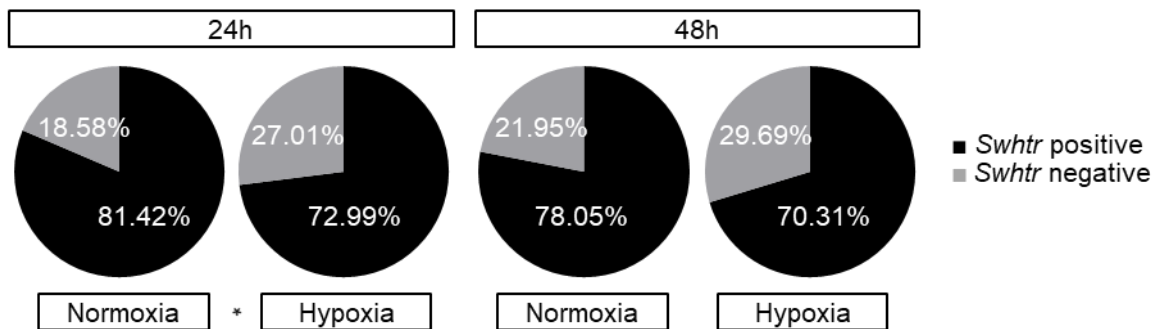

Fig. 25: Decrease of *Swhtr* expressing cells upon hypoxic stress. Quantification of *Swhtr* positive and negative HL-1 cardiomyocytes in smFISH staining after 24h and 48h of normoxic (21% O_2) or hypoxic (1% O_2) conditions (n > 63). Significant difference between percentages assessed by binomial test. * < 0.05

Indeed, less cells expressing *Swhtr* were found after 24h hypoxia as compared to the normoxic control (Fig. 25). Together with the dispersion of the lncRNA and its decreased RNA level it can be concluded that under standard conditions *Swhtr* mainly resides bound to the chromatin where it is more stable. High variation between different samples in the NKX2-5 RIP (Fig. 22) might point to a transient interaction of the lncRNA and the protein under stressed conditions. Once the interaction is lost, *Swhtr* seems to be degraded, consistent with lncRNA stability showing negative correlation with absence of an Intron-Exon structure and localization within the nucleus (Clark et al., 2012).

To investigate whether this behavior is consistent *in vivo*, *Swhtr* smFISH staining under standard and stressed conditions was repeated in murine primary cardiomyocytes isolated from neonatal mouse hearts of wild type, *Swhtr* Rescue and *Swhtr* Null mice.
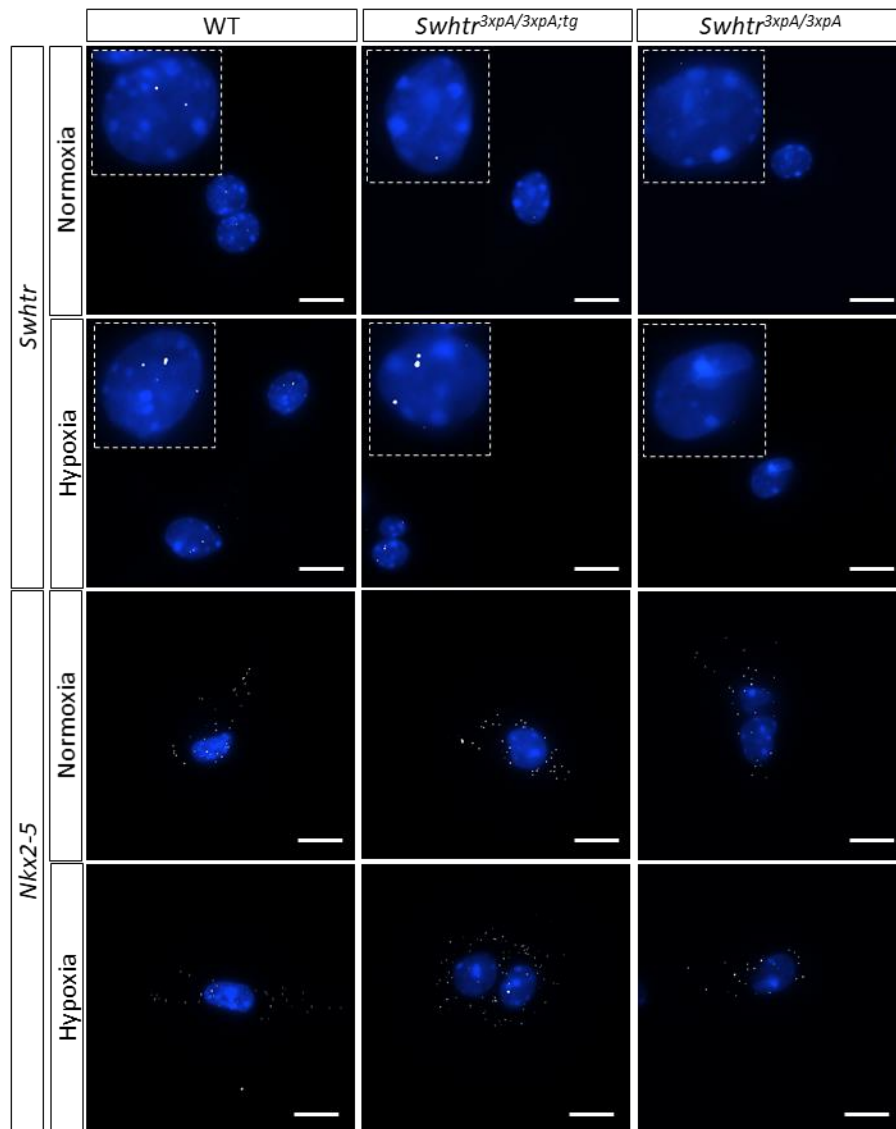
Fig. 26: _Swhtr_ disperses under stress in primary cardiomyocytes. Representative pictures of smFISH staining of _Swhtr_ and _Nkx2-5_ in primary cardiomyocytes of the indicated genotype after 24h of Normoxia (21% O_2) or Hypoxia (1% O_2). The white line represents 10 μm.

As expected, in the WT cells, two dots were visible in the majority of the cells, while in the _Swhtr_ Rescue cardiomyocytes only one dot could be found on average. Absence of any _Swhtr_ signal in the _Swhtr Null_ cardiomyocytes further validated high specificity of the _Swhtr_ smFISH probes (Fig. 26,27). Consistent with the previous experiments, the _Swhtr_ signal disperses into multiple dots in the WT but also in the _Swhtr_ Rescue cardiomyocytes, explaining how random integration of the rescue construct can rescue the _Swhtr_ phenotype.
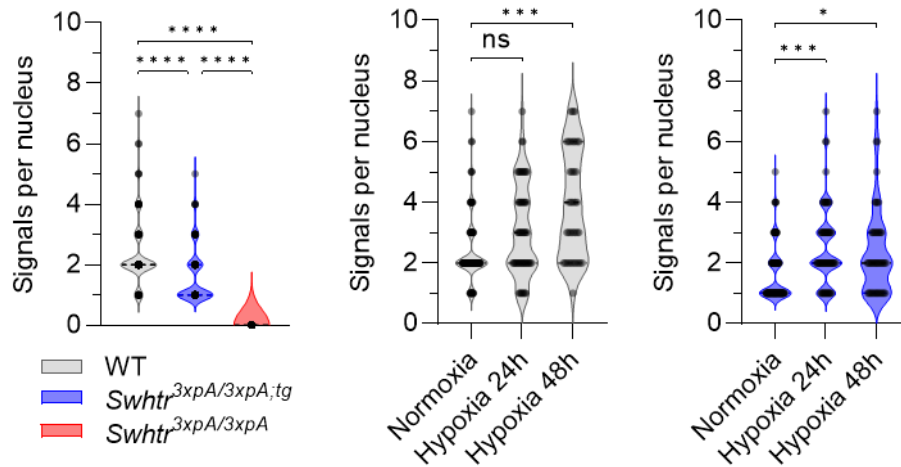
Fig. 27: Quantification of *Swhtr* dispersion under stress in primary cardiomyocytes. Signals per nucleus in smFISH staining of *Swhtr* in primary cardiomyocytes of the indicated genotype after 24h or 48h of normoxic (21% O_2) or hypoxic (1% O_2) conditions (n > 50; except WT Hypoxia 48h n=37). Color legend applies to all graphs. Statistical significance tested by Two-way ANOVA. ns = not significant; * < 0.05; *** < 0.001; **** < 0.0001.

Together, the data suggests that *Swhtr* RNA, possibly bound to NKX2-5, disperses from its locus upon stress to exert an auxiliary function in the regulation of gene programs required for cardiac stress response, in particular for tissue adaptation after myocardial injury. This data implicates for the first time a lncRNA that can act in *trans* at multiple loci, despite under standard conditions mainly residing at its own locus of origin.

Conservation in the *Nkx2-5* locus between mouse and human

For the human *NKX2-5* locus no lncRNAs have been described yet. However, the promoter element which gives rise to *Swhtr* and *IRENE-SS* is highly conserved between human and mouse. Analysis of the sequence of the core promoter element revealed 91.51% sequence conservation. Hence, publicly available RNA-sequencing datasets of cardiomyocytes differentiated from human induced pluripotent stem cells (iPSCs) were obtained from ENCODE and analyzed for RNAs expressed within the vicinity of *NKX2-5*. Structurally, the human locus resembles that in the murine system, with one convergently and two divergently expressed RNAs. Derived from the sequencing reads and structural similarity to the murine lncRNAs, they were termed *IRENE-SS*, *IRENE-DIV* and *SWHTR* (Fig. 28).
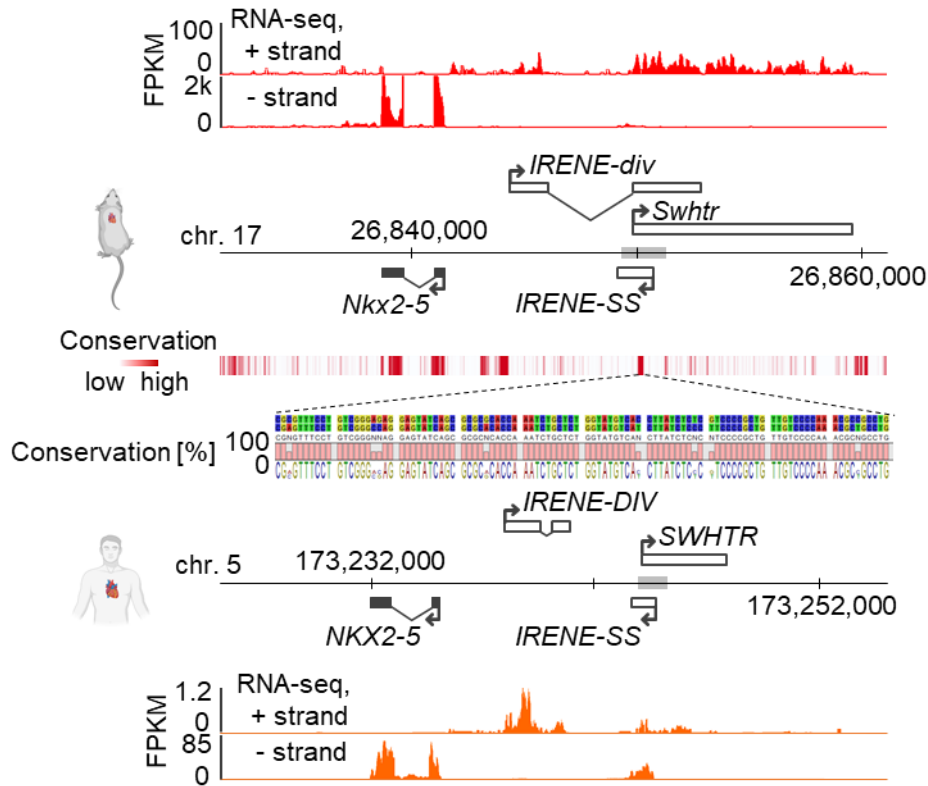
Fig. 28: The *Nkx2-5* locus is structurally conserved between mouse and human. Comparison of strand-specific RNA-sequencing from murine E9.5 hearts (red) with publicly available ENCODE RNA-sequencing from human iPSC derived cardiomyocytes (ENCFF182KCL, ENCFF571PYD; orange) with conservation data. Heat map derived from WashU Epigenome Browser (Placental PhastCons 60-way). The alignment was generated using genomic sequences from mouse and human in CLC software and the figure presents a segment of the core promoter element, demonstrating high sequence conservation.

Unlike in the murine system, in the human system no overlap between *IRENE-DIV* and *SWHTR* could be detected by PCR amplification from cDNA obtained from human iPSC derived cardiomyocytes. To further validate these RNAs, RNA-sequencing read derived sequences were extracted to assess coding potential of the three RNAs together with known coding and non-coding RNAs. CPAT analysis (Wang et al., 2013) validated the absence of coding potential in all three RNAs, while correctly labeling the known coding and non-coding control RNAs (Fig. 29A). Furthermore, human iPSCs were differentiated into cardiomyocytes (iPSCMs) to check on-set of the *NKX2-5* associated lncRNAs in the cardiac system, as comparable to the murine *Nkx2-5* associated lncRNAs. Successful differentiation of iPSCs was validated by monitoring of beating cells under the

microscope. Upon qRT-PCR analysis, induction of these lncRNAs could be detected, compared to undifferentiated iPSCs (Fig. 29B).
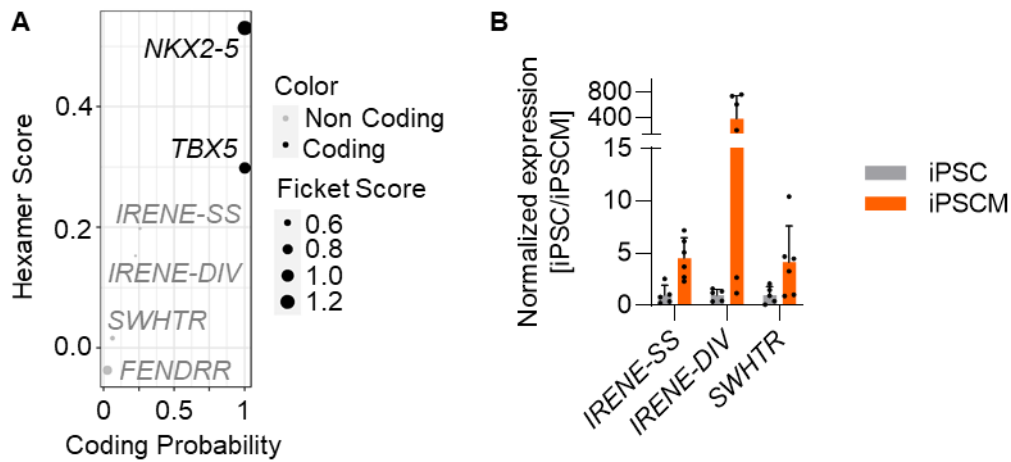
Fig. 29: Coding potential and expression of human NKX2-5 associated RNAs. (**A**) CPAT analysis demonstrating absence of coding potential in *IRENE-SS*, *IRENE-DIV* and *SWHTR* compared to known control RNAs. (**B**) On-set of *NKX2-5* associated RNAs in iPSC derived cardiomyocytes (iPSCM) compared to the undifferentiated condition (Number of n indicated by dots; n=5-6).

Despite poor sequence conservation within the *SWHTR* lncRNA itself, strong conservation of the promoter element and structural conservation of the *NKX2-5* locus, suggests similarities between human and mouse. Whether or not the human *SWHTR* interacts with NKX2-5 and exerts a similar cardioprotective role in disease and injury remains to be studied.

Discussion

In this work a marginally characterized lncRNA that was termed *Swhtr* was studied in the context of cardiac stress and disease. It could be shown that *Swhtr* is expressed abundantly in adult cardiomyocytes, following the expression pattern of its neighboring transcription factor coding gene *Nkx2-5* (Fig. 4). Genetic ablation of *Swhtr* does not lead to any overt defect in neither embryonic development (Rogala et al., 2022) nor the adult heart under standard conditions (Fig. 5,6). Upon stress, presence of *Swhtr* becomes important, as *Swhtr Null* mice show reduced survival after LAD ligation induced AMI and fail to develop compensatory cardiac hypertrophy as compared to the WT (Fig. 9,10,13). Despite localization of *Swhtr* RNA to its own locus of transcription, this effect can be rescued by re-expressing *Swhtr* from an exogenous locus via random integration clearly showing that the RNA itself bears a function (Fig. 10,13).

The role of *Swhtr* in cardiac hypertrophy and survival after injury

The promoter of *Swhtr* has been described previously as a GATA4 bound enhancer element active in cardiac progenitor cells and the developing heart (Lien et al., 1999; Wu et al., 2006). The enhancer element is active in hearts following MI and furthermore, sorted cell populations with this active enhancer showed increased expression of embryonic programs following MI (Deutsch et al., 2018). Here, the importance of the RNA that is transcribed from this genetic element and its involvement in early stress response and downstream compensatory hypertrophic remodeling has been shown, in accordance with *Swhtr* expressing cells reverting to a more embryonic expression profile. The importance of distinction between a hyperplastic or hypertrophic origin of difference in thickness of the IVS becomes apparent as over-abundance of *XNkx2-5* in in *xenopus* embryos and injection of *Nkx2-5* into zebrafish embryos can lead to hyperplastic hearts (Chen and Fishman, 1996; Cleaver et al., 1996). Absence of *Swhtr* dependent *Nkx2-5* deregulation accompanied by cardiomyocyte specific expression of *Swhtr*, however, pointed towards a hypertrophy-based phenotype. Cardiac hypertrophy is a typical compensatory mechanism that the intact cardiac tissue utilizes to maintain blood supply of the body after cardiac injury such as AMI (Rubin et al., 1983) and is associated with the recurrence to fetal gene programs, especially in regard to energy metabolism

(Taegtmeyer et al., 2010). In this work, this was demonstrated by increased cell size of the cardiomyocytes in the IVS of WT and *Swhtr* rescue animals whereas parameters assessing the function of the heart remain unchanged. Decreased IVS size of Swhtr Null hearts could be an indication for cell death, however, presence of comparable amounts of cells within the IVS as verified by number of nuclei points against it. Accordingly, *Swhtr Null* mice showed a decrease in cell size of cardiomyocytes in the IVS accompanied by a minor non-significant reduced ejection fraction. Considering the echocardiography measurements, the elevated early lethality places a compositional bias on the data, as it can be expected that death was the most drastic manifestation of a phenotype that when not completely penetrant can under certain circumstances be compensated for. Yet, it can be hypothesized that the animals depicting the most drastic phenotype, which is early lethality, showed the most drastic manifestation of heart failure following absence of compensatory hypertrophy. Hypertrophy usually occurs after some form of stress. Mild stress, such as exercise, as an example for adaptive hypertrophy or injury or disease as precursor of maladaptive hypertrophy. Distinction of the two types of cardiac hypertrophy is typically discussed controversially with low to medium coherence, however, typically adaptive hypertrophy is usually associated with mild stress and maladaptive hypertrophy with pathological stimuli (Shimizu and Minamino, 2016). Nevertheless, even initial occurrence of adaptive hypertrophy has been shown to transition into a maladaptive response should pathological stress persist (Oldfield et al., 2020). One of the main pathways discussed in association with cardiac hypertrophy is the phosphoinositide 3-kinase (PI3-K) pathway, that regulates metabolic substrate utilization and function of cardiomyocytes (Matsui et al., 2003). Its main inducing pathway, the AGE-RAGE signaling pathway, that by itself is also associated with cardiac stress response (Hou et al., 2014; Shang et al., 2010), was revealed by KEGG-pathway analysis to be among the most affected *Swhtr* dependent pathways in *ex vivo* cardiac slice stress assays (Fig. 20) coherent with failed induction of hypertrophic remodeling. In these cardiac slice stress assays a complete lack of recovery after cardiac stress, induced by hypoxia, could be detected upon absence of *Swhtr*, demonstrated by deregulation of 464 genes in *Swhtr Null* slices as compared to deregulation of 9 genes in WT slices (Fig. 18). Together with GO-term analysis revealing *Swhtr* dependent deregulation of immune response, calcium

signaling, cardiac muscle function and extracellular matrix deposition (Fig. 19) and increased early lethality of *Swhtr Null* mice after AMI, it can be concluded that an initial *Swhtr* dependent hypertrophic response has an inert cardioprotective function for short-term survival following acute myocardial injury.

Swhtr binds to NKX2-5 to regulate cardiac stress response

The *cis* located transcription factor coding gene divergent from *Swhtr*, *Nkx2-5*, is well conserved among vertebrates and best known for its role in embryonic development of the heart (Tanaka et al., 1999). As mice deficient for *Nkx2-5* suffer from early embryonic lethality, its role in an adult context has been studied poorly, despite its abundance. However, human heart diseases, such as dilated cardiomyopathy and sudden cardiac death, are often associated with mutations within the human *NKX2-5* locus (Sveinbjornsson et al., 2018). Strikingly, 98.71% of *Swhtr* dependent deregulated genes were also occupied by NKX2-5, as overlay of these genes with NKX2-5 ChIP-sequencing data revealed (Fig. 21). The heart of zebrafish is characterized by high regenerative capacity as compared to the mammalian heart (Poss et al., 2002). Interestingly, *nkx2-5* and especially the transcriptional programs regulated by *nkx2-5* are essential for repair of the adult myocardium of zebrafish after injury, indicating an important role for *Nkx2-5* in stress response and compensatory mechanisms after injury (de Sena-Tomás et al., 2022). Overexpression of a WT and a dominant negative variant in mice showed a protective role of *Nkx2-5* against doxorubicin induced stress (Toko et al., 2002). Furthermore, over-expression of human *NKX2-5* in mice led to specific deregulation of *Nkx2-5* target genes in the heart but not in the skeletal muscle, consequently only in the presence of *Swhtr* (Takimoto et al., 2000). The group concludes that an additional co-factor that *Nkx2-5* might require is present in the heart but not in the skeletal muscle. This suggests that *Swhtr* might be one of these co-factors. Accordingly, *Swhtr* physically interacts with NKX2-5 protein, showing high variation, but slight increase of that physical interaction under stress conditions (Fig. 22). Absence of a phenotypical manifestation of *Swhtr* depletion in the hearts of embryos, further suggests a stress-specific requirement of *Swhtr* and NKX2-5 interaction while during embryonic development NKX2-5 seems to be able to act in a *Swhtr* independent manner.

LncRNAs in the *Nkx2-5* locus

Within the *Nkx2-5* locus two other lncRNAs that were described as eRNAs have been characterized before. Namely, these are *IRENE-SS* and *IRENE-div*, of which the latter partially overlaps with *Swhtr*, without sharing the same TSS (Salamon et al., 2020). All data about *IRENE-SS* or *IRENE-div* that is discussed in the following were obtained from the cited publication of Salomon et al. While full-length *Swhtr* has been verified by 5' and 3' RACE-PCR (Rogala et al., 2022), the group only identified the full length of *IRENE-div* by 5' RACE-PCR and northern blotting, proposing two versions of *IRENE-div* which are the immature and the mature transcript. However, these two versions stem from presence of a 2 kb northern blot band and one over 9 kb. Neither of these are shown with a negative control nor are the probes used specified. In combination with the proposed 3' end of the lncRNA not matching the RNA-sequencing tracks the group shows and no further validation of the group to prove the existence of *IRENE-div$_{immature}$*, this raised questions about the existence of the lncRNA in the proposed form. Furthermore, none of the cDNA sequences for either of these lncRNAs has been deposited, making it necessary to contact the group personally to receive the sequences. According to this sequence, *IRENE-div* spans 3717 bp, not including the intron. Adding the intron, *IRENE-div$_{immature}$* would span 7411 bp. Both of these lengths do not agree with the observed northern blot bands. As the used probes are not given, it can only be speculated, however, it is more likely that in the case of the over 9kb large band the group referred to as *IRENE-div$_{immature}$*, they detected either *Swhtr* itself, or an alternative variant of *Swhtr* with the included first exon of *IRENE-div*.

To prove the existence of a lncRNA overlapping with *Swhtr*, with a different TSS than the one verified for *Swhtr*, for this work specific qRT-PCR primers taking advantage of the Intron-Exon structure unique for *IRENE-div* were designed and used to distinguish between the two lncRNAs, unlike in their paper, hereby proving existence of an alternative transcript overlapping with *Swhtr*. Efficient knockdown of *Swhtr* using ASO GapmeRs directed towards the 3' end in combination with qRT-PCR primers directed more towards the 5' end proves the existence of the long version that does not contain *IRENE-div$_{immature}$*. However, strongly decreased levels of *IRENE-div* by using these ASOs suggests that the shorter 3717 bp version of *IRENE-div* either is a strongly underrepresented splice variant

or, that it does not exist in the proposed form. Lack of 3' RACE PCR for *IRENE-div* rather points towards the latter, leading to the conclusion that a second splice variant exists for *Swhtr*, most likely spanning over its entire length plus the *IRENE-div* first exon, yielding a 10.288 bp long *Swhtr* isoform, that will be referred to as *Long Swhtr* (*lSwhtr*) in the following.

The function of these two isoforms most likely differ from one another, or the RNA domain required for the observed *Swhtr* dependent phenotype is located exclusively on the only *Swhtr* or the second *lSwhtr* exon, as insertion of the stop signal generated a truncated 2130 bp lncRNA that remains unaffected in its expression level in *Swhtr null* mice compared to the full length in WT (Fig. 3,14). *IRENE-div* and *IRENE-SS* are proposed to have opposed functions on *Nkx2-5* expression. While *IRENE-SS* knockdown decreased *Nkx2-5* expression levels, *IRENE-div* knockdown increased *Nkx2-5* expression levels in HL-1 cardiomyocytes. In partial accordance with these results, a hypoxia induced decrease of *IRENE-SS* and *Nkx2-5* could be seen in this work, although only a correlation and not a causation can be stated as hypoxia induces a vast genetic response and many different genes and networks (Fig. 17). As the group used qRT-PCR primers detecting *Swhtr* as well as *IRENE-div/lSwhtr* their findings of *Nkx2-5* increase should stand true for both lncRNAs, however, despite efficient knockdown of *Swhtr* and *IRENE-div/lSwhtr*, no effect on *Nkx2-5* could be reproduced with two independent ASOs (Fig. 16,17). As *Swhtr* remains at its locus of production under standard conditions and the *Swhtr null* stress-induced phenotype can be rescued by re-expression of the lncRNA from an exogenous locus, transcriptional regulation of *Swhtr* on *Nkx2-5* can be excluded.

Alternative TSS are found for many genes and are discussed controversially. This discussion reaches from alternative TSS being completely nonadaptive and only a consequence of imprecise transcriptional initiation (Akerberg et al., 2019), whereas, in some examples, the alternative transcripts are shown to be mutually exclusive or to have distinct roles, such as stress response. In the example of *Pvt1* lncRNA, two alternative TSS produce separate isoforms *Pvt1a* and *Pvt1b*. Upon genotoxic stress, p53 can activate *Pvt1b* isoform, which regulates the neighboring coding gene *Myc* by p53-independent repression (Olivero et al., 2020), while *Pvt1a* is known to increase *Myc* levels (Tseng et al., 2014). Whether *Swhtr* and *lSwhtr* have separate functions or are just arise

81

from incorrect transcriptional initiation will have to be determined in further experiments, for example by deletion of the alternative TSS and thereby *lSwhtr* knockout with *Swhtr* remaining intact. However, this work clearly shows a cardioprotective role embedded in the RNA of *Swhtr* and maybe also *lSwhtr* second exon.

Conservation of the *Nkx2-5* locus and therapeutic potential

Nkx2-5 is important for heart development, specifically formation of the ventricular structure (Tanaka *et al.*, 1999). Its conservation and importance for embryonic heart development has been shown across mammalian species, down to fish (Tu et al., 2009) but even for function of the invertebrate heart an *Nkx2-5* homologue, namely *tinman*, is known. Importance of this cardiac transcription factor is demonstrated by at least partial functional conservation shared between evolutionary distant species such as mice and fruit flies both of which show *miR-1* dependent *Cdc42* regulation of *tinman*/*Nkx2-5* (Qian et al., 2011). This demonstrates at least a possibility for partial conservation of *Swhtr* dependent function of NKX2-5 across species. Strong conservation of the enhancer element contained within the *Swhtr* TSS that serves as part of its promoter across placental mammalian species suggests presence of a homologous lncRNA. Conservation of this genetic element was strong between human and mice and presence of a lncRNA transcribed from this genetic element could be found in RNA-sequencing from human iPSC derived cardiomyocytes (Fig. 28). This RNA-sequencing data showed strong similarities of arrangement of the murine and human *Nkx2-5*/*NKX2-5* locus, while on-set of *SWHTR*, *IRENE-DIV* and *IRENE-SS* lncRNAs in cardiomyocytes compared to iPSCs could be verified (Fig. 29). Exact annotation of these lncRNAs will need to be verified in further experiments, however, PCR from iPSC derived cardiomyocyte cDNA did not reveal an overlap of *SWHTR* and *IRENE-DIV* so far. Activity of the promoter of *Swhtr* after myocardial infarction (Deutsch et al., 2018) in combination with its strong conservation between human and mouse leads to the hypothesis that the lncRNA loci might share a cardioprotective role in disease settings. Whether this potential cardioprotective role is exerted by binding of *SWHTR* to NKX2-5 cannot be excluded, but given the low sequence conservation of the lncRNA, a different mechanism of action can be possible.

As previously stated, an initial response that subsequently triggers development of cardiac hypertrophy is beneficial for short-term survival of mice after AMI, however, this reaction is coupled to persisting of the arterial blockage in the case of the LAD-ligation mouse model (Kolk et al., 2009). Human AMI is typically followed by intense medical care by appropriate medication to lower the risk of formation of more thromboses or lysis of the present ones, regulation of heartrate and in some cases surgical intervention to restore blood-flow in the form of reperfusion therapy. This medical intervention leads to low mortality rates of 3-7% caused by primary AMI (Gardarsdottir *et al.*, 2022; Rosengren *et al.*, 2001). As persistent cardiac hypertrophy, however, is usually associated with an elevated risk for subsequent CVDs (Guidry et al., 1999; Spencer et al., 2002; Steg et al., 2004), blocking of the initial response leading to cardiac hypertrophy in the combination with reperfusion therapy might lead to a beneficial outcome for long-term survival and quality of life. An attempt to decrease post-AMI hypertrophy in patients by implantation of bone marrow-derived stem cells showed a beneficial effect of decreased end-diastolic wall thickness on contractility of the heart (Rolf et al., 2011). One field of RNA therapeutics focuses on ASO-based drug discovery that already led to FDA approved drugs for various different diseases, such as Fomivirsen (Vitravene™) targeting CMV mRNA and by such significantly reducing progression of AIDS in HIV-positive patients (Vitravene Study, 2002) or Mipomersen (Kynamro™) targeting apo-B-100 mRNA which aids in establishment of favorable low-density lipoprotein levels in patients suffering from homozygous familial hypercholesterolemia (Raal et al., 2010). Should *SWHTR* prove to be responsible for initiation of cardiac hypertrophy in human AMI patients an ASO-based therapeutic approach to reduce *SWHTR* induced hypertrophy in combination with reperfusion therapy to minimize short-term mortality might lead to increased long-term survival and quality of life improvements.

Fendrr synergizes with WNT signaling to regulate fibrosis related genes during lung development via its RNA:dsDNA Triplex Element

Introduction

Fendrr lncRNA is essential for embryonic development and viability

Fetal-lethal non-coding developmental regulatory RNA / Foxf1 adjacent non-coding developmental regulatory RNA (*Fendrr*) is a 2397 nt lncRNA transcribed divergently from the transcription factor coding gene *Forkhead box F1* (*Foxf1*), most likely sharing bidirectional promoter elements. It is expressed in the nascent lateral plate mesoderm (LPM) of developing embryos showing co-expression with its neighboring gene in caudal regions of LPM while it's downregulated upon maintained *Foxf1* expression in the differentiating splanchnic mesoderm (Grote et al., 2013). At later embryonic stages *Fendrr* is also expressed in the frontonasal region of the face, the aorta gonad mesonephros region and, most importantly, in the respiratory system (Sauvageau et al., 2013) (Fig. 30). Replacement of the first exon with a transcriptional stop cassette (*FendrrNull*) caused embryonic lethality and malformed embryos suffering from omphalocele due to reduced thickness of the body wall and hypoplastic hearts that accumulate blood (Fig. 30). In these *FendrrNull* mutants, early embryonic lethality is explained through malfunction of the heart due to deregulation of transcription factors important for cardiac lineage commitment. The lethality and the embryonic malformations can be rescued by re-expression of the lncRNA from an exogenous locus via random BAC integration (Fig. 30), showing that the phenotype is caused by loss of the RNA and also that the RNA itself holds a function. Accordingly, *Fendrr* is able to bind target promoters *in vitro* and tethers histone modification complexes to them *in vivo* (Grote et al., 2013).

Using a different knock-out approach, effectively replacement of the *Fendrr* locus with a lacZ reporter (*Fendrr*[-/-]), a different group shows a different phenotype, but confirmed *Fendrr* being essential for survival. At E12.5 the group observes overtly normal embryos that do, however, display hypoplasia of lungs at E13.5 (Fig. 30) and severe alveolar

defects, weak mesenchymal and smooth muscle layers in mucosa and esophagus and defects of the interventricular septum of the heart at E18.5 (Sauvageau et al., 2013).

Presence of malformations in tissues that do not express *Fendrr* demonstrates importance of the lncRNA for lineage commitment and differentiation. In adult animals expression remains to be most abundant in the lung. *Fendrr*[-/-] embryos do not show the same lethality compared to *FendrrNull* embryos, however, muscle contraction is severely impaired, which in combination with alveolar defects and decreased pulmonary vasculature leads to inability to breathe and thus perinatal death within the first 5 hours after birth (Sauvageau et al., 2013). Despite the differences of the two mouse models, it is clear that loss of *Fendrr* has detrimental effects on the murine embryo development and is essential for life.
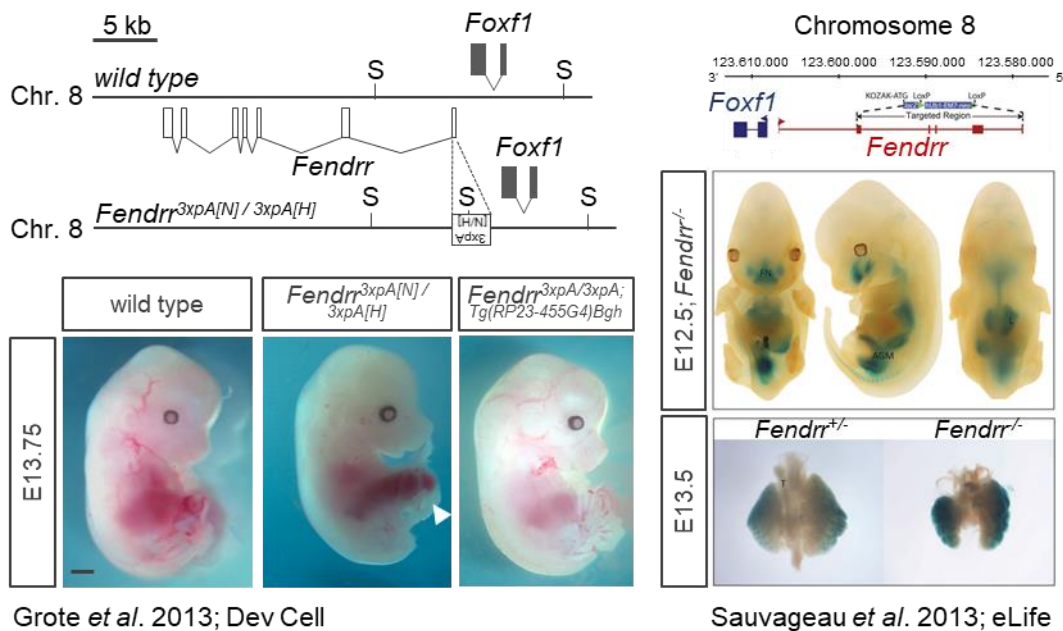
Fig. 30: Importance of *Fendrr* for embryonic development demonstrated by two different knock-out approaches. Homozygous insertion of a transcriptional stop signal (*Fendrr*[3xpA[N]/3xpA[H]]) leads to embryonic lethality that can be rescued by re-expression of *Fendrr* from an exogenous locus by random BAC integration into the *Fendrr*[3xpA[N]/3xpA[H]] background (*Fendrr*[3xpA/3xpA;Tg(RP23-455G4)Bgh]) demonstrated by Grote *et al*., 2013 on the left. The black line represents 1 mm, the white arrowhead points towards the omphalocele. Exchange of the *Fendrr* locus by a lacZ reporter (*Fendrr*[-/-]) allows visualization of *Fendrr* activity in blue and leads to malformed hypoplastic lungs at E13.5 and postnatal lethality demonstrated by Sauvageau *et al*., 2013 on the right. This figure was generated by modifying figures taken from the indicated original publications (Grote et al., 2013; Sauvageau et al., 2013).

Partial requirement of the *FendrrBox*, an RNA:dsDNA triplex formation domain

While loss of *Fendrr* leads to a detrimental phenotype and embryonic lethality, deregulation of genes is only mild. This points towards *Fendrr* being involved in fine-tuning of biological processes, especially in regard to cardio-pulmonary development and that loss of the lncRNA in turn destabilizes entire gene regulatory networks (Grote and Herrmann, 2013; Grote et al., 2013; Sauvageau et al., 2013). Towards the 3'-end of the RNA it contains a low-complexity, UC-rich region of 39 nt that is predicted to form RNA:dsDNA triplexes and tether histone modification complexes, such as PRC2, to target promoters to regulate gene expression by placement of activating or repressing histone methylation marks that can influence target genes in a long-term manner (Grote and Herrmann, 2013). This triplex formation domain has been termed *FendrrBox* (Fig. 31).
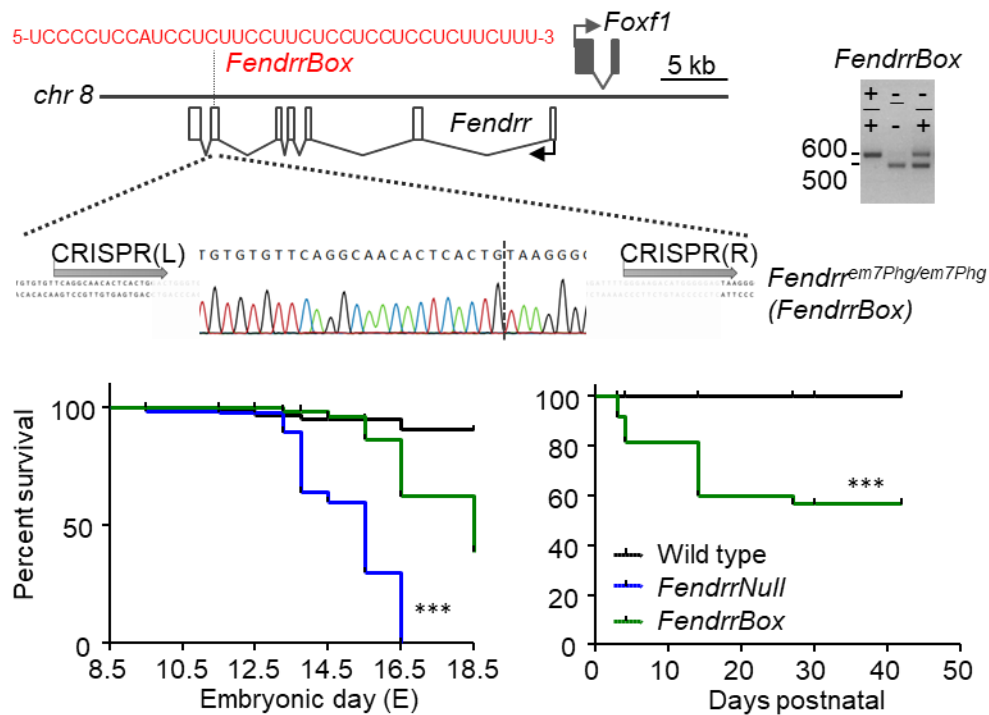
Fig. 31: **The *FendrrBox* element is partially required for *Fendrr* function.** Schematic of localization and deletion (99bp) of the *FendrrBox* element and Kaplan-Meier survival graphs of embryos (lower left) and postnatal mice (right) of the indicated genotype. *** p > 0.0001 by log-rank (Mantel-Cox) test. Figure modified from Ali, Rogala et al., 2023.

Deletion of the *FendrrBox* does not lead to the same embryonic lethality as *FendrrNull*. Only around 50% of the embryos die pre-natal and 60% of the animals that are born,

survive for at least 40 days (Ali, Rogala et al., 2023). This demonstrates that the *FendrrBox* serves a function that is partially required for *Fendrr* function, making *FendrrBox* mutation only incompletely genetically penetrant. As *FendrrNull* mutation is completely genetically penetrant and lethal at day E16.5, lungs of embryos were cultivated *ex vivo* from day E12.5 until E16.5 and used for expression profiling. While the expression profiles of both *Fendrr* mutants show similarities, there are many genes specific for either mutant. However, a subset of 60 deregulated genes is shared between *FendrrBox* and *FendrrNull* with significant correlation in respect to their Log$_2$ fold changes. Overlapping deregulated genes of the two *Fendrr* mutants are mainly associated with lung fibrosis and extracellular matrix deposition (Fig. 32; Ali, Rogala et al. 2023).
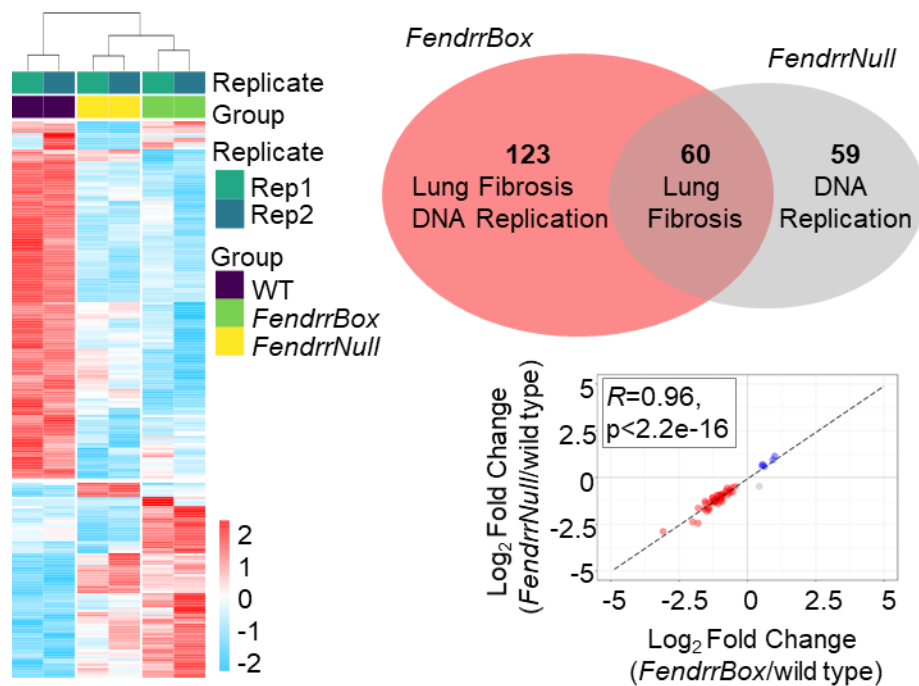
Fig. 32: ***Ex vivo* expression profiling of E16.5 *FendrrNull* and *FendrrBox* lungs.** Heat map of WT, *FendrrNull* and *FendrrBox* lungs show similarities between the two Fendrr mutations compared to WT (left). VENN-diagram depicting deregulated genes and their GO-enrichment analysis (upper right) and scatter dot plot depicting correlation (Pearson correlation coefficient test) between Log$_2$ fold changes of the 60 deregulated genes shared between *FendrrNull* and *FendrrBox* lungs (lower right). Figure modified from Ali, Rogala et al., 2023.

Of those 60 overlapping genes, 20 have significant *FendrrBox* binding domains within their promoter regions (Appendix Fig. 8; Ali, Rogala et al. 2023). Whether or not the *FendrrBox* directly regulates these target genes, however, has not been studied. This work aims to identify the nature and determining factors of this possible regulation.

Early lung development and its link to fibrosis

The first main stage of development of the lung, the pseudoglandular stage, starts at day E9.5 of the murine embryonic development, mainly on the differentiation and cellular level, while substantial structural development continues until day P30 after birth. In humans postnatal structural changes take around one year, while alveolarization can take several years (Shi et al., 2009). *Fendrr* and its most studied target *Foxf1* are mainly expressed in endothelial cells and mesenchymal cells in the developing and adult lung (Curras-Alonso et al., 2023; Negretti et al., 2021; Travaglini et al., 2020). Mesenchymal cells make up for the majority of the cells within the developing lung (Negretti et al., 2021). At day E12.0 around 75% of all mesenchymal cells present are fibroblast cell types such as *Wnt2+* fibroblasts and myofibroblasts. At early developmental stages the vast majority of fibroblasts represent *Wnt2+* fibroblasts whereas a shift towards prevalence of myofibroblasts occurs at post-natal stages. During development the contribution of fibroblasts to the mesenchymal cell niche decreases, however at day P17 the composition of the mesenchyme is still 50:50 between fibroblasts and other cell types demonstrating the importance of fibroblasts for the integrity and function of the lung (Fig. 33; Negretti et al. 2021).

Many of the cellular processes involved in lung development are also key players in lung repair mechanisms and lung fibrosis. Fibroblast growth factors (FGFs) for example regulate the transition from the pseudoglandular developmental stage to later stages, initiating lung branching morphogenesis and promoting cell differentiation (Yang et al., 2021). In a lung fibrosis context FGFs are implicated to have antifibrotic properties (Ghanem et al., 2022; Yang et al., 2021). Additionally, WNT signaling is one of the most intensely studied signaling pathway involved in stem cell maintenance and differentiation and implicated in organogenesis in general. During lung development different WNT isoforms are expressed in various cell types of the three germ layers. Those WNT proteins are involved in differentiation, proliferation and communication between mesenchymal and epithelial cells (Hussain et al., 2017; Rajagopal et al., 2008; Shu et al., 2002). In humans suffering from idiopathic pulmonary fibrosis (IPF) and in IPF mouse models, the WNT pathway is activated in epithelial but also mesenchymal cells and in some cases administration of WNT inhibitors can attenuate fibrosis demonstrating a pro-fibrotic effect

of WNT signaling (Akhmetshina et al., 2012; Konigshoff et al., 2008; Konigshoff et al., 2009). Another important signaling molecule is bone morphogenetic protein (BMP), mostly BMP-4. It is also implicated in cell differentiation, proliferation and survival. Its function has been mainly shown to be important for epithelial and endothelial tissues. Moreover, it is required for proper morphogenesis of proximal-distal patterning of the lung (Eblaghie et al., 2006; Lu et al., 2001; Weaver et al., 1999).
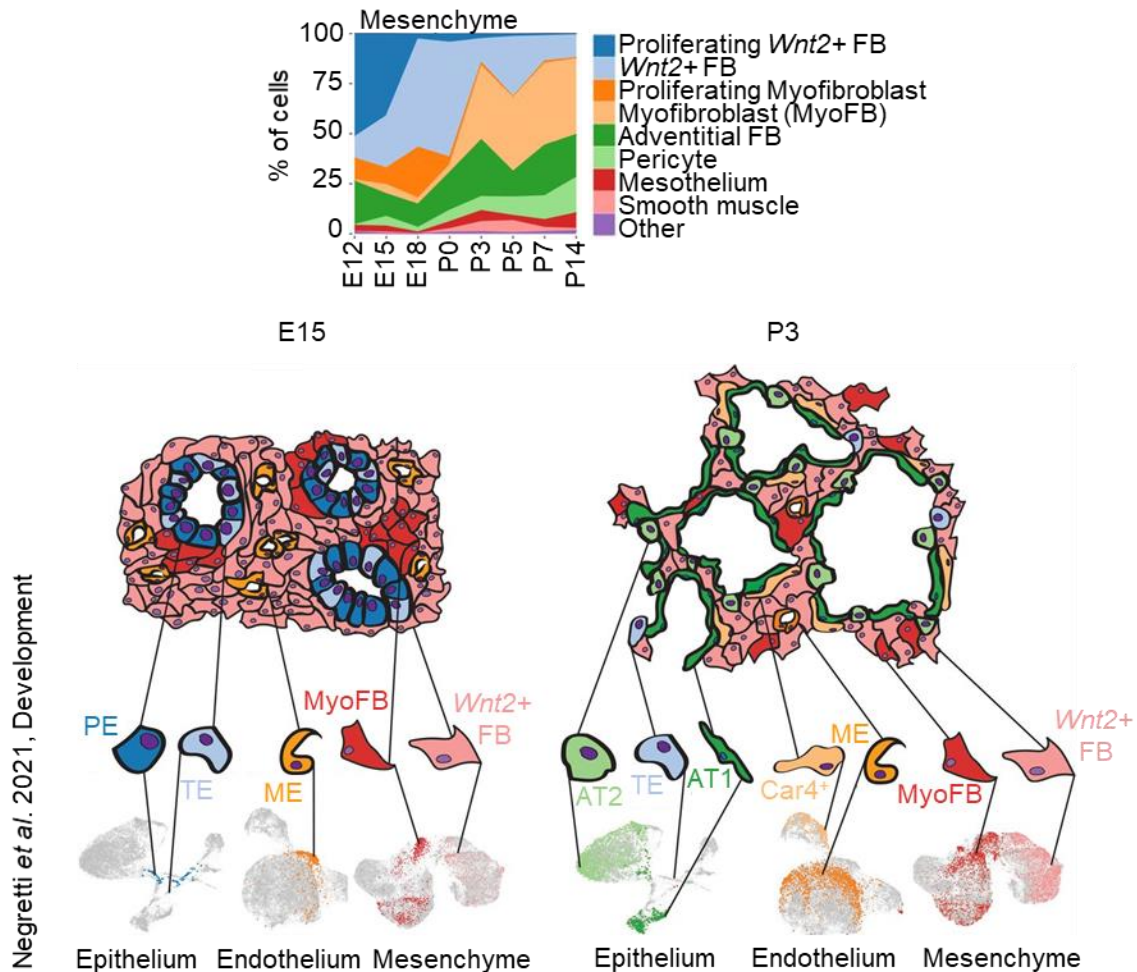
Fig. 33: Cellular plasticity of the developing murine lung. Single-cell sequencing analysis of lungs of the indicated stages showing mesenchymal cell composition (top). Schematic of E15 and P3 lung composition indicating presence and developmental plasticity of the major cell types derived from the indicated germ layer identified by single-cell sequencing (bottom). PE=Primordial Epithelium, TE=Transitional Epithelium, ME=Microvascular Endothelium, FB=Fibroblast, AT2=Alveolar Type 2 cells, AT1=Alveolar Type 1 cells. This figure was generated by modifying figures taken from the indicated original publication (Negretti et al., 2021).

Aim of the study

The lung is mainly comprised of fibroblast cells that are not only important for its development, integrity and function, but also in a disease context. These fibroblasts derive from the mesenchymal niche during development. In lung fibrosis, for example in disease settings like IPF fibroblasts but also mesenchymal progenitor cells play a key role. The involved mechanisms include excessive deposition of extracellular matrix and a resulting positive feedback loop promoting pathological remodeling of the lung tissue (Parker et al., 2014) or providing a self-renewing pro-fibrotic cell niche that supports the disease progression (Xia et al., 2017).

In mouse and human opposing functions for *Fendrr/FENDRR* have been suggested for lung and heart tissue, which leads to the conclusion of additional factors being required for its function. In a transaortic constriction cardiac fibrosis model Fendrr was upregulated and its depletion alleviated the fibrotic phenotype, demonstrating a pro-fibrotic function of the lncRNA (Gong et al., 2020). In a human IPF context, *FENDRR* is downregulated and over-expression of *FENDRR* in a bleomycin-induced lung fibrosis mouse model alleviated fibrosis suggesting *FENDRR* as an anti-fibrotic lncRNA in the human system (Huang et al., 2020). However, the murine and human transcripts do not display high sequence similarity and might even reside in different compartments of the cell. While *Fendrr* is nuclear localized (Grote et al., 2013), *FENDRR* has been proposed as a cytoplasmic localized lncRNA (Huang et al., 2020). Partial requirement of the *FendrrBox* and the rescuable embryonic lethality induced by loss of *Fendrr* argue towards an at least partially RNA based mechanism involving RNA:dsDNA triplex formation (Ali et al., 2023; Grote et al., 2013).

This study aims to identify the effect of *Fendrr* on its fibrosis related target genes in particular in regard to the *FendrrBox* region and possible requirements for additional cues required for specific, cell-type specific *Fendrr* function in fibroblasts.

Results

Fibrosis related *FendrrBox* target genes are not affected by exogenous *Fendrr* over-expression

The lncRNA *Fendrr* contains a low complexity region with the potential for DNA:RNA triplex formation, termed *FendrrBox*. Through different depletion mouse models like blocking of transcription by insertion of a stop signal (Grote et al., 2013) or complete deletion by replacement of the locus with a lacZ reporter gene (Sauvageau et al., 2013) it could be shown that *Fendrr* is essential for life. Mutants with an inserted transcriptional stop signal (*FendrrNull*) start dying at day E13.5 and display complete embryonic lethality at day e16.5. Mutants with a deletion of the *FendrrBox* element, however, exhibit incomplete genetic penetrance with slightly increased embryonic lethality and dramatically increased lethality after birth compared to WT. Involvement of *Fendrr* in lung development (Sauvageau et al., 2013) has been shown previously along involvement of mutations in the human *FENDRR* locus in lung disease (Szafranski et al., 2013). Hence, embryonic lungs of E12.5 WT, *FendrrNull* and *FendrrBox* embryos were cultivated *ex vivo* until E16.5 to avoid a compositional bias caused by premature death of *FendrrNull* animals. These lungs were used for RNA sequencing that revealed that 60 genes are deregulated in *FendrrNull* and *FendrrBox* alike. Pathway analysis revealed that the overlapping, deregulated genes are mainly associated with lung fibrosis while 20 of them showed potential *FendrrBox* triplex target sites at their promoter (Ali, Rogala et al. 2023). In the following these genes will be referred to as **F**ibrosis **R**elated *FendrrBox* **T**arget genes (FRFT genes).
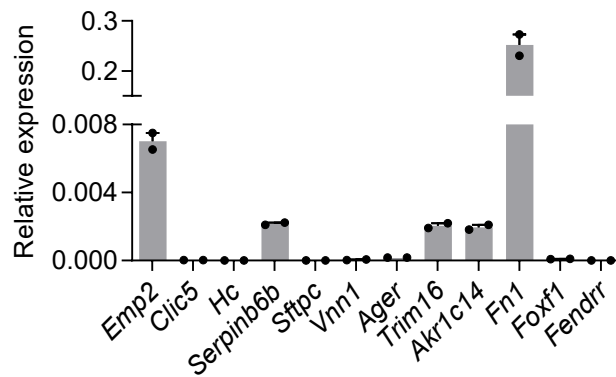
Fig. 34: RNA level of *Fendrr*, *Foxf1* and the top 10 scoring *FendrrBox* target genes in NIH3T3 fibroblasts. (n=2)

The 10 highest scoring FRFT genes were investigated more closely. Due to its location on the X Chromosome *S100g* was omitted. NIH3T3 cells, an embryonic fibroblast cell line, were chosen as a model system as the cells were known to be a suitable transfection host and were hence the prevalent model system in the field. Furthermore, these cells expressed *Fendrr* at very low levels making them an adequate model system for studying over-expression of *Fendrr* with and without the *FendrrBox*. 5 of the FRFT genes (*Emp2*, *Serpinb6b*, *Trim16*, *Akr1c14* and *Fn1*) were abundantly expressed and were subsequently studied further (Fig. 34).

To investigate a possible effect of *Fendrr* on the FRFT genes, first, a plasmid over-expression approach was used. For this purpose, an over-expression plasmid containing the full *Fendrr* sequence downstream of a human β-actin promoter was modified to express the same *FendrrBox* deletion as in the original mutant animals. A plasmid over-expressing a GFP was used as negative control. Cell morphology was monitored microscopically before the cells were harvested for analysis of expression patterns using qRT-PCR. Morphologically, the cells appeared comparable to ctrl transfected cells (Appendix Fig. 9). Furthermore, despite very strong upregulation of *Fendrr* RNA levels, the FRFT genes remained unaffected (Fig. 35).
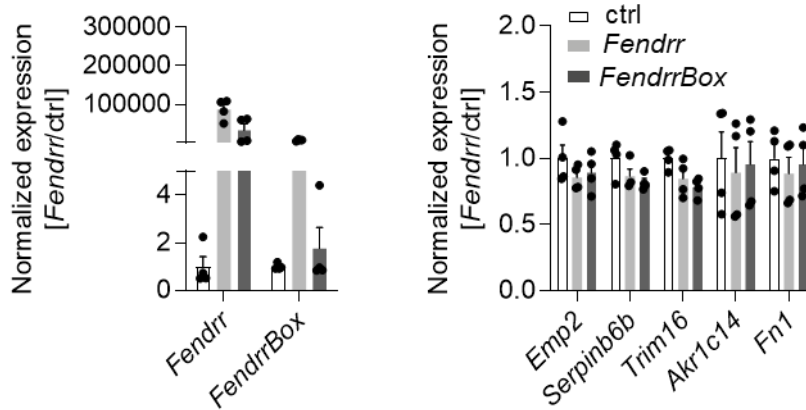
Fig. 35: RNA level of *Fendrr* and FRFT genes after exogenous over-expression of *Fendrr* or *FendrrBox* deletion mutant. Figure legend applies to both graphs. No statistically significant differences were detected in FRFT genes after *Fendrr* over-expression as assessed by t-test analysis (n=4).

Reaction to fibrosis related signaling pathways of fibrosis related *FendrrBox* target genes is unaltered by exogenous over-expression of *Fendrr*

Gene regulatory networks that are not required for homeostasis but rather involved in biological processes, such as fibrosis, however, may need additional stimuli to be affected. As the FRFT genes are enriched in fibrosis, it can be hypothesized that they are regulated by signaling pathways involved in fibrosis and that several or one of them acts in concert with *Fendrr*. Developmental pathways respectively are often shown to be reactivated in the context of disease. Some of these commonly reactivated pathways in fibrosis are Fibroblast Growth Factor (FGF), Bone Morphogenic Protein 4 (BMP-4) and Wingless/Integrase-1 (WNT) signaling; which are also required for proper embryonic development and could hence explain increased embryonic lethality of mutant mice.

Lung fibrosis is associated with unusual activation of fibroblasts, a process that involves FGF signaling (Guzy, 2020). An opposing function has been shown for BMP-4, which can decrease activation of lung fibroblasts and their differentiation into myofibroblasts and hence attenuates lung fibrosis (Guan et al., 2022). WNT signaling was shown to be upregulated in patients suffering from pulmonary fibrosis. Furthermore, its downstream targets have been implicated in the involvement of fibroblast trans-differentiation, fibrosis associated inflammation, extracellular matrix deposition and fibrotic changes in general (Baarsma and Konigshoff, 2017). Thus, these three pathways were stimulated in NIH3T3 cells in combination with exogenous *Fendrr* over-expression.
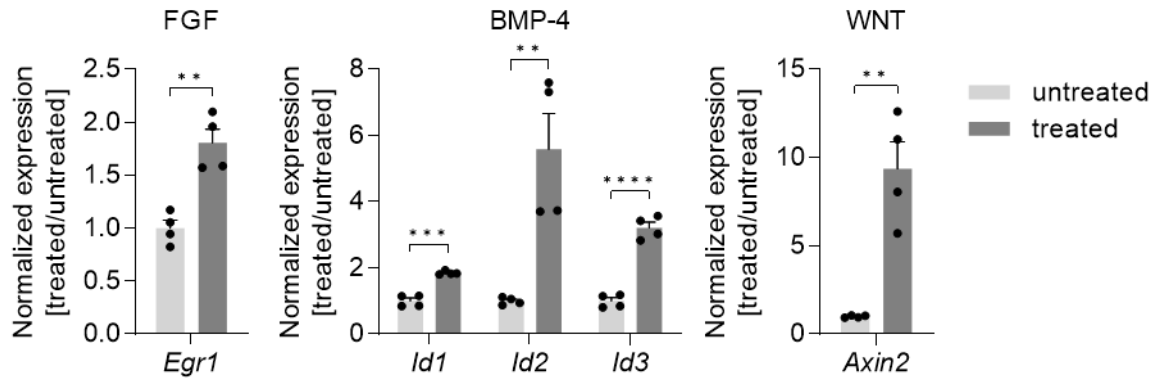
Fig. 36: RNA levels of known target genes of the respective pathway after treatment. Statistical significance assessed by t-test analysis. * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001 (n=4).

In the case of FGF and BMP-4 NIH3T3 cells were stimulated directly by adding the recombinant protein to the cell culture media, whereas for WNT stimulation an inhibitor of Glycogen Synthase Kinase 3 (GSK-3), a known repressor of the WNT pathway, was added to the cell culture medium. To validate successful activation by the treatment the RNA level of known target genes of the respective pathway were measured by qRT-PCR. For FGF a well-studied target gene is *Egr1* (Vasudevan et al., 2015), while BMP-4 is known to induce mRNA expression of Id proteins, *Id1*, *Id2* and *Id3* respectively (Yang et al., 2013). WNT regulates itself within a negative feedback loop that gets activated by WNT mediated induction of *Axin2* transcription (Jho et al., 2002). All these positive control genes were significantly upregulated after the respective treatment, verifying the inductions of the pathways (Fig. 36).

To investigate a possible *FendrrBox* dependent reaction of the FRFT genes in response to the additional stimulus of these signaling pathways, treatment of the NIH3T3 cells was combined with exogenous over-expression of *Fendrr* or the *FendrrBox* mutant. Again, morphology of the cells was monitored microscopically before harvest of the cells for qRT-PCR analysis. They appeared comparable to the GFP transfected control and depicted ordinary fibroblast morphology (Appendix Fig. 10). In terms of RNA levels, the FRFT genes do react to the treatment. However, the *Fendrr* over-expressing cells as well as the *FendrrBox* over-expressing cells reacted to the treatment comparable to control transfected cells (Fig. 37).
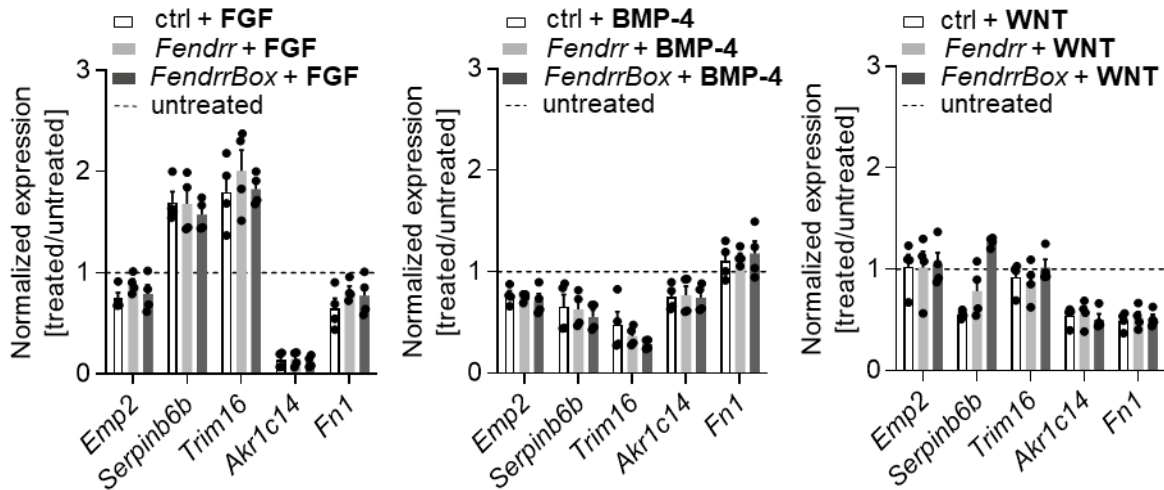
Fig. 37: RNA levels of FRFT genes in response to FGF, BMP-4 or WNT pathway stimulation in addition to exogenous *Fendrr* or *FendrrBox* over-expression. (n=4)

Fibrosis related *FendrrBox* target genes are not affected by endogenous *Fendrr* over-activation

For further investigation, RNA levels of *Fendrr* and its potential target genes were examined in greater detail using MLg cells, due to their lung origin, but also NIH3T3 cells for previously stated reasons. The very low expression levels of *Fendrr* and *Foxf1* in NIH3T3 cells allowed for investigation of reaction of FRFT genes in absence and after on-set of *Fendrr*. MLg cells did indeed express *Fendrr* and *Foxf1* at a higher level compared to NIH3T3 cells as shown by qRT-PCR analysis on a logarithmic scale demonstrating the presence of *Fendrr* in lung fibroblasts (Fig. 38).

Fig. 38: RNA levels of *Fendrr*, *Foxf1* and top 10 FRFT genes in MLg and NIH3T3 cells. The dotted line in NIH3T3 represents the *Foxf1* RNA level. (n=2)

Exogenous plasmid over-expression does not always lead to reliable results, with lncRNA genes respectively. Some lncRNA genes exhibit RNA based mechanisms while in some cases RNA is merely a byproduct of the act of transcription that harbors the function of the locus. In the case of *Fendrr* it is unlikely that the RNA is dispensable for its function as the insertion of a stop signal leads to com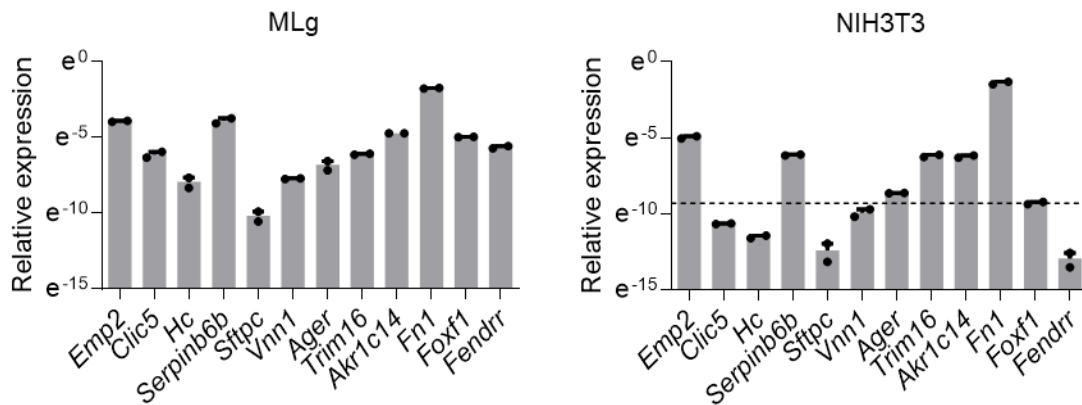plete embryonic lethality and deletion of the triplex formation domain leads to a strong phenotypic manifestation respectively. However, triplex formation relies heavily on RNA localization and structure. Exogenous expression from a plasmid might produce non-functional lncRNA molecules. Additionally, the RNA might be mis-localized and thus being unable to exert its function despite very strong over-expression. To over-activate endogenous *Fendrr*, a catalytically inactive dCas9 fused to activating domains (SAM) was guided to the *Fendrr* promoter using three pooled sgRNAs. The sgRNAs were designed to be less than 100 bp from the *Fendrr* TSS due to its location in close proximity to *Foxf1* TSS which might share some bidirectional promoter elements (Fig. 39).
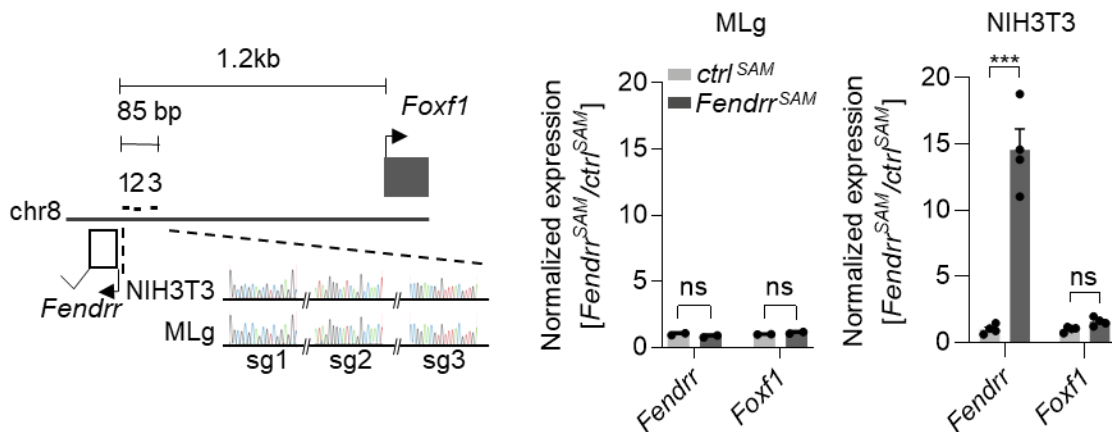
Fig. 39: dCas9 (SAM) mediated over-activation approach of endogenous *Fendrr* in MLg and NIH3T3 cells. On the left a schematic of the approach is shown together with Sanger Sequencing of the targeted promoter region in MLg and NIH3T3 cells. On the right qRT-PCR data showing *Fendrr* and *Foxf1* RNA levels after transfection with the *Fendrr* specific sgRNAs (*Fendrr^SAM*) or a control sgRNA (*ctrl^SAM*) in MLg (n=2) and NIH3T3 (n=4) cells. Statistical significance assessed by t-test analysis. ns = not significant, *** < 0.001

These specific sgRNAs were able to exclusively over-activate *Fendrr* 15-fold without significant changes in *Foxf1* only in NIH3T3 cells but not MLg cells. The *Fendrr* promoter region was amplified from NIH3T3 and MLg genomic DNA verifying the presence of the sgRNA binding sites in both cell types (Fig. 39). For future experiments, NIH3T3 cells

were used. As a previously identified known target gene of *Fendrr*, expression of *Foxf1* was set a cut-off for FRFT genes, gaining *Ager* as an additional target to be studied (Fig. 38). However, even after over-activation of endogenous *Fendrr*, no effect on the FRFT genes could be observed (Fig. 40).
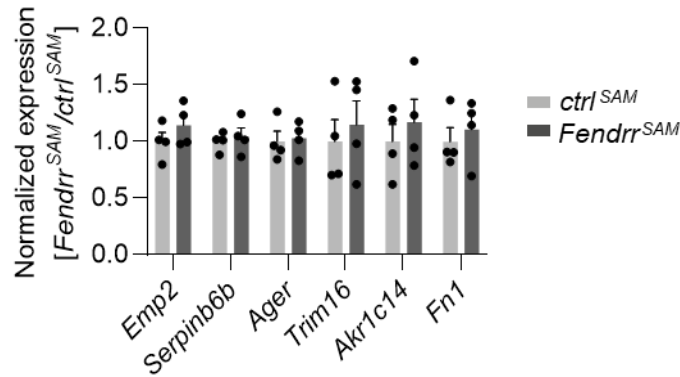
Fig. 40: qRT-PCR analysis of FRFT genes after *Fendrr^SAM* or *ctrl^SAM*. No statistically significant differences could be detected. Statistical significance assessed by t-test analysis (n=4).

Fendrr regulates fibrosis related genes via its RNA:dsDNA triplex formation domain in concert with the WNT signaling pathway

Over-expression of *Fendrr* from an exogenous locus did not lead to a change in expression of the FRFT genes with or without external stimuli. For reasons discussed above that might not be sufficient. Thus, stimulation of the fibrosis related signaling pathways was combined with over-activation of endogenous *Fendrr*. Morphologically, the cells were comparable to the control cells (Appendix Fig. 10). Only minor, but non-significant changes in gene expression could be detected in some of the targets after stimulation of the FGF signaling pathway, while there were no changes after BMP-4 signaling stimulation. Strikingly, after WNT stimulation all measured FRFT genes except *Serbinb6b* showed an altered reaction to the treatment compared to control transfected cells (Fig. 41).
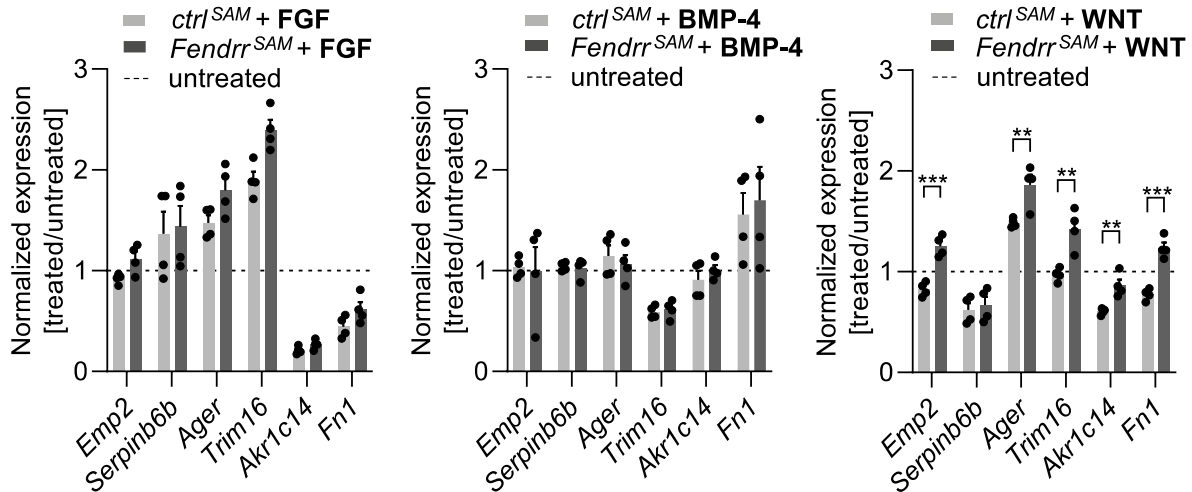
Fig. 41: RNA levels of FRFT genes in response to FGF, BMP-4 or WNT pathway stimulation in addition to endogenous *Fendrr* over-activation. The dotted line represents the untreated ctrl cells. Statistics given when significant as assessed by t-test analysis. ** < 0.01, *** < 0.001 (n=4)

To investigate *FendrrBox* dependency of this effect, the *FendrrBox* was deleted in NIH3T3 cells. For this purpose, a plasmid containing both sgRNAs to introduce the same deletion as in the mouse model and a red fluorescent mScarlet reporter was generated. This plasmid was co-transfected with a plasmid containing a Cas9-T2A-GFP construct to monitor the presence of both plasmids by double-fluorescence. Through FACS-sorting single cell clonal colonies could be obtained and expanded after verification of genotype by PCR (Fig. 42).
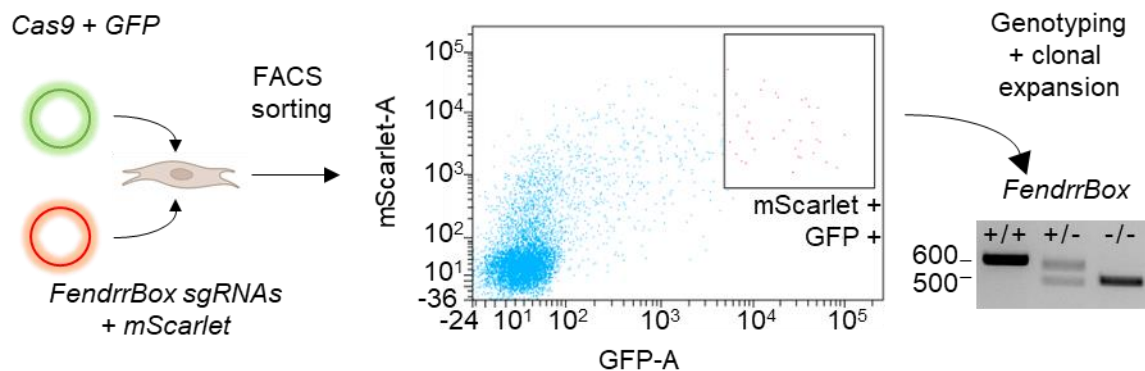
Fig. 42: Schematic of the experimental procedure to obtain mutant NIH3T3 cells. Also depicted is the final gating of the FACS showing the very strict gating strategy sorting only the top fluorescent cell population and a representative agarose gel image showing WT (+/+), heterozygous (+/-) and homozygous (-/-) deletion mutants.

During the FACS sort, GFP fluorescence depicted the typically observed pattern of single fluorescent cells, while only a very low number of cells showed single fluorescence for mScarlet. This happened most likely due to the different size of the plasmids and resulting variability in uptake efficiency. Thus, only the top cluster of double-fluorescent cells were sorted (Fig. 42). After clonal expansion, morphological differences could be detected. NIH3T3 cells are known to be a heterogeneous population (Rahimi et al., 2022), hence cells were selected for depicting typical fibroblast morphology comparable to the parental line (Fig. 43). Hereby, genetic variation was mitigated considerably, although it could not be abolished.
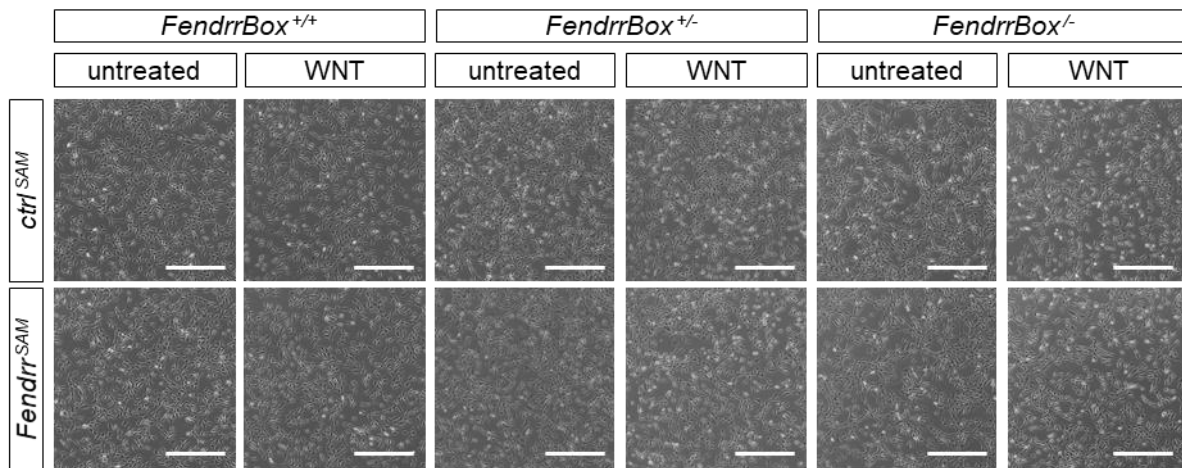

Fig. 43: Representative images of clonal NIH3T3 cell lines of the indicated genotype after transfection with and without WNT activation. The white line represents 500 μm.

The resulting experimental population consisted of 3 WT (*FendrrBox*$^{+/+}$), 3 heterozygous deletion mutants (*FendrrBox*$^{+/-}$) and 5 homozygous deletion mutants (*FendrrBox*$^{-/-}$). Even WT clonal cells showed high variability, however after activation of *Fendrr* in combination with WNT pathway stimulation, the same trend as in the original population could be detected. In the clonal cells, changes of the FRFT genes compared to control transfected cells were not significant, which is most likely due to the clonal differences. On the other hand, neither the heterozygous *FendrrBox* mutants nor the homozygous *FendrrBox* mutants showed any altered reaction to stimulation of the WNT pathway after over-activation of *Fendrr* as compared to the control transfected cells (Fig. 44). This demonstrates the *FendrrBox* dependency of *Fendrr* mediated regulation of the FRFT genes in response to WNT signaling.
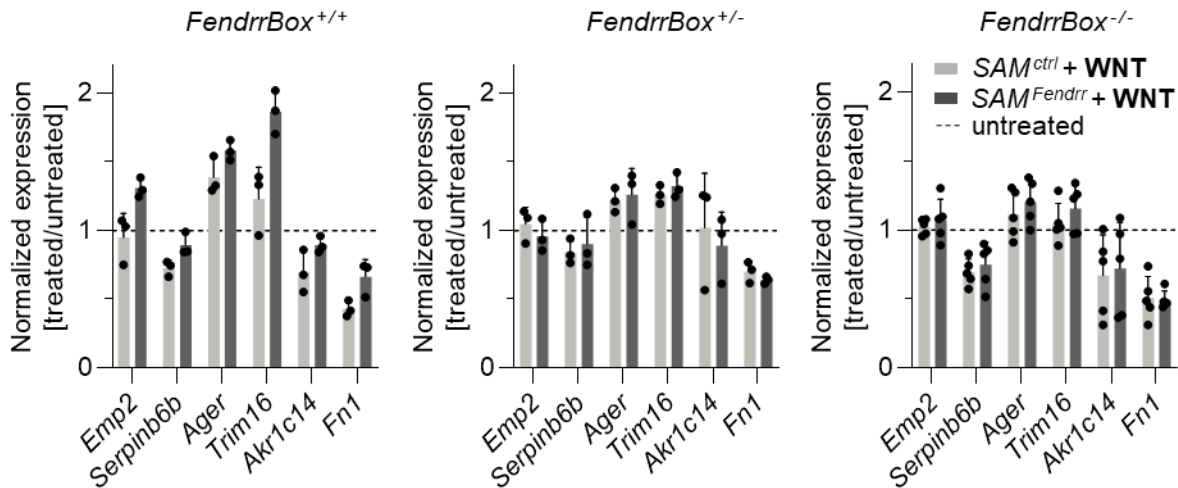
Fig. 44: qRT-PCR Analysis of WNT dependent reaction of FRFT genes to activation of WNT signaling in combination with *Fendrr* over-activation in WT (+/+) , heterozygous *FendrrBox* mutants (+/-) and homozygous *FendrrBox* mutants (-/-). Note that despite clonal variations the WT sorted NIH3T3 showed behavior comparable to the non-sorted population while neither the heterozygous nor the homozygous mutants show altered reaction to WNT stimulation in response to *Fendrr* over-activation compared to the control.

The *FendrrBox* can form RNA:dsDNA triplexes with the promoters of its target genes *in vitro*

The next step was to investigate the triplex formation potential of the *FendrrBox* with the predicted target. Circular dichroism spectroscopy is a method typically used to examine structure or conformational changes of proteins by detection of differences in light-absorbance and scattering between different samples. However, this technique can also be utilized to detect conformational properties of DNA or RNA, such as secondary structure but also triplex formation (Kypr et al., 2009). Furthermore, stability of RNA:DNA triplexes can be used to determine its presence. As RNA:DNA triplex formation typically occurs using energetically unfavorable Hoogsteen bonds as compared to the highly stable Watson-Crick bonds that the DNA double helix uses, the RNA:DNA triplex typically depicts decreased stability as compared to the double helix and a resulting biphasic melting transition. In a thermal melting assay, this effect can be demonstrated by the presence of two distinct melting points of a triplex with the first melting point presenting dissociation of Hoogsteen bonds followed by dissociation of Watson-Crick base pairs at the second melting transition.

Fig. 45: Triplex formation potential of the *FendrrBox* with known negative and positive *FendrrBox* targets. CD spectra (left graph) and thermal melting assays (right graph) for the *FendrrBox* RNA with dsDNA element (red) and the dsDNA element alone (black) for *TGFβ2* as a negative control and the previously identified target *Foxf1* as a positive control. RNA sequence and DNA sequences used are indicated. Watson-Crick base pairing is indicated with | and the Hoogsteen base pairing is indicated with ●.

Using *TGFβ2* DNA with no predicted *FendrrBox* triplex formation potential while having been shown to form a triplex with another lncRNA, *Meg3* (Mondal et al., 2015), and the previously described *FendrrBox* target *Foxf1*, specificity of the occurrence of these effects could be verified (Fig. 45). To investigate the triplex formation potential with the FRFT genes, CD spectra and thermal melting assays were conducted using the predicted *FendrrBox* triplex formation domains within the gene promoter regions. Despite variations between regions, all the investigated regions showed the typical peaks at ~220nm within their CD spectra and more or less prominently the remaining features of an increased peak at ~280 nm and a transition at ~260 nm as well as a slight negative peak at ~240 nm (Fig. 46). Together with the clear presence of a shift of melting temperatures in the thermal melting assays, this demonstrates positive triplex formation potential of the *FendrrBox* with the predicted FRFT genes presenting them as direct *FendrrBox* targets and validating the previous *ex vivo* and *in vitro* data.
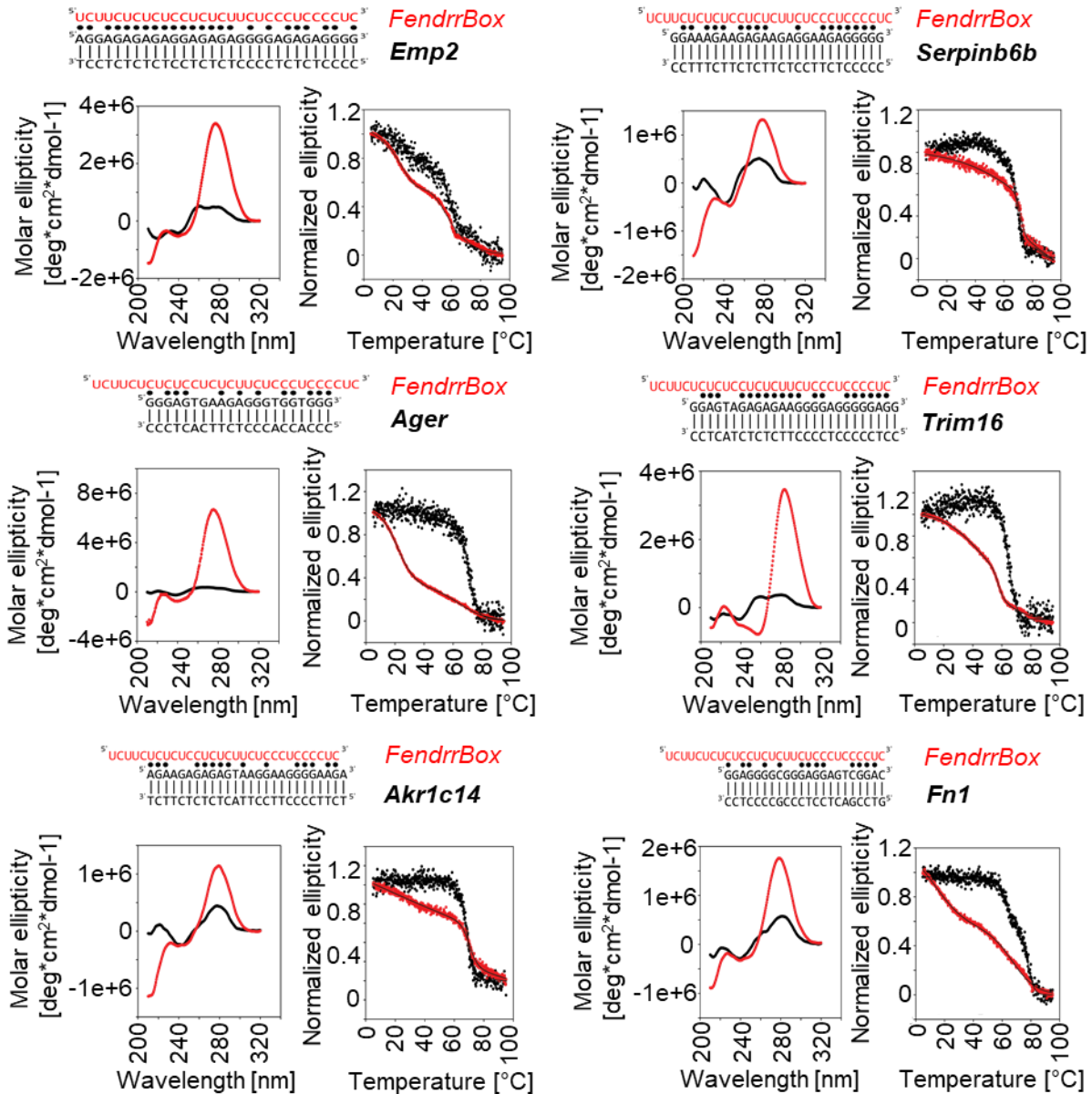
Fig. 46: Triplex formation potential of the *FendrrBox* with predicted triplex formation sites of FRFT genes. CD spectra (left graph) and thermal melting assays (right graph) for the *FendrrBox* RNA with the FRFT dsDNA elements (in red) and the dsDNAs alone (in black). RNA sequence and DNA sequences used are indicated. Watson-Crick base pairing is indicated with | and Hoogsteen base pairing is indicated with ●.

FENDRR, FOXF1 and FRFT genes in the human lung

Despite first being described in mice (Grote et al., 2013; Sauvageau et al., 2013), the human orthologue *FENDRR* has been studied in much greater detail and in various disease settings, such as IPF (Huang et al., 2020) but also cancer (Zheng et al., 2021). However, despite structural similarities of the locus, sequence conservation of the murine

and human lncRNA was found to be low (Fig. 47), as typically observed for lncRNA orthologues in different species (Hezroni et al., 2015; Johnsson et al., 2014). The exact function of *FENDRR* in human embryonic development is not clear at this moment, however, structural similarity of *Fendrr* and *FENDRR* loci (Fig. 47) in combination with similarity of the tissue of their highest expression and a similar disease relevance, allows the hypothesis of involvement of the murine and human lncRNA in similar processes.
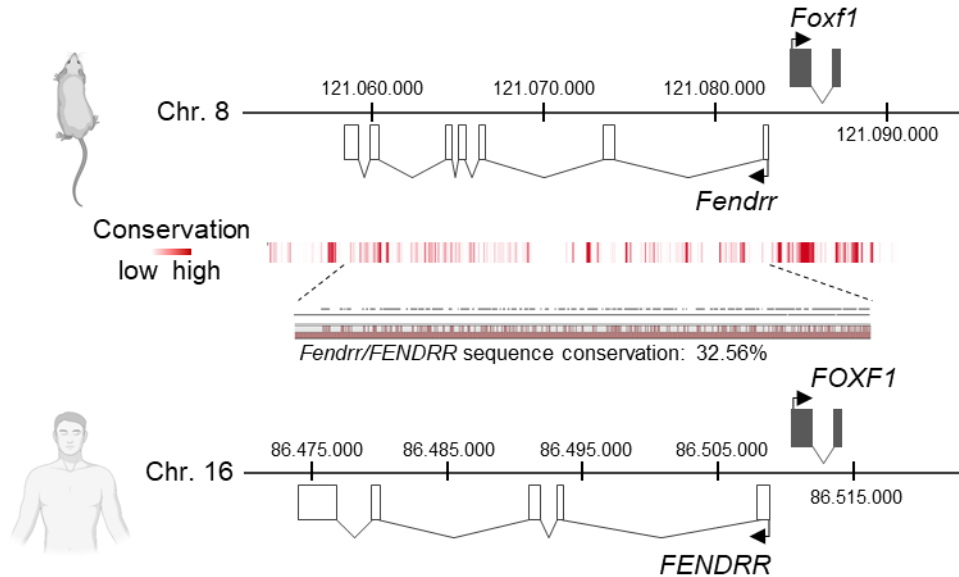

Fig. 47: Conservation of murine *Fendrr* and human *FENDRR*. Schematic of the loci depicting only the major annotated isoform of the lncRNA for simplicity. Heat map derived from WashU Epigenome Browser (Placental PhastCons 60-way). The alignment was generated using cDNA sequences from mouse and human in the CLC software.

Fendrr has been shown to be expressed most prominently in the lung, but also in colon, liver, spleen, brain and testes in the adult mouse (Sauvageau et al., 2013) while *FENDRR* has been shown to have its highest expression in lung followed by urinary bladder, gall bladder, esophagus, colon prostate and intestine (Consortium, 2013; Fagerberg et al., 2014). While a publicly available scRNA-sequencing dataset from murine lungs was not sequenced deep enough to allow expression pattern analysis for lncRNAs, expression of *Foxf1* showed similar patterns for the human version of the mRNA (Curras-Alonso et al., 2023; Travaglini et al., 2020). This suggested a similarity of an expression of *Fendrr* and *FENDRR*, although the exact cell type specificity of *Fendrr* will have to be evaluated further. Within the human lung, most of the cells expressing *FENDRR* were found in the endothelial cell cluster, however, it was also found considerably in the mesenchymal

fibroblast niche (Fig. 48). Expression of *FOXF1* and *FENDRR* in human lung fibroblasts was further validated by analysis of public datasets from RNA-sequencing of human adult fibroblasts (Strand-specific ENCODE polyA plus strand-specific RNA-sequencing data derived from human primary lung fibroblasts; ENCFF173XVB; ENCFF711PVV; ENCFF664JII; ENCFF209PVO) in which both transcripts could be detected along with all human orthologues of the investigated murine FRFT genes (data not shown). As *FENDRR* has been implicated to be involved in IPF (Adams et al., 2020; Huang et al., 2020; Morse et al., 2019) regulation of fibrosis related target genes might be a conserved function between human and mouse.

To further investigate possibility of similar targets for *FENDRR in vivo*, co-expression of *FENDRR* with the human orthologues of the murine FRFT genes was investigated in scRNA-sequencing data that was publicly available (Travaglini et al., 2020). As a known interactor of *FENDRR*, *FOXF1* was used as a positive control, revealing high co-expression with *FENDRR* in endothelial, but also mesenchymal cells. *EMP2*, *SERPINB6*, *AGER*, *TRIM16*, *FN1* all showed at least some co-expression with *FENDRR* in endothelial and mesenchymal cells (Fig. 48, more in Appendix Fig. 12).

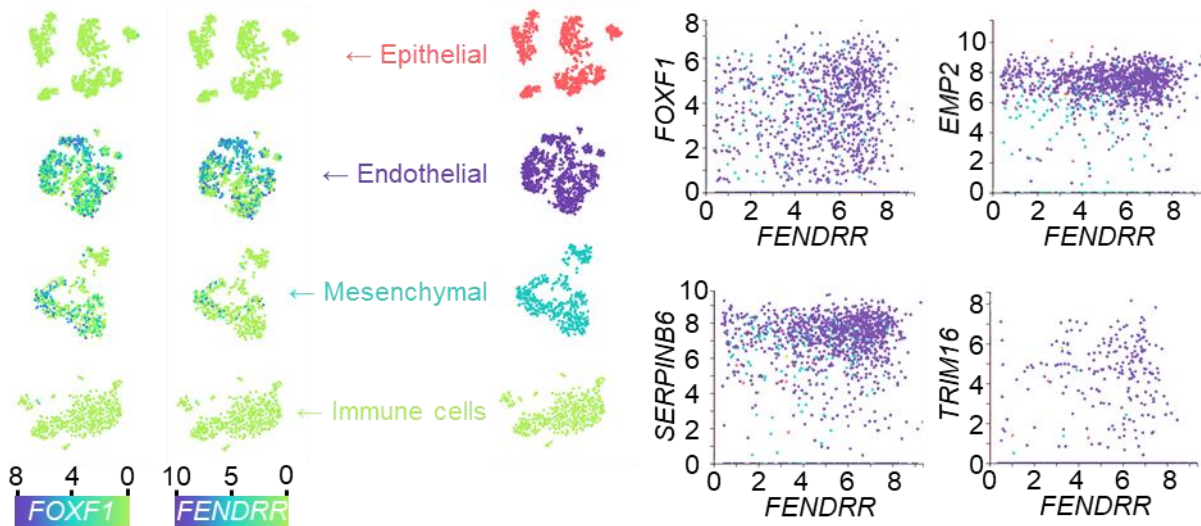**Fig. 48: Co-expression of *FENDRR* with FRFT genes.** Publicly available scRNA-sequencing data from human lungs (Travaglini *et al.*, 2020) depicting epithelial, endothelial, mesenchymal and immune cells with respective expression of *FOXF1* and *FENDRR* indicating expression level by color (left) and dot plots depicting cells co-expressing *FENDRR* with *FOXF1* and selected FRFT genes indicating the cell type by color (right).

Furthermore, the remaining genes of the top 10 predicted FRFT genes that were not expressed in NIH3T3 cells (Fig. 34,38) have been investigated for co-expression with *FENDRR*. While *CLIC5* showed a similar co-expression pattern with *FENDRR* compared to the other genes, *SFTPC* and *C5*, which was found to be the human orthologue to murine *Hc*, were co-expressed with *FENDRR* only sporadically and almost exclusively in endothelial cells. *VNN1* did not show any significant co-expression with *FENDRR*, suggesting it as either a secondary target, or a murine specific *Fendrr* target (Appendix Fig. 12).

In this work, the effect of *Fendrr* on the FRFT genes was investigated in fibroblast cells due to the fibrosis relevance. Whether *FENDRR* shows a similar regulation of fibrosis related genes as *Fendrr* and whether it serves a function in fibroblasts or endothelial cells or in both, will be an exciting field of study for further experiments, to help understand similarities and differences between the two orthologue lncRNAs. Furthermore, investigation of *Fendrr* behavior in endothelial cells and communication between fibroblasts and endothelial cells in relation to *Fendrr* will strengthen the understanding of *Fendrr* function in development and during disease.

Discussion

In this work the effect of *Fendrr* and particularly its RNA:dsDNA triplex formation domain the *FendrrBox* was investigated in fibroblasts. Mesenchymal cells, such as fibroblasts are the most abundant cell type in the developing and adult lung and are next to endothelial cells known to express *Foxf1* and *Fendrr* (Curras-Alonso et al., 2023; Negretti et al., 2021; Travaglini et al., 2020). In these cells *Fendrr* does not regulate its target genes alone but only in concert with activated WNT signaling (Fig. 41). The *FendrrBox* can form triplexes with target promoter DNA *in vitro* (Fig. 46) and accordingly, regulation of the target genes can be alleviated by removal of the *FendrrBox* (Fig. 44).

WNT dependent regulation of fibrosis related genes by the *FendrrBox*

The ability of *Fendrr* to bind target promoter DNA *in vitro* has been shown before. It can bind histone modification complexes to direct them to its target genes (Grote and Herrmann, 2013; Grote et al., 2013). *Fendrr* function is partially dependent on the *FendrrBox*, however, absence does not result in complete lethality as in the *FendrrNull* or *Fendrr*$^{-/-}$ mice (Ali et al., 2023; Grote et al., 2013; Sauvageau et al., 2013). This proves a partial requirement for the *FendrrBox in vivo*. In fibroblast cells presence of *Fendrr* can increase the expression of fibrosis related genes, indicating a function in fibrosis. This increased FRFT gene expression could be seen only upon co-stimulation with a WNT-signaling activating treatment (Fig. 41). WNT-signaling is one of the major signaling pathways required for various different biological processes and is known for its important role during embryonic development. Especially, its role in mesenchymal-to-endodermal crosstalk during lung development, but also cell differentiation and lung morphology. Mutation of WNTs leads to mice that fail to survive, despite some die post-natal, because of lung-malformations and failed or misplaced cell differentiation into the wrong cell identity (Aros et al., 2021; Li et al., 2002; Pongracz and Stockley, 2006).

Fibrosis usually takes place upon injury and is responsible to close up the wound and replace non-viable tissue. Cell migration, extracellular matrix deposition and altered metabolic requirements are all processes involved in fibrosis and use similar pathways as embryonic development does. Normal fibrosis is obligatory for wound healing,

however, when fibrosis becomes mis-controlled and pathogenic, organ failure and death can be the consequence. Canonical and non-canonical WNT signaling and its targets are involved in these processes (Buhling et al., 2004; Pongracz and Stockley, 2006). While FGF and BMP-4 signaling are mentioned frequently in association with lung development and fibrosis (Eblaghie et al., 2006; Yang et al., 2021), *Fendrr*-mediated activation of FRFT genes could only be identified in conjunction with activated WNT signaling (Fig. 41) indicating not only a synergistic but also specific role of the lncRNA within the WNT-mediated lung development in the mesenchymal cells and possibly its reaction to fibrotic stimuli. While *Fendrr* is shown to be a suppressor of target genes in the precursor cells of the heart (Grote et al., 2013), its role as a positive co-regulator of WNT signaling in the developing lung establishes that *Fendrr* can be a suppressor or activator depending on the context. This highlights the cross-talk between lncRNAs and signaling pathways and the versatility with which lncRNAs can function.

Fendrr/*FENDRR* conservation and implication in a disease context

Conservation of *Fendrr* in humans has been studied most frequently in the context of cancer (Li et al., 2015; Xu et al., 2014). In gastric cancer for example, *FENDRR* is differentially expressed in the tumor and is associated with increased tumor invasion and more aggressive and metastatic tumors. Low expression of *FENDRR* in gastric cancer correlates with poor prognosis. Furthermore, in endothelial cells associated with gastric cancer *FENDRR* acts by regulating FN1, which was also detected as one of the FRFT genes, and secreted matrix metalloproteinase, one of the prime targets of WNT signaling in the developing lung (Tamamura et al., 2005; Xu et al., 2014). This implies a similar importance of *FENDRR* in the lung, possibly with similar target genes in mouse and human.

Patients suffer from IPF show increased *FENDRR* levels in lung tissue and is associated with anti-fibrotic properties (Huang et al., 2020). Single-cell RNA-sequencing analysis from human patient and healthy lung explants revealed increased *FENDRR* levels in endothelial cells but also fibroblasts (Adams et al., 2020). Analysis of pathways involved in IPF revealed a high contribution of pathways utilized during lung development (Shi et al., 2009), WNT-signaling being an excellent example (Baarsma and Konigshoff, 2017).

Furthermore, *FENDRR* has been shown to have the ability to regulate β-catenin levels in fibroblasts and reduced fibrotic lesions in an asbestos-induced fibrosis mouse model (Senavirathna et al., 2021). The exact mode of action of *FENDRR* is discussed controversially, as much as even its location. Some groups find it in the cytoplasm whereas other groups find it in the nucleus and others in the cytoplasm (Huang et al., 2020; Wang et al., 2022). In mouse, *Fendrr* located mainly to the nucleus (Grote et al., 2013) with only a small fraction remaining in the cytoplasm (Appendix Fig. 11). In the murine system, it has been shown that the *FendrrBox* is partially required in order for the RNA to exert its function as deletion of only the *FendrrBox* region leads to decreased survival of the mice (Ali et al., 2023), which indicates a requirement for nuclear *Fendrr* RNA to bind DNA targets (Grote and Herrmann, 2013). In this work, no triplex forming domain could be detected in human *FENDRR* (data not shown), however, recently another group stated to have found a triplex formation domain using LongTarget algorithm and reported physical interaction of the lncRNA with DNA target regions. They show downregulation of *FENDRR* in pulmonary artery endothelial cells upon stress and propose an m6A mediated degradation of *FENDRR* which promotes expression of target genes (Wang et al., 2022). Hence, similar mechanisms of *Fendrr* and *FENDRR* cannot be excluded.

Plasmid-overexpression versus CRISPRa

Plasmid-overexpression of genes of interest has advantages and disadvantages. It is a quick and easy method to achieve mostly reliable over-expression of the gene of interest to very high quantities compared to the control situation, especially with the use of a strong promoter contained in the plasmid. The selectivity of the single isoform of the RNA which is expressed from the plasmid can be an advantage, but also a disadvantage. It allows to identify varying functions of isoforms, but diminishes the potential of detecting functions of unknown isoforms for example. In the context of coding genes, presence of an open reading frame and possibility to detect presence of the protein after overexpression makes this method a nice tool to study the effect of over-abundance of a protein or on-set thereof. This is different in regard to lncRNAs.

Especially, lncRNAs that exert an RNA-based function, such as direct binding of proteins or DNA, are highly dependent on correct localization and secondary structure of the RNA. While exogenous overexpression from a plasmid is an artificial process in a different locus, using a foreign promoter on only one selected isoform, overactivation using CRISPRa takes place at the *Fendrr* locus, using the endogenous promoter, leading to overexpression of *Fendrr* in its native form. For that reason, localization of *Fendrr* after overexpression using the two different methods of plasmid-overexpression and CRISPRa was examined. As expected CRISPRa in NIH3T3 cells shows the same localization pattern of *Fendrr* as compared to MLg cells, in which *Fendrr* is active by default. In contrast, plasmid-overexpression shows a more ectopic pattern in which *Fendrr* is distributed equally in the cytoplasm and nucleus (Appendix Fig. 11).

Another important contributing factor to RNA-based lncRNA function is the integrity of its secondary structure. This is especially apparent as even RNAs with little to no sequence conservation among species can show high conservation in terms of secondary structure and exert similar mechanisms in different species (Johnsson et al., 2014). Secondary structure is one of the main contributors to RNA:dsDNA triplex formation, as the triplex formation domain has to be presented in a free RNA region that is not concealed by the structure of the RNA (Li et al., 2016). Inclusion or exclusion of exons, usage of alternative poly(A)-signals or inefficient splicing can all have a big impact on RNA secondary structure. In the case of *Fendrr*, only partial requirement of the *FendrrBox* implies a second RNA:dsDNA triplex independent function for the lncRNA. This indicates a requirement of different isoforms and potential masking of the *FendrrBox* in order to exert *FendrrBox*-independent functions. It has not been determined exhaustively how many different isoforms there are for *Fendrr*/*FENDRR*, although the UCSC database implies 5 isoforms for mouse and 7 for humans (Kent et al., 2002). The nature and mechanism of *FendrrBox*-independent functions of *Fendrr* will have to be determined in further experiments as well as whether and how the *FendrrBox* is masked in order to serve different functions.

Initial characterization of the novel lncRNA *lncFsd2* in relation to its coding *cis*-located gene *Fsd2*

Introduction

The protein-coding gene *Fsd2*

Fibronectin Type III And SPRY Domain Containing 2 (*Fsd2*) is protein-coding gene located on chromosome 7 in the mouse genome and chromosome 15 in the human genome. However, *Fsd2*/*FSD2* is not only conserved between mice and humans, but itself and its paralogue proteins can be detected in many organisms with a chambered heart (Fig. 49). Despite this very high conservation the protein-coding gene remains poorly studied.
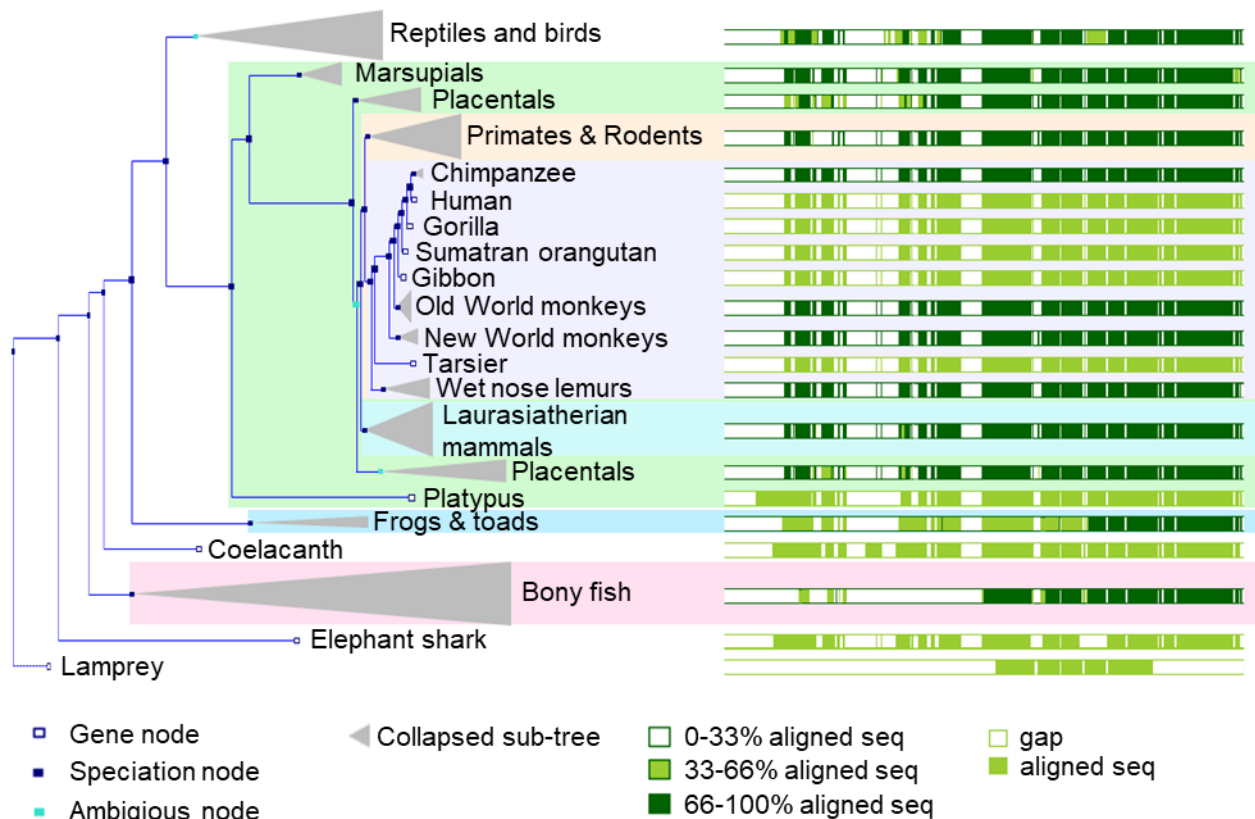

Fig. 49: Phylogenetic tree of FSD2 depicting protein conservation among species. This tree was generated using the Ensembl database (Cunningham et al., 2022).

Expression of *Fsd2* can be detected specifically in heart and skeletal muscle of mice and pigs (Benson et al., 2017; Lim et al., 2016). Interestingly, the gene has two paralogues, *Fibronectin Type III And SPRY Domain Containing 1* (*Fsd1*), *Fibronectin Type III And*

SPRY Domain Containing 1 Like (*Fsd1l*) and *Cardiomyopathy Associated 5* (*Cmya5*). While not much is known about *Fsd1l*, *Fsd1* seems to have specific expression in brain tissue (Carim-Todd et al., 2001). *Cmya5* on the other hand is expressed specifically in striated muscle, namely the cardiac and skeletal muscle (Durham et al., 2006), comparable to *Fsd2*. It has been proposed that the entire genomic clusters surrounding the two genes is the result of conserved chromosomal duplication, as both clusters contain paralogues of contained genes in a similar organization and are conserved even in teleost fish. The *Fsd2* locus, furthermore, harbors a divergent lncRNA downstream of *Fsd2*, that is conserved between human and mouse. However, lack of a divergent lncRNA in the *Cmya5* locus implies that the lncRNA evolved after the chromosomal duplication took place. Despite the difference in size, as FSD2 spans 716 amino acids and CMYA5 3739 amino acids, FSD2 and the CMYA5 C-terminus share the same protein domain architecture (Benson et al., 2017). Furthermore, both proteins co-localize with Ryanodine Receptor 2 in the sarcomeric structures of the muscle Z-lines (Fig. 50) and form a multi-protein complex (Benson et al., 2017).
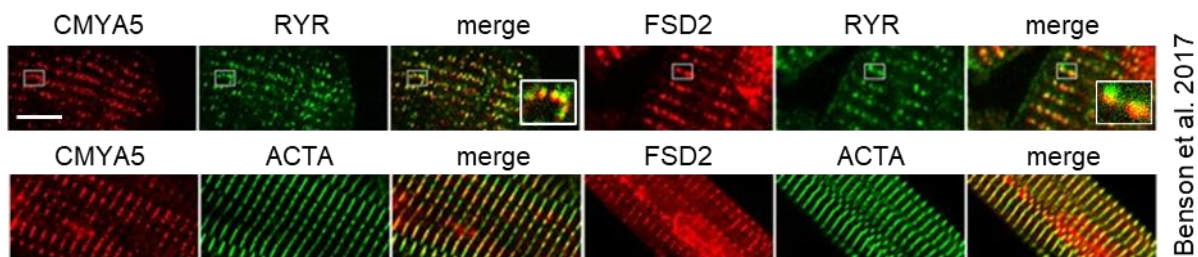

Fig. 50: Co-localization of CMYA5, RYR and FSD2 within the Z-lines (ACTA) of isolated cardiomyocytes. Immunofluorescence of the indicated protein at high magnifications showing co-localization in the merged image. The white line represents 5 μm. This figure was generated by modifying a figure from Benson et al. 2017.

Ryanodine receptors are a known to be an important component of the sarcoplasmic reticulum as they mediate calcium-induced sarcoplasmic calcium release. The three main ryanodine receptors are Ryanodine receptor 1, 2 and 3 (RYR1, RYR2, RYR3) encoded by three independent genes, *Ryr1*, *Ryr2* and *Ryr3*. While *Ryr1* is mainly expressed in skeletal muscle, *Ryr2* is the cardiac ryanodine receptor and *Ryr3* more ectopic, but mostly associated with the brain. Interestingly, the organs with the highest demand for ryanodine receptor function are also the ones with expression of either *Fsd2* or either of its paralogues. Together with the known co-localization of CMYA5, FSD2 and RYR2 in order

to form the myospryn complex (Benson et al., 2017) and the high conservation of the proteins but also the entire genomic clusters, an interaction of these proteins can be hypothesized.

Ryr2 is essential for life. Mutant mice lacking *Ryr2* show embryonic lethality as early as day E10 (Takeshima et al., 1998), but even inducing acute loss of approximately 50% RYR2 protein by cardiac-specific inducible knockout is sufficient to induce cardiac disease such as arrhythmia leading to heart failure and sudden cardiac death (Bround et al., 2012). In human patients suffering from cardiovascular diseases such as catecholaminergic polymorphic ventricular tachycardia type 1 and arrhythmogenic right ventricular dysplasia type 2 mutations in RYR2 are believed to be causative (Gomez and Richard, 2004; Priori and Chen, 2011; Priori et al., 2001).

In contrast, neither *Cmya5* nor *Fsd2* knockout seems to have a comparably strong effect on the heart. The International Mouse Phenotyping Consortium reports abnormal heart morphology and heart weight as well as increased heart rate and shortened RR interval for homozygous *Fsd2* mutant mice lacking Exon 3, while viability of the mice seems unaffected and they are born following standard mendelian ratios. Prediction data furthermore implies *FSD2* involvement in cardiac diseases, in particular those associated with arrhythmia (Dickinson et al., 2017; Groza et al., 2023). Although no further detailed data is available at this moment, this indicates a function of *Fsd2* in the cardiac system, especially in regard to cardiac conduction, which is plausible considering its co-localization with RYR2 and CMYA5.

Cmya5 mutant mice display similar cardiac phenotypes with hearts that are increased in size and weight and show impairment of systolic ventricular function (Lu et al., 2022; Tsoupri et al., 2021). Absence of CMYA5 completely abolishes dyad structure and localization to the Z-disks of RYR2 and FSD2 and most interestingly, neither FSD2 nor re-expression of the CMYA5 C-terminus can rescue that effect completely (Lu et al., 2022). Lu et al. conclude that CMYA5 N-terminus is mainly responsible for structural organization of the complex, however, an N-terminal CMYA5 rescue is missing in the paper despite presence of N-terminus over-expression in WT cardiomyocytes. Furthermore, the reported phenotypes for *Cmya5* mutant mice closely resemble that reported for *Fsd2* mutants and their association with human disease is comparable.

Hence, a partially redundant effect of the two proteins due to their high similarity cannot be excluded, especially, as ablation of neither single one of the genes impacts viability and health of the mice as strongly as loss of *Ryr2*.

Aim of the study

The *Fsd2* locus contains a divergently expressed lncRNA, annotated in the database as *2900076A07Rik*. Nothing is known about the lncRNA so far, however, its close proximity to *Fsd2* might be indicative of a functional relationship between the genes, especially in regard to divergent lncRNAs often regulating their *cis*-neighbors (Mele et al., 2017). Database derived, unstudied RNAs and predicted RNAs often can be incorrectly annotated or even false positives. This work aims to characterize the unstudied lncRNA to verify its presence and annotation. Typical characteristics of the lncRNA, such as conservation, expression pattern and specificity will be evaluated. Especially, the relationship between *2900076A07Rik* and its *cis*-located coding gene *Fsd2* will be investigated using knockdown and CRISPR-Cas9 based approaches. As not much is known about *Fsd2* either, it will be interesting to see, whether the *Fsd2* divergent lncRNA interacts with the coding gene in order to exert a function.

Results

lncFsd2 is a conserved lncRNA convergent to *Fsd2*

The protein-coding gene *Fibronectin type III and SPRY domain containing 2* (*Fsd2*) has been annotated, however, it has been described poorly. The protein is known to be expressed in the heart and skeletal muscle (Benson *et al.*, 2017; Lim *et al.*, 2016), but its function remains yet to be discovered. FSD2 protein co-localizes with the essential cardiac ryanodine receptor 2 (RYR2) and Cardiomyopathy Associated 5 (CMYA5) together forming the myospryn complex (Benson et al., 2017). While for *Cmya5* a malformed heart and impaired function and even embryonic lethality in the case of *Ryr2* was described, the data for loss of *Fsd2* seems more unclear.
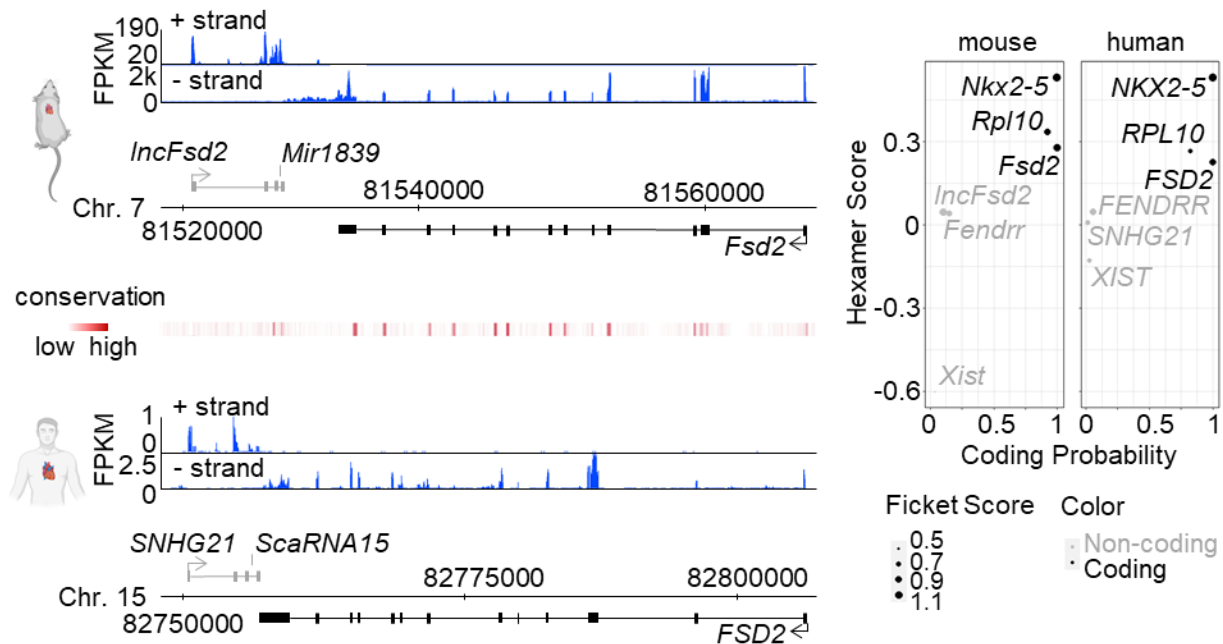

Fig. 51: *Fsd2* **and its convergent lncRNA are conserved between mouse and human.** Strand specific RNA sequencing of murine embryonic heart (E9.5) and publicly available ENCODE data from human iPSC derived cardiomyocytes (ENCFF182KCL, ENCFF571PYD) with conservation data. Conservation heat map derived from WashU Epigenome Browser (Placental PhastCons 60-way; mm10). For simplicity only one of the annotated 5 *SNHG21* isoforms is shown (left). CPAT analysis of murine and human *Fsd2* convergent RNAs compared to known coding and non-coding RNAs revealing absence of coding potential for *lncFsd2* and *SNHG21* (right).

Downstream of *Fsd2* a convergent RNA is transcribed. In the human system it has been annotated as *Small nucleolar host gene 21* (*SNHG21*), as it harbors a microRNA (miRNA), *ScaRNA15*, within its last intron. In the murine system, the RNA has been

annotated as *2900076A07Rik*, harboring the miRNA *Mir1839* within its last intron. The murine RNA will be referred to as *lncFsd2* in the following (Fig. 51). The sequence for *lncFsd2* was obtained by PCR amplification from cDNA derived from murine hearts, demonstrating two splice variants, of which the major one was annotated in the graph (Fig. 51) and the minor one lacks exon 2. Absence of coding potential was verified using CPAT analysis for the human and murine RNAs compared to known coding and non-coding genes (Fig. 51). Typically, lncRNAs evolve faster than mRNAs and thus display poor sequence conservation. However, *lncFsd2* and *SNHG21* show sequence similarity of 49.87%, which is considerably high for a lncRNA, with the 3' end being the lncRNA feature with the highest conservation. Using three different miRNA prediction analysis tools, no significant interaction between the miRNAs contained within the lncRNAs could be detected with *Fsd2*, hence this work focused on characterization of the lncRNA first (Agarwal et al., 2015; Chen and Wang, 2020; Rehmsmeier et al., 2004).

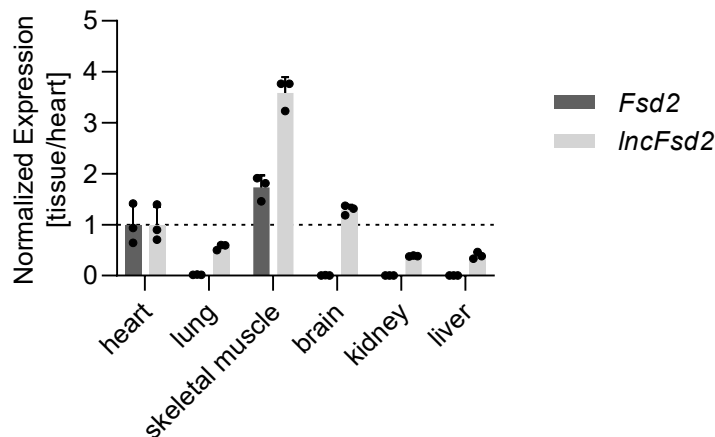*lncFsd2* is expressed ubiquitously as compared to the cardiomyocyte specific *Fsd2*

Fig. 52: Tissue specificity of *Fsd2* and *lncFsd2*. qRT-PCR analysis of *Fsd2* and *lncFsd2* in different tissues from independent adult mice, normalized to RNA levels in the heart (n=3).

So far, *Fsd2* has been described to be expressed in the heart and skeletal muscle (Benson et al., 2017; Lim et al., 2016). To further characterize the expression of the lncRNA in relation to its neighboring coding gene, different tissues from adult mice were

harvested for qRT-PCR analysis of RNA levels of *Fsd2* and *lncFsd2* compared to their respective expression in the heart.

One characteristic of lncRNAs compared to mRNAs is their high tissue specificity compared to mRNAs (Mattioli et al., 2019). In the case of *lncFsd2*, however, its neighboring coding gene was found to be highly specific for muscle tissue, while the lncRNA was found to be ubiquitously expressed with highest expression in heart, skeletal muscle and brain (Fig. 52). In this work, a possible interaction between *lncFsd2* and *Fsd2* was investigated in more detail in the murine system. While in the human system, *FSD2* and *SNHG21* were already annotated to be overlapping (Fig. 51), according to the public annotation of *Fsd2* the mRNA was terminated leaving almost 3 kb of non-transcribed region between the two RNAs (database annotation depicted in Fig. 51). However, in RNA sequencing tracks of the heart (Fig. 51) but also in RNA sequencing tracks derived from different cell types of the heart it could be observed that *Fsd2* seemed to contain a long 3'-UTR overlapping with the lncRNA last two exons (Appendix Fig. 13). Furthermore, this dataset in combination with a publicly available RNA-sequencing dataset from skeletal muscle (GSE123879) did not show expression of *lncFsd2* when the mRNA was present. Publicly available Poly(A) sequencing data (GSE111134) showed the presence of three distinct polyadenylation signals for *Fsd2*, corresponding to the long 3' UTR. To investigate whether the signal was true and the polyadenylation signals were indeed used, 3'-RACE PCR was conducted using RNA isolated from E12.5 hearts, testing for the polyadenylation signal corresponding to the annotated *Fsd2* 3' end but also one that was located within the *lncFsd2* 3' end and one between lncFsd2 exon 2 and 3. This experiment showed positive signals for all three investigated polyadenylation signals demonstrating usage of alternative *Fsd2* polyadenylation signals *in vivo* (Appendix Fig. 13). Combined with the absence of *lncFsd2* in the RNA-sequencing datasets mentioned above, this led to the hypothesis of negative regulation and mutual exclusivity of the two RNAs which was further investigated.
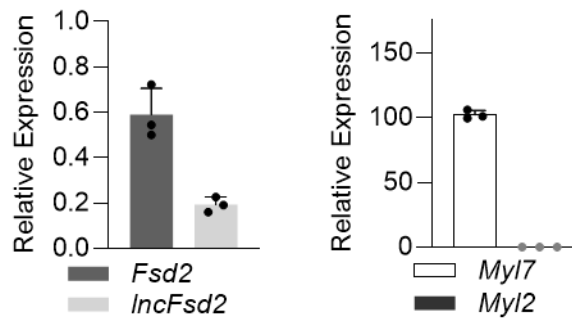
Fig. 53: HL-1 cardiomyocytes are an atrial cell line that expresses both *Fsd2* and *lncFsd2*. qRT-PCR expression data from HL-1 cardiomyocytes showing relative expression of *Fsd2* and *lncFsd2* as well as the atrial marker *Myl7* and ventricular marker *Myl2* (n=3).

HL-1 cardiomyocytes are a murine, immortal cardiomyocyte-like cell line are widely used to investigate cardiomyocyte behavior and function *in vitro* (Claycomb et al., 1998). Under the assumption that they would express only *Fsd2*, expression levels of *Fsd2* and *lncFsd2* were assessed by qRT-PCR, however, counterintuitively *lncFsd2* was also found to be expressed. In order to contract and exert their proper function, muscle cells require the myosin motor proteins, which are specific for different types of muscle. *Myosin light polypeptide 7* (*Myl7*) was previously identified to be typically expressed in atrial cardiomyocytes while *Myosin light polypeptide 2* (*Myl2*) was found to be expressed in ventricular cardiomyocytes. HL-1 cardiomyocytes were derived from an atrial tumor and should hence represent more atrial cardiomyocytes (Claycomb et al., 1998). As they clearly also expressed the atrial marker *Myl7*, it was further investigated whether expression of the lncRNA differs between the heart atria and ventricles. For this purpose, murine hearts were isolated and atria and ventricles separated surgically to assess RNA levels by qRT-PCR. The relative expression of the atrial marker *Myl2* and the ventricular marker *Myl7* were assessed as quality control of the obtained biopsies, revealing successful separation of the heart (Fig. 53).
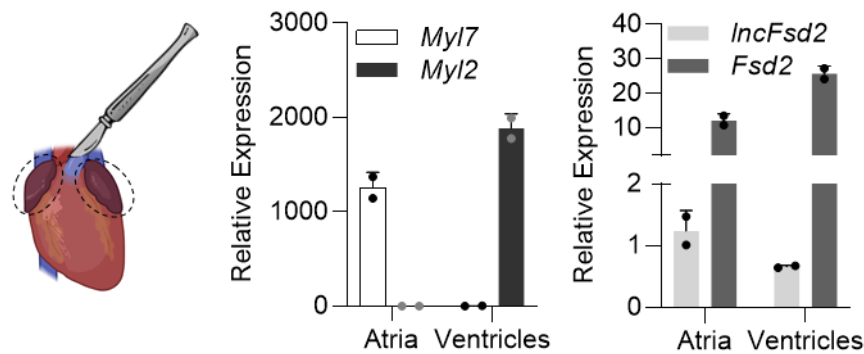
Fig. 54: *lncFsd2* **and** *Fsd2* **are both expressed abundantly in the atria and ventricles of adult murine hearts.** Schematic depiction of dissection of the heart into atria and ventricles (left) and qRT-PCR analysis of atrial and ventricular markers and *lncFsd2* and *Fsd2* (n=2).

While the expression of *Fsd2* was higher in ventricles as compared to atria, the expression of *lncFsd2* was only slightly reduced while still being present in both tissues (Fig. 54). This experiment was conducted, however, using RNA from complete tissue, leaving the possibility of the *lncFsd2* signal being caused by other cell types of the heart. The four main cell types present in the heart were found to be cardiomyocytes, endothelial cells, fibroblasts and immune cells. In order to examine the expression of the lncRNA in cardiomyocytes and non-cardiomyocytes, two different methods were used to obtain an enriched cardiomyocyte population from the embryonic and postnatal heart. The hearts of E12.5 embryos were dissociated into single cells and subsequently FACS sorted using an APC conjugated antibody for cardiomyocyte specific cardiac troponin (*cTnT*) into the CTNT positive population representing cardiomyocytes and the CTNT negative population representing non-cardiomyocytes (Fig. 55). The absence of specific markers for typical non-cardiomyocytes served as quality control for the sorted cell-populations. *Cadherin 5* (*Cdh5*) that was found to encode for a protein mediating cell adhesion at endothelial cell junctions was used as endothelial marker. *Discoidin domain receptor 2* (*Ddr2*), encodes for a tyrosine kinase receptor binding collagen to be activated, and was used as a marker for fibroblasts. For immune cells, protein tyrosine phosphatase receptor type C (*Ptprc*), which was found to be a transmembrane enzyme regulating key functions such as B- and T-cell receptor signaling and exclusive for cells of hematopoietic lineages, was utilized. Enrichment of the cardiomyocyte specific *Nkx2-5* mRNA in CTNT positive

cells over CTNT negative cells further validated the enrichment of the cardiomyocyte population (Fig. 55).

The postnatal hearts were dissociated similarly, however, the non-cardiomyocytes were labelled magnetically and separated from the cardiomyocytes by passing them through a column attached to a strong magnet to collect the enriched cardiomyocytes and subsequently flush out the non-cardiomyocyte population. The quality control was conducted in the same manner as for the embryonic protocol using the same markers showing enrichment of *Nkx2-5* in the cardiomyocyte population in addition to absence of non-cardiomyocyte marker genes (Fig. 55).
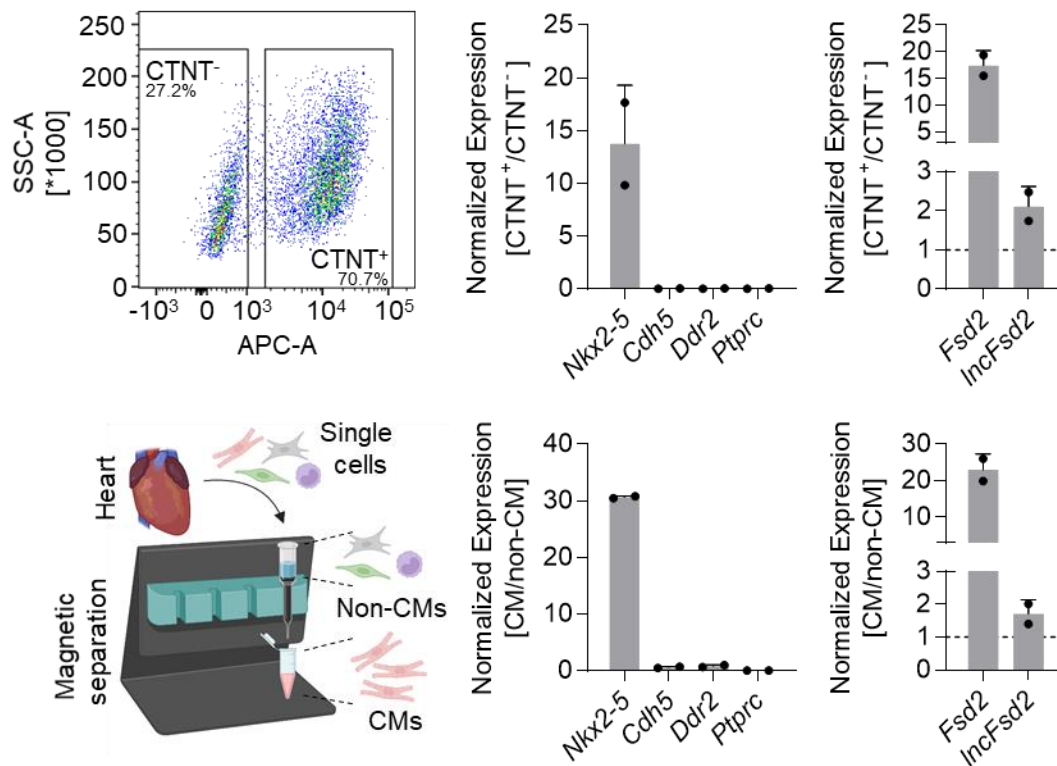

Fig. 55: *Fsd2* is specific for cardiomyocytes while *lncFsd2* is expressed comparably in cardiomyocytes and non-cardiomyocytes. Representation of cardiomyocyte isolation (left; upper left: FACS blot of sorted CTNT⁺ and CTNT ⁻ cells; lower left: Schematic of magnetic separation) with corresponding qRT-PCR analysis of known marker genes and the two genes of interest (middle and right). Relative expression levels normalized to expression of non-CMs (n=2).

In both, the embryonic but also the postnatal heart, *Fsd2* was strongly enriched in the cardiomyocyte population presenting it as a gene specific for muscle cells in the heart. However, even though only slightly, expression of *lncFsd2* was also increased in

cardiomyocytes as compared to non-cardiomyocytes (Fig. 55). While this does not exclude the possibility of negative regulation, it argued against a strong "on-off" effect of mutual exclusivity of the two genes and pointed more towards a subtle effect.

lncFsd2 and *Fsd2* expression do not overtly affect each other

Neuro-2A cells are a glioblastoma derived immortal cell line that is known to be a suitable transfection host. As the cell line does not express *Fsd2* while expressing *lncFsd2* abundantly (information derived from publicly available RNA-sequencing GSM1097466 and later validated by qRT-PCR), it was used to investigate, whether on-set of *Fsd2* might have an effect on *lncFsd2* expression. By targeted mutation of the Cas9 catalytic domains, an inactive Cas9 protein was generated that can bind DNA, while lacking the ability to induce double strand breaks. This dCas9 was fused to VP64, p65 and HSF1 activator domains creating the Synergistic Activation Mediator (SAM) capable of targeted activation of genes of interest by use of single guide RNAs (sgRNAs) designed to target its promoter (Konermann et al., 2015). Using this system, expression a gene from its endogenous locus could be induced, as compared to plasmid over-expression, where the RNA would be generated from an exogenous locus. This could be used to mimic on-set of Fsd2 expression in a cellular system that should not express the gene under normal circumstances, and investigate whether *lncFsd2* expression was influenced by it. By pooling three sgRNAs designed to target the *Fsd2* promoter, *Fsd2* could be overactivated 600-fold as compared to the cells transfected with the SAM construct and a control sgRNA. Despite strong *Fsd2* activation, no effect on *lncFsd2* expression could be observed (Fig. 66).
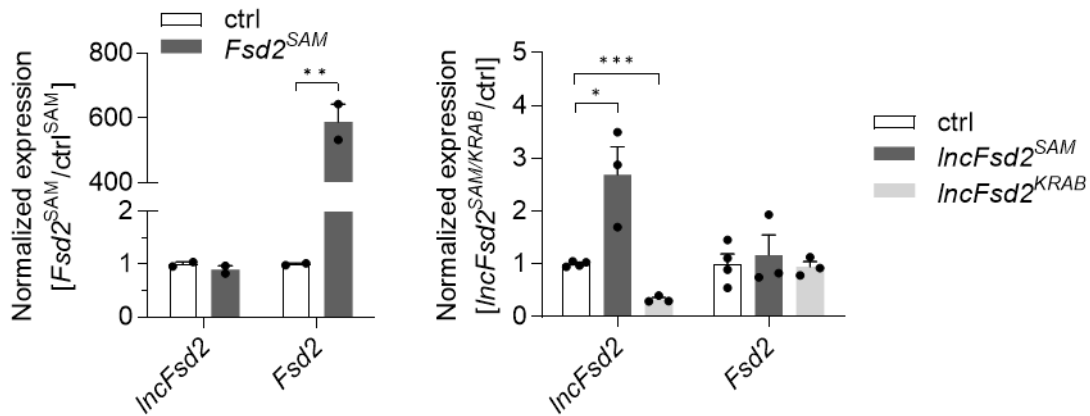
Fig. 66: No effect of dCas9 mediated *Fsd2* or *lncFsd2* activation or repression on each other in Neuro-2A cells. qRT-PCR analysis after 48h of dCas9 mediated activation (SAM) or repression (KRAB) of *Fsd2* or *lncFsd2* in Neuro-2A cells (number of n indicated by dots). Statistical significance tested by t-test analysis and given when significant. * < 0.05, ** < 0.01, *** < 0.001

Fusion of dCas9 to a strong repressive KRAB-MeCP2 domain led to an efficient expression inhibition complex (Yeo et al., 2018). As *Fsd2* was not expressed, repression was only attempted for *lncFsd2*, as well as over-activation. Activation of *lncFsd2* proved to be problematic, most likely due to its abundancy in Neuro-2A cells. Out of 9 tested gRNAs, only 2 were able to lead to a slight 2-3-fold over-activation when combined. Repression using 3 pooled sgRNAs, in contrast, worked efficiently, reducing *lncFsd2* levels by 70-80%. Neither slight up-regulation nor strong downregulation of the lncRNA had an effect on the neighboring coding gene (Fig. 66). This could be, among the absence of an effect, due to the fact that *Fsd2* was already switched off and could not be reduced further or lack of co-factors to induce *Fsd2* expression. To study the effect of *lncFsd2* on *Fsd2* in an *in vitro* system that recapitulated the *in vivo* situation more accurately, HL-1 cardiomyocytes were used. As these cells were not a suitable transfection host, all efforts to use the dCas9 system for over-activation and repression were in vain. Anti-sense oligonucleotide (ASO) mediated knockdown of *lncFsd2* was used as an alternative. 2 out of 4 tested ASOs led to knockdown of 50% or more, however, with no overt effect on *Fsd2* expression (Fig. 67).
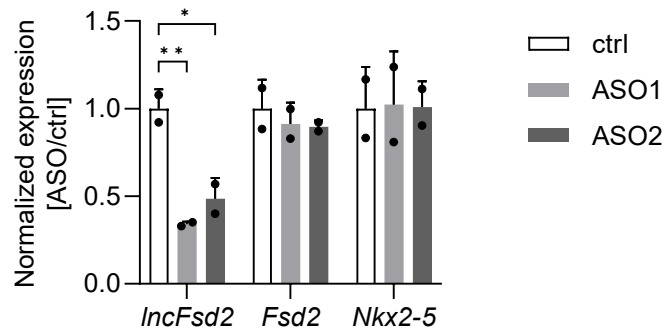
Fig. 67: No effect of ASO mediated *lncFsd2* knockdown on *Fsd2* expression in HL-1 cardiomyocytes. qRT-PCR analysis of HL-1 cardiomyocytes 48h after transfection with lncFsd2 or control ASOs normalized to the control (n=2) Statistical significance tested by t-test analysis and given when significant. * < 0.05, ** < 0.01

FSD2 is localized in the cytoplasm and nucleus of cardiac tissue

HL-1 cardiomyocytes are a widely used experimental system to investigate cardiomyocytes *in vitro*, and were hence used for initial characterization. However, HL-1 cells were found to be a cell line that while resembling expression of primary neonatal cardiomyocytes quite well, the expression patterns were varying from adult cardiac tissue. Expression of major cardiac transcription factors, however, was found to mimic that of adult cardiac tissue well. Furthermore, while organized sarcomeric structures being present, the cells were found to be heterogenous in their population and only a minority depicted this kind of organized structure (Onodi et al., 2022). As FSD2 was found to be involved in the myospryn complex that organizes ryanodine receptor clusters in striated muscle tissue (Benson et al., 2017) it was hypothesized that the protein requires organized structures in order to function. Hence, localization of the protein was investigated in HL-1 cardiomyocytes and adult cardiac tissue by immunofluorescence, to determine similarities and differences and investigate possible limitations of HL-1 cardiomyocytes for this project.
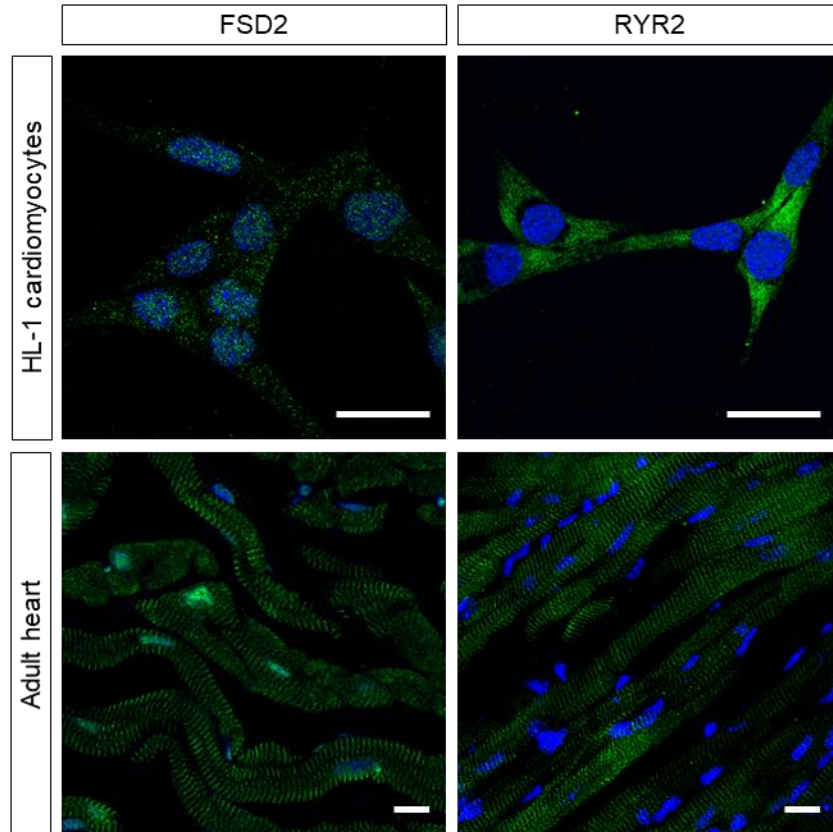
Fig. 68: Localization of FSD2 and RYR2 in HL-1 cardiomyocytes and adult murine cardiac tissue. Representative pictures of immunofluorescence of HL-1 cardiomyocytes and cryosections of 8-week-old adult mouse hearts depicting the indicated protein in green (Alexa-488) and DAPI in blue. The white line represents 50 μm.

While both FSD2 and RYR2 were highly organized in striated structures in cardiac tissue, HL-1 cardiomyocytes did not resemble that degree of organization. Interestingly, RYR2 which localized to the sarcomeric reticulum, was found to be localized exclusively in the cytoplasm, while FSD2 in both samples was present in the cytoplasm as well as in the nucleus, which has not been reported by any other publications (Fig. 68). Upon further examination of the other publications, FSD2 staining unlike the other staining did not show unstained spots where presumably nuclei of the cells were located, despite no definite statement can be made due to missing nucleus counter-stains. In some of the presented images, however, FSD2 positive nuclei can be seen, further validating authenticity of this finding (Benson et al., 2017; Lu et al., 2022). Keeping the differences between the primary tissue and the cell culture in mind, whenever possible, future experiments were conducted *in vivo*, using cardiac tissue.

Loss of *lncFsd2* leads to a mild increase of FSD2

To study the relationship of *lncFsd2* and *Fsd2 in vivo*, *lncFsd2* mutant mESCs to generate *lncFsd2* knock-out mice were produced using the CRIPSR-Cas9 system. In a first attempt, a strong transcriptional SV40-MAZ-bGH poly(A) signal was inserted into the first *lncFsd2* intron to generate a transcriptional stop mutant. This was achieved by transfecting one gRNA together with a single-strand DNA repair template containing homology arms flanking the gRNA cutting site and the stop signal inducing a double strand break and subsequent integration of the mutation by homology directed repair. Integration of the stop signal was verified by genotyping PCR and positive clones were expanded. As *lncFsd2* was also expressed in mESCs, successful knockout could be validated by qRT-PCR from those mutant cells, however, despite insertion of the stop signal with no mutations, in the two independent homozygous clones (*lncFsd2*^{stop/stop}) expression of *lncFsd2* was reduced, but not lost completely (Fig. 69).
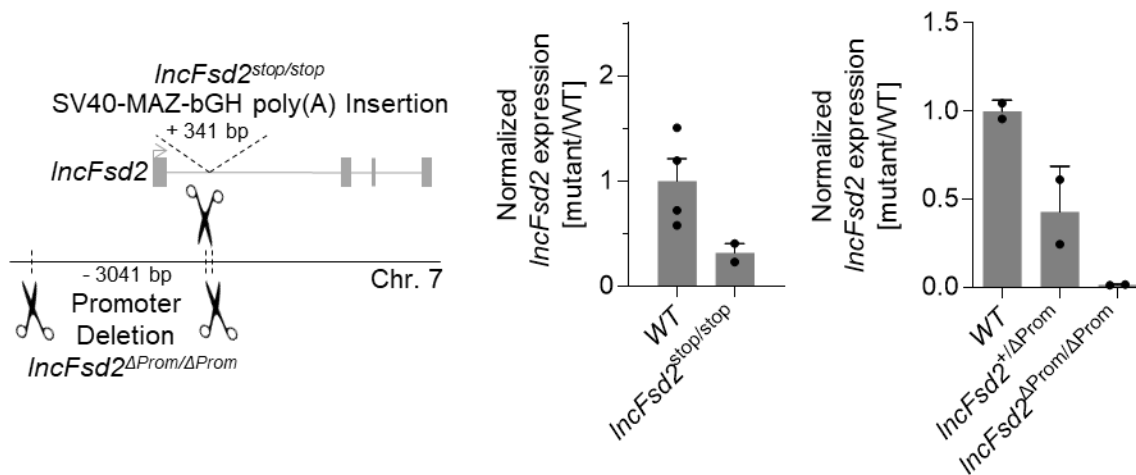

Fig. 69: *lncFsd2* promoter deletion leads to complete loss of *lncFsd2* expression in mESCs. Schematic depiction of CRISPR-Cas9 induced targeted mutations (left) and corresponding qRT-PCR of WT and mutant mESCs (number of n indicated by dots). Each dot represents an independent clone.

To create a mutant with no leftover expression of *lncFsd2*, the promoter of the lncRNA and its first exon was removed by transfection of two gRNAs and subsequent deletion of the sequence by non-homologous end joining. The first pair of gRNAs did not lead to mESC clones harboring the desired promoter deletion, whereas the second pair of gRNAs yielded two heterozygous (*lncFsd2*^{+/ΔProm}) and two homozygous (*lncFsd2*^{ΔProm/ΔProm}) clones that were expanded and tested for *lncFsd2* expression. In the

heterozygous clones, expression was reduced by 50%, in accordance with loss of one allele, and in homozygous clones, expression of *lncFsd2* was lost completely (Fig. 69). The two expanded *lncFsd2^ΔProm/ΔProm^* clones were used for *in vitro* generation of mutant blastocysts and subsequent re-transfer to pseudo-pregnant foster animals (George et al., 2007). However, despite several attempts with both clones no animals or embryos could be generated. As right before the experiment a new batch of mESCs was received, non-transfected wild type cells were used for the same procedure, to assess whether the problem stemmed from the mESCs themselves. From several attempts only one life animal could be generated, leading to the conclusion that the batch of mESCs was not suitable to create mice.

As Prof. Dr. Thomas Boettger from the department of Cardiac Development and Remodeling at the Max-Planck-Institute Bad Nauheim was interested in the same lncRNA and was in the possession of adult *lncFsd2* knock-out mice, harboring a poly(A) signal within the first exon, a co-operation was established. The *lncFsd2* mutant mice were born following standard mendelian ratios and with no overt phenotype. In the following these *lncFsd2* mutant mice will be referred to as *lncFsd2^-/-^*. Investigation of the *lncFsd2^-/-^* mice within Prof. Dr. Boettger's laboratory was directed towards skeletal muscle function, showing no overt defect and no de-regulation in micro-arrays. Snap-frozen heart and skeletal muscle tissue from 8-week-old adult male mice was received and used for qRT-PCR and Western Blot analysis.
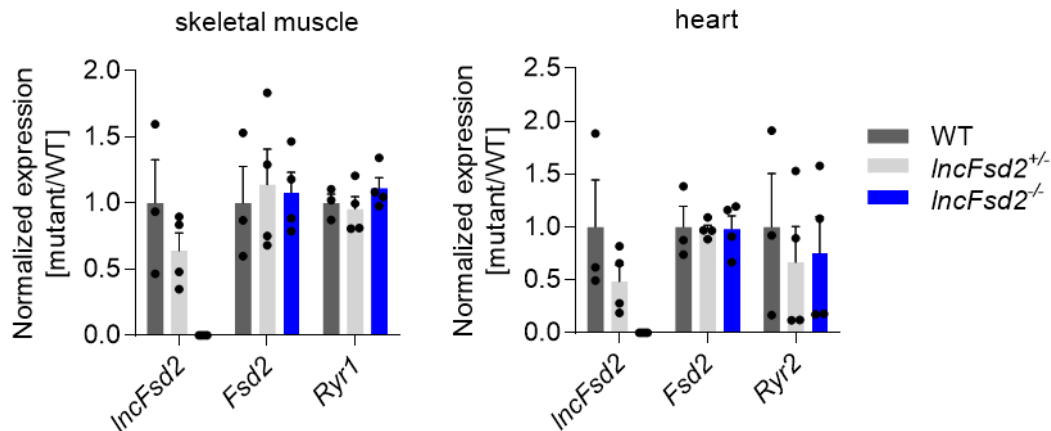

Fig. 70: No differences in *Fsd2* or ryanodine receptor mRNA levels in the skeletal muscle or the heart of *lncFsd2^-/-^* mice. qRT-PCR analysis of samples from WT, *lncFsd^+/-^* or *lncFsd2^-/-^* mice. (WT n=3, *lncFsd2^+/-^* n=4, *lncFsd2^-/-^* n=4)

qRT-PCR showed a 50% reduction of *lncFsd2* in the heterozygous mutant mice (*lncFsd*[+/-]) and complete loss in *lncFsd2*[-/-] mice, demonstrating efficiency of the stop signal as well as specificity of the qRT-PCR primers (Fig. 70). While *Fsd2* showed high variation in skeletal muscle, no deregulation could be determined neither in skeletal muscle nor in the heart. FSD2 has only been shown to co-localize with RYR2, the heart specific ryanodine receptor (Benson et al., 2017), however, it could not be excluded that in the skeletal muscle, FSD2 might interact with RYR1, the skeletal muscle specific ryanodine receptor. Hence mRNA levels of both receptors were assessed in the respective tissue. In the skeletal muscle samples, *Ryr1* remained stable within and between the genotypes. In contrast, *Ryr2* showed high variation in the heart, even between samples from the same genotype. However, no differences between genotypes could be detected (Fig. 70).
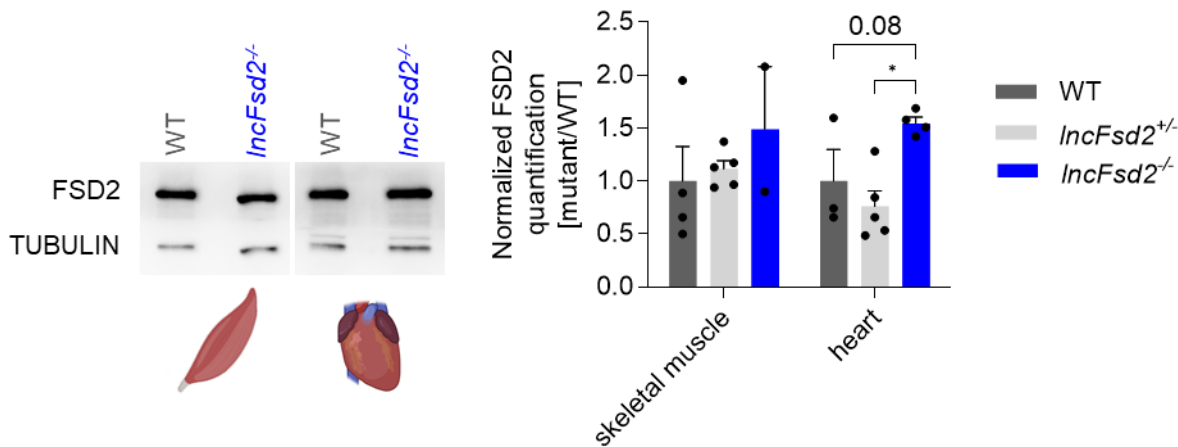

Fig. 71: FSD2 is mildly increased in *lncFsd2*[-/-] heart tissue. Representative Western Blot images from skeletal muscle and heart derived from WT and *lncFsd*[-/-] mice (left) and corresponding protein quantification of all genotypes and samples normalized to the respective WT tissue (right; number of n indicated by dots). Statistical significance assessed by t-test analysis. * < 0.05

In the skeletal muscle derived from lncFsd2 mutant mice, no differences in FSD2 protein levels could be detected. In the heart, on the other hand, a slight increase in FSD2 protein in *lncFsd2*[-/-] hearts was detectable compared to *lncFsd*[+/-] hearts. No significant difference between WT and *lncFsd2*[-/-] hearts could be detected, however, one sample varied strongly from the other two and the p-value was close to 0.05 indicating that the one WT sample might be an outlier (Fig. 71).

FSD2 dual cytoplasm/nucleus localization remains under loss of *lncFsd2*

As the observed effect on FSD2 protein amount in the heart was mild and did not stem from a change in mRNA levels (Fig. 72), it was hypothesized that stability of the protein might be increased upon loss of *lncFsd2*. The protein seemed to be able to shuttle between cytoplasm and nucleus for reasons unknown so far (Fig. 68). In the cytoplasm the protein is part of a complex structure (Benson et al., 2017) and as such might be more stable as compared to remaining in the nucleus. To investigate a possible effect of the lncRNA on localization of the protein, first localization of the lncRNA was determined. For this purpose, three different cell lines were used, NIH3T3 cells which were an embryonic fibroblast cell line with no *Fsd2* expression and HL-1 cardiomyocytes that expressed *Fsd2* as well as C2C12 myoblast cells that did not express *Fsd2* but could be differentiated into myotubes that did express *Fsd2*. Using these four cell types it could be investigated whether localization of the lncRNA might change in presence of *Fsd2*.
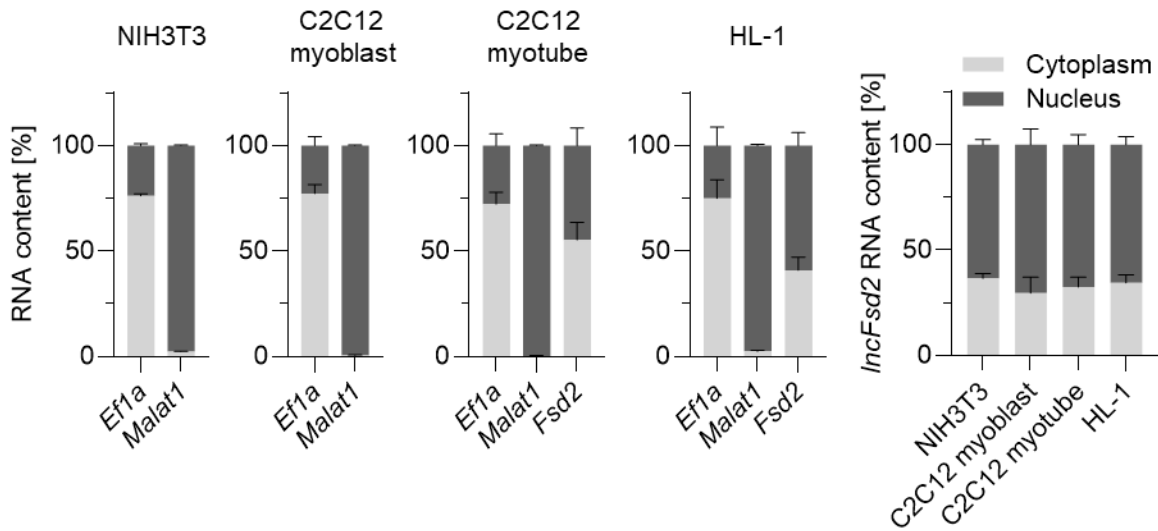

Fig. 72: *lncFsd2* does not change its localization based on *Fsd2* presence. Fractionation of NIH3T3 cells, C2C12 myoblasts, C2C12 myotubes and HL-1 cardiomyocytes into cytoplasm and nucleus. The four graphs on the right show the cytoplasmic control mRNA *Ef1a* and nuclear control lncRNA *Malat1* as well as *Fsd2* in cell types where it was expressed. The left four graphs share the label of the y-axis. Localization of lncFsd2 within the cell types depicted in the graph on the right. Color legend applies to all graphs. (n=3)

Translation elongation factor 1 alpha (*Ef1a*) mRNA was used as cytoplasmic quality control as all mRNAs have to be exported to the cytoplasm for translation. As quality control for the nucleus the well-studied *Metastasis-associated lung adenocarcinoma transcript 1* (*Malat1*) lncRNA, that was found to be highly enriched in the nucleus

(Hutchinson et al., 2007) was used. The marker genes behaved as they were expected with the majority of *Ef1a* mRNA being located in the cytoplasm and *Malat1* exclusively detectable in the nucleus, proving the efficiency of the used protocol (Fig. 72). *LncFsd2* RNA content was stable at around 30% of the RNA present in the cytoplasm while around 70% of the RNA remained within the nucleus. No major difference in RNA localization could be detected among the cell types or in relation to *Fsd2* expression (Fig. 72).

FSD2 was shown to be co-localized with the myospryn complex within the dyad structure of Z-lines in cardiac muscle tissue (Benson et al., 2017; Lu et al., 2022). Furthermore, dual localization of FSD2 in the cytoplasm and nucleus of the cell implies the possibility of the protein to shuttle in response to unknown stimuli. To examine whether *lncFsd2* was involved in that process and could influence localization of FSD2 protein, immunofluorescence stainings from the heart and skeletal muscle of WT and *lncFsd2*$^{-/-}$ mice were obtained. Interestingly, in the skeletal muscle FSD2 could not be found in the nucleus, neither in WT nor in mutant samples, while in the hearts of both genotypes the protein could be detected in both compartments (Fig. 73). Absence of FSD2 staining within the nuclei of skeletal muscle proved the specificity of the antibody, despite absence of a predicted nuclear localization signal (Nguyen Ba et al., 2009) within the FSD2 protein sequence as compared to known nuclear and cytoplasmic proteins (Appendix Fig. 14). However, no difference between the genotypes raised the question whether *lncFsd2* interacted with *Fsd2* at all. The nuclear staining within the heart did seem to be decreased in *lncFsd2*$^{-/-}$, however, quantification was not trivial due to the efficiency of the antibody. In conclusion, it can only be said that ectopic loss of *lncFsd2* does not abolish organization of FSD2 within the dyad structure of Z-lines of striated muscles completely and shuttling of the protein between the cytoplasm and nucleus remains possible.
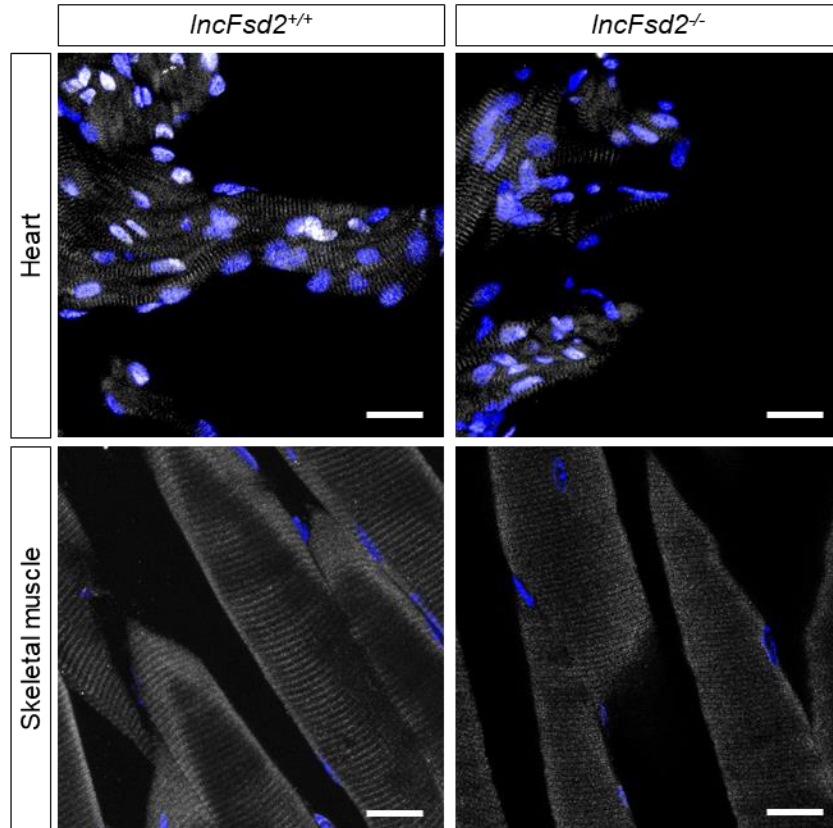
Fig. 73: FSD2 only shows dual localization in the heart with no overt regard to *lncFsd2* expression. Immunofluorescence of 8-week-old adult WT (*lncFsd2*$^{+/+}$) and *lncFsd2*$^{-/-}$ mouse cryosections obtained from heart and skeletal muscle depicting FSD2 in grey (Alexa-647) and DAPI in blue. The white line represents 50 μm.

lncFsd2 possibly interacts with *Fsd2*

To identify, whether *lncFsd2* was able to bind the FSD2 protein, RNA-immunoprecipitation was performed. For an initial test, HL-1 cardiomyocytes were used, despite not representing FSD2 organization accurately. However, all efforts to visualize FSD2 in the protein lysate of HL-1 cardiomyocytes failed, hence pull-down was not performed. Instead, adult heart tissue was used. In the process of obtaining single cells and crosslinking, all RNA was lost from the samples, evident by absence of detectable RNA in the input samples. Thus, RNA-immunoprecipitation was performed using native heart tissue.
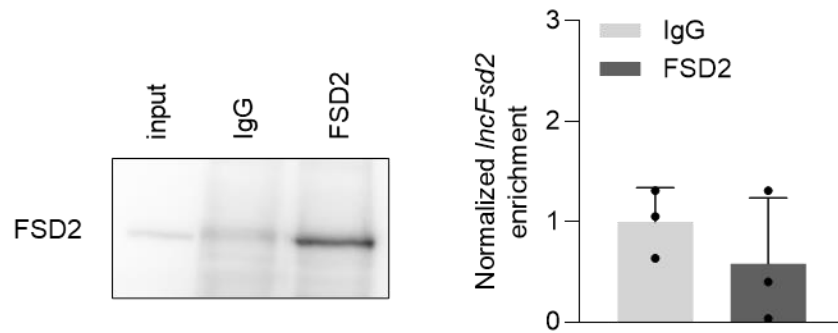
Fig. 74: *lncFsd2* does not co-precipitate with FSD2 in native RNA-immunoprecipitation from adult heart tissue. Representative Western Blot showing enrichment of FSD2 over input in FSD2 immunoprecipitation but not IgG control (left) and qRT-PCR quantification of *lncFsd2* enrichment (left, n=3).

Despite successful pulldown of FSD2 protein, no enrichment of *lncFsd2* could be detected (Fig. 74). One reason for this could be that *lncFsd2* did not interact with FSD2 protein, but without cross-linking of the samples no definitive statement was possible. During lysis of the heart tissue, for which homogenization was necessary, transient or weak interactions could have been broken and thus would not be detectable in this experimental setup. Another possibility would be that *lncFsd2* does not interact with FSD2 protein but rather *Fsd2* RNA. All RNAs are transcribed within the nucleus, however, mRNAs need to be exported to the cytoplasm in order to be translated. Considering *Fsd2* is an mRNA, only a surprisingly small fraction of *Fsd2* RNA was found within the cytoplasm, especially in HL-1 cells (Fig. 72), where FSD2 protein seems to be present, however, in a different organization and undetectable by Western Blot (Fig. 68, data not shown). Furthermore, analysis of RNA-RNA interaction (Mann et al., 2017) revealed a possible interaction of the *lncFsd2* 3'-end with the *Fsd2*-3'UTR even upon using the annotated *Fsd2* version with no overlap to *lncFsd2* as input sequence. Thus, *lncFsd2* RNA was pulled down using biotinylated probes designed against the RNA in order to investigate co-precipitation of *Fsd2* RNA. As a negative control, probes designed against *Swhtr* were used. Enrichment of *lncFsd2* exclusively using the *lncFsd2* but not the *Swhtr* probes and vice versa validated the specificity of the probes. Only a mild enrichment of *Fsd2* in the *lncFsd2* pulldown could be measured compared to the *Swhtr* pulldown (Fig. 75). Despite high variation, and only mild enrichment, *Fsd2* did precipitate with *lncFsd2* compared to *Swhtr*.

The mild effect in HL-1 cardiomyocytes might be due to a difference in FSD2 organization and an associated different behavior of *Fsd2* as compared to cardiac tissue *in vivo*.
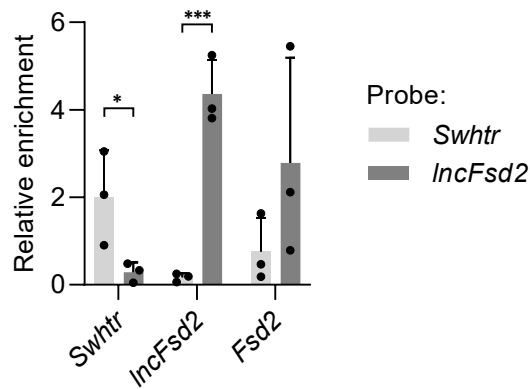

Fig. 75: *Fsd2* mildly co-precipitates with *lncFsd2* in HL-1 cardiomyocytes. qRT-PCR analysis of *Swhtr* (light grey) and *lncFsd2* (dark grey) RNA pulldown (n=3). Statistical significance tested by t-test analysis and given when significant. * < 0.05, *** < 0.001

Whether this observation pointed towards a mechanism relevant *in vivo*, could not be determined within the scope of this work. However, interaction of *lncFsd2* and *Fsd2* would be one explanation of the conspicuous localization of the mRNA and a resulting mild increase in FSD2 protein in absence of the lncRNA. In following experiments, it will have to be determined whether *lncFsd2* is responsible for the nuclear retention of *Fsd2* mRNA.

Discussion

In this work the relationship of an under-studied protein-coding gene *Fsd2* and its neighboring convergent lncRNA *lncFsd2* was characterized. Both genes were found to be highly conserved among placental animals with especially the lncRNA's 3' end showing notable sequence similarity in regard to being a non-coding RNA (Fig. 51). Natural antisense transcripts are an abundant class of non-coding RNAs and are frequently well conserved with their coding sense transcripts and often share a partial 3' overlap. In many cases they show negative correlation in terms of expression with their coding neighbors due to negative regulation *in vitro* and *in vivo* (Modarresi et al., 2012) but they can also be involved with the choice of alternative polyadenylation signals of the mRNA. These convergent lncRNAs are generally believed to often act in *cis* (Gil and Ulitsky, 2020), which would point towards an interaction of *lncFsd2* and *Fsd2*, that might be conserved between mouse and human due to high structural and partial sequence-conservation of the *Fsd2*/*FSD2* locus (Fig. 51).

Interaction of *lncFsd2* and FSD2

The expression pattern of *Fsd2* was confirmed to be specific for heart and skeletal muscle tissue as shown previously (Benson *et al.*, 2017; Lim *et al.*, 2016) and furthermore, within the heart shown to be strongly enriched in cardiomyocytes (Fig. 52,55). *LncFsd2*, on the other hand, was found to be expressed in all tissues examined and comparably expressed in cardiomyocytes and other cell types present in the heart, being only slightly enriched in cardiomyocytes (Fig. 52,55). Strikingly, on protein level, a minor increase in FSD2 protein could be detected in the heart, but not skeletal muscle, of *lncFsd2*$^{-/-}$ mice. In heart tissue, FSD2 has been shown to co-localize with RYR2 and form the myospryn complex (Fig. 50, 58). This complex was shown to be involved in the organization of ryanodine receptor clusters, whereas FSD2 alone lacks the ability to facilitate the assembly of the clusters on its own, leaving the exact function of FSD2 within the complex in the unclear (Lu et al., 2022). It can, however, be speculated that it holds a similar function within the skeletal muscle, as ryanodine complexes are essential for function of the tissue, the predominant ryanodine receptor being RYR1. Interestingly, FSD2 showed dual localization in the cytoplasm and the nucleus in heart but not skeletal muscle (Fig.

73). Proteins can harbor varying functions in the cytoplasm and nucleus, as well documented on the example of Argonaute RISC Catalytic Component 2 (AGO2), which was mainly studied for its involvement in RNA cleavage and silencing, however, it was also found to be imported into the nucleus where it participated in transcriptional regulation (Ross and Kassir, 2014). Interestingly, AGO2 much like FSD2 lacked a nuclear localization signal (Appendix Fig. 14) despite being shown to shuttle between cytoplasm and nucleus (Sharma et al., 2016). This dual localization in the heart and exclusive cytoplasmic localization within the skeletal muscle seemed unaltered by the absence of *lncFsd2*, although a minor change would not have been noticed within the scope of this experiment (Fig. 73).

Interaction of *lncFsd2* and *Fsd2*

In cells with no canonical expression of *Fsd2*, no regulation on the transcriptional level of either of these genes on each other could be found by dCas9 mediated activation or inactivation of both genes (Fig. 66). By ASO mediated *lncFsd2* knockdown in HL-1 cardiomyocytes, that resemble the expression pattern of primary cardiomyocytes (Fig. 67), while not mimicking structural organization of the protein (Fig. 67) no *Fsd2* deregulation on the transcriptional level could be detected. Due to significant differences of HL-1 cardiomyocytes to primary cardiomyocytes in terms of structural organization and transcriptome (Onodi et al., 2022), interaction between the two genes was investigated in transcriptional *lncFsd2* stop mutants, with no overt deregulation of *Fsd2* (Fig. 70).

In this study, it was found that around half of the *Fsd2* mRNA present in the cell, is retained within the nucleus. This holds true for HL-1 cardiomyocytes as well as C2C12 myotubes (Fig. 72). Nuclear retention of mRNAs is an additional mechanism of gene regulation and quality control, that has recently emerged and is often associated with quick response to stimuli, such as stress. Two frequently discussed features that promote nuclear retention of mRNAs are retained introns or extended poly(A)-tails or 3'-UTRs and an increased association with lncRNAs (Bahar Halpern et al., 2015; Wegener and Muller-McNicoll, 2018). The mechanism behind the nuclear retention of *Fsd2* has yet to be elucidated, however, as the qRT-PCR primers utilized for *Fsd2* detection are exon-spanning, and the RNA-sequencing does not show any signs of intron retention (Fig. 51, Appendix Fig. 13),

unlike the known intron-retaining mRNA *Taf1d* (Ninomiya et al., 2020), that was used as a control to exclude the possibility of the datasets not being sequenced deep enough to detect intron retention (data not shown). The more likely possibilities are that *Fsd2* localization is influenced by its *cis*-located lncRNA *lncFsd2* by a 3'-UTR based mechanism. This is supported by the 2 additional poly(A)-signals that are used *in vivo* (Appendix Fig. 13). The created overlap of the mRNA 3'-UTR and the lncRNA sequence in combination with the anti-sense divergent transcription, creates two transcripts with high affinity to each other due to their compatible sequences. Despite high variability, *Fsd2* seems to co-precipitate with *lncFsd2* arguing for the involvement of *lncFsd2* in *Fsd2* localization (Fig. 75).

Hypothesis for further experiments

This project is unlike the work on *Swhtr* and *Fendrr* in a more preliminary stage. However, the previously unstudied lncRNA *2900076A07Rik* here referred to as *lncFsd2* has been initially characterized in terms of expression pattern, localization, sequence and possible interaction with its *cis*-located coding neighbor *Fsd2*. The currently proposed mechanism based on the obtained data is that while *lncFsd2* does not interact with FSD2 protein, it might interact with *Fsd2* mRNA to promote nuclear retention. This would explain the mild effect reflected in protein quantity (Fig. 71). Whether the binding of *lncFsd2* to *Fsd2* is causative for *Fsd2* nuclear retention or whether the binding is coincidental due to 3'-UTR/lncRNA sequence similarity, will have to be elucidated further, by subcellular fractionation of *Fsd2* expressing cells with depleted *lncFsd2* expression, or by subcellular fractionation of cardiomyocytes derived from *lncFsd2*$^{-/-}$ mice.

Despite prediction analysis of the miRNAs contained within *lncFsd2* and *SNHG21* did not show any significant indication for interaction with *Fsd2*, further experiments aimed towards studying their function will be another step towards understanding *lncFsd2* and *SNHG21*. *Mir1836* does not seem essential as blocking the entire *lncFsd2* transcription does not lead to a strong phenotype, but that does not exclude a function for the miRNA, which might also be required in the context of stress. Furthermore, for a more unbiased approach mass spectrometry following *lncFsd2* pulldown might lead to interesting FSD2 independent protein interactors.

Conclusion

The studies conducted on *Swhtr*, *Fendrr*, and *lncFsd2* within the scope of this work have significantly advanced our understanding of the impact of lncRNAs on gene regulation in the cardiopulmonary system, especially in regard to RNA-based functions. These investigations have shed light on the intricate mechanisms through which lncRNAs interact with protein-coding genes and signaling pathways to influence cellular processes and contribute to cardiovascular and pulmonary diseases.

The study on *Swhtr* provided insights into its role in cardiac stress and disease. The findings demonstrated that *Swhtr* is abundantly expressed in adult cardiomyocytes and plays a crucial role in initiating a hypertrophic response in the heart under stress conditions. The interaction between *Swhtr* and *Nkx2-5*, a neighboring transcription factor coding gene, suggests a stress-specific requirement for their interaction. Under standard conditions, loss of the lncRNA does not lead to any overt defects, while *Swhtr* remains at its locus of production as evident from smFISH staining. This possibly serves as a deposit that can be released upon stress. Mice lacking *Swhtr* depicted increased lethality after acute myocardial infarction and fail to develop cardiac hypertrophy as a compensatory response to loss of cardiac tissue and hence flexibility of the left ventricle. Furthermore, *Swhtr* disperses upon external stress-stimuli, such as hypoxia, while its interaction with NKX2-5 increases. Genes deregulated in dependence of *Swhtr* show significant occupation of NKX2-5 within their gene bodies, demonstrating an auxiliary effect of the lncRNA and the transcription factor in response to stress. This data proposes *Swhtr* as one possible co-factor required for NKX2-5 function under specific conditions. Despite highly abundant expression of *Nkx2-5* in the adult heart, not much was known about its function beyond cardiac development, although *NKX2-5* is frequently implied in human cardiac disease. This work elucidated the role *Nkx2-5* plays in the adult heart especially in the context cardiac injury and the requirement of *Swhtr* as a co-factor in the regulation of stress-response genes.

The study also examined other lncRNAs within the *Nkx2-5* locus, indicating the need for further investigation into their functions and their interplay with *Swhtr*, although it can be concluded that *IRENE-div*, a lncRNA proposed to partially overlap with *Swhtr*, is an

underrepresented long *Swhtr* isoform with alternative TSS. Structural conservation of the *Nkx2-5* locus between mouse and human and validated presence of a lncRNA potentially orthologue to *Swhtr*, which gets activated in hiPSC derived cardiomyocytes as compared to their undifferentiated state, suggested *Swhtr* as a potential therapeutic target.

Similarly, the investigation of *Fendrr* elucidated its role in fibrosis-related gene regulation. *Fendrr* is expressed in endothelial cells and fibroblasts of the lung tissue, however due to its implied importance for the regulation of fibrotic gene programs, this work focused on *Fendrr* function in fibroblast cells. The study highlighted the importance of *Fendrr*, and explicitly the lncRNA's RNA:dsDNA triplex formation element, the *FendrrBox*, in collaboration with activated WNT signaling in terms of regulating fibrosis associated target genes. Fibrosis is a critical process in wound healing, but dysregulation can lead to pathological fibrosis and organ failure. The findings presented in this work provided insights into the specific involvement of *Fendrr* in WNT-mediated lung development and fibrotic stimuli. Mutant mice lacking the *FendrrBox* show a similar yet less severe phenotype, demonstrating partial requirement for the *FendrrBox* and thus a second *FendrrBox*-independent function of *Fendrr*. This work investigated the impact of *Fendrr* and its RNA:dsDNA triplex formation element on fibrosis related target genes in embryonic fibroblast cells. The *FendrrBox* was shown to bind target gene promoters *in vitro* and fine-tune their expression in association with WNT-signaling. Other *FendrrBox*-independent functions remain targets to be studied in the future, however, this work highlights the importance of the *FendrrBox* for the fine-tuning of gene programs responding to WNT-signaling. This explains the lethality of the mutant mice, despite only minor changes in gene expression can be found. While minor changes in only one gene might not be reflected in a phenotype, the lack of regulatory fine-tuning can lead to destabilization of entire gene regulatory networks and thus disease and lethality. Furthermore, the conservation of *Fendrr*/*FENDRR* and similar implication of *FENDRR* in human diseases, such as idiopathic pulmonary fibrosis, suggested a potentially conserved mechanism. Fibrosis related *FendrrBox* target genes identified in mice were furthermore found to be co-expressed with *FENDRR* in human lungs. Despite no RNA:dsDNA triplex formation element could be detected in human *FENDRR* within the

scope of this work and the utilized algorithm (TriplexDomainFinder), another study used a different algorithm (LongTarget) and reported a functional RNA:dsDNA triplex formation domain. Furthermore, the lncRNA has been shown to be upregulated in the lungs of patients suffering from IPF. Together with *FENDRR* being associated with anti-fibrotic properties in human lung tissue, this might point towards a conserved function of *Fendrr/FENDRR* in the regulation of fibrosis-associated gene programs. Whether *Fendrr/FENDRR* have similar, or opposing functions in different cell types has to be examined in future experiments.

Regarding *lncFsd2*, the characterization of its relationship with the protein-coding gene *Fsd2* focused on potential involvement in gene regulation and localization processes. *Fsd2* is an understudied protein-coding gene, that produces a protein which forms a complex with RYR2 and CMYA5 in striated muscle tissue and thus might be involved in calcium-induced calcium release of the sarcoplasmic reticulum. Although no significant transcriptional regulation of *Fsd2* by *lncFsd2* was observed, neither *in vitro* nor *in vivo*, the data presented in this work highlighted the need for further investigations to fully understand the precise mechanisms of interaction and the functions of *lncFsd2*. A subtle increased protein amount could be detected in *lncFsd2* mutant mouse hearts but not skeletal muscle. As the protein showed dual localization between cytoplasm and nucleus only in the heart, a difference in response to loss of *lncFsd2* was examined but could not be detected. Immunoprecipitation experiments revealed potential binding of *lncFsd2* to *Fsd2* mRNA and absence of *lncFsd2* interaction with FSD2 protein. Subcellular fractionation revealed that *lncFsd2* predominantly resides in the nucleus, while *Fsd2* mRNA does not show the typical cytoplasmic localization of an mRNA and is instead localized to 50% in the nucleus and cytoplasm. RNA-sequencing data did not reveal any overt retained introns that would explain the high rate of nuclear retention. Nuclear retention of RNAs is one way of protein dosage control, thus a change in mRNA localization might be explanatory for a subtle change of protein quantity. Together with predominant nuclear localization of the lncRNA and its physical interaction with the *Fsd2* mRNA, this might indicate involvement of *lncFsd2* in nuclear retention of the mRNA, but requires further validation.

Furthermore, lncFsd2 contains a miRNA, *Mir1839*, within its last exon. Within the scope of this work, a *Mir1839*- dependent impact on *Fsd2* was not investigated further, as miRNA target prediction did not yield an interaction between *Mir1839* and *Fsd2*. Absence of a strong phenotype in mice lacking expression of *lncFsd2* and thus also *Mir1839* excludes the miRNA to be essential, miRNAs can have various functions in gene regulation and are often implicated in disease. *Fsd2*-independent functions of *lncFsd2* or *Mir1839* and potentially involvement in stress reaction, are further topics to be examined in the future.

The role of *Swhtr* in cardiac stress and disease highlights how lncRNAs can play critical roles in specific physiological conditions. It initiates a hypertrophic response in the heart under stress conditions, and its physical interaction with NKX2-5, reveals stress-specific regulatory requirements. This study showcases the significance of lncRNAs as co-factors in the form of direct protein-binders in regulating gene expression under specific circumstances. *Fendrr* involvement in fibrosis-related gene regulation sheds light on how lncRNAs contribute to developmental and pathological processes. The RNA:dsDNA triplex formation element collaborates with WNT signaling to fine-tune the expression of fibrosis-associated genes. Dysregulation of these genes can lead to serious health conditions, emphasizing the importance of lncRNAs in maintaining proper gene expression programs by direct binding of target gene promoter regions. The characterization of *lncFsd2* and its relationship with the protein-coding gene *Fsd2* provides insights into the regulatory mechanisms involving lncRNAs. Although no significant transcriptional regulation of *Fsd2* by *lncFsd2* was observed, possible lncRNA involvement in nuclear retention of an mRNA and subsequent fine-tuning of protein quantity reveals how lncRNA can add additional layers beyond direct transcriptional regulation to help adjust protein expression.

In conclusion, studying *Swhtr*, *Fendrr*, and *lncFsd2* significantly contributed to our knowledge of lncRNA-mediated gene regulation in the cardiopulmonary system. By unraveling the roles of these lncRNAs in stress responses, fibrosis-related gene regulation, and gene localization processes, these investigations have advanced our understanding of the complex mechanisms underlying cardiovascular and pulmonary

maintenance, stress reaction and diseases. Further exploration of the molecular interactions, signaling pathways, and functions of lncRNAs in these contexts will deepen our understanding of gene regulation. The three lncRNAs highlighted in this work demonstrate the intricate and diverse roles of lncRNAs in genetic regulation. They contribute to our understanding of how lncRNAs influence various cellular processes and disease states, providing potential avenues for future therapeutic targets and further exploration of lncRNA functions in different biological contexts. Continued research in this field holds great promise for understanding the foundation of a healthy organism, while potentially also improving the diagnosis, treatment, and management of cardiovascular and pulmonary disorders.

Literature

Adams, T.S., Schupp, J.C., Poli, S., Ayaub, E.A., Neumark, N., Ahangari, F., Chu, S.G., Raby, B.A., DeIuliis, G., Januszyk, M., et al. (2020). Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. Sci Adv *6*, eaba1983. 10.1126/sciadv.aba1983.

Agarwal, V., Bell, G.W., Nam, J.W., and Bartel, D.P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. Elife *4*. 10.7554/eLife.05005.

Akazawa, H., and Komuro, I. (2005). Cardiac transcription factor Csx/Nkx2-5: Its role in cardiac development and diseases. Pharmacol Ther *107*, 252-268. 10.1016/j.pharmthera.2005.03.005.

Akerberg, B.N., Gu, F., VanDusen, N.J., Zhang, X., Dong, R., Li, K., Zhang, B., Zhou, B., Sethi, I., Ma, Q., et al. (2019). A reference map of murine cardiac transcription factor chromatin occupancy identifies dynamic and conserved enhancers. Nat Commun *10*, 4907. 10.1038/s41467-019-12812-3.

Akhmetshina, A., Palumbo, K., Dees, C., Bergmann, C., Venalis, P., Zerr, P., Horn, A., Kireva, T., Beyer, C., Zwerina, J., et al. (2012). Activation of canonical Wnt signalling is required for TGF-beta-mediated fibrosis. Nat Commun *3*, 735. 10.1038/ncomms1734.

Ali, T., Rogala, S., Krause, N.M., Bains, J.K., Melissari, M.T., Wahrisch, S., Schwalbe, H., Herrmann, B.G., and Grote, P. (2023). Fendrr synergizes with Wnt signalling to regulate fibrosis related genes during lung development via its RNA:dsDNA triplex element. Nucleic Acids Res. 10.1093/nar/gkad395.

Aros, C.J., Pantoja, C.J., and Gomperts, B.N. (2021). Wnt signaling in lung development, regeneration, and disease progression. Commun Biol *4*, 601. 10.1038/s42003-021-02118-w.

Arun, G., Aggarwal, D., and Spector, D.L. (2020). MALAT1 Long Non-Coding RNA: Functional Implications. Noncoding RNA *6*. 10.3390/ncrna6020022.

Baarsma, H.A., and Konigshoff, M. (2017). 'WNT-er is coming': WNT signalling in chronic lung diseases. Thorax *72*, 746-759. 10.1136/thoraxjnl-2016-209753.

Bahar Halpern, K., Caspi, I., Lemze, D., Levy, M., Landen, S., Elinav, E., Ulitsky, I., and Itzkovitz, S. (2015). Nuclear Retention of mRNA in Mammalian Tissues. Cell Rep *13*, 2653-2662. 10.1016/j.celrep.2015.11.036.

Benson, M.A., Tinsley, C.L., Waite, A.J., Carlisle, F.A., Sweet, S.M.M., Ehler, E., George, C.H., Lai, F.A., Martin-Rendon, E., and Blake, D.J. (2017). Ryanodine receptors are part of the myospryn complex in cardiac muscle. Sci Rep *7*, 6312. 10.1038/s41598-017-06395-6.

Bodmer, R. (1993). The gene tinman is required for specification of the heart and visceral muscles in Drosophila. Development *118*, 719-729. 10.1242/dev.118.3.719.

Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., and Rastan, S. (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell *71*, 515-526. 10.1016/0092-8674(92)90519-i.

Bround, M.J., Asghari, P., Wambolt, R.B., Bohunek, L., Smits, C., Philit, M., Kieffer, T.J., Lakatta, E.G., Boheler, K.R., Moore, E.D., et al. (2012). Cardiac ryanodine receptors control heart rate and rhythmicity in adult mice. Cardiovasc Res *96*, 372-380. 10.1093/cvr/cvs260.

Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafreniere, R.G., Xing, Y., Lawrence, J., and Willard, H.F. (1992). The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell *71*, 527-542. 10.1016/0092-8674(92)90520-m.

Bryzghalov, O., Makalowska, I., and Szczesniak, M.W. (2021). lncEvo: automated identification and conservation study of long noncoding RNAs. BMC Bioinformatics 22, 59. 10.1186/s12859-021-03991-2.

Buhling, F., Rocken, C., Brasch, F., Hartig, R., Yasuda, Y., Saftig, P., Bromme, D., and Welte, T. (2004). Pivotal role of cathepsin K in lung fibrosis. Am J Pathol 164, 2203-2216. 10.1016/S0002-9440(10)63777-7.

Cai, X., and Cullen, B.R. (2007). The imprinted H19 noncoding RNA is a primary microRNA precursor. RNA 13, 313-316. 10.1261/rna.351707.

Carim-Todd, L., Escarceller, M., Estivill, X., and Sumoy, L. (2001). Characterization of human FSD1, a novel brain specific gene on chromosome 19 with paralogy to 9q31. Biochim Biophys Acta 1518, 200-203. 10.1016/s0167-4781(01)00178-6.

Chaumeil, J., Waters, P.D., Koina, E., Gilbert, C., Robinson, T.J., and Graves, J.A. (2011). Evolution from XIST-independent to XIST-controlled X-chromosome inactivation: epigenetic modifications in distantly related mammals. PLoS One 6, e19040. 10.1371/journal.pone.0019040.

Chen, J.N., and Fishman, M.C. (1996). Zebrafish tinman homolog demarcates the heart field and initiates myocardial differentiation. Development 122, 3809-3816. 10.1242/dev.122.12.3809.

Chen, Y., and Wang, X. (2020). miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Res 48, D127-D131. 10.1093/nar/gkz757.

Clark, M.B., Johnston, R.L., Inostroza-Ponta, M., Fox, A.H., Fortini, E., Moscato, P., Dinger, M.E., and Mattick, J.S. (2012). Genome-wide analysis of long noncoding RNA stability. Genome Res 22, 885-898. 10.1101/gr.131037.111.

Claycomb, W.C., Lanson, N.A., Jr., Stallworth, B.S., Egeland, D.B., Delcarpio, J.B., Bahinski, A., and Izzo, N.J., Jr. (1998). HL-1 cells: a cardiac muscle cell line that contracts and retains phenotypic characteristics of the adult cardiomyocyte. Proc Natl Acad Sci U S A 95, 2979-2984. 10.1073/pnas.95.6.2979.

Cleaver, O.B., Patterson, K.D., and Krieg, P.A. (1996). Overexpression of the tinman-related genes XNkx-2.5 and XNkx-2.3 in Xenopus embryos results in myocardial hyperplasia. Development 122, 3549-3556. 10.1242/dev.122.11.3549.

Clemson, C.M., Hutchinson, J.N., Sara, S.A., Ensminger, A.W., Fox, A.H., Chess, A., and Lawrence, J.B. (2009). An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol Cell 33, 717-726. 10.1016/j.molcel.2009.01.026.

Concordet, J.P., and Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res 46, W242-W245. 10.1093/nar/gky354.

Consortium, G.T. (2013). The Genotype-Tissue Expression (GTEx) project. Nat Genet 45, 580-585. 10.1038/ng.2653.

Costa, M.W., Guo, G., Wolstein, O., Vale, M., Castro, M.L., Wang, L., Otway, R., Riek, P., Cochrane, N., Furtado, M., et al. (2013). Functional characterization of a novel mutation in NKX2-5 associated with congenital heart disease and adult-onset cardiomyopathy. Circ Cardiovasc Genet 6, 238-247. 10.1161/CIRCGENETICS.113.000057.

Cunningham, F., Allen, J.E., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Austine-Orimoloye, O., Azov, A.G., Barnes, I., Bennett, R., et al. (2022). Ensembl 2022. Nucleic Acids Res 50, D988-D995. 10.1093/nar/gkab1049.

Curras-Alonso, S., Soulier, J., Defard, T., Weber, C., Heinrich, S., Laporte, H., Leboucher, S., Lameiras, S., Dutreix, M., Favaudon, V., et al. (2023). An interactive murine single-cell atlas of the lung responses to radiation injury. Nat Commun *14*, 2445. 10.1038/s41467-023-38134-z.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. Gigascience *10*. 10.1093/gigascience/giab008.

de Sena-Tomás, C., Aleman, A.G., Ford, C., Varshney, A., Yao, D., Harrington, J.K., Saúde, L., Ramialison, M., and Targoff, K.L. (2022). Activation of Nkx2.5 transcriptional program is required for adult myocardial repair. Nature Communications *13*, 2970. 10.1038/s41467-022-30468-4.

Deutsch, M.A., Doppler, S.A., Li, X., Lahm, H., Santamaria, G., Cuda, G., Eichhorn, S., Ratschiller, T., Dzilic, E., Dressen, M., et al. (2018). Reactivation of the Nkx2.5 cardiac enhancer after myocardial infarction does not presage myogenesis. Cardiovasc Res *114*, 1098-1114. 10.1093/cvr/cvy069.

Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L., Wong, M.D., White, J.K., Meehan, T.F., Weninger, W.J., Westerberg, H., Adissu, H., et al. (2017). Corrigendum: High-throughput discovery of novel developmental phenotypes. Nature *551*, 398. 10.1038/nature24643.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21. 10.1093/bioinformatics/bts635.

Durham, J.T., Brand, O.M., Arnold, M., Reynolds, J.G., Muthukumar, L., Weiler, H., Richardson, J.A., and Naya, F.J. (2006). Myospryn is a direct transcriptional target for MEF2A that encodes a striated muscle, alpha-actinin-interacting, costamere-localized protein. J Biol Chem *281*, 6841-6849. 10.1074/jbc.M510499200.

Eblaghie, M.C., Reedy, M., Oliver, T., Mishina, Y., and Hogan, B.L. (2006). Evidence that autocrine signaling through Bmpr1a regulates the proliferation, survival and morphogenetic behavior of distal lung epithelial cells. Dev Biol *291*, 67-82. 10.1016/j.ydbio.2005.12.006.

Fagerberg, L., Hallstrom, B.M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., Tahmasebpoor, S., Danielsson, A., Edlund, K., et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics *13*, 397-406. 10.1074/mcp.M113.035600.

Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet *9*, 102-114. 10.1038/nrg2290.

Freese, N.H., Norris, D.C., and Loraine, A.E. (2016). Integrated genome browser: visual analytics platform for genomics. Bioinformatics *32*, 2089-2095. 10.1093/bioinformatics/btw069.

Furlan, G., Gutierrez Hernandez, N., Huret, C., Galupa, R., van Bemmel, J.G., Romito, A., Heard, E., Morey, C., and Rougeulle, C. (2018). The Ftx Noncoding Locus Controls X Chromosome Inactivation Independently of Its RNA Products. Mol Cell *70*, 462-472 e468. 10.1016/j.molcel.2018.03.024.

Gardarsdottir, H.R., Sigurdsson, M.I., Andersen, K., and Gudmundsdottir, I.J. (2022). Long-term survival of Icelandic women following acute myocardial infarction. Scand Cardiovasc J *56*, 114-120. 10.1080/14017431.2022.2075561.

George, M.R., Duan, Q., Nagle, A., Kathiriya, I.S., Huang, Y., Rao, K., Haldar, S.M., and Bruneau, B.G. (2019). Minimal in vivo requirements for developmentally regulated cardiac long intergenic non-coding RNAs. Development *146*. 10.1242/dev.185314.

George, S.H., Gertsenstein, M., Vintersten, K., Korets-Smith, E., Murphy, J., Stevens, M.E., Haigh, J.J., and Nagy, A. (2007). Developmental and adult phenotyping directly from mutant embryonic stem cells. Proc Natl Acad Sci U S A *104*, 4455-4460. 10.1073/pnas.0609277104.

Ghanem, M., Justet, A., Hachem, M., Boghanim, T., Vadel, A., Jaillet, M., Mailleux, A., and Crestani, B. (2022). Involvement of FGF21 in pulmonary fibrosis. ERJ Open Research *8*, 108. 10.1183/23120541.Lsc-2022.108.

Gil, N., and Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. Nat Rev Genet *21*, 102-117. 10.1038/s41576-019-0184-5.

Gomez, A.M., and Richard, S. (2004). Mutant cardiac ryanodine receptors and ventricular arrhythmias: is 'gain-of-function' obligatory? Cardiovasc Res *64*, 3-5. 10.1016/j.cardiores.2004.07.018.

Gong, L., Zhu, L., and Yang, T. (2020). Fendrr involves in the pathogenesis of cardiac fibrosis via regulating miR-106b/SMAD3 axis. Biochem Biophys Res Commun *524*, 169-177. 10.1016/j.bbrc.2020.01.062.

Goudarzi, M., Berg, K., Pieper, L.M., and Schier, A.F. (2019). Individual long non-coding RNAs have no overt functions in zebrafish embryogenesis, viability and fertility. Elife *8*. 10.7554/eLife.40815.

Grote, P., and Herrmann, B.G. (2013). The long non-coding RNA Fendrr links epigenetic control mechanisms to gene regulatory networks in mammalian embryogenesis. RNA Biol *10*, 1579-1585. 10.4161/rna.26165.

Grote, P., Wittler, L., Hendrix, D., Koch, F., Wahrisch, S., Beisaw, A., Macura, K., Blass, G., Kellis, M., Werber, M., and Herrmann, B.G. (2013). The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. Dev Cell *24*, 206-214. 10.1016/j.devcel.2012.12.012.

Groza, T., Gomez, F.L., Mashhadi, H.H., Munoz-Fuentes, V., Gunes, O., Wilson, R., Cacheiro, P., Frost, A., Keskivali-Bond, P., Vardal, B., et al. (2023). The International Mouse Phenotyping Consortium: comprehensive knockout phenotyping underpinning the study of human disease. Nucleic Acids Res *51*, D1038-D1045. 10.1093/nar/gkac972.

Guan, R., Yuan, L., Li, J., Wang, J., Li, Z., Cai, Z., Guo, H., Fang, Y., Lin, R., Liu, W., et al. (2022). Bone morphogenetic protein 4 inhibits pulmonary fibrosis by modulating cellular senescence and mitophagy in lung fibroblasts. Eur Respir J *60*. 10.1183/13993003.02307-2021.

Guidry, U.C., Evans, J.C., Larson, M.G., Wilson, P.W., Murabito, J.M., and Levy, D. (1999). Temporal trends in event rates after Q-wave myocardial infarction: the Framingham Heart Study. Circulation *100*, 2054-2059. 10.1161/01.cir.100.20.2054.

Guzy, R. (2020). Fibroblast Growth Factor Inhibitors in Lung Fibrosis: Friends or Foes? Am J Respir Cell Mol Biol *63*, 273-274. 10.1165/rcmb.2020-0156ED.

Han, P., Li, W., Lin, C.H., Yang, J., Shang, C., Nuernberg, S.T., Jin, K.K., Xu, W., Lin, C.Y., Lin, C.J., et al. (2014). A long noncoding RNA protects the heart from pathological hypertrophy. Nature *514*, 102-106. 10.1038/nature13596.

Hezroni, H., Koppstein, D., Schwartz, M.G., Avrutin, A., Bartel, D.P., and Ulitsky, I. (2015). Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. Cell Rep *11*, 1110-1122. 10.1016/j.celrep.2015.04.023.

Hou, X., Hu, Z., Xu, H., Xu, J., Zhang, S., Zhong, Y., He, X., and Wang, N. (2014). Advanced glycation endproducts trigger autophagy in cadiomyocyte via RAGE/PI3K/AKT/mTOR pathway. Cardiovasc Diabetol *13*, 78. 10.1186/1475-2840-13-78.

Huang, C., Liang, Y., Zeng, X., Yang, X., Xu, D., Gou, X., Sathiaseelan, R., Senavirathna, L.K., Wang, P., and Liu, L. (2020). Long Noncoding RNA FENDRR Exhibits Antifibrotic Activity in Pulmonary Fibrosis. Am J Respir Cell Mol Biol *62*, 440-453. 10.1165/rcmb.2018-0293OC.

Hussain, M., Xu, C., Lu, M., Wu, X., Tang, L., and Wu, X. (2017). Wnt/beta-catenin signaling links embryonic lung development and asthmatic airway remodeling. Biochim Biophys Acta Mol Basis Dis *1863*, 3226-3242. 10.1016/j.bbadis.2017.08.031.

Hutchinson, J.N., Ensminger, A.W., Clemson, C.M., Lynch, C.R., Lawrence, J.B., and Chess, A. (2007). A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. BMC Genomics *8*, 39. 10.1186/1471-2164-8-39.

Jho, E.H., Zhang, T., Domon, C., Joo, C.K., Freund, J.N., and Costantini, F. (2002). Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. Mol Cell Biol *22*, 1172-1183. 10.1128/MCB.22.4.1172-1183.2002.

Jiang, L., Shao, C., Wu, Q.J., Chen, G., Zhou, J., Yang, B., Li, H., Gou, L.T., Zhang, Y., Wang, Y., et al. (2017). NEAT1 scaffolds RNA-binding proteins and the Microprocessor to globally enhance pri-miRNA processing. Nat Struct Mol Biol *24*, 816-824. 10.1038/nsmb.3455.

Johnsson, P., Lipovich, L., Grander, D., and Morris, K.V. (2014). Evolutionary conservation of long non-coding RNAs; sequence, structure, function. Biochim Biophys Acta *1840*, 1063-1071. 10.1016/j.bbagen.2013.10.035.

Kasahara, H., Bartunkova, S., Schinke, M., Tanaka, M., and Izumo, S. (1998). Cardiac and extracardiac expression of Csx/Nkx2.5 homeodomain protein. Circ Res *82*, 936-946. 10.1161/01.res.82.9.936.

Keniry, A., Oxley, D., Monnier, P., Kyba, M., Dandolo, L., Smits, G., and Reik, W. (2012). The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. Nat Cell Biol *14*, 659-665. 10.1038/ncb2521.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res *12*, 996-1006. 10.1101/gr.229102.

Kolk, M.V., Meyberg, D., Deuse, T., Tang-Quan, K.R., Robbins, R.C., Reichenspurner, H., and Schrepfer, S. (2009). LAD-ligation: a murine model of myocardial infarction. J Vis Exp. 10.3791/1438.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583-588. 10.1038/nature14136.

Konigshoff, M., Balsara, N., Pfaff, E.M., Kramer, M., Chrobak, I., Seeger, W., and Eickelberg, O. (2008). Functional Wnt signaling is increased in idiopathic pulmonary fibrosis. PLoS One *3*, e2142. 10.1371/journal.pone.0002142.

Konigshoff, M., Kramer, M., Balsara, N., Wilhelm, J., Amarie, O.V., Jahn, A., Rose, F., Fink, L., Seeger, W., Schaefer, L., et al. (2009). WNT1-inducible signaling protein-1 mediates pulmonary fibrosis in mice and is upregulated in humans with idiopathic pulmonary fibrosis. J Clin Invest *119*, 772-787. 10.1172/JCI33950.

Koressaar, T., Lepamets, M., Kaplinski, L., Raime, K., Andreson, R., and Remm, M. (2018). Primer3_masker: integrating masking of template sequence with primer design software. Bioinformatics *34*, 1937-1938. 10.1093/bioinformatics/bty036.

Koressaar, T., and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. Bioinformatics *23*, 1289-1291. 10.1093/bioinformatics/btm091.

Kypr, J., Kejnovska, I., Renciuk, D., and Vorlickova, M. (2009). Circular dichroism and conformational polymorphism of DNA. Nucleic Acids Res *37*, 1713-1725. 10.1093/nar/gkp026.

Lagarde, J., Uszczynska-Ratajczak, B., Carbonell, S., Perez-Lluch, S., Abad, A., Davis, C., Gingeras, T.R., Frankish, A., Harrow, J., Guigo, R., and Johnson, R. (2017). High-throughput annotation of full-length long noncoding RNAs with capture long-read sequencing. Nat Genet *49*, 1731-1740. 10.1038/ng.3988.

Laird, P.W., Zijderveld, A., Linders, K., Rudnicki, M.A., Jaenisch, R., and Berns, A. (1991). Simplified mammalian DNA isolation procedure. Nucleic Acids Res *19*, 4293. 10.1093/nar/19.15.4293.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359. 10.1038/nmeth.1923.

Lawrence, M., Huber, W., Pages, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput Biol *9*, e1003118. 10.1371/journal.pcbi.1003118.

Lee, S., Kopp, F., Chang, T.C., Sataluri, A., Chen, B., Sivakumar, S., Yu, H., Xie, Y., and Mendell, J.T. (2016). Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. Cell *164*, 69-80. 10.1016/j.cell.2015.12.017.

Li, C., Xiao, J., Hormi, K., Borok, Z., and Minoo, P. (2002). Wnt5a participates in distal lung morphogenesis. Dev Biol *248*, 68-81. 10.1006/dbio.2002.0729.

Li, D., Hsu, S., Purushotham, D., Sears, R.L., and Wang, T. (2019). WashU Epigenome Browser update 2019. Nucleic Acids Res *47*, W158-W165. 10.1093/nar/gkz348.

Li, D., Purushotham, D., Harrison, J.K., Hsu, S., Zhuo, X., Fan, C., Liu, S., Xu, V., Chen, S., Xu, J., et al. (2022). WashU Epigenome Browser update 2022. Nucleic Acids Res *50*, W774-W781. 10.1093/nar/gkac238.

Li, Q., Wu, C., Song, G., Zhang, H., Shan, B., Duan, Y., and Wang, Y. (2015). Genome-Wide Analysis of Long Noncoding RNA Expression Profiles in Human Xuanwei Lung Cancer. Clin Lab *61*, 1515-1523. 10.7754/clin.lab.2015.150323.

Li, Y., Syed, J., and Sugiyama, H. (2016). RNA-DNA Triplex Formation by Long Noncoding RNAs. Cell Chem Biol *23*, 1325-1333. 10.1016/j.chembiol.2016.09.011.

Lien, C.L., Wu, C., Mercer, B., Webb, R., Richardson, J.A., and Olson, E.N. (1999). Control of early cardiac-specific transcription of Nkx2-5 by a GATA-dependent enhancer. Development *126*, 75-84. 10.1242/dev.126.1.75.

Lim, K.S., Lee, K.T., Lee, S.W., Chai, H.H., Jang, G., Hong, K.C., and Kim, T.H. (2016). Genomic structure, expression and association study of the porcine FSD2. Mol Biol Rep *43*, 1011-1018. 10.1007/s11033-016-4029-4.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. 10.1186/s13059-014-0550-8.

Lu, F., Ma, Q., Xie, W., Liou, C.L., Zhang, D., Sweat, M.E., Jardin, B.D., Naya, F.J., Guo, Y., Cheng, H., and Pu, W.T. (2022). CMYA5 establishes cardiac dyad architecture and positioning. Nat Commun *13*, 2185. 10.1038/s41467-022-29902-4.

Lu, M.M., Yang, H., Zhang, L., Shu, W., Blair, D.G., and Morrisey, E.E. (2001). The bone morphogenic protein antagonist gremlin regulates proximal-distal patterning of the lung. Dev Dyn *222*, 667-680. 10.1002/dvdy.1231.

Mann, M., Wright, P.R., and Backofen, R. (2017). IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions. Nucleic Acids Res *45*, W435-W439. 10.1093/nar/gkx279.

Matsui, T., Nagoshi, T., and Rosenzweig, A. (2003). Akt and PI 3-kinase signaling in cardiomyocyte hypertrophy and survival. Cell Cycle *2*, 220-223.

Mattioli, K., Volders, P.J., Gerhardinger, C., Lee, J.C., Maass, P.G., Mele, M., and Rinn, J.L. (2019). High-throughput functional analysis of lncRNA core promoters elucidates rules governing tissue specificity. Genome Res *29*, 344-355. 10.1101/gr.242222.118.

McCown, P.J., Corbino, K.A., Stav, S., Sherlock, M.E., and Breaker, R.R. (2017). Riboswitch diversity and distribution. RNA *23*, 995-1011. 10.1261/rna.061234.117.

Mele, M., Mattioli, K., Mallard, W., Shechner, D.M., Gerhardinger, C., and Rinn, J.L. (2017). Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. Genome Res *27*, 27-37. 10.1101/gr.214205.116.

Modarresi, F., Faghihi, M.A., Lopez-Toledano, M.A., Fatemi, R.P., Magistri, M., Brothers, S.P., van der Brug, M.P., and Wahlestedt, C. (2012). Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. Nat Biotechnol *30*, 453-459. 10.1038/nbt.2158.

Mondal, T., Subhash, S., Vaid, R., Enroth, S., Uday, S., Reinius, B., Mitra, S., Mohammed, A., James, A.R., Hoberg, E., et al. (2015). MEG3 long noncoding RNA regulates the TGF-beta pathway genes through formation of RNA-DNA triplex structures. Nat Commun *6*, 7743. 10.1038/ncomms8743.

Morse, C., Tabib, T., Sembrat, J., Buschur, K.L., Bittar, H.T., Valenzi, E., Jiang, Y., Kass, D.J., Gibson, K., Chen, W., et al. (2019). Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. Eur Respir J *54*. 10.1183/13993003.02441-2018.

Negretti, N.M., Plosa, E.J., Benjamin, J.T., Schuler, B.A., Habermann, A.C., Jetter, C.S., Gulleman, P., Bunn, C., Hackett, A.N., Ransom, M., et al. (2021). A single-cell atlas of mouse lung development. Development *148*. 10.1242/dev.199512.

Nguyen Ba, A.N., Pogoutse, A., Provart, N., and Moses, A.M. (2009). NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. BMC Bioinformatics *10*, 202. 10.1186/1471-2105-10-202.

Nian, M., Lee, P., Khaper, N., and Liu, P. (2004). Inflammatory cytokines and postmyocardial infarction remodeling. Circ Res *94*, 1543-1553. 10.1161/01.RES.0000130526.20854.fa.

Ninomiya, K., Adachi, S., Natsume, T., Iwakiri, J., Terai, G., Asai, K., and Hirose, T. (2020). LncRNA-dependent nuclear stress bodies promote intron retention through SR protein phosphorylation. EMBO J *39*, e102729. 10.15252/embj.2019102729.

Oldfield, C.J., Duhamel, T.A., and Dhalla, N.S. (2020). Mechanisms for the transition from physiological to pathological cardiac hypertrophy. Can J Physiol Pharmacol *98*, 74-84. 10.1139/cjpp-2019-0566.

Olivero, C.E., Martinez-Terroba, E., Zimmer, J., Liao, C., Tesfaye, E., Hooshdaran, N., Schofield, J.A., Bendor, J., Fang, D., Simon, M.D., et al. (2020). p53 Activates the Long Noncoding RNA Pvt1b to Inhibit Myc and Suppress Tumorigenesis. Mol Cell *77*, 761-774 e768. 10.1016/j.molcel.2019.12.014.

Onodi, Z., Visnovitz, T., Kiss, B., Hambalko, S., Koncz, A., Agg, B., Varadi, B., Toth, V.E., Nagy, R.N., Gergely, T.G., et al. (2022). Systematic transcriptomic and phenotypic characterization of human and murine cardiac myocyte cell lines and primary cardiomyocytes reveals serious limitations and low resemblances to adult cardiac phenotype. J Mol Cell Cardiol *165*, 19-30. 10.1016/j.yjmcc.2021.12.007.

Paralkar, V.R., Taborda, C.C., Huang, P., Yao, Y., Kossenkov, A.V., Prasad, R., Luan, J., Davies, J.O., Hughes, J.R., Hardison, R.C., et al. (2016). Unlinking an lncRNA from Its Associated cis Element. Mol Cell *62*, 104-110. 10.1016/j.molcel.2016.02.029.

Parker, M.W., Rossi, D., Peterson, M., Smith, K., Sikstrom, K., White, E.S., Connett, J.E., Henke, C.A., Larsson, O., and Bitterman, P.B. (2014). Fibrotic extracellular matrix activates a profibrotic positive feedback loop. J Clin Invest *124*, 1622-1635. 10.1172/JCI71386.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature *379*, 131-137. 10.1038/379131a0.

Pongracz, J.E., and Stockley, R.A. (2006). Wnt signalling in lung development and diseases. Respir Res *7*, 15. 10.1186/1465-9921-7-15.

Poss, K.D., Wilson, L.G., and Keating, M.T. (2002). Heart regeneration in zebrafish. Science *298*, 2188-2190. 10.1126/science.1077857.

Priori, S.G., and Chen, S.R. (2011). Inherited dysfunction of sarcoplasmic reticulum Ca2+ handling and arrhythmogenesis. Circ Res *108*, 871-883. 10.1161/CIRCRESAHA.110.226845.

Priori, S.G., Napolitano, C., Tiso, N., Memmi, M., Vignati, G., Bloise, R., Sorrentino, V., and Danieli, G.A. (2001). Mutations in the cardiac ryanodine receptor gene (hRyR2) underlie catecholaminergic polymorphic ventricular tachycardia. Circulation *103*, 196-200. 10.1161/01.cir.103.2.196.

Qian, L., Wythe, J.D., Liu, J., Cartry, J., Vogler, G., Mohapatra, B., Otway, R.T., Huang, Y., King, I.N., Maillet, M., et al. (2011). Tinman/Nkx2-5 acts via miR-1 and upstream of Cdc42 to regulate heart function across species. J Cell Biol *193*, 1181-1196. 10.1083/jcb.201006114.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842. 10.1093/bioinformatics/btq033.

R Core Team (2022). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

Raal, F.J., Santos, R.D., Blom, D.J., Marais, A.D., Charng, M.J., Cromwell, W.C., Lachmann, R.H., Gaudet, D., Tan, J.L., Chasan-Taber, S., et al. (2010). Mipomersen, an apolipoprotein B synthesis inhibitor, for lowering of LDL cholesterol concentrations in patients with homozygous familial hypercholesterolaemia: a randomised, double-blind, placebo-controlled trial. Lancet *375*, 998-1006. 10.1016/S0140-6736(10)60284-X.

Rahimi, A.M., Cai, M., and Hoyer-Fender, S. (2022). Heterogeneity of the NIH3T3 Fibroblast Cell Line. Cells *11*. 10.3390/cells11172677.

Rajagopal, J., Carroll, T.J., Guseh, J.S., Bores, S.A., Blank, L.J., Anderson, W.J., Yu, J., Zhou, Q., McMahon, A.P., and Melton, D.A. (2008). Wnt7b stimulates embryonic lung growth by coordinately increasing the replication of epithelium and mesenchyme. Development *135*, 1625-1634. 10.1242/dev.015495.

Rehmsmeier, M., Steffen, P., Hochsmann, M., and Giegerich, R. (2004). Fast and effective prediction of microRNA/target duplexes. RNA *10*, 1507-1517. 10.1261/rna.5248604.

Robinson, J.T., Thorvaldsdottir, H., Turner, D., and Mesirov, J.P. (2023). igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). Bioinformatics *39*. 10.1093/bioinformatics/btac830.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26. 10.1038/nbt.1754.

Rogala, S., Ali, T., Melissari, M.-T., Währisch, S., Schuster, P., Sarre, A., Boettger, T., Rogg, E.-M., Kaur, J., Krishnan, J., et al. (2022). The lncRNA *Sweetheart* regulates compensatory cardiac hypertrophy after myocardial injury. bioRxiv, 2022.2011.2014.516395. 10.1101/2022.11.14.516395.

Rolf, A., Assmus, B., Schachinger, V., Rixe, J., Mollmann, S., Mollmann, H., Dimmeler, S., Zeiher, A.M., Hamm, C.W., and Dill, T. (2011). Maladaptive hypertrophy after acute myocardial infarction positive effect of bone marrow-derived stem cell therapy on regional remodeling measured by cardiac MRI. Clin Res Cardiol *100*, 983-992. 10.1007/s00392-011-0330-3.

Rosengren, A., Spetz, C.L., Koster, M., Hammar, N., Alfredsson, L., and Rosen, M. (2001). Sex differences in survival after myocardial infarction in Sweden; data from the Swedish National Acute Myocardial Infarction Register. Eur Heart J *22*, 314-322. 10.1053/euhj.2000.2368.

Ross, J.P., and Kassir, Z. (2014). The varied roles of nuclear argonaute-small RNA complexes and avenues for therapy. Mol Ther Nucleic Acids *3*, e203. 10.1038/mtna.2014.54.

Ruan, X., Li, P., Chen, Y., Shi, Y., Pirooznia, M., Seifuddin, F., Suemizu, H., Ohnishi, Y., Yoneda, N., Nishiwaki, M., et al. (2020). In vivo functional analysis of non-conserved human lncRNAs associated with cardiometabolic traits. Nat Commun *11*, 45. 10.1038/s41467-019-13688-z.

Rubin, S.A., Fishbein, M.C., and Swan, H.J. (1983). Compensatory hypertrophy in the heart after myocardial infarction in the rat. J Am Coll Cardiol *1*, 1435-1441. 10.1016/s0735-1097(83)80046-1.

Salamon, I., Serio, S., Bianco, S., Pagiatakis, C., Crasto, S., Chiariello, A.M., Conte, M., Cattaneo, P., Fiorillo, L., Felicetta, A., et al. (2020). Divergent Transcription of the Nkx2-5 Locus Generates Two Enhancer RNAs with Opposing Functions. iScience *23*, 101539. 10.1016/j.isci.2020.101539.

Sauvageau, M., Goff, L.A., Lodato, S., Bonev, B., Groff, A.F., Gerhardinger, C., Sanchez-Gomez, D.B., Hacisuleyman, E., Li, E., Spence, M., et al. (2013). Multiple knockout mouse models reveal lincRNAs are required for life and brain development. Elife *2*, e01749. 10.7554/eLife.01749.

Schaub, M.C., Hefti, M.A., Zuellig, R.A., and Morano, I. (1998). Modulation of contractility in human cardiac hypertrophy by myosin essential light chain isoforms. Cardiovasc Res *37*, 381-404. 10.1016/s0008-6363(97)00258-7.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676-682. 10.1038/nmeth.2019.

Senavirathna, L.K., Liang, Y., Huang, C., Yang, X., Bamunuarachchi, G., Xu, D., Dang, Q., Sivasami, P., Vaddadi, K., Munteanu, M.C., et al. (2021). Long Noncoding RNA FENDRR Inhibits Lung Fibroblast Proliferation via a Reduction of beta-Catenin. Int J Mol Sci *22*. 10.3390/ijms22168536.

Shang, L., Ananthakrishnan, R., Li, Q., Quadri, N., Abdillahi, M., Zhu, Z., Qu, W., Rosario, R., Toure, F., Yan, S.F., et al. (2010). RAGE modulates hypoxia/reoxygenation injury in adult murine cardiomyocytes via JNK and GSK-3beta signaling pathways. PLoS One *5*, e10092. 10.1371/journal.pone.0010092.

Sharma, N.R., Wang, X., Majerciak, V., Ajiro, M., Kruhlak, M., Meyers, C., and Zheng, Z.M. (2016). Cell Type- and Tissue Context-dependent Nuclear Distribution of Human Ago2. J Biol Chem *291*, 2302-2309. 10.1074/jbc.C115.695049.

Shi, W., Xu, J., and Warburton, D. (2009). Development, repair and fibrosis: what is common and why it matters. Respirology *14*, 656-665. 10.1111/j.1440-1843.2009.01565.x.

Shimizu, I., and Minamino, T. (2016). Physiological and pathological cardiac hypertrophy. J Mol Cell Cardiol *97*, 245-262. 10.1016/j.yjmcc.2016.06.001.

Shu, W., Jiang, Y.Q., Lu, M.M., and Morrisey, E.E. (2002). Wnt7b regulates mesenchymal proliferation and vascular development in the lung. Development *129*, 4831-4842. 10.1242/dev.129.20.4831.

Silva, A.C., Pereira, C., Fonseca, A., Pinto-do, O.P., and Nascimento, D.S. (2020). Bearing My Heart: The Role of Extracellular Matrix on Cardiac Development, Homeostasis, and Injury Response. Front Cell Dev Biol *8*, 621644. 10.3389/fcell.2020.621644.

Spencer, F.A., Meyer, T.E., Gore, J.M., and Goldberg, R.J. (2002). Heterogeneity in the management and outcomes of patients with acute myocardial infarction complicated by heart failure: the National Registry of Myocardial Infarction. Circulation *105*, 2605-2610. 10.1161/01.cir.0000017861.00991.2f.

Steg, P.G., Dabbous, O.H., Feldman, L.J., Cohen-Solal, A., Aumont, M.C., Lopez-Sendon, J., Budaj, A., Goldberg, R.J., Klein, W., Anderson, F.A., Jr., and Global Registry of Acute Coronary Events, I. (2004). Determinants and prognostic impact of heart failure complicating acute coronary syndromes: observations from the Global Registry of Acute Coronary Events (GRACE). Circulation *109*, 494-499. 10.1161/01.CIR.0000109691.16944.DA.

Sveinbjornsson, G., Olafsdottir, E.F., Thorolfsdottir, R.B., Davidsson, O.B., Helgadottir, A., Jonasdottir, A., Jonasdottir, A., Bjornsson, E., Jensson, B.O., Arnadottir, G.A., et al. (2018). Variants in NKX2-5 and FLNC Cause Dilated Cardiomyopathy and Sudden Cardiac Death. Circ Genom Precis Med *11*, e002151. 10.1161/CIRCGEN.117.002151.

Szafranski, P., Dharmadhikari, A.V., Brosens, E., Gurha, P., Kolodziejska, K.E., Zhishuo, O., Dittwald, P., Majewski, T., Mohan, K.N., Chen, B., et al. (2013). Small noncoding differentially methylated copy-number variants, including lncRNA genes, cause a lethal lung developmental disorder. Genome Res *23*, 23-33. 10.1101/gr.141887.112.

Taegtmeyer, H., Sen, S., and Vela, D. (2010). Return to the fetal gene program: a suggested metabolic link to gene expression in the heart. Ann N Y Acad Sci *1188*, 191-198. 10.1111/j.1749-6632.2009.05100.x.

Takeshima, H., Komazaki, S., Hirose, K., Nishi, M., Noda, T., and Iino, M. (1998). Embryonic lethality and abnormal cardiac myocytes in mice lacking ryanodine receptor type 2. EMBO J *17*, 3309-3316. 10.1093/emboj/17.12.3309.

Takimoto, E., Mizuno, T., Terasaki, F., Shimoyama, M., Honda, H., Shiojima, I., Hiroi, Y., Oka, T., Hayashi, D., Hirai, H., et al. (2000). Up-regulation of natriuretic peptides in the ventricle of Csx/Nkx2-5 transgenic mice. Biochem Biophys Res Commun *270*, 1074-1079. 10.1006/bbrc.2000.2561.

Tamamura, Y., Otani, T., Kanatani, N., Koyama, E., Kitagaki, J., Komori, T., Yamada, Y., Costantini, F., Wakisaka, S., Pacifici, M., et al. (2005). Developmental regulation of Wnt/beta-catenin signals is required for growth plate assembly, cartilage integrity, and endochondral ossification. J Biol Chem *280*, 19185-19195. 10.1074/jbc.M414275200.

Tanaka, M., Chen, Z., Bartunkova, S., Yamasaki, N., and Izumo, S. (1999). The cardiac homeobox gene Csx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. Development *126*, 1269-1280. 10.1242/dev.126.6.1269.

Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform *14*, 178-192. 10.1093/bib/bbs017.

Tichon, A., Gil, N., Lubelsky, Y., Havkin Solomon, T., Lemze, D., Itzkovitz, S., Stern-Ginossar, N., and Ulitsky, I. (2016). A conserved abundant cytoplasmic long noncoding RNA modulates repression by Pumilio proteins in human cells. Nat Commun *7*, 12209. 10.1038/ncomms12209.

Toko, H., Zhu, W., Takimoto, E., Shiojima, I., Hiroi, Y., Zou, Y., Oka, T., Akazawa, H., Mizukami, M., Sakamoto, M., et al. (2002). Csx/Nkx2-5 is required for homeostasis and survival of cardiac myocytes in the adult heart. J Biol Chem *277*, 24735-24743. 10.1074/jbc.M107669200.

Travaglini, K.J., Nabhan, A.N., Penland, L., Sinha, R., Gillich, A., Sit, R.V., Chang, S., Conley, S.D., Mori, Y., Seita, J., et al. (2020). A molecular cell atlas of the human lung from single-cell RNA sequencing. Nature *587*, 619-625. 10.1038/s41586-020-2922-4.

Tseng, Y.Y., Moriarity, B.S., Gong, W., Akiyama, R., Tiwari, A., Kawakami, H., Ronning, P., Reuland, B., Guenther, K., Beadnell, T.C., et al. (2014). PVT1 dependence in cancer with MYC copy-number increase. Nature *512*, 82-86. 10.1038/nature13311.

Tsoupri, E., Kostavasili, I., Kloukina, I., Tsikitis, M., Miliou, D., Vasilaki, E., Varela, A., Nakos-Bimpos, M., Davos, C., Mavroidis, M., et al. (2021). Myospryn deficiency leads to impaired cardiac structure and function and schizophrenia-associated symptoms. Cell Tissue Res *385*, 675-696. 10.1007/s00441-021-03447-2.

Tu, C.T., Yang, T.C., and Tsai, H.J. (2009). Nkx2.7 and Nkx2.5 function redundantly and are required for cardiac morphogenesis of zebrafish embryos. PLoS One *4*, e4249. 10.1371/journal.pone.0004249.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3--new capabilities and interfaces. Nucleic Acids Res *40*, e115. 10.1093/nar/gks596.

Valencia, A., and Burgess, J.H. (1969). Arterial hypoxemia following acute myocardial infarction. Circulation *40*, 641-652. 10.1161/01.cir.40.5.641.

Vasudevan, H.N., Mazot, P., He, F., and Soriano, P. (2015). Receptor tyrosine kinases modulate distinct transcriptional programs by differential usage of intracellular pathways. Elife *4*. 10.7554/eLife.07186.

Venkata Subbaiah, K.C., Hedaya, O., Wu, J., Jiang, F., and Yao, P. (2019). Mammalian RNA switches: Molecular rheostats in gene regulation, disease, and medicine. Comput Struct Biotechnol J *17*, 1326-1338. 10.1016/j.csbj.2019.10.001.

Vitravene Study, G. (2002). A randomized controlled clinical trial of intravitreous fomivirsen for treatment of newly diagnosed peripheral cytomegalovirus retinitis in patients with AIDS. Am J Ophthalmol *133*, 467-474. 10.1016/s0002-9394(02)01327-2.

Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.P., and Li, W. (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res *41*, e74. 10.1093/nar/gkt006.

Wang, X., Li, Q., He, S., Bai, J., Ma, C., Zhang, L., Guan, X., Yuan, H., Li, Y., Zhu, X., et al. (2022). LncRNA FENDRR with m6A RNA methylation regulates hypoxia-induced pulmonary artery endothelial cell pyroptosis by mediating DRP1 DNA methylation. Mol Med *28*, 126. 10.1186/s10020-022-00551-z.

Weaver, M., Yingling, J.M., Dunn, N.R., Bellusci, S., and Hogan, B.L. (1999). Bmp signaling regulates proximal-distal differentiation of endoderm in mouse lung development. Development *126*, 4005-4015. 10.1242/dev.126.18.4005.

Wegener, M., and Muller-McNicoll, M. (2018). Nuclear retention of mRNAs - quality control, gene regulation and human disease. Semin Cell Dev Biol *79*, 131-142. 10.1016/j.semcdb.2017.11.001.

Wu, S.M., Fujiwara, Y., Cibulsky, S.M., Clapham, D.E., Lien, C.L., Schultheiss, T.M., and Orkin, S.H. (2006). Developmental origin of a bipotential myocardial and smooth muscle cell precursor in the mammalian heart. Cell *127*, 1137-1150. 10.1016/j.cell.2006.10.028.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (N Y) *2*, 100141. 10.1016/j.xinn.2021.100141.

Xia, H., Gilbertsen, A., Herrera, J., Racila, E., Smith, K., Peterson, M., Griffin, T., Benyumov, A., Yang, L., Bitterman, P.B., and Henke, C.A. (2017). Calcium-binding protein S100A4 confers mesenchymal progenitor cell fibrogenicity in idiopathic pulmonary fibrosis. J Clin Invest *127*, 2586-2597. 10.1172/JCI90832.

Xu, T.P., Huang, M.D., Xia, R., Liu, X.X., Sun, M., Yin, L., Chen, W.M., Han, L., Zhang, E.B., Kong, R., et al. (2014). Decreased expression of the long non-coding RNA FENDRR is associated with poor prognosis in gastric cancer and FENDRR regulates gastric cancer cell metastasis by affecting fibronectin1 expression. J Hematol Oncol *7*, 63. 10.1186/s13045-014-0063-7.

Yang, J., Li, X., Li, Y., Southwood, M., Ye, L., Long, L., Al-Lamki, R.S., and Morrell, N.W. (2013). Id proteins are critical downstream effectors of BMP signaling in human pulmonary arterial smooth muscle cells. Am J Physiol Lung Cell Mol Physiol *305*, L312-321. 10.1152/ajplung.00054.2013.

Yang, L., Zhou, F., Zheng, D., Wang, D., Li, X., Zhao, C., and Huang, X. (2021). FGF/FGFR signaling: From lung development to respiratory diseases. Cytokine Growth Factor Rev *62*, 94-104. 10.1016/j.cytogfr.2021.09.002.

Yeo, N.C., Chavez, A., Lance-Byrne, A., Chan, Y., Menn, D., Milanova, D., Kuo, C.C., Guo, X., Sharma, S., Tung, A., et al. (2018). An enhanced CRISPR repressor for targeted mammalian gene regulation. Nat Methods *15*, 611-616. 10.1038/s41592-018-0048-5.

Yin, Y., Yan, P., Lu, J., Song, G., Zhu, Y., Li, Z., Zhao, Y., Shen, B., Huang, X., Zhu, H., et al. (2015). Opposing Roles for the lncRNA Haunt and Its Genomic Locus in Regulating HOXA Gene Activation during Embryonic Stem Cell Differentiation. Cell Stem Cell *16*, 504-516. 10.1016/j.stem.2015.03.007.

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Giron, C.G., et al. (2018). Ensembl 2018. Nucleic Acids Res *46*, D754-D761. 10.1093/nar/gkx1098.

Zhang, B., Arun, G., Mao, Y.S., Lazar, Z., Hung, G., Bhattacharjee, G., Xiao, X., Booth, C.J., Wu, J., Zhang, C., and Spector, D.L. (2012). The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. Cell Rep *2*, 111-123. 10.1016/j.celrep.2012.06.003.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137. 10.1186/gb-2008-9-9-r137.

Zheng, Q., Zhang, Q., Yu, X., He, Y., and Guo, W. (2021). FENDRR: A pivotal, cancer-related, long non-coding RNA. Biomed Pharmacother *137*, 111390. 10.1016/j.biopha.2021.111390.

Zhu, L.J. (2013). Integrative analysis of ChIP-chip and ChIP-seq dataset. Methods Mol Biol *1067*, 105-124. 10.1007/978-1-62703-607-8_8.
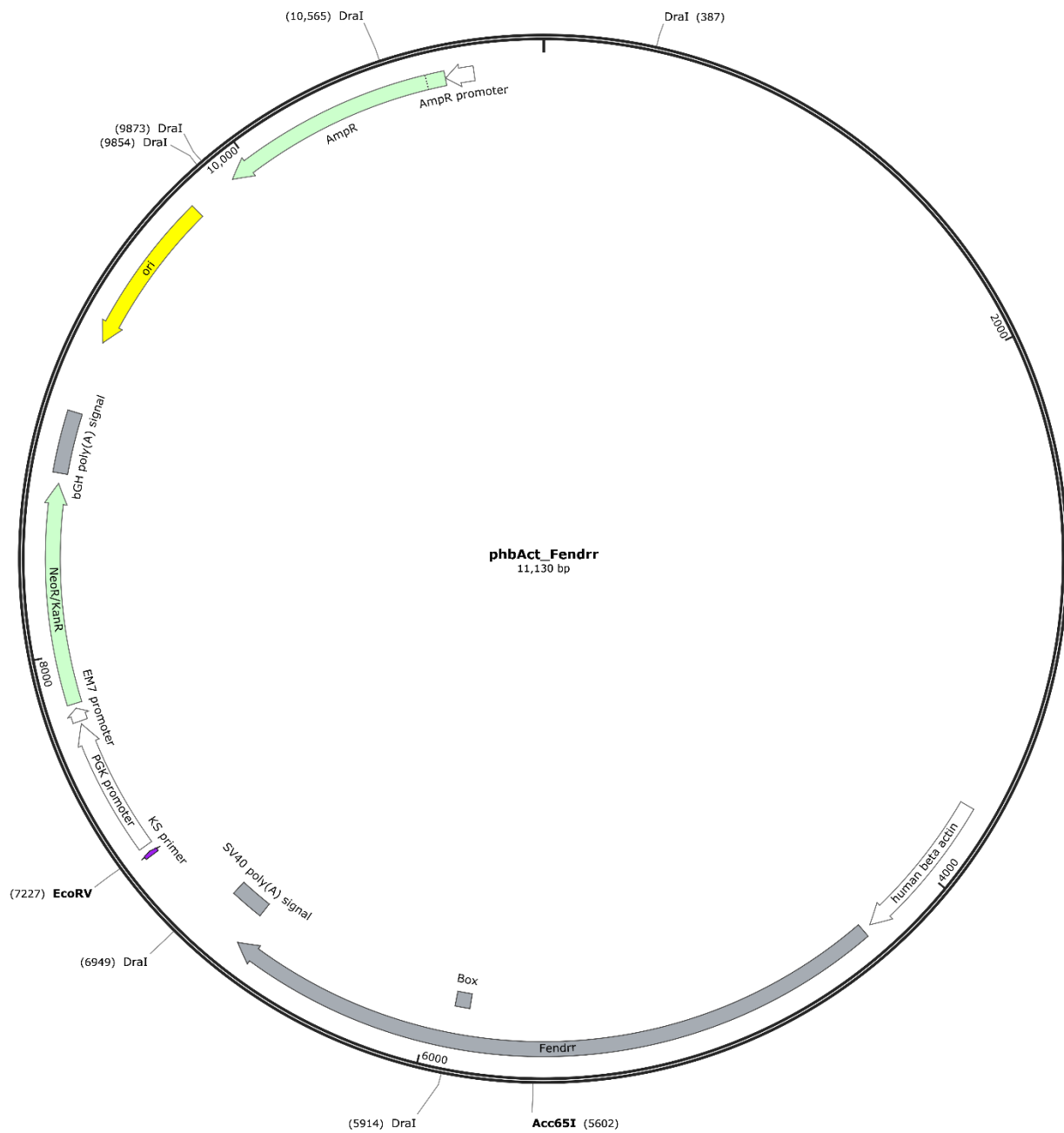
Appendix

Material and Methods
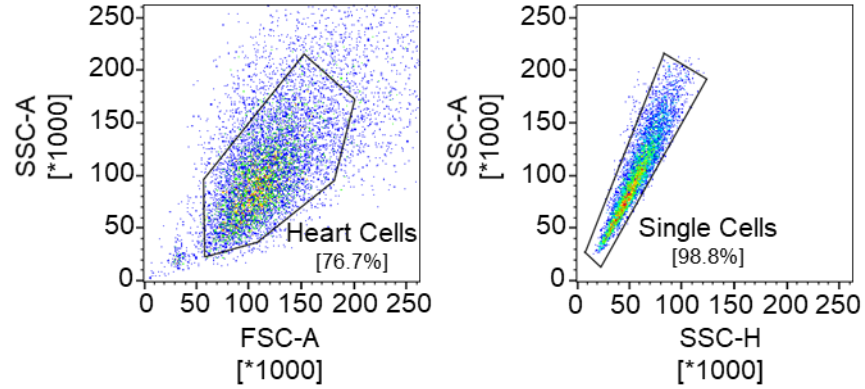
Appendix Fig. 1: Gating strategy for cardiac single cells for FACS sorting. Gates were set to remove debris, clusters and doublets and obtain single cells for the final fluorescence sorting.
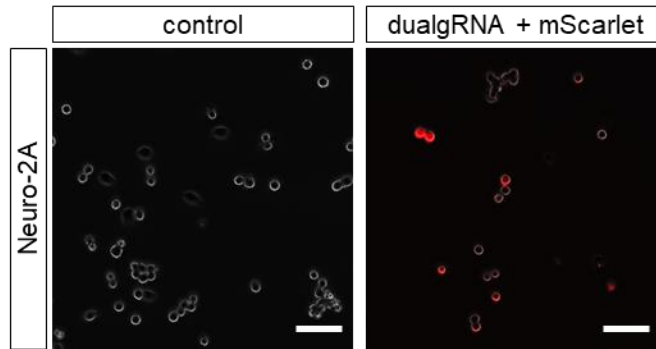

Appendix Fig. 2: Presence of mScarlet red fluorescence in Neuro-2A cells transfected with dualgRNA + mScarlet plasmid. The picture was taken 48 hours after Lipofectamine aided transfection of 1 µg plasmid. The white line represents 100 µm.

Appendix Fig. 3: *Fendrr*-**overexpression plasmid.** The plasmid was received as a kind gift from Dr. Phillip Grote (Transgenic Core Facility, Georg-Speyer-Haus, Frankfurt am Main). The part later removed for overexpression of *Fendrr* lacking the *FendrrBox* is marked with 'Box'. All restriction sites used for modification as described in the Material and Methods are shown. The full-length sequence was determined by Sanger-Sequencing (Microsynth Seqlab) and illustrated using SnapGene Viewer®.

Swhtr
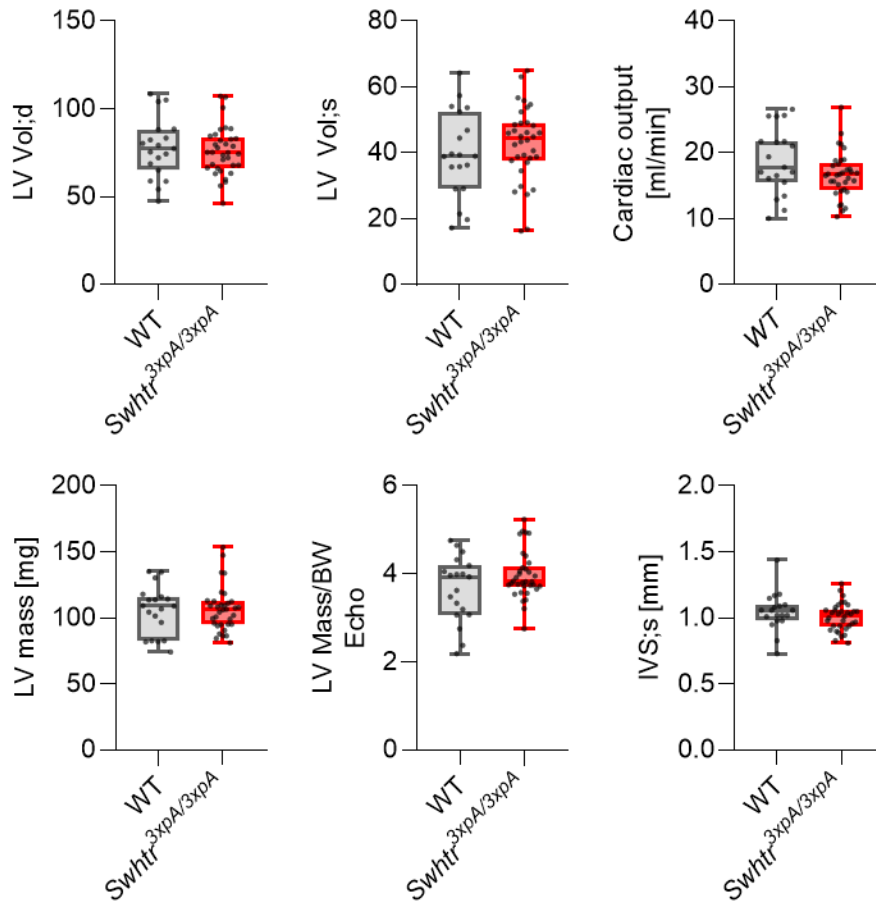
Appendix Fig. 4: Additional heart parameters of WT and *Swhtr^{3xpA/3xpA}* mice. Implicated parameters assessed by echocardiography of 8-week-old adult mice (WT n=19; *Swhtr^{3xpA/3xpA}* n=34). Statistical significance was tested by Two-way ANOVA. No statistically significant differences could be detected.
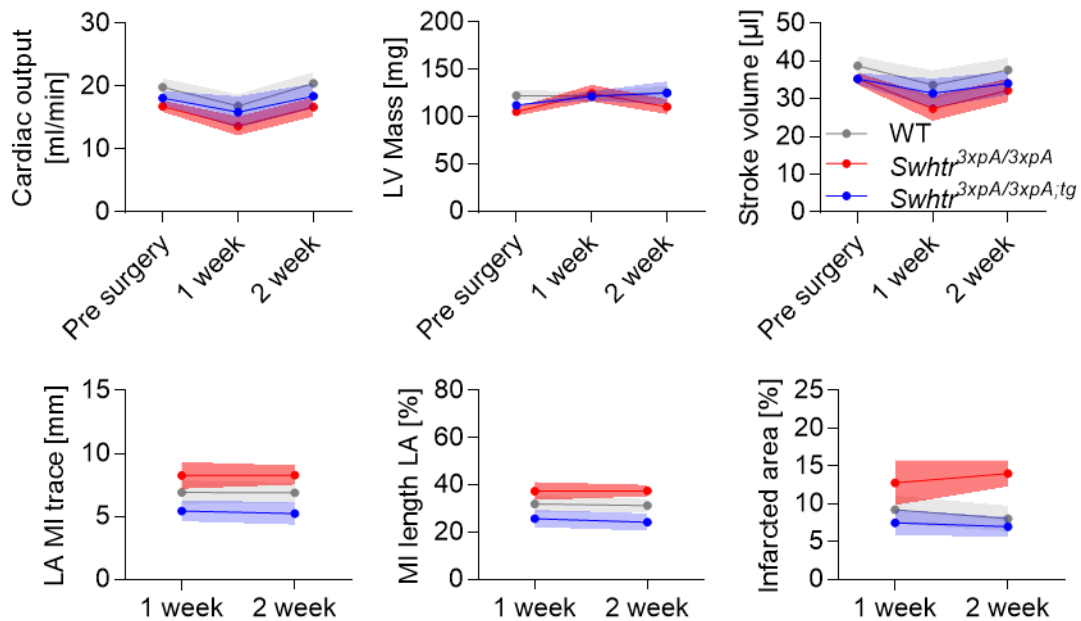
Appendix Fig. 5: Additional heart parameters of WT and *Swhtr* mutant mice before and after AMI. Implicated parameters assessed by echocardiography of mice before and after AMI (n=9 per genotype). Statistical significance was tested by Two-way ANOVA. No statistically significant differences could be detected.
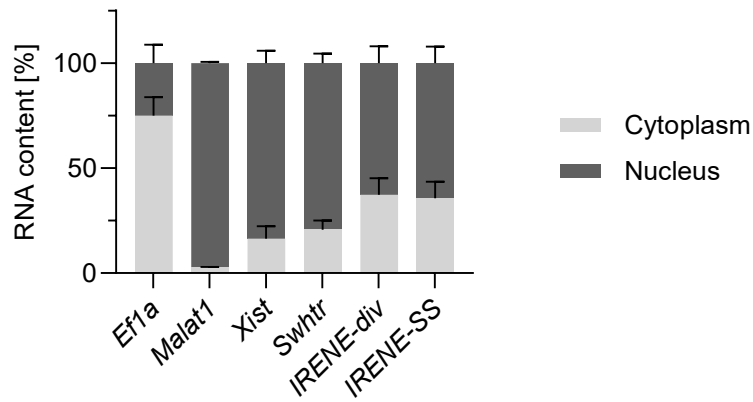
Appendix Fig. 6: Subcellular fractionation of HL-1 cardiomyocytes. RNA localization of the indicated transcript was assessed by qRT-PCR (n=3).
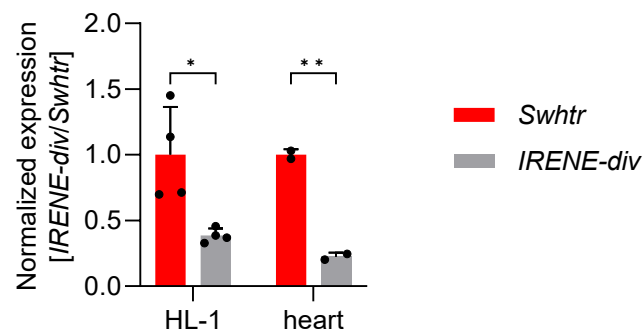
Appendix Fig. 7: Expression level of *Swhtr* and *IRENE-div* in HL-1 cardiomyocytes and heart tissue. RNA levels normalized to *Swhtr* demonstrate *Swhtr* as the predominant isoform. Statistical significance tested by t-test analysis. * < 0.05, ** < 0.01

Fendrr
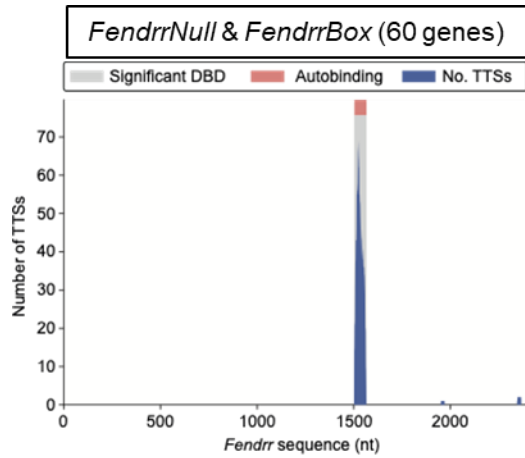
Rank	Gene	TTSs count	TTS cov.	LFC
1	*Emp2*	50	0.37	-1.07
2	*Clic5*	16	0.08	-1.16
3	*Hc*	5	0.05	-1.80
4	*Serpinb6b*	4	0.0450	-0.92
5	*Sftpc*	7	0.0330	-0.64
6	*Vnn1*	3	0.0460	-1.88
7	*Ager*	3	0.0440	-1.17
8	*S100g*	3	0.0330	-2.78
9	*Trim16*	3	0.0290	-1.18
10	*Akr1c14*	1	0.0250	-1.07
11	*Fn1*	1	0.0210	-0.68
12	*Rtkn2*	1	0.0210	-1.39
13	*Icam1*	1	0.0210	-0.92
14	*Timp3*	1	0.0200	-1.26
15	*Lmo7*	1	0.0200	-0.95
16	*Scd1*	1	0.0200	-3.08
17	*Phgdh*	1	0.0200	0.99
18	*Hnrnpdl*	1	0.0200	0.53
19	*Hbb-y*	1	0.0200	2.80
20	*Eif2s3y*	1	0.0200	25.64

Appendix Fig. 8: Potential direct targets of the *FendrrBox*. Triplex prediction analysis of the 60 shared dysregulated genes identifies 20 genes with a potential *Fendrr* triplex interacting site at their promoter. DBD = DNA Binding Domain on RNA, TTS = triple target DNA site (top). List of the 20 *Fendrr* target genes that depend on the *Fendrr* triplex and have a *Fendrr* binding site at their promoter. Figure taken from Ali, Rogala et al., 2023.
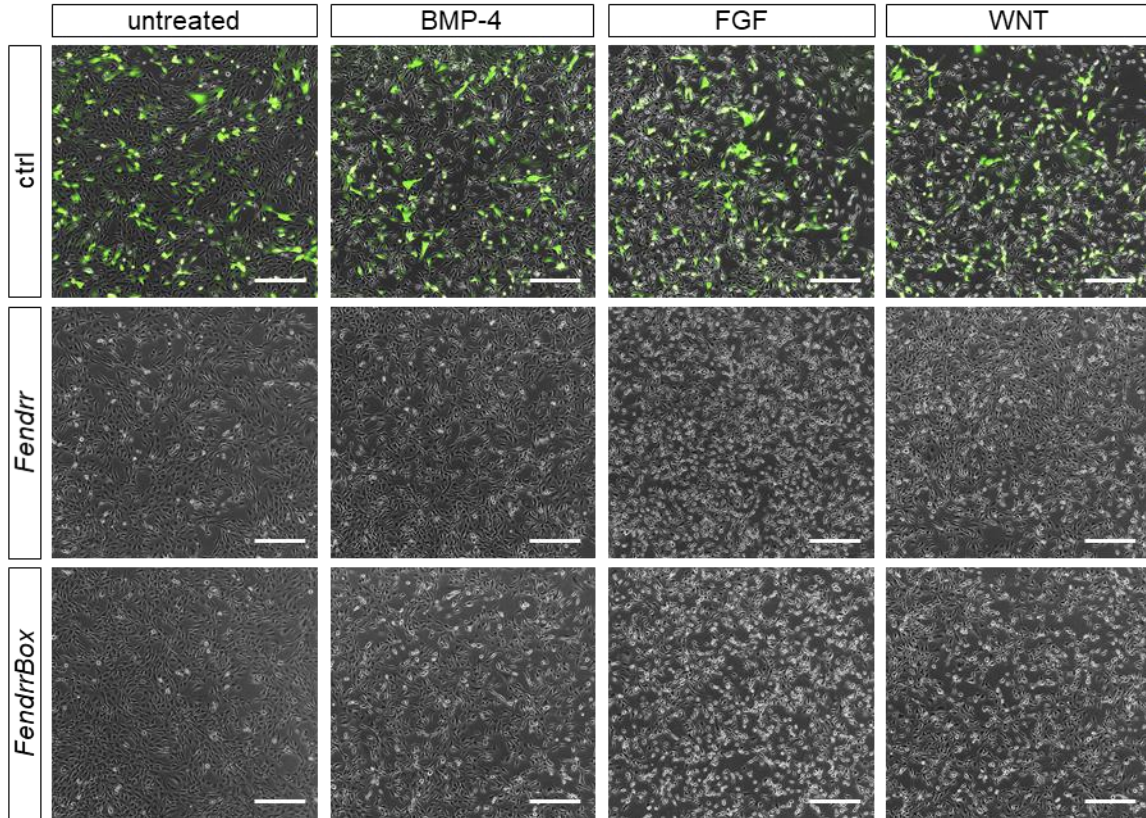
Appendix Fig. 9: Morphology of NIH3T3 after plasmid over-expression. Representative images of cells transfected with the indicated plasmid, treated with a ctrl (untreated), BMP-4, FGF or WNT. The white line represents 500 μm.
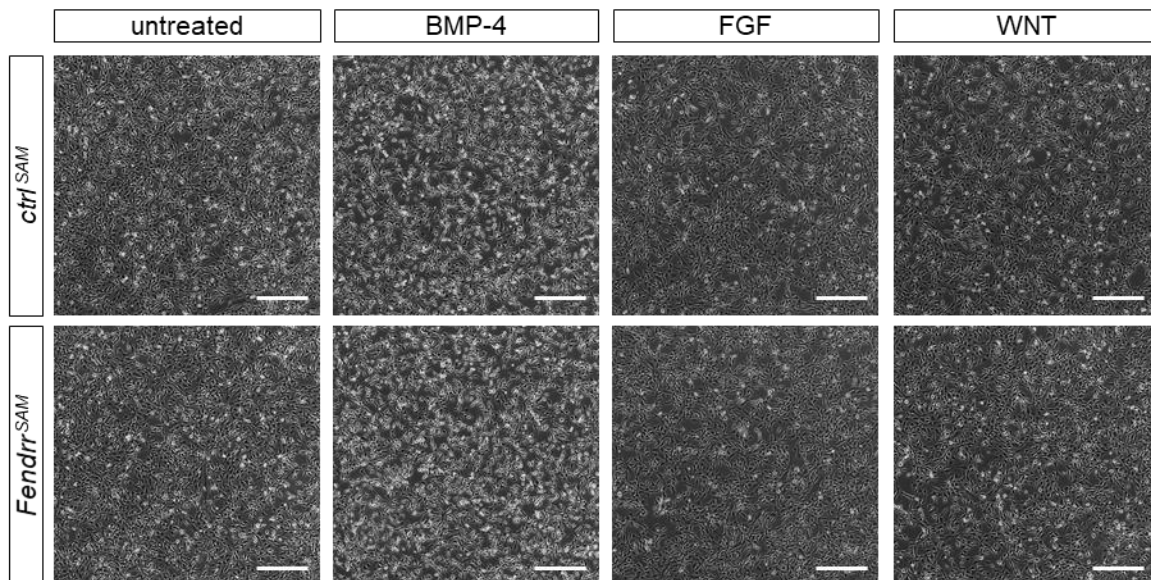

Appendix Fig. 10: Morphology of NIH3T3 after CRISPRa. Representative images of cells transfected with the indicated gRNAs, treated with a ctrl (untreated), BMP-4, FGF or WNT. The white line represents 500 μm.
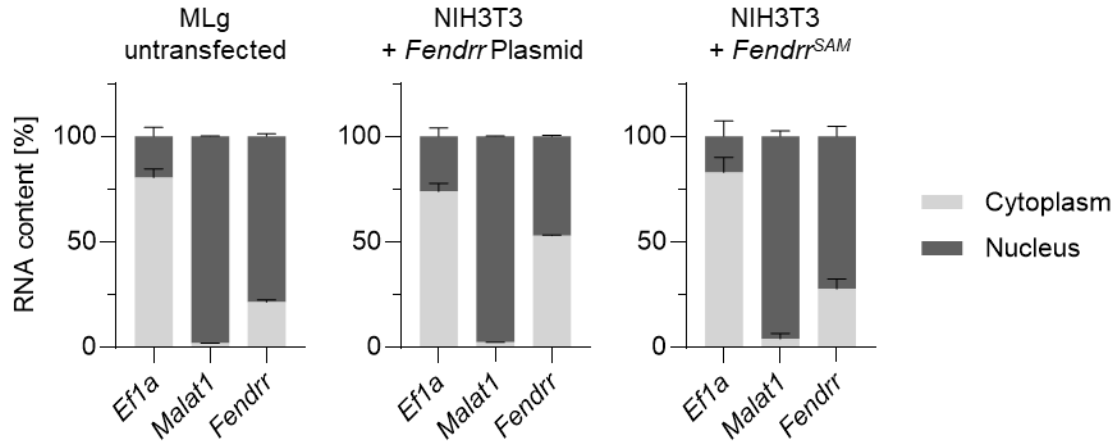
Appendix Fig. 11: Subcellular fractionation of untransfected MLg cells and transfected NIH3T3 cells.
Note that *Fendrr* localization in CRISPRa cells resembles the untransfected MLg WT situation while *Fendrr*
from the over-expression plasmid is located 50-50 in cytoplasm and nucleus (n=2).
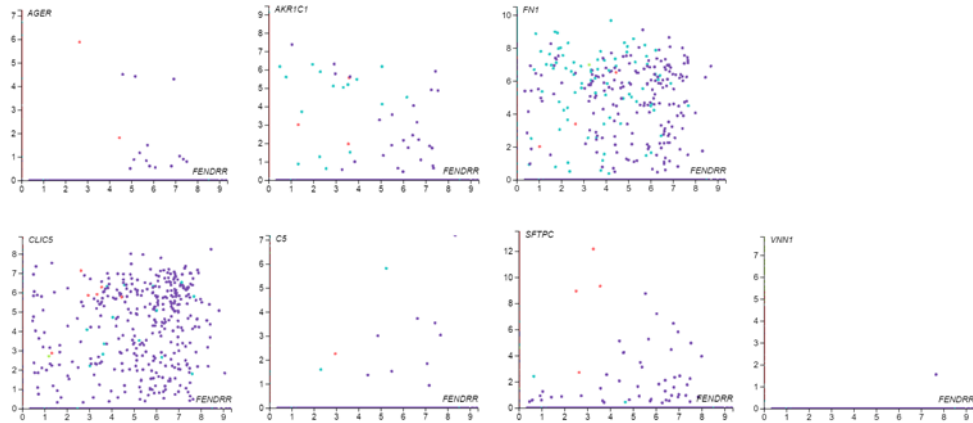
Appendix Fig. 12: Co-expression of *FENDRR* with remaining FRFT genes. Publicly available scRNA-
sequencing data from human lungs (Travaglini *et al.*, 2020) depicting dot plots of cells co-expressing
FENDRR with remaining FRFT genes indicating the cell type by color. Cyan = fibroblast, purple =
endothelial.

LncFsd2
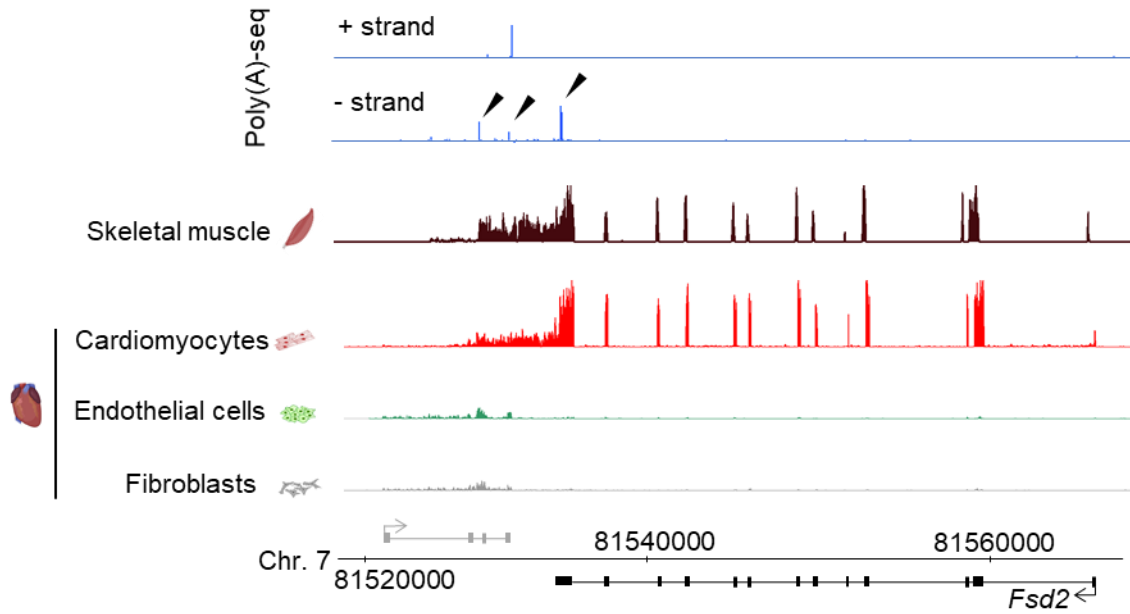
Appendix Fig. 13: *Fsd2* **polyadenylation-signals and 3'-UTR**. Publicly available strand-specific Poly(A)-sequencing (GSE111134) data showing three distinct polyadenylation-signals for *Fsd2* marked with arrowheads (top), RNA-sequencing tracks from skeletal muscle (GSE123879) and three major cell types of the heart.
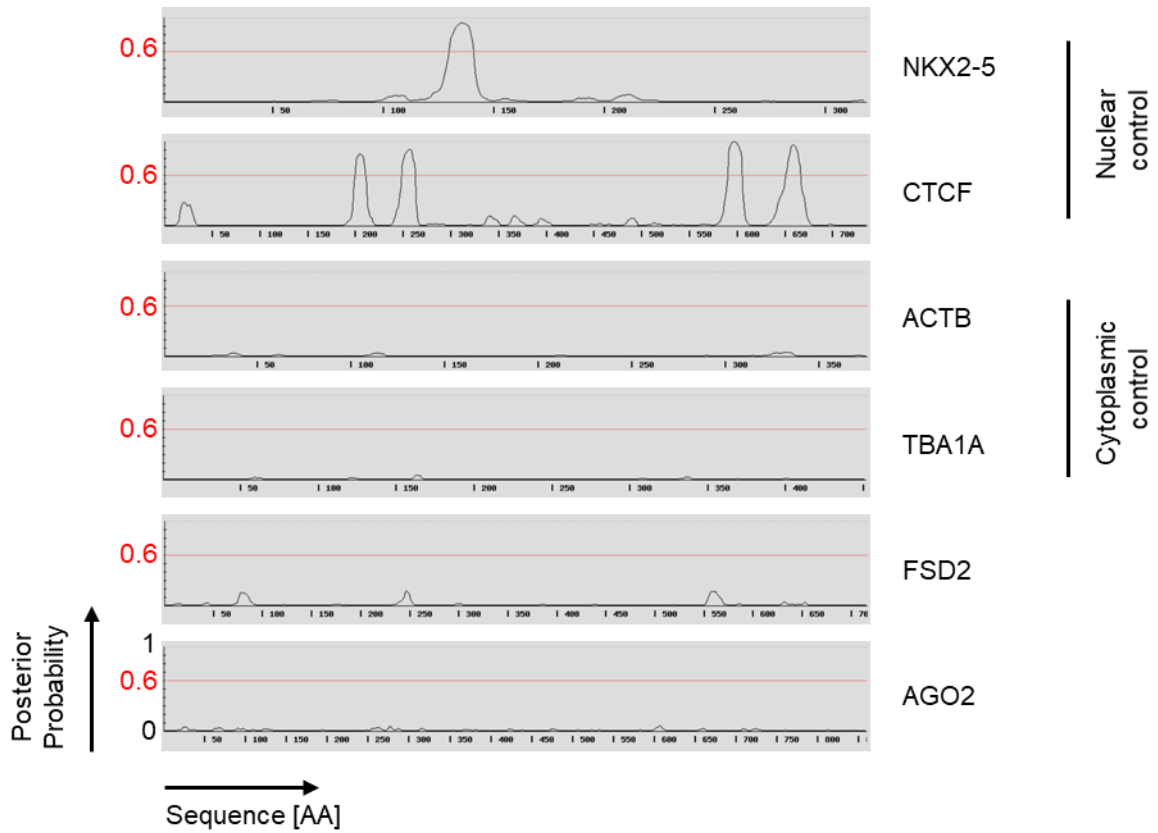
Appendix Fig. 14: Nuclear localization signal (NLS) prediction of FSD2, AGO2 and known nuclear and cytoplasmic proteins. NLStradamus (Nguyen Ba et al., 2009) derived NLS prediction of the indicated proteins. The threshold is marked in red. Axis descriptions given in the bottom left apply for all graphs.

Acknowledgements

It's nothing new that writing a dissertation involves walking a long, rocky road. Some days this path is easier to walk, others you feel like you have to fight against a high slope and stormy winds and every once in a while, you are overrun by an avalanche of stones. These are the moments when you have to fight your way out, through perseverance and persistence; but they are also the moments when you are grateful for every helping hand that is extended to you. I, too, had the unspeakable good fortune not to have had to tread this path alone. First and foremost, I would like to thank my supervisor Phillip. He not only gave me the opportunity to work on the exciting projects, which are described in more detail in the following work, but also always stood by me with advice and support, while he still gave me enough freedom to develop my scientific independence, to express my own ideas and to drive the projects forward on my own account. Even beyond the work he was always up for a good talk, being the heart and soul of the laboratory.

In the following I would like to mention Michaela. Thank you very much for the support from the university in the form of being the first supervisor of this thesis. In addition, the many discussions, including those at the poster sessions at the SFB meetings, have helped me to present this thesis in a form that fills me with pride. Similarly, Jaya deserves mentioning, as a referee of this thesis, but also for the scientific discussions and support – especially in regard to the *Swhtr* project.

I have been lucky enough to be able to complete my work on this thesis in a laboratory with generally positive atmosphere, with co-workers that built each other up when one fell and created an environment of learning and understanding the foundation of life together. In particular I would like to mention my fellow doctoral candidate Hsuan, who has been not only a comrade in suffering, but a real friend, always ready for technical discussions but also moral support. Postdoctoral researcher Tamer was a great source of knowledge and always there to help me fill in the gaps left by my inexperience in terms of bioinformatics. Not least, the other members of the lab take the crown. Technicians Peggy, Nina, Matthias and Felix, who were always there with good talk and open ears in science and beyond while greatly helping to push the projects to a new level. Beyond that I would like to thank the students visiting the lab for rotations or Master projects, for being patient while I developed my abilities in terms of supervision, for always spreading

CURRICULUM VITAE   SANDRA ROGALA

EDUCATION

2018-present	PhD student at the Institute for Cardiovascular Regeneration (AG Grote) and later Georg-Speyer-Haus Frankfurt (Transgenic Core Facility) PhD program of the Transregio SFB "ncRNA in the cardiovascular system"
2018	Research assistant at Goethe-University Frankfurt (AG Kössl)
2015-2017	Goethe-University Frankfurt Master 'Physical Biology of Cells and Cell Interactions' Thesis: "Inflammation Associated Extracellular Vesicles Promote Tumor Initiation from Neural Stem Cells"
2012-2015	Goethe-University Frankfurt Bachelor 'Biosciences' Thesis: "The Effect of the Insecticide Pymetrozin on Neuronal Activity in the Hearing Organ of *Mecopoda elongata*"
2003-2012	Karl-Rehbein-Grammar school Hanau
1999-2003	Friedrich-Ebert-Primary school in Hanau – Klein-Auheim

QUALIFICATIONS & COMPETENCES

Technical excellence	Genome engineering of mouse stem cells
	Immunohistochemistry
	CRISPR/Cas9 genome manipulation
	Molecular Biology
	Tissue culture (Cell lines/primary cells)
Teaching	Supervising Bachelor student internships
	Supervising Master Students
Others	Proficient in oral and written English
	Presentation of projects at conferences in oral and Poster format

AWARDS

2022	Aegean Conference Travel Award in recognition of an excellent research contribution

PUBLICATIONS

*Ali T, ***Rogala S**, Krause NM, Bains JK, Melissari MT, Währisch S, Schwalbe H, Herrmann BG, Grote P. *Fendrr* synergizes with Wnt signalling to regulate fibrosis related genes during lung development via its RNA:dsDNA triplex element. Nucleic Acids Res. 2023 Jul 7;51(12):6227-6237. doi: 10.1093/nar/gkad395. PMID: 37207329; PMCID: PMC10325902.
*equal contribution

Rogala S, Ali T, Melissari MT, Währisch S, Schuster P, Sarre A, Boettger T, Rogg EM, Kaur J, Krishnan J, et al. (2022). The lncRNA *Sweetheart* regulates compensatory cardiac hypertrophy after myocardial injury. *bioRxiv*, 2022.2011.2014.516395. 10.1101/2022.11.14.516395. (under revision in Nature Communications)

List of common Abbreviations

Abbreviation	Meaning
3xpA	Triple Polyadenylation Signal
Ager	*Advanced Glycosylation End-Product Specific Receptor*
Akr1c14	*Aldo-Keto Reductase Family 1, Member C14*
AMI	Acute myocardial infarction
ASO	Anti-sense oligonucleotide
BMP-4	Bone Morphogenetic Protein 4
Cas9	CRISPR-associated protein 9
CD	Circular dichroism
CM	cardiomyocyte
CPAT	Coding Potential Assessment Tool
ctrl	control
CVD	Cardiovascular disease
DFO	Deferoxamine
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dsDNA	Double-strand deoxyribonucleic acid
Ef1a	*Elongation factor 1-alpha*
Emp2	*Epithelial Membrane Protein 2*
et al.	et alii (Latin: "and others")
FACS	Fluorescence-Activated Cell Sorting
FB	Fibroblast
FBS	Fetal Bovine Serum
Fendrr	*Fetal-lethal noncoding developmental regulatory RNA / Foxf1 adjacent non-Coding developmental regulatory RNA*
FGF	Fibroblast Growth Factor
Fn1	*Fibronectin 1*
Foxf1	*Forkhead box F1*
FPKM	Fragments per kilobase million
FRFT	Fibrosis related *FendrrBox* Target
Fsd2	*Fibronectin Type III And SPRY Domain Containing 2*
IgG	Immunoglobulin G
IVS	Interventricular septum
kb	kilobase

LAD ligation	Left anterior descending artery ligation
LNA	Locked nucleic acid
lncRNA	Long non-coding RNA
Malat1	*Metastasis Associated Lung Adenocarcinoma Transcript 1*
MI	Myocardial infarction
mRNA	Messenger RNA
Nkx2-5	*NK2 Homeobox 5*
ns	not significant
nt	nucleotide
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
Poly(A)	Polyadenylation
qRT-PCR	Quantitative real-time polymerase chain reaction
RACE	Rapid amplification of cDNA-ends with polymerase chain reaction
RIP	RNA-immunoprecipitation
RNA	Ribonucleic acid
rpm	Rotations per minute
Ryr	*Ryanodine*
SAM	Synergistic Activation Mediator
SDS	Sodium dodecyl sulfate
Serpinb6b	*Serpin Family B Member 6*
sgRNA	Single guide ribonucleic acid
smFISH	Single-molecular fluorescent *in situ* hybridization
Swhtr	*Sweetheart RNA*
TBS	Tris-buffered saline
tg	transgene
Trim16	*Tripartite Motif Containing 16*
UTR	Untranslated region
WGA	Wheat Germ Agglutinin
WNT	Wingless/Integrated
WT	Wild type