Molecular mechanisms underlying the development of Atrial Fibrillation

Matilde Stefanini

The Queen's College

University of Oxford

Master of Science by Research

Hilary 2018

ABSTRACT

Atrial fibrillation (AF) is the most common type of clinical arrhythmia, affecting roughly 1% of the population. While not fatal *per se*, it is associated with a significant increase in the risk of stroke, accounting for around a quarter of all strokes. AF does not have a clear cause, but it appears to be a progressive remodelling of the atrial tissue characterised by extensive electrical and structural modifications. Its main risk factors are age, hypertension, heart failure, obesity, diabetes mellitus and thyroid dysfunction. The overall goal of this project is to shed light on the molecular mechanisms that promote the development of AF.

The first part of the project focuses on the role of neuronal Nitric Oxide Synthase (nNOS) and dystrophin in AF. Previous studies reported that both proteins are reduced in the atrial tissues of patients with persistent AF. In addition, $nNOS^{-/-}$ mice showed an increased propensity to develop AF in response to pacing. Given that nNOS is physically associated with dystrophin at the membrane in atrial myocytes, I investigated whether a mouse lacking dystrophin (*mdx*) could be used as a model for AF development.

In the second part of the project, I investigated atrial samples from patients in sinus rhythm undergoing coronary revascularisation with the aim of identifying early changes linked to the development of post-operative AF (poAF). PoAF occurs in one-third of patients within a week of undergoing cardiac surgery. Although poAF is usually self-limiting, patients are more likely to develop

persistent AF on follow-up. I adopted a whole-transcriptome approach, profiling the mRNAs and microRNAs in samples of patients who developed poAF, and compared them to samples from patients who stayed in sinus rhythm postoperatively. Differences between the two groups may bring valuable insights on the atrial substrate that subtends the development of AF and uncover new therapeutic targets.

TABLE OF CONTENTS

ABSTRACT 1

1. INTRODUCTION 8

1.1. CLINICAL MANIFESTATIONS OF ATRIAL FIBRILLATION 8

1.1.1. Molecular alterations underlying Atrial Fibrillation 11

1.1.1.1. Transcriptomics of Atrial Fibrillation 12

1.1.1.2. The role of Nitric Oxide in Atrial Fibrillation 19

1.2. SYSTEMIC AND MOLECULAR EFFECTS OF DYSTROPHIN ABSENCE 21

1.2.1. The Dystrophin-associated glycoprotein complex 21

1.2.2. Dystrophinopathies 28

1.2.2.1. Duchenne Muscular Dystrophy 28

1.2.2.2. Becker Muscular Dystrophy 30

1.2.2.3. X-Linked Dilated Cardiomyopathy 30

1.2.3. Mouse models of Duchenne Muscular Dystrophy 32

1.2.3.1. Molecular features of the mdx mouse model 35

1.2.4. Clinical and molecular manifestations of Duchenne Muscular Dystrophy 40

1.2.4.1. Role of nitric oxide 42

1.2.4.2. Electrocardiographic features 44

1.2.5. Considerations on DMD research 47

1.3. PROJECT STRUCTURE 49

2. THE MDX MOUSE AS A MODEL FOR ATRIAL FIBRILLATION 51

2.1. INTRODUCTION 51

2.2. METHODS 52

2.2.1. Mouse models 52

2.2.2. Electrocardiographic records 55

2.2.3. Atrial Fibrillation inducibility protocol 56

2.2.4. Echocardiography 59

2.2.5. Protein extraction 62

2.2.6. Western Blots 63

2.2.7. Immunohistochemistry 64

2.2.8. NOS activity measurement 65

2.3. RESULTS 67

2.3.1. Mouse model 67

2.3.2. Electrocardiographic characteristics of mdx mice 69

2.3.3. mdx mice do not have increased susceptibility to AF 74

2.3.4. Echocardiographic characterisation of the mdx mouse 78

2.3.5. mdx hearts have reduced nNOS protein content....81

2.3.6. Increased NOS activity compensates for protein reduction in the atria of mdx mice....85

2.4. DISCUSSION....87

2.4.1. Cardiac characterisation of mdx mice....88

2.4.1.1. Electrocardiography....88

2.4.1.2. Echocardiography....88

2.4.2. The development of AF in mdx mice....91

3. TRANSCRIPTOMIC ANALYSIS OF POST-OPERATIVE ATRIAL FIBRILLATION .. 95

3.1. INTRODUCTION....95

3.2. METHODS....97

3.2.1. Patients....97

3.2.2. RNA extraction....98

3.2.3. mRNA profiling with microarray....99

3.2.4. miRNA profiling with small RNA-Sequencing....100

3.3. RESULTS....101

3.3.1. RNA profiling with microarray....101

3.3.1.1. First batch of samples....101

3.3.1.2. Sample size calculation 105
3.3.1.3. Second batch of samples 108
3.3.1.4. Combined analysis 111
3.3.2. Small RNA profiling with RNA-Sequencing 117
3.3.2.1. Sample size calculation 122
3.4. DISCUSSION 126
4. APPENDIX: R SCRIPTS 131
4.1. MICROARRAY ANALYSIS 131
4.1.1. Quality control of first batch of samples 131
4.1.2. Illumina_GEX_preprocessing_v1.0.R 137
4.1.3. Exploratory_QC_plot_functions_v1.3.R 140
4.1.4. Analysis of the first batch of samples 142
4.1.5. Quality control and analysis of second batch of samples 147
4.1.6. Combined microarray analysis 151
4.2. SMALL RNA-SEQUENCING 155
4.2.1. Quality control of RNA-Seq 155
4.2.2. Analysis of RNA-Seq 158
REFERENCES 162

1. INTRODUCTION

1.1. CLINICAL MANIFESTATIONS OF ATRIAL FIBRILLATION

Atrial fibrillation (AF) is the most common type of arrhythmia, affecting roughly 1% of the total population. It manifests as fast, irregular electrical activity of the atria, which is detected through standard electrocardiography (ECG) as irregular ventricular depolarisations (R waves) in the absence of conventional P waves. It is generally characterised by a progressive development, whereby paroxysmal AF episodes slowly become more frequent and long-lasting until the arrhythmia becomes persistent. While not life-threatening *per se*, AF causes a sharp increase in the risk of stroke, which is partially prevented by therapy with anticoagulants (Schotten et al., 2011).

Post-operative AF (poAF) is the most common complication of cardiac surgery, manifesting in approximately one-third of patients during the first week after surgery. The development of poAF is associated with prolonged hospital stay and increased stroke risk and mortality (Aranki et al., 1996; Mathew et al., 2004; Whitlock et al., 2014). The main independent predictors of poAF are age, male sex, and the presence of cardiovascular risk factors, such as heart failure, hypertension, obesity and diabetes (Danelich et al., 2014; Zakkar et al., 2015). The current guidelines for management of poAF recommend amiodarone or beta-blockers, and anti-coagulant therapy if AF persists for more than 48 hours.

However, while it is well established that "AF begets AF" there is no evidence that "sinus rhythm (SR) begets SR" as early recovery of SR with anti-arrhythmic drugs in patients with AF has no impact on the subsequent burden of AF or on patients outcome (Carlsson et al., 2003; Hohnloser et al., 2000; Opolski et al., 2004; Van Gelder et al., 2002; Wyse et al., 2002).

A number of mechanisms have been reported to trigger poAF, including perioperative sympathetic activation, hypoxia, inflammation, oxidative stress and electrolyte imbalance (Zakkar et al., 2015). However, patients who experience self-limiting postoperative AF are more likely to develop persistent AF thereafter (Açil et al., 2007; Lee et al., 2014), suggesting that a subclinical atrial cardiomyopathy may precede and subtend both conditions. To date, published studies have investigated single pathways of interest, but none has adopted a systematic approach to the discovery of local (atrial) factors that underlie the future development of AF.

Previous work proved that there are detectable differences between the atrial myocardium of patients who develop poAF and those who remain in sinus rhythm (SR) (Antoniades et al., 2012; Kim et al., 2008; Reilly et al., 2011). Samples of the right atrial appendage collected at the time of surgery from patients developing poAF show increased superoxide production from NOX2-containing NADPH oxidases, as well as increased atrial protein content of the $p22^{phox}$ and NOX2 subunits (Reilly et al., 2011). This is in line with previously reported results showing increased NADPH-oxidase activity in right atrial samples from patients who develop poAF after cardiac surgery (Antoniades et

al., 2012; Kim et al., 2008). These findings suggest that the pre-operative levels of atrial reactive oxygen species may be implicated in the new onset of poAF. Interestingly, the NAPDH-oxidase upregulation is present to a much lesser extent in samples from long-standing persistent AF, suggesting that alterations detected in later stages of the disease may not mirror early changes, and therefore may be poor therapeutic targets.

A number of additional parameters, measured pre-operatively, are reported to be associated with the development of poAF, such as increased levels of mono-amino oxidase activity (Anderson et al., 2014), brain natriuretic peptide (Cai et al., 2014), N-Terminal pro-brain natriuretic peptide (Iskesen et al., 2011), left atrial dysfunction (Haffajee et al., 2011), left ventricular dysfunction (Banach et al., 2007; Melduni et al., 2011), high heart-rate variability (Kinoshita et al., 2011), and elevated Na^+/K^+-ATPase activity (Tran et al., 2009).

All these findings are consistent with the notion that atrial changes may precede AF, thereby validating the prospective whole-transcriptome approach that I intend to use in this project.

1.1.1. Molecular alterations underlying Atrial Fibrillation

The myocardium of patients with AF progressively undergoes both electrical and structural remodelling. Such changes create a favourable environment for disorganised impulse conduction that prevents the synchronous and functional contraction of the atria. These features are not mutually exclusive and cooperate to permanently remodel the myocardium.

A hallmark of structural remodelling in AF is the increase of atrial fibrosis (Dzeshka et al., 2015), where the fibrotic tissue prevents the correct conduction from the sinoatrial node and may cause the emergence of re-entry loops. Atrial samples of patients with permanent AF show increased levels of pro-fibrotic molecules such as Angiotensin II (Goette et al., 2000; Li et al., 2001; Tsai et al., 2008) and Transforming Growth Factor (TGF) β (Brand and Schneider, 1996; Verheule et al., 2004).

Alterations in ion currents play a pivotal role in the development of AF. The atria of patients with AF are characterised by a decrease in L-type Ca^{2+}-channels and I_{CaL} density (Brundel et al., 2001; Van Wagoner et al., 1999), as well as various changes in potassium currents (Nattel et al., 2007). This determines the overall shortening of the action potential duration and refractory period in AF, which is a common feature of patients with AF (Schotten et al., 2011).

AF is also characterised by altered Ca^{2+} homeostasis in the myocardium, with increased release of Ca^{2+} from the sarcoplasmic reticulum via leaky Ryanodine Receptor channels (Neef et al., 2010; Vest et al., 2005). In turn, higher

intracellular Ca^{2+} favours the activation of Ca-activated proteases, such as calpain, which further contribute to the structural remodelling of the atrial tissue (Brundel et al., 2002; Goette et al., 2002). High intracellular Ca^{2+} levels induce long-term adaptations in human and animal models of AF, characterised by the activation of the calcineurin-nuclear factor of activated T-cell (NFAT) signalling pathway, which promotes myocyte hypertrophy (Bukowska et al., 2006; Crabtree and Olson, 2002; Lin et al., 2004).

Last, increased oxidative stress is reported to contribute to the development of AF (Mihm et al., 2001), with lower glutathione (Carnes et al., 2007) and increased reactive oxygen species (Youn et al., 2013).

1.1.1.1. Transcriptomics of Atrial Fibrillation

A number of recent studies have explored the global transcriptional profile of patients with AF.

The most rigorous study is from Barth and colleagues, who profiled atrial mRNAs in the left atrial appendage (LAA) of patients with AF and in patients in SR. Their work identified a general downregulation of atrial-specific transcripts in AF atrial samples, and, to a lesser extent, upregulation of ventricular-specific transcripts, as well as an increase in transcripts involved in metabolic activities (Barth et al., 2005). The top hits from their analysis are reported in Figure 1.

Affymetrix ID	Gene	Qualitative Change Microarray (pmAF vs SR)	Fold-Change Microarray	Fold-Change RT-PCR	Validation Status
214316_x_at	CALR	Down	1.3	2.6	+
204312_x_at	CREB1	Down	1.5	1.4	+
202765_s_at	FBN1	Down	1.4	2.0	+
216944_s_at	ITPR1	Down	1.3	1.0	−
208351_s_at	MAPK1	Down	1.3	1.7	+
210555_s_at	NFATc3	Down	1.4	2.1	+
209239_at	NFKB1	Down	1.2	2.0	+
205026_at	STAT5B	Down	1.2	1.7	+
229606_at	PPP3CA	Down	1.5	1.0	−
215253_s_at	DSCR1	Up	1.8	1.9	+
213453_x_at	GAPDH	Up	1.2	1.2	+
632_at	GSK3A	Up	1.2	0.7	−
228127_at	KCNK3	Up	1.2	2.6	+

Mean values of fold-changes were compared between all SR and pmAF tissue samples (same samples as for array experiments).

Figure 1: Top hits from differential gene expression analysis between AF and SR samples

The original table summarising the validated findings from (Barth et al., 2005).

A more recent study investigated the transcriptional changes in the LAA of 239 patients, 32 of who had no previous history of AF (noAF) (Deshmukh et al., 2015). The group of patients with a history of AF was then split into patients in SR at the time of surgery (AF/SR), and patients in AF at the time of surgery (AF/AF). All the comparisons were then carried out between the noAF, AF/SR and AF/AF group, with the assumption that transcripts altered in the AF/SR group would reflect AF susceptibility, while transcripts altered in the AF/AF group

would reflect AF persistence. To me, this grouping (AF/SR vs. AF/AF) may not reflect different disease states and may only pick up short-term transcriptional changes elicited by fast atrial activity, while failing to identify the large-scale, long-term alterations that characterise the myocardium of patients in AF. A comparison of noAF vs. AF (i.e., AF/SR + AF/AF) is not reported in the manuscript, and would be of great interest given the large sample size of this study. Regardless, the comparison of noAF with the AF/AF group highlights a number of differentially expressed transcripts (Figure 2), but the paper does not elaborate any further on the meaning of these changes and how they may relate to the pathophysiology of AF. A pathway analysis would be needed in order to draw meaningful conclusions on this dataset.

Lamirault and colleagues also ran a microarray study on right atrial appendages of AF vs. SR patients (Lamirault et al., 2006). Their top hits, validated with qPCR, are reported in Figure 3.

From a quick comparison of the top hits reported in the aforementioned three papers (Figure 1, Figure 2, Figure 3), it appears that there is no overlap of any kind between them. None of the significantly disregulated transcripts appears in list of top hits reported by the others. This observation hints to the fact that transcriptomic studies in the context of AF are still at an early stage, and no consensus on the transcriptional profile of AF has been reached yet.

AF/AF vs. NoAF (AF Susceptibility & Activity)				
Illumina ID	**Gene Name**	**p-value**	**FDR<0.05 adjusted p-value**	**Fold Change**
ILMN_1772612	ANGPTL2	2.36E-23	2.78E-19	4.3
ILMN_1668411	FHL2	5.49E-23	3.24E-19	4.4
ILMN_2355831	FHL2	5.62E-22	2.21E-18	4.1
ILMN_1666385	CALM3	3.43E-21	1.01E-17	3.8
ILMN_1733998	DHRS9	9.11E-21	2.15E-17	3.8
ILMN_1678643	FHL2	2.78E-18	5.47E-15	3.5
ILMN_1729980	RNF216	1.77E-16	2.98E-13	3.1
ILMN_2367239	RCAN1	3.56E-16	5.25E-13	3.2
ILMN_1783805	PNMA3	5.99E-16	7.85E-13	3.1
ILMN_1696675	CES2	1.04E-15	1.23E-12	3.2
ILMN_1793549	PTPN4	5.27E-15	5.65E-12	2.9
ILMN_2384181	DHRS9	4.10E-14	4.03E-11	2.9
ILMN_2305544	DBI	8.49E-14	7.71E-11	3.0
ILMN_2301722	PDE8B	1.04E-13	8.75E-11	3.0
ILMN_2404049	RBM38	1.56E-13	1.23E-10	0.40

Figure 2: Top differentially expressed transcripts in AF/AF vs. noAF

The original table summarising the top differentially expressed transcripts between the two groups (Deshmukh et al., 2015).

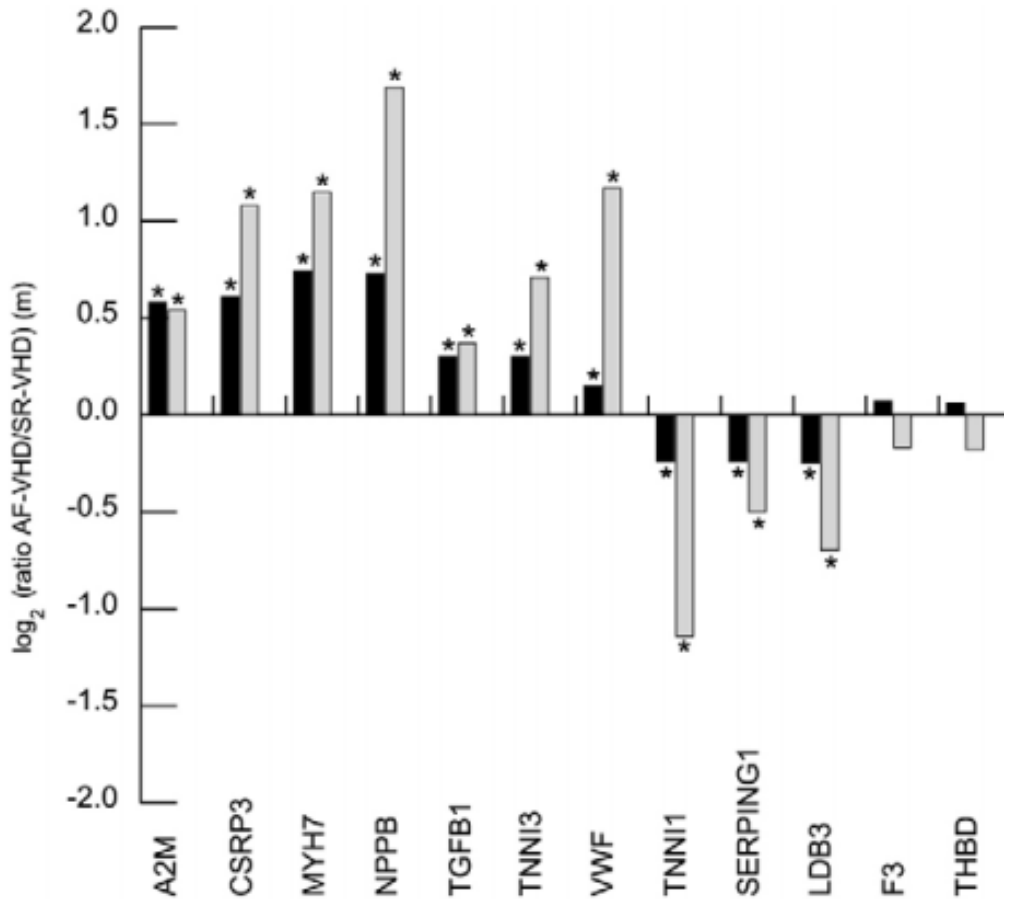


Figure 3: Top hits from the Lamirault paper

Differential gene expression between AF and SR measured via microarray (black bar) and qPCR (grey bar). * indicates statistical significance. (Lamirault et al., 2006)

While the transcriptomic studies mentioned above worked with atrial samples, the work by Lin and colleagues used blood samples collected from the Framingham Heart Study, comparing 2126 participants with no history of AF (noAF) to 177 participants with AF (prevalent AF), and with a third group of 143 patients who developed AF after the blood sample was taken (incident AF) (Lin et al., 2014). This study identified 7 transcripts over-expressed in the

prevalent AF group (Figure 4), but no transcript was significantly differentially expressed in the incident AF group. These altered genes have an exceptional potential as biomarkers for the diagnosis of AF. However, this finding should be approached with caution as the cohorts of patients were poorly matched for sex (women 58% vs. 36% vs. 39%) and co-morbidities (diabetes mellitus 15% vs. 32% vs 26%, myocardial infarction 1% vs. 23% vs. 5%; heart failure 3% vs. 25% vs. 6%), even though this was later corrected for during the statistical analysis.

Most significant transcripts associated with prevalent AF (FDR<0.05).

Transcript ID	Gene Symbol	Average Expression		Effect size	SE*	*P* value	FDR*
		AF (n=177)	No AF (n=2269)				
2364677	*PBX1*	6.85	6.71	0.17	0.03	2.8×10^{-7}	0.005
3712922	*C17orf39*	5.68	5.56	0.11	0.02	6.5×10^{-6}	0.037
3527514	*PNP*	7.22	7.11	0.14	0.03	7.1×10^{-6}	0.037
3804358	*C18orf10*	7.51	7.37	0.17	0.04	8.2×10^{-6}	0.037
3507710	*SLC7A1*	5.65	5.58	0.07	0.02	1.3×10^{-5}	0.046
3568534	*SPTB*	5.92	5.76	0.17	0.04	1.6×10^{-5}	0.047
2849469	*ANKH*	7.34	7.26	0.10	0.02	1.9×10^{-5}	0.048

*SE: Standard error; FDR: false discovery rate [32].

Figure 4: Top hits from the Framingham Heart Study paper

(Lin et al., 2014)

The paper by Cooley and colleagues offers an analysis of differentially expressed microRNAs paired with a sound statistical analysis (Cooley et al., 2012). Using human right atrial appendage (RAA) samples, they identified 47 differentially

expressed microRNAs (Figure 5), with miR-146b-5p and miR-21 among the upregulated in AF, and the miR-133 and miR-30 families among the downregulated.

miRNA	FC	Direction	*P* Value	False Discovery Rate
hsa-miR-146b-5p	6.5546	up	9.375E-06	0.001588
hsa-miR-142-3p	3.7638	up	0.0018188	0.023099
hsa-miR-21	3.5135	up	4.194E-05	0.00343
hsa-miR-1202	3.4125	up	0.0045582	0.037691
hsa-miR-142-5p	2.5911	up	0.000645	0.013653
hsa-miR-21*	2.554	up	0.0010058	0.016877
hsa-miR-483-5p	2.3163	up	0.0039893	0.03494
hsa-miR-223	2.2202	up	0.0010793	0.017133
hsa-miR-1308	2.193	up	0.005793	0.045274
hsa-miR-224	1.8652	up	0.0003876	0.010826
hsa-miR-16	1.7736	up	0.0003339	0.0106
hsa-miR-377	1.5647	up	0.0022217	0.025967
hsa-miR-337-5p	1.547	up	0.0044363	0.037691
hsa-miR-1290	1.5468	up	0.0031016	0.031217
hsa-miR-198	1.5362	up	0.0007135	0.014498
hsa-miR-490-3p	3.0632	down	8.251E-06	0.001588
hsa-miR-378*	2.2356	down	0.0001676	0.007738
hsa-miR-133a	2.1906	down	0.000259	0.009398
hsa-miR-143*	2.1808	down	8.515E-05	0.004325
hsa-miR-22*	2.1315	down	0.0016117	0.020993
hsa-miR-378	1.9828	down	0.003299	0.032229
hsa-miR-133b	1.9674	down	0.0027509	0.029114
hsa-miR-30a*	1.9245	down	0.0002356	0.009382
hsa-miR-29c*	1.9139	down	0.0038026	0.03404
hsa-miR-24-1*	1.9079	down	0.0033752	0.032351
hsa-miR-197	1.8873	down	0.0013192	0.018112
hsa-miR-30c	1.8584	down	0.0016087	0.020993
hsa-miR-139-5p	1.858	down	4.726E-05	0.00343
hsa-miR-125b-2*	1.8441	down	7.338E-05	0.004324
hsa-miR-30e	1.8362	down	0.0004928	0.011379
hsa-miR-99b	1.8246	down	7.661E-05	0.004324
hsa-miR-99a	1.7935	down	0.0020262	0.025105
hsa-miR-181c	1.771	down	0.0002401	0.009382
hsa-miR-331-3p	1.7535	down	5.939E-06	0.001588
hsa-miR-374b	1.7212	down	0.0004149	0.010826
hsa-miR-628-5p	1.7153	down	0.0045447	0.037691
hsa-miR-128	1.7018	down	0.0012997	0.018112
hsa-miR-181d	1.6812	down	0.0010299	0.016877
hsa-miR-145	1.6682	down	0.0031339	0.031217
hsa-miR-203	1.633	down	0.0050888	0.041033
hsa-miR-367	1.5954	down	0.005777	0.045274
hsa-let-7c	1.5817	down	3.237E-05	0.003289
hsa-miR-484	1.5752	down	0.0022945	0.025967
hsa-miR-490-5p	1.5532	down	0.0008714	0.015325
hsa-miR-30b	1.5447	down	0.0012841	0.018112
hsa-miR-149	1.5282	down	0.0004262	0.010826
hsa-miR-30a	1.5061	down	0.0037241	0.03404

miRNAs that were differentially expressed between VHD-RA-SR and VHD-RA-AF were defined by FC >1.5, false discovery rate <0.05, $P < 0.01$. miRNAs described as 'up' are higher in VHD-RA-AF than in VHD-RA-SR, whereas 'down' means expression is higher in the SR group.

Figure 5: Differentially expressed microRNAs from the Cooley paper

(Cooley et al., 2012)

A more recent study profiled microRNAs in the blood and right atrial tissues of AF patients (McManus et al., 2015). They identified miRNA-21 as downregulated in both the plasma and RAA of patients with AF (vs. SR), and miRNA-150 as downregulated in the plasma of patients with AF. The finding on miRNA-21 is in sharp contrast with what reported in the article by Cooley and colleagues (i.e., upregulation of miRNA-21 in AF). It is difficult to hypothesise why this may be.

The evidence from a few more studies (Chiang et al., 2015; Kharlap et al., 2006; Kim et al., 2005; Kim et al., 2003; Liu et al., 2014; Ruan et al., 2015; Su et al., 2018; Xiao et al., 2011; Xu et al., 2016) was not included as they appear to use p-values that have not been corrected for multiple comparison testing, and therefore are not good indicators of statistical significance. The work by Nishi and colleagues was also excluded as they do not report any statistics with their microarray results (Nishi et al., 2013).

1.1.1.2. The role of Nitric Oxide in Atrial Fibrillation

In the context of AF, a molecule of interest is the second messenger nitric oxide (NO). NO is synthesised by Nitric Oxide Synthases (NOS) through the conversion of Arginine into Citrulline. There are three isoforms of NOS: neuronal (nNOS or NOS1), endothelial (eNOS or NOS3) and inducible (iNOS or NOS2). eNOS is found in cardiac myocytes and endothelial cells where NO regulates β3-adrenergic receptor signalling (Moens et al., 2010) and flow-mediated

vasodilation (Shaul, 2002), respectively. iNOS is expressed in macrophages and other inflammatory cells where it contributes to pathogen clearance by forming reactive nitrogen compounds, and nNOS is primarily expressed in neurones and both in skeletal and cardiac myocytes (Simon et al., 2014; Stamler and Meissner, 2001).

In cardiac and skeletal myocytes, nNOS is physically anchored to dystrophin and other proteins of the dystrophin-associated protein complex (Lai et al., 2009; Lai et al., 2013a). Lack of dystrophin in the skeletal muscle, as observed in patients with Duchenne muscular dystrophy (see 1.2), leads to a strong reduction in nNOS (Brenman et al., 1995). This also seems to be the case in the cardiac muscle, although it is less well characterised (Stamler and Meissner, 2001).

Previous studies have shown that human AF is characterised by a reduction of dystrophin and nNOS in the atrial myocardium (Reilly et al., 2016). Moreover, mice genetically lacking nNOS display an increased susceptibility to AF (Reilly et al., 2016). Here, I sought to clarify whether a lack of dystrophin would cause a pro-arrhythmic phenotype in the *mdx* mouse.

1.2. SYSTEMIC AND MOLECULAR EFFECTS OF DYSTROPHIN ABSENCE

1.2.1. The Dystrophin-associated glycoprotein complex

Dystrophin is encoded on human chromosome X, spanning over 2 million basepairs on the human genome. Mutations in this gene cause a relatively common form of inherited muscular dystrophy, called Duchenne Muscular Dystrophy (DMD).

The dystrophin gene is composed by 86 exons, which, through alternative splicing, can combine to generate different isoforms of the protein. In addition, there are a number of different promoters from which the transcription of the mRNA can start. The most commonly active are the M, B and P promoters, giving rise respectively to the muscle, brain and Purkinje isoforms of the protein (Figure 1). These three isoforms differ only in an alternative first exon, and all generate the full-length 427 KDa isoform of dystrophin. This isoform has three main domains: an N-terminal domain which binds actin, a central rod domain characterised by spectrin-like repeats, and C-terminal domain which contains the binding sites for various proteins of the Dystrophin-Associated Glycoprotein Complex (DAGC).

Transcription from alternative promoters located inside the dystrophin gene gives rise to a variety of shorter isoforms, which are called Dp260, Dp140,

Dp116 and Dp71, according to their reported molecular weight (Muntoni et al., 2003).

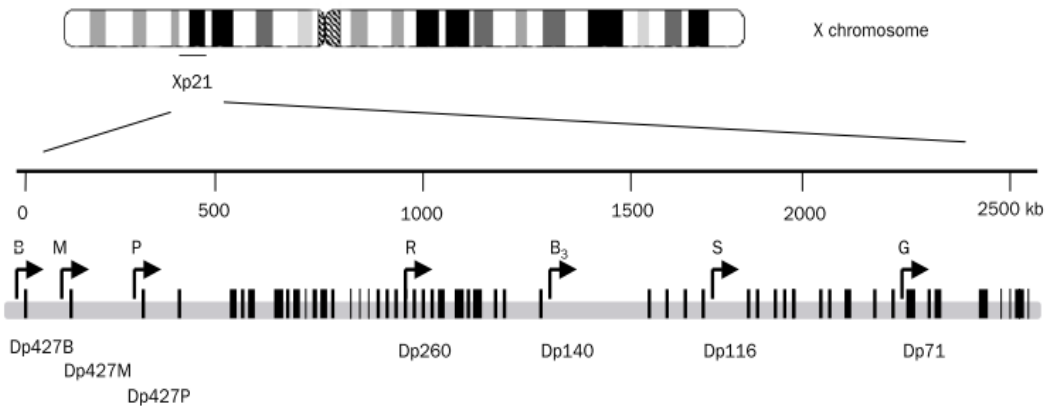


Figure 6: Arrangement of alternative promoters and exons of the dystrophin gene.

(Adapted from Muntoni et al., 2003).

Dystrophin links the intracellular actin cytoskeleton to integral membrane proteins, which in turn function as a bridge with the extracellular matrix. This protein complex is often referred to as DAGC (Rando, 2001) (Figure 2) and has a key structural role in maintaining membrane integrity, helping to anchor the cell to the extracellular matrix. This is particularly important for contractile cells, like myocytes, which depend on this anchoring to produce force and resist to mechanical stress. Indeed, the skeletal muscles of DMD patients exhibit signs of damage and repair, and *mdx* mice (i.e., mouse model lacking dystrophin, see 1.2.3) show increase exercise-induced damage to the sarcolemmal membrane (Brussee et al., 1997). Additionally, the DAGC assembles proteins with key

signalling activities, such as neuronal Nitric Oxide Synthase (nNOS or NOS1), which are able to exert their function in the local membrane environment. For instance, the nitric oxide produced by nNOS regulates the tone of the vessels which bring oxygenated blood to the muscles (Björnberg et al., 1990; Brock et al., 1998; Hickner et al., 1997; Hussain, 1998).

The main components of the DAGC are:

- **Dystroglycan**: a key protein of the DAGC, dystroglycan directly links the intracellular actin cytoskeleton with proteins of the extracellular matrix (Ibraghimov-Beskrovnaya et al., 1993). Alterations in dystroglycan glycosylation, either via direct mutation of the gene or of modifying proteins, have been associated with human muscular dystrophies (Dinçer et al., 2003; Godfrey et al., 2007; Hara et al., 2011; Hayashi et al., 2001). Although the molecular mechanisms are not fully understood, the defective glycosylation of dystroglycans seems to prevent or alter its binding affinity to key components of the basal membrane, such as laminin and agrin. This, in turn, leads to developmental and progressive abnormalities in affected individuals. It should be noted that dystroglycan gene deletion is embryonically lethal, indicating a necessary role for this protein very early in the mouse development (Williamson et al., 1997) (OMIM 128239).

- **Sarcoglycan**: sarcoglycans are integral membrane proteins whose function has not been fully elucidated. Although the sarcoglycan-null

mouse maintains the correct localisation of dystrophin at the membrane (Hack et al., 1998), animals develop a form of dystrophy that is similar to DMD (Barresi et al., 2000; Hack et al., 1999). Interestingly, one of the proteins that appear to be delocalised and downregulated when sarcoglycans are absent is nNOS. However, there is no clear mechanism explaining the pathogenesis of this type of dystrophy (Angelini et al., 2014).

- **Dystrobrevin**: has binding sites for dystrophin and syntrophin, and is reportedly associated with sarcoglycan and nNOS (Blake, 2002; Roberts, 2001). Mice lacking alpha-dystrobrevin develop skeletal and cardiac myopathies (Grady et al., 1999).

- **Syntrophin**: provides anchoring sites for a number of accessory proteins, including neuronal Nitric Oxide Synthase (nNOS) (Butler et al., 1992; Campbell and Kahl, 1989; Kramarcy et al., 1994). Its main role is thought to be associated with the development of functional synapses. The alpha1 isoform is prevalently expressed in cardiac and skeletal muscle, and a mutation in this isoform has been causally related to a long-QT syndrome (Ueda et al., 2008). The mouse model deficient for alpha1 syntrophin shows defective muscular regeneration and decreased tolerance to osmotic shock, which could be contributing factors in the disease progression of DMD. However, the authors did not investigate the cardiac phenotype of these mice, and is therefore difficult to define the

specific role of syntrophin in the cardiac muscle (Hosaka et al., 2002; Yokota et al., 2014).

- **NOS**: various proteins in the DAGC have been reported to bind NOS, including syntrophin, dystrophin and sarcoglycans. The association of both endothelial (eNOS) and neuronal NOS with the DAGC has been reported in different cell types. The association of eNOS has been indirectly proven, as it appears to bind to caveolin-3, a further protein associated to the DAGC (García-Cardeña et al., 1997). The association between eNOS and the DAGC has been reported in endothelial cells, but has not been investigated in myocytes (Palladino et al., 2013). Conversely, nNOS is able to bind directly to dystrophin though interaction with alpha-helices located in the central rod domain of the protein (Lai et al., 2013b). A further protein playing a role in anchoring nNOS to the membrane is caveolin-3, a membrane-associated protein present in skeletal and cardiac cells (Brenman et al., 1995; Sander et al., 2000; Thomas et al., 1998; Thomas et al., 2003). nNOS displays cell type-specific localisation: while in the skeletal muscle and atrial myocardium nNOS co-localises with dystrophin at the sarcolemma, in the ventricular myocardium it is predominantly located to the sarcoplasmic reticulum (Johnson et al., 2012).

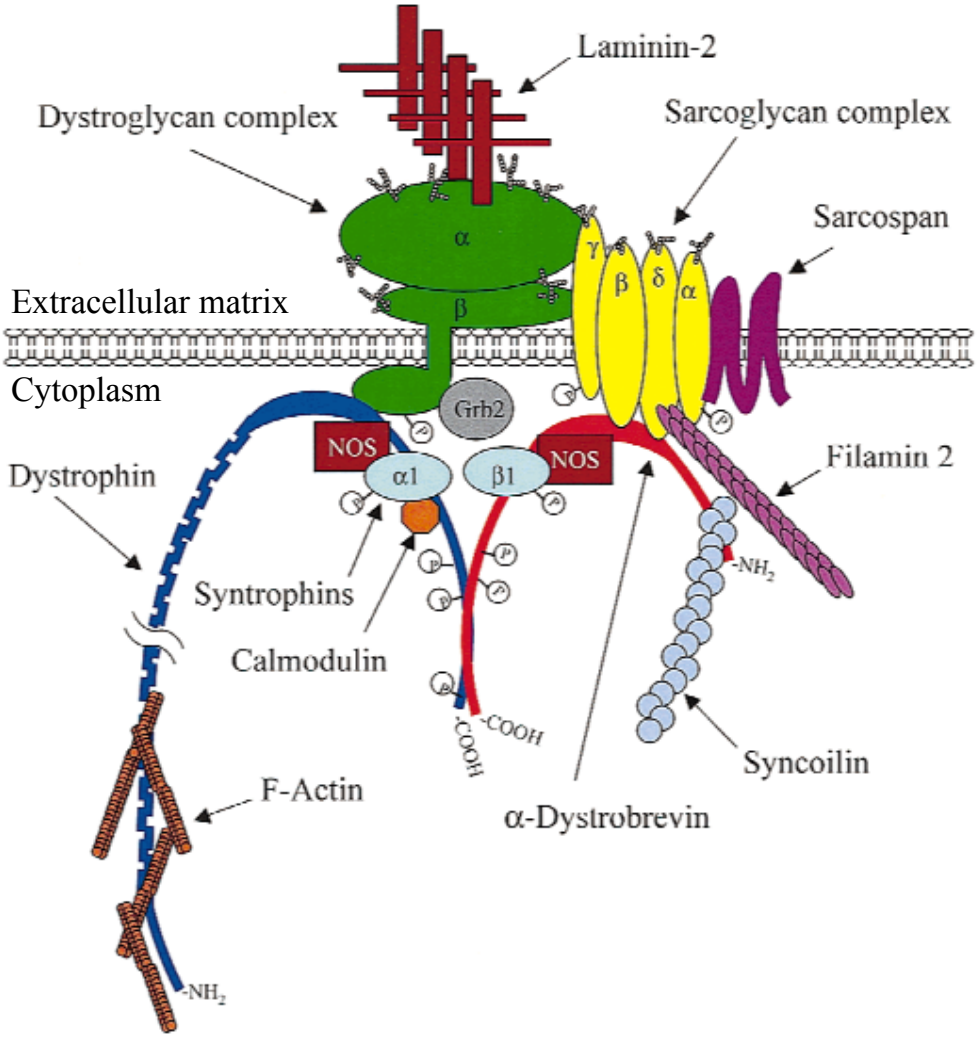


Figure 7: The dystrophin-associated glycoprotein complex

Dystrophin associates with a number of proteins at the plasma membrane level, forming an interlinked macromolecular network. Members of the dystrophin-associated glycoprotein complex include Nitric Oxide Synthases, Dystroglycan, Dystrobrevin and Syntrophins (Rando, 2001).

It is important to note that the DAGC does not only serve a fundamental structural function, but is also a centre for intracellular signalling. Proteins such as calmodulin and nNOS are associated with the complex, and are able to exert their signalling function at the membrane (Rando, 2001). When the complex is disrupted, as is the case in muscular dystrophies, disease progression is not confined to the lack of structural support to the plasma membrane, but also to an alteration of the signalling environment. A notable example is the association of calmodulin and calcium-calmodulin-dependent kinase protein II (CaMKII) to the DAGC. CaMKII is reported to phosphorylate both dystrophin and syntrophin, and thereby reduce their association in *in vitro* studies (Madhavan and Jarrett, 1994, 1999). In light of the binding of nNOS to both syntrophin and dystrophin, the CamKII pathway could be an interesting signalling pathway to further investigate.

1.2.2. Dystrophinopathies

1.2.2.1. Duchenne Muscular Dystrophy

DMD is an X-linked genetic disease caused by various different mutations in the dystrophin gene. These mutations prevent the normal translation of the protein, causing the absence or severe reduction of dystrophin in cells. Over a thousand different mutations have been described, and this is one of the features contributing to the heterogeneity of clinical phenotype in patients (OMIM 310200).

A meta-study on published datasets has estimated that the global incidence of DMD is 1-3 affected individuals in 10,000 males (Mah et al., 2014).

The disease manifests during childhood as progressive muscular weakness and wasting, with respiratory insufficiency and cardiomyopathy, and patients are wheelchair-bound by puberty. Additional symptoms that have a more variable manifestation are mental retardation and abnormalities of the retina. The variability of these phenotypic manifestations is likely to be related to (a) the use of different tissue-specific dystrophin promoters, and (b) the compensation by shorter isoforms in these tissues. Indeed, mutations associated with the loss of Dp71, the smallest dystrophin isoform, are invariably associated with mental retardation (Muntoni et al., 2003). It is not yet clear why no compensation is apparent in the cardiac and skeletal muscle, but it can be hypothesised that smaller isoforms are not able to sustain the mechanical stress derived from contraction.

To date, no effective cure is available for treating DMD patients. Pharmacological treatment with glucocorticoids delays the progress of the disease, but the prognosis is invariably fatal within the patient's third or fourth decade (Bushby et al., 2010; Fairclough et al., 2013).

Female carriers are generally asymptomatic, but can exhibit a spectrum of clinical signs. This is likely due to skewed X-chromosome inactivation, where by chance the majority of the active X-chromosomes carry the affected dystrophin allele (Giliberto et al., 2014; Viggiano et al., 2013). Although severe skeletal muscular disease is present in only a minority of cases, female carriers are at higher risk of developing cardiomyopathy (Florian et al., 2016; Iwase et al., 2010). For this reason, the American Academy of Paediatrics recommends a complete cardiac evaluation every 5 years for female carriers of the disease (Imbornoni et al., 2014; Surgery, 2005). Signs of cardiomyopathy, either hypertrophic or dilated, are not detectable in girls under the age of 20, but become apparent in 20-40% of older individuals when probed with echocardiography or magnetic resonance imaging (Nolan et al., 2003). Other manifestations include mild muscular weakness, cognitive impairment and increased levels of serum creatine kinase when compared to healthy females (Mercier et al., 2013).

1.2.2.2. Becker Muscular Dystrophy

In some instances, mutations in the dystrophin gene do not alter the overall reading frame of the protein, and give rise to truncated forms of dystrophin, which retain some degree of function. In these cases the phenotype of affected individuals tends to be milder, and the disease is classified as Becker Muscular Dystrophy (BMD). The age of onset of BMD tends to be higher than DMD, with many patients reporting the first muscular symptoms when they are in their 30s or 40s (Beynon and Ray, 2008) (OMIM 300376).

It is to be noted that BMD and DMD are not sharply separated, but should instead be regarded as a continuum, and there are some intermediate phenotypes among patients (Gillard et al., 1989; Koenig et al., 1989; Monaco et al., 1988).

1.2.2.3. X-Linked Dilated Cardiomyopathy

In some rare cases, the absence of dystrophin manifests almost exclusively with a progressive myocardial disease, and the disease is therefore called X-Linked Dilated Cardiomyopathy (XLDC) (OMIM 302045). It is not clear why in these cases the cardiac muscle is affected whereas skeletal muscles are spared. However, patients affected by XLDC show increased levels of serum creatine kinase, suggesting skeletal muscle damage at a sub-symptomatic level (Muntoni et al., 2003).

In some instances, XLDC is caused by mutations in the promoter region of the muscular isoform of dystrophin. There is evidence of upregulation of the Brain and Purkinje isoforms of dystrophin exclusively in the skeletal (but not in the cardiac) muscle of these patients. This suggests the presence of a compensatory mechanism by which isoforms B and P would be able to compensate for the absence of the M isoform (Muntoni et al., 1993; Neri et al., 2012; Obler et al., 2010).

1.2.3. Mouse models of Duchenne Muscular Dystrophy

Dystrophin is highly conserved among vertebrates and invertebrates. This has allowed the use of various animal models to research the pathophysiology of DMD, including *D. Melanogaster*, *C. Elegans*, zebrafish, mouse and dog. As in many areas of research, mouse models are the most widely used in studying DMD. In particular, the *mdx* mouse has been the first mouse model of DMD. This C57BL/10 strain originated from a spontaneous nonsense mutation in exon 23 of the murine dystrophin gene, preventing the translation of the dystrophin protein with an X-linked transmission pattern (Bulfield et al., 1984; Sicinski et al., 1989).

The phenotype of the *mdx* mouse is milder than the one displayed by DMD patients. Cycles of necrosis, inflammation and regeneration occur in the skeletal muscles of *mdx* mice from 1 month of age onwards, but the animals retain motor functions until late in their life, exhibiting in general only a mild phenotype (Mosqueira et al., 2013). When compared to wild-type mice, in addition to muscular symptoms, *mdx* mice display only a 15-20% reduction in lifespan and a slight increase in the incidence of sarcoma (Chamberlain et al., 2007). Although *mdx* mice do not display overt cardiac symptoms, molecular and functional alterations of the heart muscles are present upon investigation (Stuckey et al., 2012).

An interesting feature of this animal model, which could in part explain why the symptoms are milder than the human, is the overexpression of utrophin in skeletal muscles (Pons et al., 1991; Pons et al., 1994). Utrophin is a structural protein that has high sequence homology (roughly 80%) with the C-terminal domain of dystrophin (Love et al., 1989; Tinsley et al., 1992). In healthy skeletal muscle of mice and humans, utrophin is localised at the neuromuscular junction (Karpati et al., 1993; Perkins and Davies, 2002). Reports on the localisation of utrophin in cardiac tissue are inconsistent at times. While some report a colocalisation with dystrophin (Grady et al., 1997a), others report very low expression of utrophin in the myocardium (Matsumura et al., 1992). In skeletal muscles, the lack of dystrophin seems to induce an upregulation of utrophin, with relocalisation at the sarcolemma, both in the *mdx* mouse model (Grady et al., 1997b; Matsumura et al., 1992) and in DMD patients (Karpati et al., 1993; Mizuno et al., 1993).

A recent study showed that utrophin expression in the skeletal muscle of *mdx* mice is strongly induced by exercise (Gordon et al., 2014), suggesting that transcription may be activated or enhanced when regeneration and/or repair pathways are activated. Consistent with this hypothesis, the sarcolemmal localisation of utrophin has been reported in the *mdx* mouse at 4-5 weeks, an age when a strong wave of remodelling takes place (Grady et al., 1997b).

Due to its homology with dystrophin, utrophin is also able to bind proteins of the DAGC, and therefore its upregulation in the *mdx* mouse might explain the mild phenotype of these animals. However, this idea is not consistent with the

observation that utrophin is also upregulated in some DMD patients with a severe disease progression (Karpati et al., 1993; Mizuno et al., 1993). Nevertheless, the forced overexpression of utrophin in *mdx* mice prevents the disease progression, and has therefore been investigated as a therapeutic strategy (Perkins and Davies, 2002).

A way to induce a more severe phenotype in the *mdx* mouse to more closely resemble the human pathology is to force the animals to exercise (Granchelli et al., 2000; Hara et al., 2002). An alternative strategy is to generate a double-KO model where a second gene is silenced in addition to dystrophin. Various double-KO models for DMD have been proposed during the years, targeting structural proteins (e.g., utrophin, integrin, alpha-dystrobrevin) or proteins involved in regeneration pathways (e.g., telomerase, myoD) (McGreevy et al., 2015; Mosqueira et al., 2013). Although the disease progression of double-KO mice more closely resembles that of DMD patients, it should be noted that they carry a mutation in a gene that is still functional in DMD patients. It becomes therefore difficult to relate the discoveries done in these models to the actual pathophysiology of the human disease. An exception worth mentioning is the *mdx/Cmah*$^{-/-}$ model, as the gene coding for cytidine monophosphate sialic acid hydroxylase (Cmah) is also inactive in DMD patients. This mouse model is characterised by a more severe disease progression in both skeletal and cardiac muscles (Chandrasekharan et al., 2010).

1.2.3.1. Molecular features of the mdx mouse model

Several alterations of Ca^{2+} homeostasis are reported in the heart of *mdx* mice, contributing to the cardiomyopathy that develops in aged animals. In paced atria of *mdx* mice relaxation time is increased, and force depressed, hinting at a possible defective reuptake of Ca^{2+} in the sarcoplasmic reticulum (Sapp et al., 1996). Other changes include elevated resting Ca^{2+}, abnormal Ca^{2+} transients, increased expression of Ryanodine Receptor 2 (RyR2), decreased expression and phosphorylation of phospholamban (Williams and Allen, 2007a). There are consistent reports of increased Ca^{2+} influx in isolated ventricular myocytes, possibly from L-type Ca^{2+} Channels, as well as increased mitochondrial ROS production (Jung et al., 2008; Viola et al., 2013; Williams and Allen, 2007b).

Studies in neonatal cardiomyocytes report a reduction in the density of sodium channels, and a slower Ca^{2+} channel inactivation that are present before any sign of cardiomyopathy is detectable. Such changes are maintained in adult *mdx* mice. This evidence suggests that lack of dystrophin may affect ion channel localisation at the sarcolemmal membrane, effectively priming the dystrophic heart for the development of cardiomyopathy (Koenig et al., 2011; Koenig et al., 2013).

Studies employing echocardiography and magnetic resonance imaging (MRI) proved that various cardiac measurements are altered in *mdx* mice as young as 3-months, although the studies are not always in agreement. Stuckey and colleagues report increased RV end systolic volume and lower RV ejection fraction in mice aged 3-months and older. They also report a lower LV cardiac

output from the age of 6 months, and a lower LV ejection fraction at 12 months (Stuckey et al., 2012). Spurney and colleagues report an increased LV diameter in the hearts of 10-months old *mdx* mice, as well as a thinner LV posterior wall (Spurney et al., 2008). Zhang and colleagues report increased RV end-systolic and -diastolic volumes, reduced fractional shortening, and no difference in LV parameters in 8-months old *mdx* mice (Zhang et al., 2008).

Several articles report metabolic alterations in *mdx* cardiomyocytes, although the mechanisms are still unclear. Features of this metabolic impairment include shift from glycolysis to oxidative phosphorylation (Khairallah et al., 2007), reduced energy production (Zhang et al., 2008) and mitochondrial metabolic inhibition (Viola et al., 2014).

The first report of impaired endothelial function in *mdx* mice was published in 2001. Specifically, impaired flow-induced vasodilation was detected in isolated arteries, a mechanism mediated by eNOS-released NO (Loufrani et al., 2001). Consistently, it was reported that the endothelium of *mdx* mice is compromised, with reduced angiogenic activity and NO production by eNOS. Immunoprecipitation experiments also indicated a physical association between dystrophin and eNOS in endothelial cells, although it is not clear to what extent this would affect eNOS localisation and activity in *mdx* mice (Palladino et al., 2013).

Gene expression analysis in skeletal muscle indicates strong upregulation of Serca2 and NOX2 (Camerino et al., 2014). The same upregulation of NOX2 is apparent in *mdx* cardiomyocytes, correlating with increased total ROS and impaired contractility (Gonzalez et al., 2014). Contrasting evidence is reported by Spurney and colleagues, who show upregulation of NOX4 specifically in *mdx* left ventricles, and of NOX2 in skeletal muscles only (Spurney et al., 2008). The reason behind this discrepancy is not clear and cannot be determined without further investigation.

There are different reports on the alterations of the Serca expression in the *mdx*. Although the results do not always point in the same direction, there seems to be a general consensus on the slower Ca^{2+} uptake to the sarcoplasmic reticulum in dystrophic mice (Divet and Huchet-Cadiou, 2002; Kargacin and Kargacin, 1996). This may be due to increased expression of the inhibitor protein sarcolipin (Schneider et al., 2013). Consistently, increasing Serca1 or 2 expression by means of transgenes or viral vectors in the *mdx* showed beneficial effects, including lower susceptibility to contraction-induced damage (Morine et al., 2010; Mázala et al., 2015).

Fauconnier and colleagues described the role of RyR2 and Calstabin in *mdx* mice at 1 and 6 months of age. Their work highlights the progressive nature of pathological changes in the myocardium. While at 35 days there is no difference between wild-type and *mdx* mice, at 180 the level of nNOS is reduced and iNOS increased, while iNOS also co-precipitates with RyR2. It should be noted that all

experiments were carried out in raw heart lysates, and therefore there is lack of information of site-specificity (i.e., atria *vs* ventricles). RyR2 shows increased nitrosylation with disease progression, and increased leakiness due to reduction in Calstabin 2. Most differences appear to be reversed by treatment with a compound which enhances the binding of RyR2 to Calstabin 2, hence stabilising the channel (Fauconnier et al., 2010). A similar beneficial effect is reported upon inhibition of RyR2 phosphorylation by Protein Kinase A (PKA) (Sarma et al., 2010). More recently, amelioration of Ca^{2+} homeostasis in *mdx* cardiac muscle was induced by treatment with NOX and CaMKII inhibitor, but not with NOS or PKA antagonists (Figure 3) (Kyrychenko et al., 2013).

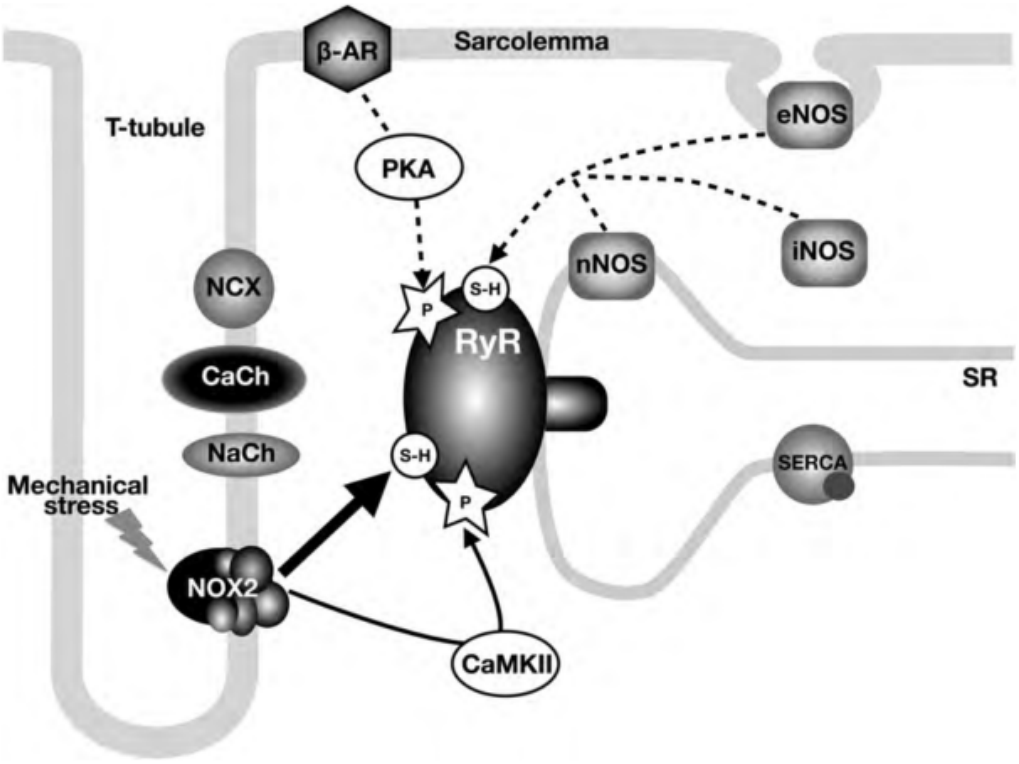


Figure 8: Signalling pathways regulating RyR in dystrophic cardiomyopathy

β-AR: beta adrenergic receptor; CaCh: Ca^{2+} channel; CaMKII: Ca^{2+}/calmodulin-dependent protein kinase II; NaCh: Na^{+} channel; NCX: sodium-calcium exchanger; eNOS: endothelial nitric oxide synthase; iNOS: inducible nitric oxide synthase; nNOS: neuronal nitric oxide synthase; NOX2: NADPH oxidase 2; RyR: ryanodine receptor; SERCA: sarco/endoplasmic reticulum Ca^{2+}-ATPase; dotted line: inhibitory effect; full line: activatory effect. (Kyrychenko et al., 2013)

1.2.4. Clinical and molecular manifestations of Duchenne Muscular Dystrophy

The first symptom of DMD is muscular weakness during childhood, characterised by a progressive decline that forces patients to start using a wheelchair before they reach puberty. The lack of dystrophin in the skeletal muscle makes the sarcolemma very prone to activity-induced damage. DMD muscle fibres are characterised by cycles of degeneration and repair, with inflammatory infiltrate. At later stages, the muscular tissue is almost completely replaced by connective and adipose tissue (Beynon and Ray, 2008).

The skeletal muscle phenotype is invariably associated with progressive cardiomyopathy, which is characterised by ventricular dilation and fibrosis in the late stages. Histological analysis of the conduction system revealed fatty acid infiltration, vacuolisation and fibrotic areas (Sanyal and Johnson, 1982). Cardiac failure, along with respiratory insufficiency, is one of the main causes of death in these patients.

Therapy and research have traditionally been focused on preventing or ameliorating the damage to skeletal muscle. Only in recent years it has emerged that cardiomyopathy is a major aspect and clinical target of DMD. In fact, a study on mice shows that therapeutically re-introducing dystrophin in the skeletal muscle alone increases the cardiac injury and dilated cardiomyopathy fivefold. It is speculated, although not directly proven, that the increase in voluntary activity consequent to the re-introduction of dystrophin in the skeletal muscles

may place additional demand on an already disease-prone myocardium (Finsterer and Cripe, 2014; Townsend et al., 2008).

At the molecular level, the absence of dystrophin seems to alter Ca^{2+} homeostasis in skeletal and cardiac muscle cells. It is not clear whether this may be due to increased membrane fragility, alteration of signalling molecules, or, more likely, to a combination of both. However, it is consistently reported that fibres from *mdx* muscles display increased intracellular Ca^{2+} responses to mechanical challenges, elevated resting Ca^{2+} levels and enhanced excitation-contraction coupling. In the long term, elevated Ca^{2+} can favour the activation of proteases and caspases, contributing to cell degeneration. Increased levels of reactive oxygen species are also reported in dystrophic heart muscle, likely due to increased expression and activity of NOX2-NADPH oxidase (Kaspar et al., 2009; Shirokova and Niggli, 2013).

Another key molecular feature of the disease is the decreased resistance of myocytes to stresses of various nature. Muscle fibres isolated from *mdx* mice are more sensitive to osmotic shock (Menke and Jockusch, 1991), and have reduced stiffness (Pasternak et al., 1995). A further, indirect proof of structural membrane alteration is the increased permeability of *mdx* muscle to dyes injected intravenously or in isolated preparations (Matsuda et al., 1995; Petrof et al., 1993). This could easily be attributed to the membrane rupture that happens when cells undergo apoptosis, but it seems that only a proportion of the dye-positive fibres co-localise with apoptotic markers (Hack et al., 1998). Interestingly, the chemical membrane sealer poloxamer has been shown to

prevent stretch-induced membrane damage and calcium overload in isolated murine myocytes (Yasuda et al., 2005). Poloxamer exerts its function by inserting directly into the lipid bilayer and promoting the tighter aggregation of other membrane lipids (Wu et al., 2004), thereby recovering the membrane structure of electro-permeabilised muscle fibres (Collins et al., 2007; Lee et al., 1992).

1.2.4.1. Role of nitric oxide

A phenomenon which may contribute to the disease progression in DMD patients is a defective functional sympatholysis in the exercising muscle. In healthy muscle cells, the nitric oxide (NO) produced by nNOS helps to counteract the vasoconstriction induced by alpha-adrenergic stimulation during exercise (Grange et al., 2001; Sander et al., 2000; Thomas et al., 1998). In DMD skeletal muscle, where nNOS is profoundly downregulated and no longer tethered to the sarcolemmal membrane by the dystrophin complex, alpha-adrenergic vasoconstriction is not counterbalanced by increased NO release, leading to local hypoxic conditions of the exercising muscles (Thomas et al., 2003).

When *mdx* mice are injected with norepinephrine, vasoconstriction is not attenuated by exercise, unlike the typical response seen in wild-type mice. A similar phenotype was recorded in $nNOS^{-/-}$ mice (Thomas et al., 1998). The same findings were confirmed in DMD patients, where oxygenation in exercising forearm muscle was reduced by 100%. Such effect was not seen in boys affected

by muscular diseases that do not alter the levels of nNOS, such as polymiositis and limb-girdle muscular dystrophy (Sander et al., 2000).

Studies on nNOS$^{-/-}$ mice clarified the role of this protein in normal muscle physiology. Indeed nNOS$^{-/-}$ mice are characterised by reduced body and muscle weight and reduced muscle force in the tibialis anterior. This phenomenon may be due to reduced oxygenation of the muscle during exercise, but the precise mechanism has not been specifically investigated (Church et al., 2011; Percival et al., 2008). Interestingly, two independent studies on double nNOS and dystrophin knock-outs found no difference to standard *mdx* mice with regard to skeletal muscle pathology (Chao et al., 1998; Crosbie et al., 1998). Additional research proposed nNOS as a regulator of activity-dependent muscle plasticity (Ito et al., 2013; Suzuki et al., 2007). This finding appears to be less relevant for DMD patients, as their ability to exercise is extremely limited from a very early age.

The central role of NO in the pathophysiology of DMD is also confirmed by the beneficial effect of interventions that increase NO availability in the muscle. A first therapeutic strategy employed NO precursor or donors, such as isosorbide dinitrate or L-arginine. These can be administered alone or in conjunction with corticosteroids or non-steroidal anti-inflammatory drugs (NSAIDs). Initial studies in *mdx* mice showed promising results (De Palma and Clementi, 2012). A pilot safety study treating DMD and BMD patients with isosorbide dinitrate in combination with ibuprofen showed good safety and tolerability. There was a

perceived improvement of quality of life measured through questionnaires; however, this may as well be indicative of a placebo effect as the trial was not blinded (D'Angelo et al., 2012).

Other experimental approaches on *mdx* mice showed amelioration of muscular and cardiac disease in response to overexpression of nNOS, either through a transgene or an adenoviral vector (Colussi et al., 2008; Thomas, 2013; Tidball and Wehling-Henricks, 2004; Wehling et al., 2001). However, such approaches in the context of DMD have not yet been applied on patients, and can only serve as proof of principle of the beneficial effects of increased NO availability in the muscle.

1.2.4.2. Electrocardiographic features

The dystrophic cardiomyopathy becomes symptomatic only at later stages of the disease, when the skeletal muscle function is already heavily compromised. However, DMD patients exhibit a wide array of electrocardiogram (ECG) abnormalities, detectable even at a very young age before the presentation of clinical cardiomyopathy symptoms. Rhythm alterations are attributed to the progressive replacement of cardiac tissue by fibrotic and adipose tissue.

Sinus tachycardia and both atrial and ventricular premature beats are the most frequently recorded anomalies (Finsterer and Stöllberger, 2003). Recently, sinus tachycardia was reported to be one of the first abnormalities that is observed in

young DMD patients, and it was shown to correlate with ventricular dysfunction (Finsterer and Cripe, 2014; Markham et al., 2006).

Sinus bradycardia, ventricular tachycardia and Atrioventricular block have also been reported in a minority of cases. Atrial arrhythmias like atrial fibrillation and flutter have been reported only anecdotally, with a total of 8 cases to date (Finsterer and Stöllberger, 2008). A number of studies used continuous ECG (Holter) recording over 24 hours to assess cardiac rhythm in DMD patients. Kirchmann et al. recorded abnormal heart rate variability and premature ventricular complexes, but no premature supraventricular beats, tachycardia, or bradyarrhythmia in a group of 39 subjects (Kirchmann et al., 2005). Lanza and colleagues report a 25% incidence of ventricular arrhythmias and abnormal heart rate variability in 60 DMD patients (Lanza et al., 2001). Similar findings were reported in a cohort of 55 patients in a study by Yotsukura (Yotsukura et al., 1995).

The incidence and severity of all types of arrhythmias appears to correlate with the disease progression in DMD patients. ECG characteristics that appear to be common to most patients are short PR interval, increased R/S amplitude and abnormally deep Q waves (Nigro et al., 1990; Perloff, 1984; Perloff et al., 1966; Perloff et al., 1967; Ronan et al., 1972; Sanyal and Johnson, 1982; Slucka, 1968; Yanagisawa et al., 1992).

A recent study reported a significant increase in resting heart rate in young (<10 yr) DMD patients, which persists after the development of the disease (Thomas, 2002).

Female carriers of DMD/BMD, aged between 6 and 16, did not show any cardiac abnormality upon examination (Nolan, 2003). However, half of DMD/BMD carriers will show signs of cardiac involvement later on in life (Finsterer and Stöllberger, 2003).

1.2.5. Considerations on DMD research

Research on DMD has been very active since the first studies in the fifties, and gives no sign of slowing down. However, there are still some major challenges and unanswered needs.

From a purely clinical point of view, there is still no effective therapy available. Even though patients can now survive up to the age of 40, thanks to corticosteroid treatment, physiotherapy and assisted ventilation, their quality of life is poor. New biologicals are in the pipeline and some are being tested in clinical trials (Fairclough et al., 2013). Much attention has been given to oligonucleotide-based therapies, such as exon-skipping strategies. These therapeutics have the great advantage of an easy synthesis that allows extensive preclinical studies, and seem to be generally well tolerated. However, the huge spectrum of mutations carried by DMD patients poses a challenge as it does not allow a one-for-all approach to drug design, but requires a different therapeutic if not for each patient, at least for each mutation.

Secondly, it seems that a disproportionate effort is put into studying an animal model that, for partially known reasons, does not represent the human disease very well. Studying the *mdx* mouse appears to be a research field *per se*, regardless of how the discoveries relate to the clinical phenotype. The studies that look for a human counterpart to their findings are the minority, and therefore caution needs to be applied when extrapolating knowledge to human DMD.

Lastly, the literature is very fragmentary. In general, each study only focuses on a single organ or cell type affected by the disease. Whereas the wealth of information accumulated on DMD patients and *mdx* mice through the years is titanic, it is very difficult to piece it together in a coherent and integrated manner. Historically, research has been focusing on skeletal muscles, as they are the most dramatically affected in DMD. Recently, scientists have started to investigate more carefully the role of dystrophin in the central nervous system and in the heart. As a consequence, there is now much better awareness of the importance of treating the cardiac symptoms alongside the skeletal ones from a very early stage.

This said, research on DMD still offers many opportunities and rewards, as new effective therapeutic strategies are urgently needed.

1.3. PROJECT STRUCTURE

My project comprises two well-defined parts. The first aims to shed light on the interaction between nNOS and dystrophin, and how the lack of one or the other may create a substrate for AF. These experiments were carried out in the *mdx* mouse, which is a genetic KO model for dystrophin.

The second part of my project aims to characterise the atrial transcriptome of patients who are in SR but develop poAF later on.

2. THE MDX MOUSE AS A MODEL FOR ATRIAL FIBRILLATION

2.1. INTRODUCTION

Our previous work has shown that long-standing AF in humans and goats is associated with a reduction of dystrophin and nNOS (Reilly et al., 2016). Moreover, mice genetically lacking nNOS display an increased susceptibility to AF (Reilly et al., 2016). nNOS is physically associated to dystrophin at the membrane, and a lack of dystrophin contributes to the reduction of nNOS protein (Brenman et al., 1995). A further molecule playing a role in the dystrophin-nNOS interaction is miR-31. This miRNA directly downregulates the expression of dystrophin and it is increased in the skeletal muscles of DMD patients and *mdx* mice (Cacchiarelli et al., 2011). The expression of miR-31 is also upregulated in the atria of patients with AF, where it targets the mRNA of both nNOS and dystrophin (Reilly et al., 2016).

Given this evidence, I wanted to clarify the functional relationship of dystrophin-nNOS in the atria, and how it relates to the development of AF. In particular, I hypothesised that a lack of dystrophin would cause a pro-arrhythmic phenotype in a mouse model.

2.2. METHODS

2.2.1. Mouse models

The C57BL/10ScSn-*Dmd*^{mdx}/J mouse, commonly known as *mdx*, carries a spontaneous mutation at position 3185 within exon 23 of the dystrophin gene located on the X chromosome, resulting in a termination codon, and is commonly used as a model for Duchenne Muscular Dystrophy (Jackson Laboratory reference 001801).

When I obtained the mice from the group of Prof. Wood (Department of Physiology, Anatomy and Genetics, University of Oxford), no reliable PCR-based method had been established in Oxford for routine genotyping of the *mdx* strain. Therefore, the colony was kept as homozygous dmd^{mdx}/dmd^{mdx}, while the wild-type counterparts (C57BL/10 mice) were purchased from a commercial supplier (Harlan). Such breeding strategy may lead to biased results due to genetic drift of the isolated colonies, thus I devised an alternative breeding scheme to obtain control mice for use in key experiments. My breeding strategy exploits the X-linked transmission of the dystrophin gene to obtain case-control mice with a shared genetic background, as illustrated in Figure 9. A male *mdx* will always pass its unaffected Y chromosome to its male offspring, while a female homozygous *mdx* will always pass an affected X chromosome to its male offspring. I mated these breeders with wild type BL10 mice, generating a male progeny bearing respectively a *mdx* and a wild-type phenotype with respect to the gene of interest (i.e., dystrophin), but with an overall genotype equally

determined by the *mdx* and BL10 original strain. The wild-type male offspring generated by mating a male *mdx* with a female BL10 will be indicated as WT-cross throughout the text, while the *mdx* males generated by mating a female *mdx* with a male BL10 will be indicated as *mdx*-cross throughout the text. Whenever possible, the starting breeding pairs for the WT-cross and *mdx*-cross colonies were two sets of *mdx* and BL10 siblings, thus ensuring maximum uniformity of the genetic background.

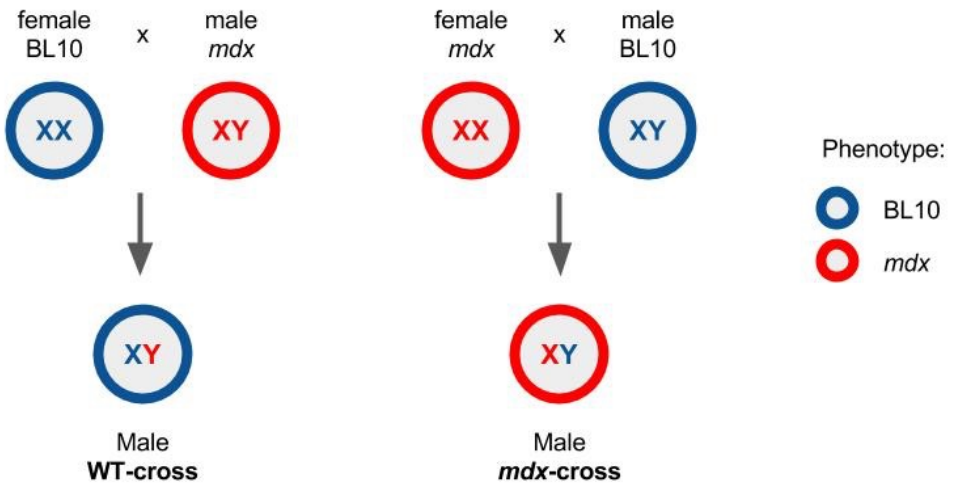


Figure 9: Schematic representation of the mouse breeding strategy

In all the experiments only male mice were used, all between 12 to 16 weeks of age. Male mice were preferred over females for two reasons. Firstly, most of the literature available on the *mdx* strain employs male mice, as Duchenne Muscular Dystrophy affects only males. Secondly, generating the WT-cross and *mdx*-cross control colonies is only possible for male mice, since females would retain the alleles of interest in Mendelian proportions, and would therefore need

genotyping by PCR or other method. The 12-16 week range was chosen as it is the earliest time point at which it is feasible to perform the AF induction protocol, as younger mice are too small to handle. In addition, *mdx* mice progressively develop cardiomyopathy (Stuckey et al., 2012), which is a known risk factor for AF (Schotten et al., 2011), so it was important to choose an early time point before the cardiac structure had undergone gross remodelling.

All the experiments involving animals were conducted in accordance with the UK Home Office Guidance on the Operation of the Animals (Scientific Procedures) Act 1986 and with institutional ethics approval.

2.2.2. Electrocardiographic records

Mice were placed in the anaesthesia induction chamber filled with 3% isoflurane in 21% oxygen. Once the animals lost consciousness they were transferred to the operating table where 1.5% isoflurane was administered through a face mask. Mice were kept on top of a heated mat set at 37 °C for the duration of the experiment. ECG parameters were measured two needle electrodes placed in the forelimbs. Traces were collected and analysed in a blind fashion using the Spike2 software. The data was analysed with one-way ANOVA and Tukey's post-hoc correction. All the ECG parameters are averaged across 3 consecutive beats.

2.2.3. Atrial Fibrillation inducibility protocol

Anaesthesia was induced as described in 2.2.2. Mice were kept under stable anaesthesia with 1.5% isoflurane through a vaporiser and placed on a mat heated at 37 °C. A trans-oesophageal octapolar catheter (1.7 Fr), which also worked as a recording probe, was inserted through the mouth of the mouse. The catheter was connected to a signal amplifier, which was in turn connected to Analog-to-Digital Converter before reaching a laptop equipped with the Spike2 software. Surface ECG traces were simultaneously recorded as described in 2.2.2. The Spike2 software controls the electrode, and visualises the signal recorder from the catheter and the surface electrode in real time. The placement of the catheter was adjusted until an electrical impulse triggered an atrial contraction (observed live on the ECG trace). An electrical stimulus was given every 60 ms (i.e., cycle-length = 60 ms), for a total of 40 stimuli. This was defined as a "burst". If the burst triggered an arrhythmic activity enough time was given for the arrhythmia to self-terminate before giving another burst. The re-appearance of the atrial P-wave was used as a marker of return to the normal sinus rhythm (Figure 10). If no arrhythmic activity arose, I waited a few seconds before giving the subsequent burst. The cycle-length of each burst was decreased by 2 ms each time, down to a cycle-length of 10 ms, for a total of 26 bursts given in each stimulation protocol. Three protocols were performed in sequence on each mouse, for a grand total of 78 bursts (scheme in Figure 11).

The traces were obtained and analysed in a blind fashion using the Spike2 software. AF was defined as an arrhythmic event with irregular, fast heart beats

lasting more than 2 s. AF probability was expressed as the ratio between the number of AF-inducing stimuli over the total number of stimuli administered to the mouse (i.e., number of stimuli/78).

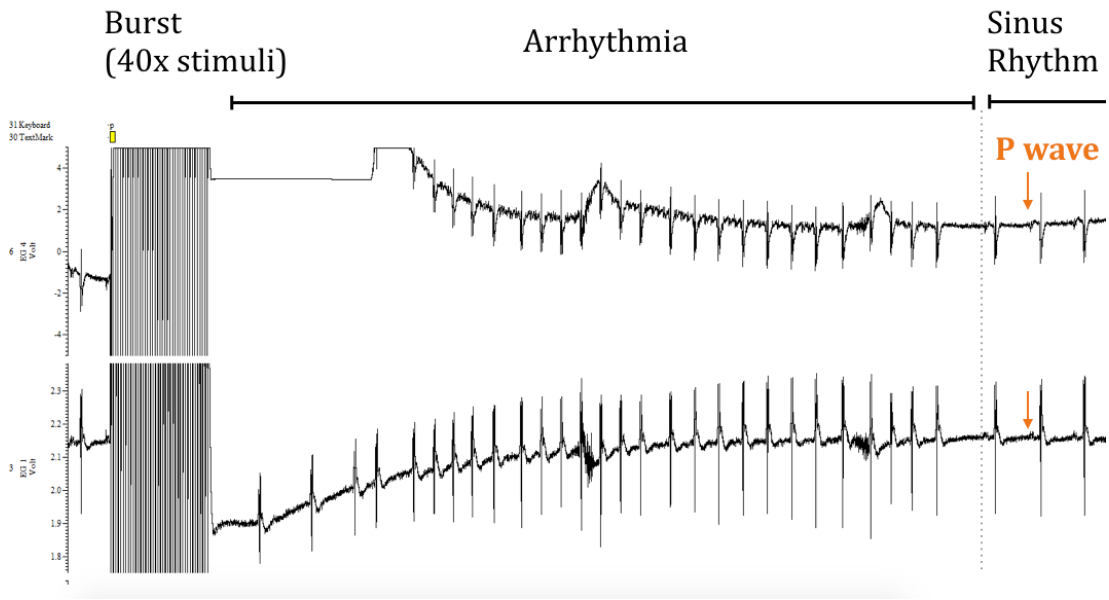


Figure 10: Representative trace of trans-oesophageal electrical stimulation

The ECG trace at the top is from the recording electrode of the octapolar catheter, while the trace at the bottom is from the surface ECG probes. A burst of 40 stimuli is given (block of black bars on the left side), which triggers an arrhythmic activity with fast atrial and ventricular contractions. The activity subsides on its own and normal rhythm returns, with stable heart beat duration, and a P-wave (orange arrow) followed by a QRS complex.

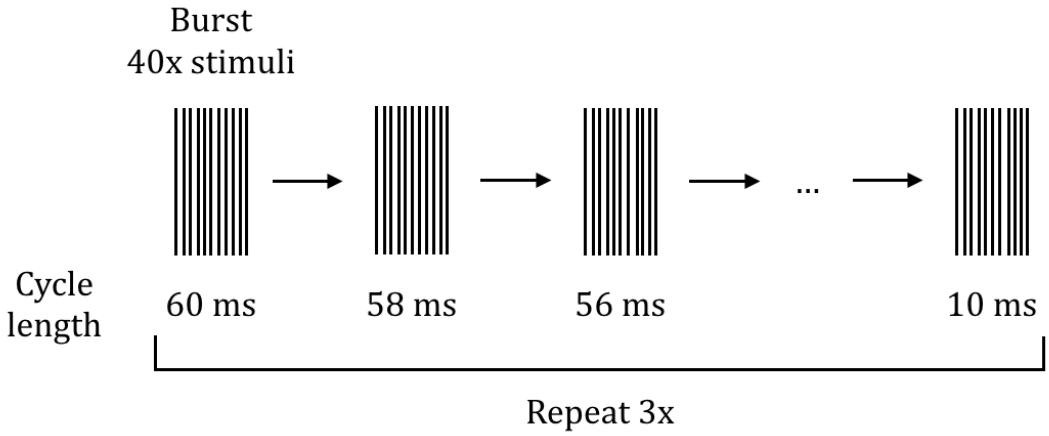


Figure 11: Scheme of the stimulation protocol

A burst of 40 electrical stimuli were given 26 times progressively decreasing the cycle length by 2 ms, from 60 ms to 10 ms. This protocol was repeated 3 times in a row on each mouse.

2.2.4. Echocardiography

Mouse echocardiography was performed with a high resolution 2D-echocardiography system (Vevo 2100, Visual Sonics, Canada) with an ultra-high frequency probe (18-38 MHz). The procedure was carried out with the help of Dr Klemen Ziberna, and the analysis with the help of Dr Parag Gajendragadkar; all investigators were blinded to the genotype of the animals, which were identified by a number.

Mice were first anaesthetised in an induction chamber with 4% isoflurane, and then moved to the imaging table (Figure 12). Mice were kept under stable anaesthesia with isoflurane through a vaporiser. The platform was heated to 38˚C and the mouse temperature monitored through a rectal probe and kept in the region of 37˚C. The eyes of the animals were protected with viscotear gel (Novartis).

The percentage of isoflurane was initially adjusted to slow the heart rate to ~350 beats/min, which allows the measurement of LV diastolic function. We carried out Doppler measurements in the four-chamber view, measuring passive (E) and active (A) transmitral flow velocities. We also measured the same parameters in tissue Doppler mode (e' and a').

The percentage of isoflurane was then reduced to bring the heart rate back to 450-500 beats/min to measure LV systolic function. We used the M-MODE in short axis view to measure wall thickness and LV diameter.

The heart rate and respiratory rates of the mice were monitored through the different phases of the experiment, as reported in Table 1. These parameters were not different between *mdx* and WT mice.

The following formulae were used to calculate the LV parameters:

- LV diastolic volume: $((7.0 / (2.4 + LVID;d)) * LVID;d^3)$
- LV systolic volume: $((7.0 / (2.4 + LVID;s)) * LVID;s^3)$
- LV mass: $1.053 * ((LVID;d + LVPW;d + IVS;d)^3 - LVID;d^3)$

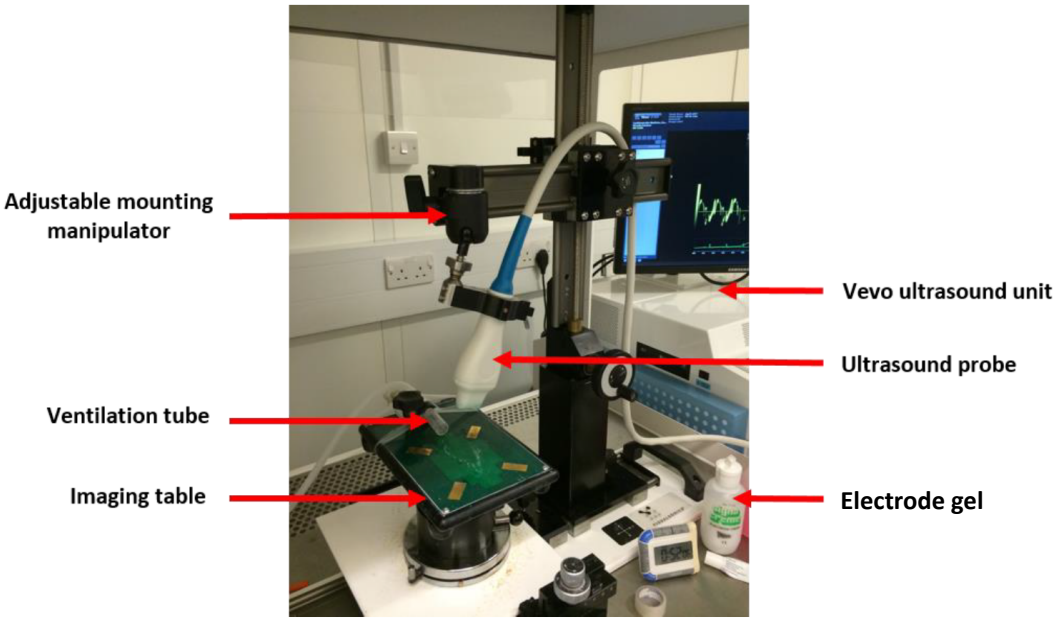


Figure 12: Echocardiography setup

Picture of the echocardiography machinery used in the experiment. The heated recording table is visible in the bottom left corner, together with the ultrasound probe and the Vevo unit in the back on the right.

Table 1: Heart rate and respiratory rate during echocardiography

	WT		***mdx***		**t-test p-value**
	mean	**SD**	**mean**	**SD**	
HR (beats per minute) during Doppler measurements in the 4-chamber view	335.1	32.1	338.3	30.0	0.7911
Respiratory rate (breaths per minute) as above	41.4	12.5	48.4	17.6	0.2413
HR (beats per minute) during tissue Doppler measurements in the 4-chamber view	336.4	29.5	336.0	34.0	0.9718
Respiratory rate (breaths per minute) as above	45.7	16.8	46.1	15.7	0.954
HR (beats per minute) during tissue Doppler measurements in the short axis view	468.3	26.0	464.6	85.9	0.8805
Respiratory rate (breaths per minute), as above	124.6	30.4	137.9	20.2	0.1849
HR during the M-Mode short axis view	476.9	18.6	487.4	19.4	0.1555
Respiratory rate (breath per minute) as above	118.5	19.0	137.3	25.6	0.0366

2.2.5. Protein extraction

12-16 weeks old male WT and *mdx* mice were culled by cervical dislocation, the chest and abdomen quickly opened and the abdominal aorta resected to completely terminate blood circulation. 5-10 ml of ice cold PBS were injected in each ventricle to eliminate residual blood. The hearts were excised, rinsed in PBS, patted dry, cut into RV, LV, RA and LA, and quickly frozen in dry ice upon collection. The tissue samples were then stored at -80 ˚C. Mouse atria and ventricles were homogenised in RIPA buffer, using the Tissue Lyser bead homogeniser (Qiagen) at 50 Hz for 2 minutes. The protein content was quantified with a BCA assay.

2.2.6. Western Blots

The protein extracts (40 ug per lane) were separated on a 4-12% Bis-Tris acrylamide gel in MOPS buffer (Life Technologies) at 180 V at room temperature. Transfer to a PVDA membrane was performed with the NuPAGE Transfer buffer (Life Technologies) with 50% Methanol, at 4 ˚C, at 280 mA for roughly 1.5 hours. After transfer, the membrane was blocked in 5% milk for 1 h at room temperature, then rinsed and cut into strips according to the molecular weight of the target proteins. The strips were incubated at 4 ˚C over night individually with antibodies against: dystrophin-1 (1:2000), nNOS (Santa Cruz A-11, 1:500), eNOS (BD 610296, 1:10000), iNOS (Abcam 49999, 1:1000), syntrophin (1:2000), and Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (1:25000). As their molecular weights are too close to be separated, the nNOS and iNOS antibodies were used on separate, replicate membranes, while eNOS was re-blotted on the same membranes after a stripping cycle. The positive control sample was a brain homogenate (0.5 ug per lane) for the membrane blotted with nNOS, and LPS-activated macrophages (1 ug per lane) for iNOS. Incubation with secondary HRP-conjugated antibodies was carried out at room temperature for 1 h. The membranes were then washed and developed with ECL Prime (GE). Developed membranes were scanned with a ChemiDoc machine (BioRad) and a densitometric analysis performed on the images with the ImageLab software.

2.2.7. Immunohistochemistry

12-16 weeks old male WT and *mdx* mice were culled by cervical dislocation, the chest and abdomen quickly opened and the abdominal aorta resected to completely terminate blood circulation. 5-10 ml of ice cold PBS were injected in each ventricle to eliminate residual blood. The hearts were excised, rinsed in PBS and patted dry. The samples were placed in a small cylindrical tin foil pouch, covered in OCT compound and snap-frozen. 0.7 mm-thick slices of tissue were cut using a cryostat, applied onto glass coverslips, and blocked with 20% Normal Goat Serum, 20% Foetal Bovine Serum in PBS. The tissue was then stained at 4 ˚C over night using an anti-Dystrophin-1 antibody (1:100), and mounted with ProLong Gold with DAPI nuclear stain (ThermoFisher).

2.2.8. NOS activity measurement

Frozen heart tissue samples were homogenised in Krebs-HEPES buffer pH 7.4 with 7.5 mM of the arginase inhibitor N^{ω}-hydroxy-nor-Arginine (nor-NOHA) (Calbiochem) and Complete protease inhibitor (Roche), using a Tissue Lyser (Qiagen) at 50 Hz for 2 minutes at 4 ˚C. Protein content was quantified with a BCA assay and the samples normalised in order to have 250 µg of protein per reaction. Substrate ([^{3}H]L-Arginine) and enzyme cofactors (FAD 1 mM, FMN 1mM, NADPH 1 mM, Tetrahydrobiopterin (BH_4) 1 mM) were added to the reaction mix. Each sample was split into two aliquots, one of which was incubated with 100 mM of the non-isoform specific NOS-inhibitor NG-nitro-L-arginine methyl ester (L-NAME). The reaction was left for 4 hours on ice, and then the samples were deproteinated with trichloroacetic acid and run on a dedicated High Pressure Liquid Chromatography (HPLC) column (de Bono et al., 2007). The addition of the radioactive arginine and the subsequent handling of the tubes were performed by Dr. Ricardo Carnicer.

A representative trace and its magnification can be seen in Figure 13. The black trace is the atrial sample without L-NAME, while the red trace is the same sample plus L-NAME. The three peaks that can be seen, from left to right, correspond to L-Citrulline, Ornithine, and L-Arginine. L-Arginine is the unconverted substrate, L-Citrulline is the product of the activity of NOS enzymes, while Ornithine is generated by active arginases that are not fully inhibited by nor-NOHA. The area of each peak is measured through the software, and then the area of the L-NAME samples is subtracted from the area

without L-NAME. The results are reported as area percentage of L-Citrulline over L-Arginine, per mg of protein used.

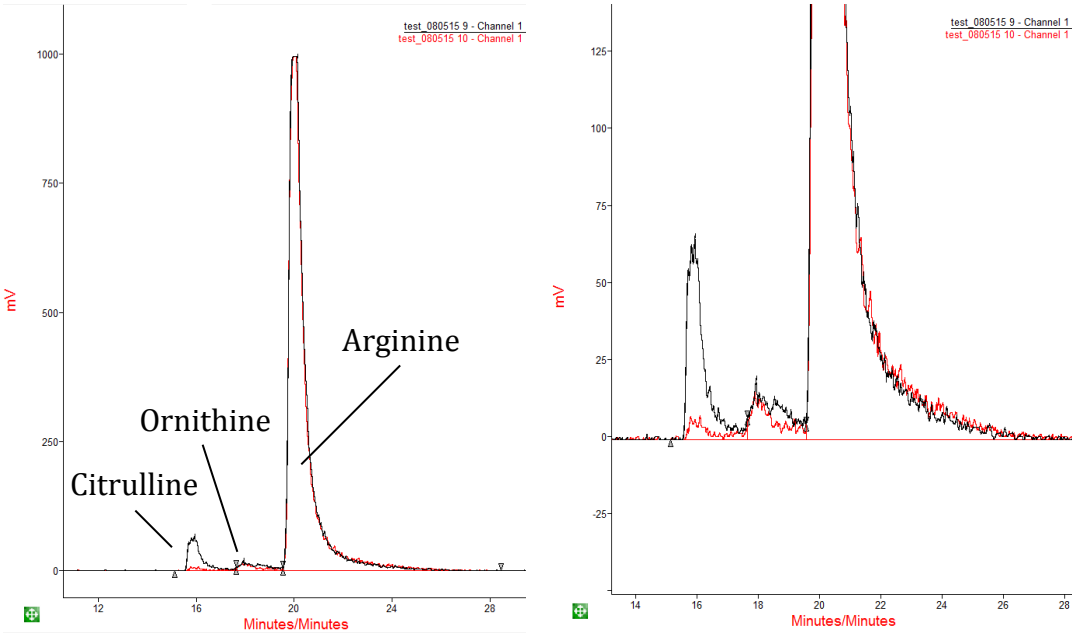


Figure 13: Representative HLPC trace of atrial samples

Left panel: full HPLC trace of a pool of 4 mouse atria (2 RA + 2 LA) from WT mice. Right panel: magnification of peaks from the trace in the left panel. In red is represented the trace obtained from samples treated with the NOS-inhibitor L-NAME, in black without L-NAME.

2.3. RESULTS

2.3.1. Mouse model

I recorded the body weight of the mice at time of culling to check for strain-specific variations. The mean body weight of BL10 mice was significantly lower from that of all the other strains (one-way ANOVA, $p < 0.0001$) (Table 2). However, there were no significant differences between the mean heart weight of the strains, suggesting that the difference in body weight was likely due to environmental factors rather than genetics. It seems reasonable to hypothesise that the difference in body weight was caused by transport-related stress, as BL10 mice were the only ones purchased from a supplier after birth, while the other strains were bred in-house. Consequently, BL10 mice were the only strain that had to endure transport and change of facility and conditions. Indeed, it was reported that C57BL/6 mice lose an average of 4% of their body weight over a 24-hour transportation (Lee et al., 2012), which may be compatible with the lower body weight observed in the BL10 mice.

Alternatively, the difference in body weight may be due to the different genetic background of the BL10 strain when compared to the *mdx*, WT-cross and *mdx*-cross strains.

Table 2: Average body and heart weight of mouse strains at 12-16 weeks

	C57BL/10	mdx	WT-cross	*mdx*-cross
Body weight (g)	28.82 (60)	**32.22****** (87)	**31.89****** (41)	**31.98****** (54)
Heart weight (mg)	139.7 (30)	141.9 (53)	141.5 (28)	142.5 (45)
Heart weight / Body weight 10^{-3}	4.802 (29)	**4.382****** (53)	**4.537****** (28)	**4.513****** (44)

The number of mice per group is indicated in brackets.

**** P< 0.0001 vs C57BL/10 by one-way ANOVA

2.3.2. Electrocardiographic characteristics of *mdx* mice

The surface electrocardiogram (ECG) in isoflurane-anaesthetised mice showed that heart rate is not significantly different between strains (Figure 14). The analysis of ECG parameters reveals that at 3 months of age *mdx* mice have a longer duration of the QRS complex when compared to BL10 (Figure 15 and Table 3). This difference was confirmed in the WT-cross and *mdx*-cross populations, indicating that it is a bona fide consequence of the lack of dystrophin, rather than a result of genetic drift between the *mdx* and BL10 colonies. The longer QRS complex registered in *mdx* mice is consistent with the literature and indicates a slower ventricular conduction in these mice, possibly due to a reduction in sodium channels, fibrosis, and other as yet unidentified causes (Bostick et al., 2008; Colussi et al., 2010; Koenig et al., 2011).

Compared with that of BL10 mice, the P wave duration was significantly shorter in both WT-cross and *mdx*-cross mice. Similarly, the PQ interval in *mdx*-cross mice was significantly shorter than in BL10 mice. These differences in P wave and PQ interval duration are probably reflecting genetic drift between the colonies. The remaining ECG parameters analysed were not significantly different between the groups.

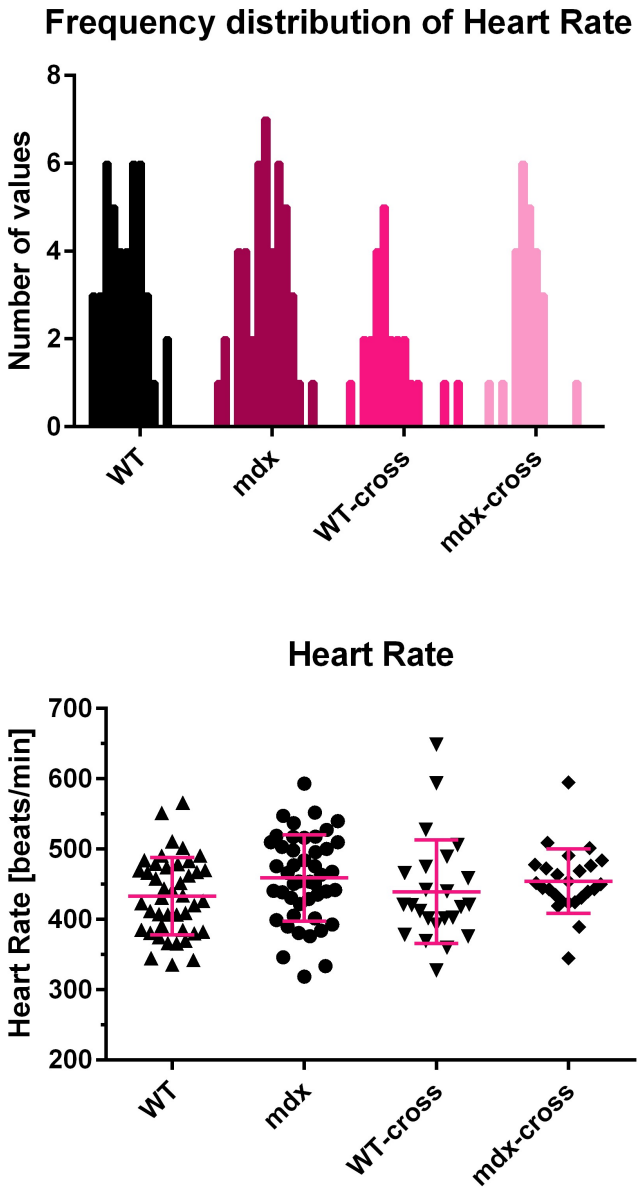
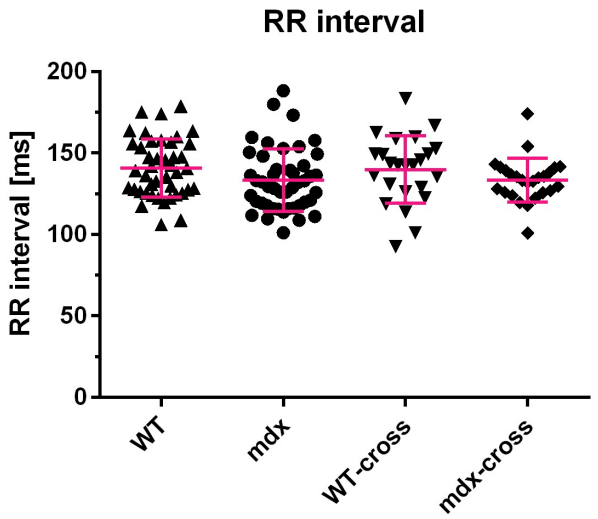
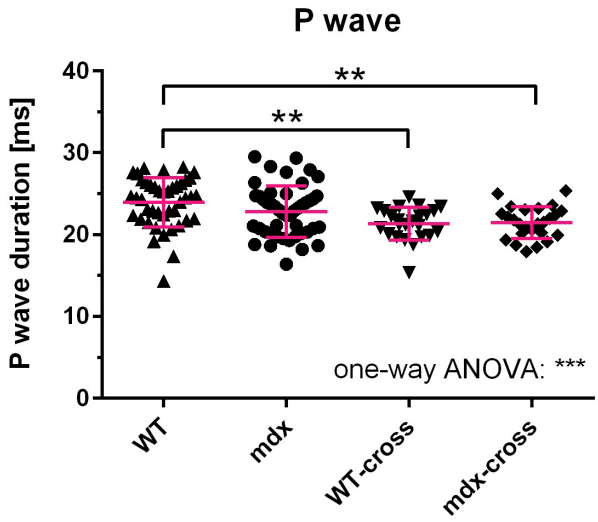


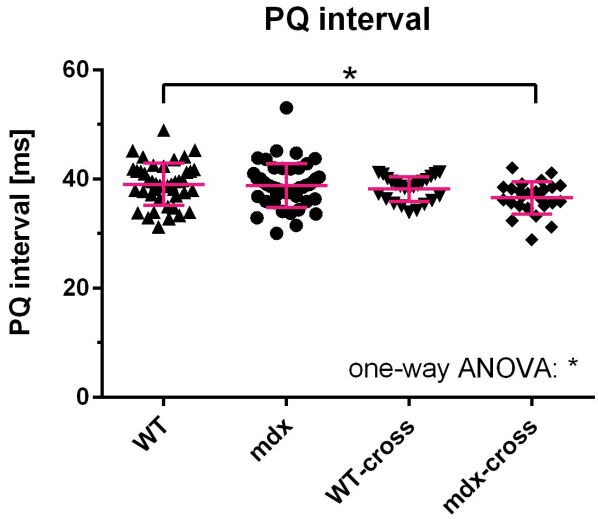
Figure 14: Heart rate of BL10 and *mdx* mice

Heart rate follows a normal distribution in every group. Heart rate is not significantly different among the groups (one-way ANOVA with Tukey's multiple comparison test).

n = 43 for the BL10 group, n = 46 for the *mdx* group, n = 24 for the WT-cross group, n = 25 for the *mdx*-cross group.


RR interval
RR interval [ms]
200
150
100
50
0
WT
mdx
WT-cross
mdx-cross


P wave
P wave duration [ms]
40
30
20
10
0
**
**
one-way ANOVA: ***
WT
mdx
WT-cross
mdx-cross


PQ interval
PQ interval [ms]
60
40
20
0
*
one-way ANOVA: *
WT
mdx
WT-cross
mdx-cross

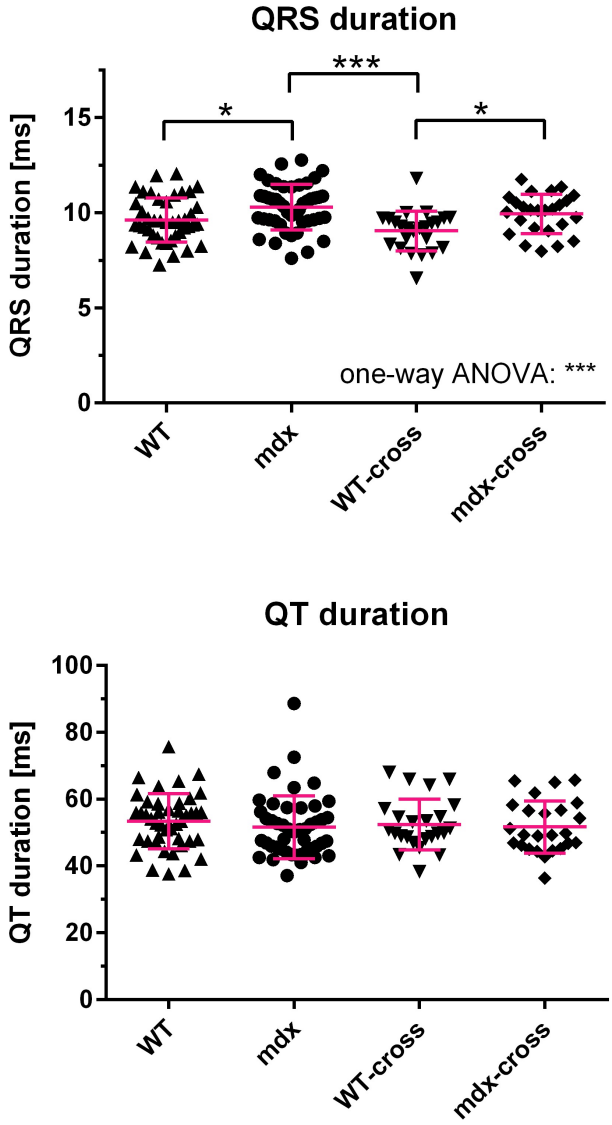


Figure 15: Electrocardiographic (ECG) parameters of BL10 and *mdx* mice

The P wave duration is significantly different between BL10 and both WT-cross and *mdx*-cross mice (one-way ANOVA with Tukey's multiple comparison test). The PQ interval is significantly different between WT and *mdx*-cross mice (one-way ANOVA with Tukey's multiple comparison test). QRS duration is significantly different between WT and *mdx* mice, between WT-cross and *mdx*-cross mice, and between *mdx* and WT-cross mice (one-way ANOVA with Tukey's

multiple comparison test). The remaining ECG parameters analysed are not significantly different between the groups. n = 43 for the BL10 group, n = 46 for the *mdx* group, n = 24 for the WT-cross group, n = 25 for the *mdx*-cross group.

Table 3: ECG parameters of the mouse strains

	WT	***mdx***	**WT-cross**	***mdx*-cross**
Heart rate (beats/min)	433.1 ± 54.9	458.8 ± 61.4	439.3 ± 73.7	454.3 ± 45.8
RR interval (ms)	140.7 ± 17.9	133.3 ± 19.2	139.8 ± 20.8	133.4 ± 13.5
P wave (ms)	23.96 ± 3.01	22.82 ± 3.14	**21.34 ± 2.01 ****	**21.47 ± 1.96 ****
PQ interval (ms)	39.01 ± 3.85	38.80 ± 4.03	38.14 ± 2.29	**36.53 ± 2.99 ***
QRS complex (ms)	9.62 ± 1.17	**10.29 ± 1.20 ***	**9.05 ± 1.04 ###**	**9.94 ± 1.04** ‡
QT interval (ms)	53.37 ± 8.21	51.58 ± 9.41	52.44 ± 7.62	51.69 ± 7.80

Mean and standard deviation of the ECG parameters measured in the different mouse strains. n = 43 for the BL10 group, n = 46 for the *mdx* group, n = 24 for the WT-cross group, n = 25 for the *mdx*-cross group.

One-way ANOVA results:

* significant difference with WT;

\# significant difference with *mdx*;

‡ significant difference to WT-cross.

2.3.3. *mdx* mice do not have increased susceptibility to AF

Since AF does not develop spontaneously in the mouse, I used an induction protocol (Verheule et al., 2004). Contrary to my initial hypothesis, mdx mice did not show a significant difference in AF inducibility measured as number of events over total stimulations (Mann-Whitney test), absolute number of events (Mann-Whitney test), duration of the fibrillation events (Mann-Whitney test), or number of responsive mice (Fisher's exact test) (Figure 16 and Figure 17).

The fact that the *mdx* mouse fails to recapitulate the differences in AF inducibility seen in $nNOS^{-/-}$ mice merited further investigation, as it may have been due to a number of causes. First, the NO released by the residual nNOS may be sufficient to prevent AF; secondly, NO availability may be maintained by means of compensation by other NOS isoforms; thirdly, the activity of the residual nNOS may be enhanced. It should be noted that these hypotheses are not mutually exclusive.

As no differences in AF inducibility were identified between the BL10 and WT-cross, or between the *mdx* and *mdx*-cross strains, in the subsequent experiments I only used BL10 and *mdx* mice. This reduced the total number of animals per experiment, and also reduced the number of groups, increasing statistical power.

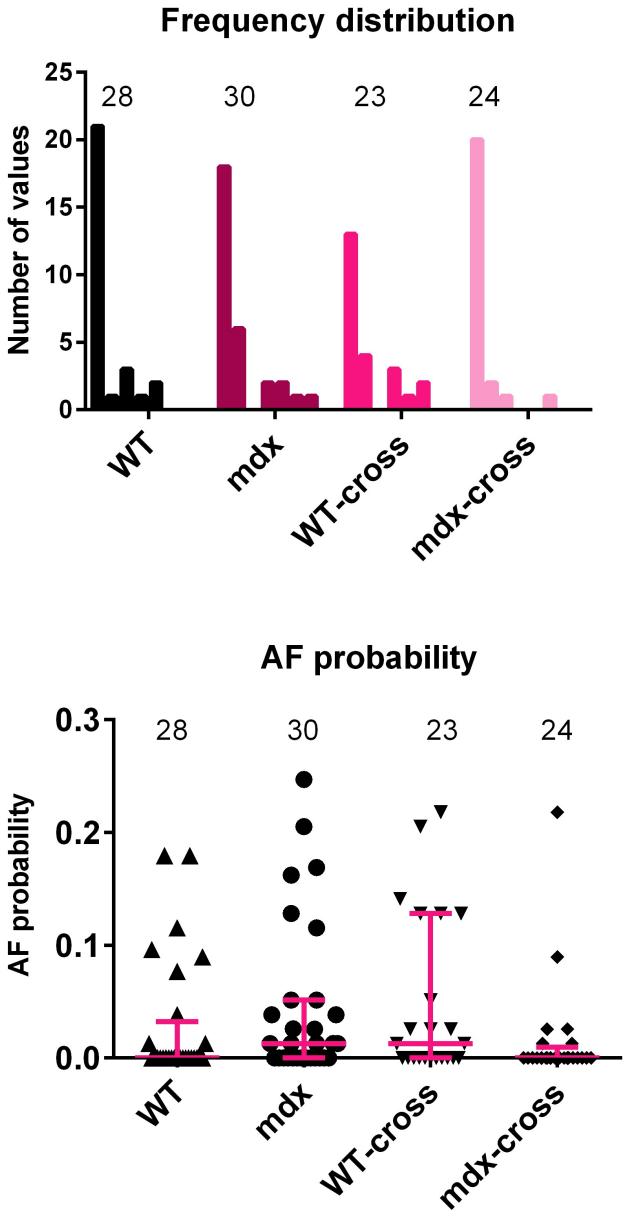
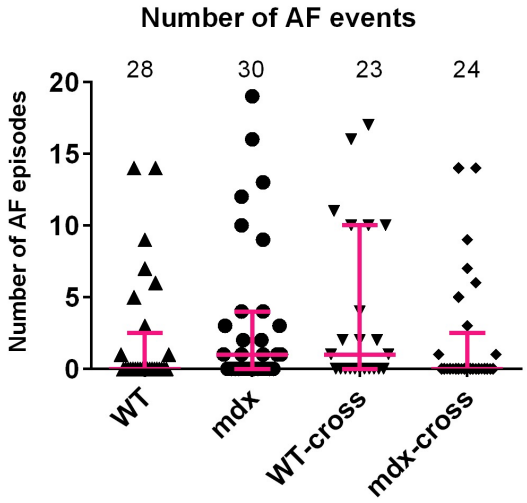
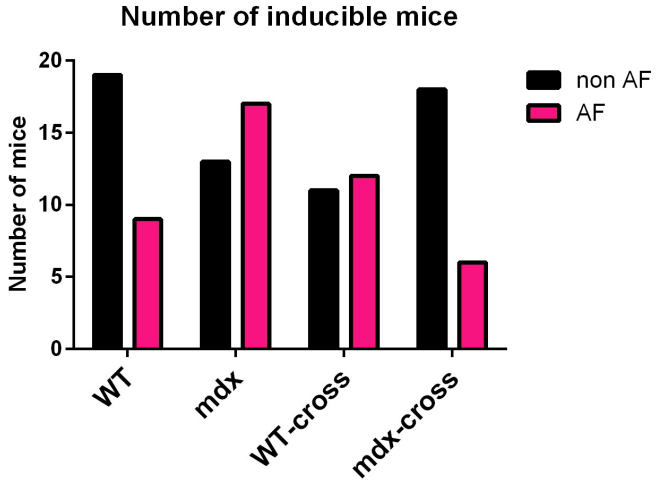


Figure 16: AF Inducibility

Top panel: frequency distribution of AF probability scores. AF probability was not normally distributed. Bottom panel: AF probability among the genotypes. The AF probability was not significantly different between the groups (Mann-Whitney test). The numbers at the top of the graph indicate the number of mice per group.


Number of AF events
28
30
23
24
Number of AF episodes
20
15
10
5
0
WT
mdx
WT-cross
mdx-cross


Number of inducible mice
Number of mice
20
15
10
5
0
non AF
AF
WT
mdx
WT-cross
mdx-cross

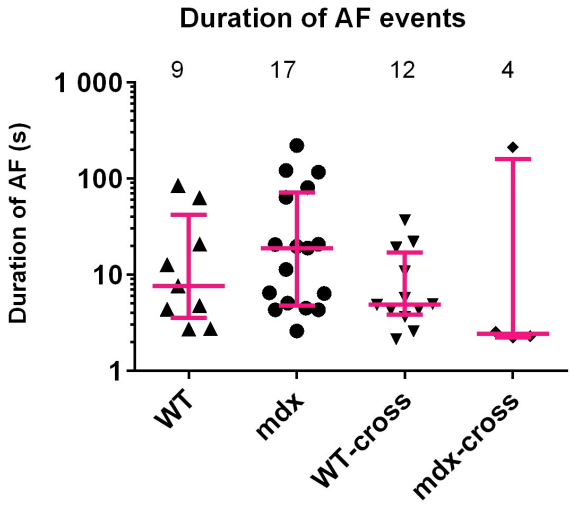

Duration of AF events
9
17
12
4
Duration of AF (s)
1 000
100
10
1
WT
mdx
WT-cross
mdx-cross

Figure 17: AF - Number of events, number of inducible mice and duration

The absolute number of AF events (Mann-Whitney test), the proportion of fibrillating mice (Fisher's exact test), and the duration of AF events (Mann-Whitney test) were not significantly different between the groups.

The numbers at the top of the graph indicate the number of mice in each group.

2.3.4. Echocardiographic characterisation of the *mdx* mouse

We proceeded to further characterise the cardiac size and function of *mdx* mice by using echocardiography. The results are summarised in Table 4. The isovolumic contraction and relaxation times (IVCT and IVRT, respectively) and the ejection time (ET) were not significantly different between *mdx* and WT mice. This suggests that the function of the myocardial muscle is not grossly compromised in 3-month old *mdx* mice.

Mdx mice have a significantly larger end-diastolic and end-systolic LV diameter and volume than WT mice (t-test $p < 0.005$). The thickness of the anterior and posterior ventricular walls is unchanged in diastole, while the posterior wall becomes thinner when stretched during systole (t-test $p < 0.05$). LV volume and mass were significantly higher in *mdx* mice (t-test $p < 0.0001$), which is to be expected as they are derived from the internal diameter. However stroke volume and cardiac output were not significantly impaired in *mdx* mice. Ejection fraction and fractional shortening were lower in *mdx* mice (t-test $p < 0.0001$), which reflects the higher end-diastolic volume of these mice.

Mdx mice have higher passive and active filling velocities (E and A waves, respectively) by pulse-wave Doppler echocardiography. The E/A ratio showed a trend towards being increased in *mdx* mice, which appeared to be driven by a greater increase in passive filling velocity (E wave). Similarly, E'/A' ratio was also higher in *mdx* mice (t-test $p < 0.005$). The E/E' ratio was slightly increased in *mdx* although the difference is not statistically significant.

Table 4: Echocardiographic measurements

	WT		*mdx*			
	mean	**SD**	**mean**	**SD**	**p-value**	**p-value summary**
IVCT (s)	14.66	3.01	14.71	2.26	0.9635	
IVRT (s)	18.31	4.52	18.01	2.85	0.8326	
ET (s)	68.46	6.93	68.75	7.30	0.9165	
E (mm/s)	498.61	71.49	641.30	81.37	< 0.0001	****
A (mm/s)	372.82	58.86	422.44	58.78	0.0345	*
E' (mm/s)	-16.89	4.03	-19.21	3.47	0.1139	
A' (mm/s)	-17.00	3.22	-14.76	3.54	0.0917	
E/A	1.35	0.22	1.54	0.29	0.0578	
E'/A'	0.99	0.19	1.35	0.32	0.0014	**
E/E'	30.85	7.00	34.14	5.98	0.1931	
LVID;d (mm)	3.18	0.12	3.38	0.21	0.0046	**
LVID;s (mm)	1.72	0.20	2.21	0.33	< 0.0001	****
LVAW;d (mm)	0.94	0.11	0.96	0.09	0.6437	
LVAW;s (mm)	1.34	0.13	1.29	0.14	0.3577	
LVPW;d (mm)	0.95	0.12	0.97	0.09	0.6418	
LVPW;s (mm)	1.51	0.15	1.39	0.14	0.0419	*
EF (%)	78.27	6.08	64.59	9.08	< 0.0001	****
FS (%)	46.08	6.06	34.81	6.80	< 0.0001	****
LV Mass (mg)	103.73	14.96	117.04	15.92	0.031	*
LV Volume;d (ul)	40.49	3.87	47.17	7.42	0.0061	**
LV Volume;s (ul)	8.78	2.53	17.03	6.74	0.0002	***
SV (ul)	31.71	4.13	30.14	4.05	0.1584	
CA (ml/min)	15.16	2.26	14.70	2.12	0.2922	

List of abbreviations:

IVCT: Isovolumic Contraction Time;

ICRT: Isovolumic Relaxation Time;

ET: Ejection Time;

A: active ventricular filling velocity;

E: passive ventricular filling velocity;

A': active ventricular filling velocity in Tissue Doppler;

E': passive ventricular filling velocity in Tissue Doppler;

d: diastolic;

s: systolic;

LVID: Left Ventricular Internal Diameter;

LVAW: Left Ventricular Anterior Wall;

LVPW: Left Ventricular Posterior Wall;

EF: Ejection Fraction;

FS: Fractional Shortening;

SV: Stroke Volume;

CA: Cardiac Output.

2.3.5. *mdx* hearts have reduced nNOS protein content

In order to shed light on the lack of difference in AF inducibility, I first wanted to measure the protein level of nNOS and of other NOS isoforms, namely endothelial NOS (eNOS) and inducible NOS (iNOS), in the atrial and ventricular myocardium of BL10 and *mdx* mice.

As expected, Western blots showed a complete absence of dystrophin in *mdx* mice (Figure 18), and this was further confirmed by immunostaining (Figure 20). Syntrophin, a protein belonging to the dystrophin complex, was significantly reduced in the right (RA) and left (LA) atrium of *mdx* mice versus BL10 mice (two-way ANOVA: $p < 0.0001$ for genotype, n.s. for sidedness, n.s. for the interaction). nNOS protein content was markedly reduced in the atria of *mdx* mice versus BL10, and was significantly higher in the LA when compared to the RA regardless of the genotype (two-way ANOVA: $p < 0.01$ for genotype, $p < 0.001$ for sidedness, $p < 0.01$ for the interaction). Additionally, individual mice with high susceptibility to AF (↑AF in Figure 18) did not have different nNOS expression from mice with low or no susceptibility to AF (↓AF in Figure 18).

eNOS did not differ between mouse strains or atrial side (two-way ANOVA: n.s. for genotype, n.s. for sidedness, n.s. for the interaction), and iNOS was undetectable. The pattern for all of the proteins analysed was replicated exactly in the ventricular tissue of these mice (Figure 18 and Figure 19).

In light of these results, it seems unlikely that iNOS expression may account for the lack of increase in AF inducibility in *mdx* mice. A possible explanation for

this phenomenon would be a hyper-active state of eNOS and/or nNOS in *mdx* mice, which would increase the NO availability in the tissue.

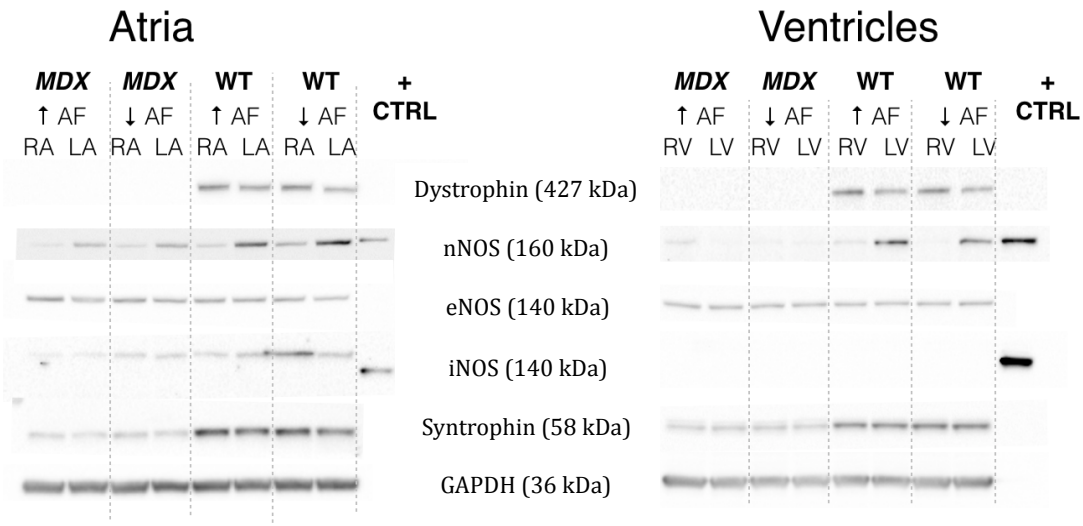


Figure 18: Expression of proteins of the dystrophin-complex in BL10 and *mdx* mice

Representative cropped images of Western Blots of left and right atrial and ventricular samples from 12-16 weeks old male *mdx* and BL10 mice (n = 8 per group). The positive control is mouse brain for the nNOS staining, and LPS-activated macrophages for iNOS staining.

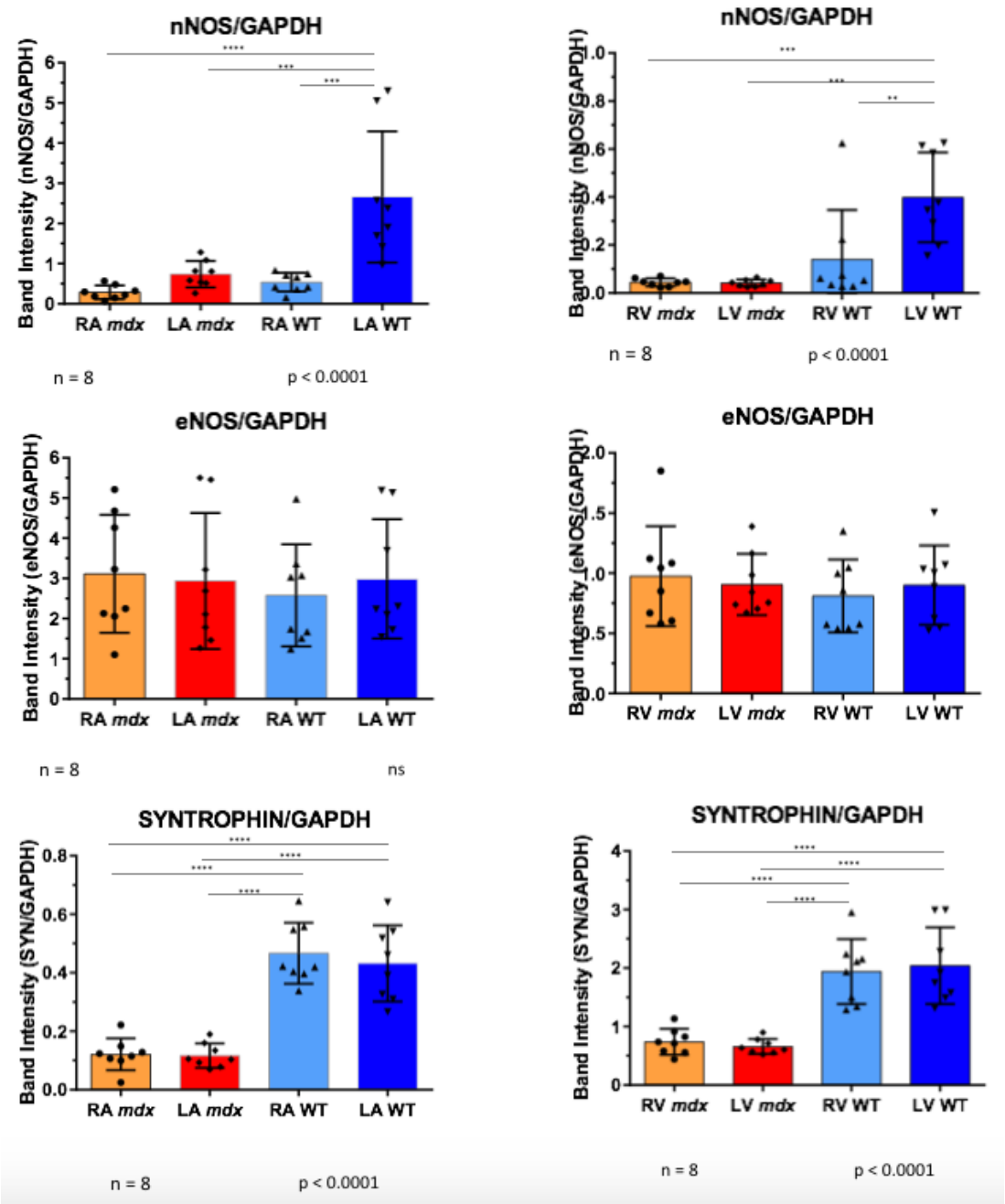


Figure 19: Quantification of protein expression from Western Blots

Histograms deriving from the densitometric analysis of the Western Blot membranes presented in Figure 18.

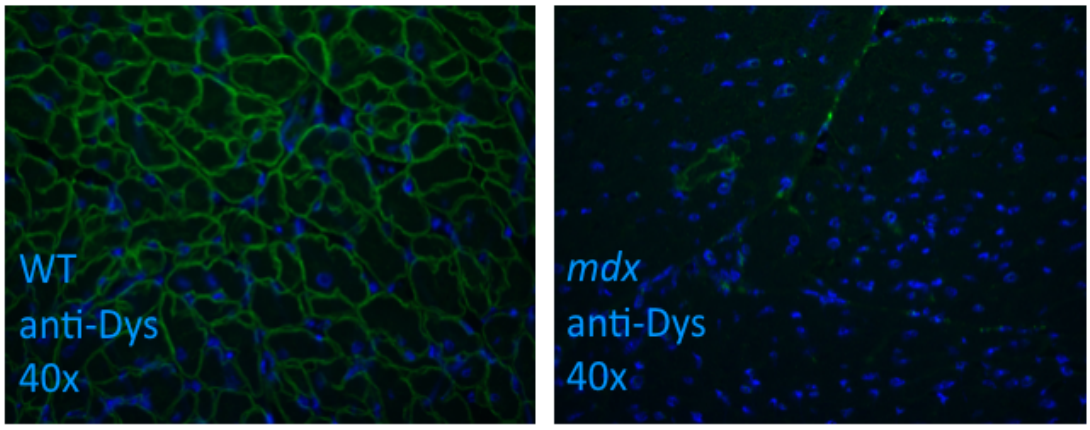


Figure 20: Immunohistochemistry of Dystrophin in cross-section of BL10 and *mdx* ventricular myocardium

The heart sections were stained with DAPI (blue) and an anti-Dystrophin primary antibody (green). While the WT heart (left) shows a clear dystrophin staining at the sarcolemmal membrane, this is absent in the *mdx* heart (right).

2.3.6. Increased NOS activity compensates for protein reduction in the atria of *mdx* mice

I then tested whether increased eNOS or the residual nNOS activity may compensate for the reduction in nNOS protein. To do so, I measured NOS activity using a radioactive assay combined with HPLC apparatus.

Interestingly, the atria of *mdx* mice showed a significantly higher Arginine to Citrulline conversion, indicating higher atrial NOS activity in these mice. No significant difference in NOS activity was measured between RA and LA regardless of the genotype (two-way ANOVA: $p < 0.05$ for genotype, n.s. for sidedness, n.s. for interaction) (Figure 21).

	BL10 RA	BL10 LA	*mdx* RA	*mdx* LA
Pool 1	1.48	1.52	2.21	2.12
Pool 2	2.15	0.9	1.9	1.96
Pool 3	1.79	1.51	1.65	1.76
Pool 4	1.31	1.38	1.58	1.87
Pool 5	1.55	1.49	2.59	1.17
Average	**1.66**	**1.36**	**1.99**	**1.78**

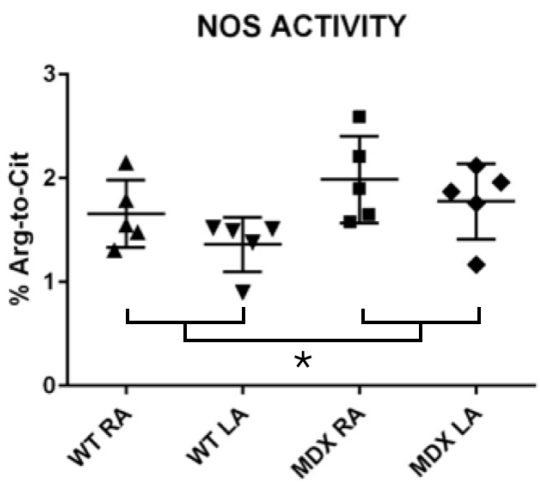


Figure 21: Nitric Oxide Synthase (NOS) Activity in the atria of WT and *mdx* mice

The table of the right contains the raw values of the Arg-to-Cit conversion rates represented in the plot on the right. 4 atria were pooled together prior to homogenisation in order to obtain measurable traces. 5 pools (a total of 20 mice) per group were used in the experiment. Statistics: two-way ANOVA.

2.4. DISCUSSION

Given that:

i) Patients with AF have reduced atrial levels of nNOS and dystrophin (Reilly et al., 2016)

ii) $nNOS^{-/-}$ mice are more susceptible than WT controls to AF induced by trans-oesophageal electrical burst stimulation (Reilly et al., 2016)

I set out to investigate whether the lack of dystrophin would increase AF inducibility in a k.o. mouse model (i.e., *mdx* mouse). In order to do so I first characterised the cardiac profile of *mdx* mice using surface ECG. Then, I induced AF through a trans-oesophageal probe. Finally, I measured the expression level and activity of NOS isoforms.

The work carried out in this thesis suggests that *mdx* mice are no more susceptible than their WT counterparts to induced AF through trans-oesophageal stimulation. This is despite the reduction of nNOS protein in the *mdx* myocardium. This seemingly contradictory evidence can be explained by an enhanced functional activity of the residual NOS enzymes, as measured by an Arg-to-Cit conversion assay.

2.4.1. Cardiac characterisation of *mdx* mice

2.4.1.1. Electrocardiography

Surface ECG of *mdx* mice revealed *mdx* mice have a longer QRS complex when compared to C57BL/10 mice (Table 3, page 73). This finding has been previously reported in literature, although the various works are not always in agreement. Two studies carried out in young (4-12 week-old), conscious mice reported a shorter or unvaried QRS complex (Chu et al., 2002; Fauconnier et al., 2010). A further study on 4-8 week-old mice anaesthetised with ketamine also reported no difference in QRS duration between *mdx* and WT controls (Sadeghi et al., 2002). Three different groups reported a longer QRS complex in 4-14 month-old mice under isoflurane anaesthesia (Koenig et al., 2011; Lai et al., 2013b; Yue et al., 2016). It is conceivable to think that a slower QRS could be characteristic of later stages of the dystrophic cardiomyopathy, and therefore not detectable in young conscious mice.

2.4.1.2. Echocardiography

The echocardiographic study showed that the key parameters IVCT, IVRT, ET, SV and CO were not significantly different between *mdx* and WT mice. This suggests that the systolic and diastolic function are not grossly compromised in 3-month old *mdx* mice.

However, a number of systolic parameters are significantly different from C57BL/10 mice, namely:

- Thinner LV posterior wall, systolic
- Increased LV diameter, both systolic and diastolic
- Increased LV volume, both systolic and diastolic, and LV mass (calculated from LV diameter values)
- Reduced EF and FS (calculated from LV diameter values)

These data indicate that *mdx* mice have bigger LVs. This could be an initial sign of systolic dysfunction, where contractility is already impaired and stroke volume shows a trend towards impairment. It is worth nothing that all the parameters calculated from the ventricular diameter can only be compared between different mouse strains under the assumption that the mice have the same heart shape. Considering that:

(i) the WT and *mdx* mouse colonies were kept separately and not properly bred together;

(ii) the body weight of *mdx* mice is higher than WT, and the heart weight shows a slight, although not significant, increase;

one could argue that the increased LVID, and the difference in calculated parameters, may simply be due to the fact that *mdx* mice are bigger and have slightly bigger hearts. This would also reconcile the differences found in other studies, which report no alterations in such parameters in 3-month old *mdx* mice (Stuckey et al., 2012).

Some diastolic function parameters were also significantly altered in *mdx* mice, namely:

- Faster passive (E) and active atrial (A) filling velocity of the LV through the mitral valve
- Increased E'/A' ratio

The increase in LV E/A and E'/A' in *mdx* mice appears to be primarily driven by an increase in E and E'. This suggest that *mdx* mice have a faster passive filling of the LV, which may be influenced by higher preload, lower end-systolic LV pressure or reduced myocardial stiffness. Without further investigation it is not possible to determine which (if any) of these factors may be at play. As mentioned above, this could also be an artefact of the larger heart size of *mdx* mice, derived from the suboptimal colony management.

In order to confidently interpret the data it would be helpful to repeat the experiment in cross-bred *mdx* and WT, eliminating the bias derived from a possible genetic drift between the isolated mouse colonies. If this experiment were to be repeated, I would also recommend the measurement of LA size, which has been shown to be a valuable indicator of diastolic dysfunction in mice (Schnelle et. al, 2018).

The literature regarding echocardiography results in *mdx* mice is somewhat inconsistent, with most investigators reporting no difference in LV diameter and

ejection fraction between *mdx* and WT mice at 2-4 months (Li et al., 2017; Stuckey et al., 2012). Wagner and colleagues report increased wall thickness and LV internal diameter in 10-14 week old *mdx* mice (Wagner et al., 2012), in accordance with my results. Parvatiyar and colleagues also report increased LV diameter, reduced factional shortening and ejection fraction, increased amplitude of E and A waves, albeit in older (10-11 months) mice (Parvatiyar et al., 2015).

2.4.2. The development of AF in *mdx* mice

My initial hypothesis was that mice lacking dystrophin (i.e., *mdx*) will have lower or absent levels of nNOS and reduced NOS activity and would therefore display increased susceptibility to AF. However, *mdx* mice are no more susceptible to AF than control BL10 mice (Figure 16). This may be caused by a number of factors: (a) nNOS is not decreased in the atria of *mdx* mice; (b) nNOS is decreased but residual levels of expression are enough to prevent the AF phenotype; (c) other NOS isoform compensate for the reduction in nNOS-derived NO; (d) the residual nNOS works at a higher rate. I sequentially investigated these hypotheses.

Western blots showed that nNOS is significantly decreased in the atria and ventricles of *mdx* mice, in keeping with the literature in the skeletal muscle (Stamler and Meissner, 2001). eNOS levels are not different from BL10 mice, and iNOS is not detectable in the heart of *mdx* mice at 3 months, ruling out the hypothesis that increased eNOS and/or iNOS may compensate at the

transcriptional level for the reduced nNOS. Other studies report the presence of iNOS in *mdx* heart at older ages, when inflammatory cells infiltrate the damaged muscle fibres. In particular, Fauconnier et. al show that iNOS is undetectable by Western blot at 35 days, while is present at 180 days (Fauconnier et al., 2010).

I then assessed NOS activity in the absence and presence of the NOS inhibitor L-NAME. Interestingly, the atria of *mdx* mice show higher levels of L-NAME inhibitable L-Arginine to Citrulline conversion, consistent with a higher NO synthesis compared with BL10. The increase in NOS activity was even more striking considering that nNOS protein content was significantly reduced in *mdx* mice and eNOS was unchanged. Thus, a compensatory increase in atrial eNOS and/or nNOS activity in *mdx* mice may explain why they do not have a higher propensity for AF.

These findings are in contrast with a previous study reporting unchanged eNOS activity and decreased nNOS activity in the myocardium of *mdx* mice (Bia et al., 1999). A number of factors may explain this difference. First, Bia and colleagues used allegedly specific nNOS and eNOS inhibitors that are not reported in the paper. After cross-referencing their article, it would appear they use inhibitory peptides from Calbiochem, which have since been discontinued and cannot be checked further. Another difference with my study is that they added calcium and calmodulin to the reaction mix. Last, they use whole heart homogenates rather than atrial preparations, and therefore atrial-specific changes in NOS activity would be greatly overshadowed by the ventricular enzymes.

It remains to be investigated what mechanism causes increased NOS activity in the *mdx*. It is tempting to speculate that nNOS alone may be responsible for this result, but the contribution of eNOS cannot currently be ruled out. Both proteins can be the target of post-translational modifications that affect their functional state, such as the activatory Akt-mediated phosphorylation and the inhibitory S-gluthationylation (Carnicer et al., 2013). Additionally, NOS activity may be enhanced by increased availability of a rate-limiting cofactor or substrate. This seems less likely as the cofactors FAD, FMN, NADPH and BH_4 were exogenously supplemented to the reaction mix. However, calcium and calmodulin were not. Future experiments should focus on screening for post-translational modification of nNOS and eNOS, and understanding whether Ca^{2+}-dependent calmodulin regulation plays a role in the process.

3. TRANSCRIPTOMIC ANALYSIS OF POST-OPERATIVE ATRIAL FIBRILLATION

3.1. INTRODUCTION

AF has a very broad spectrum of molecular alterations that can be recorded in clinical samples. However, since it is often asymptomatic in the early stages, and since obtaining human heart samples is not possible unless patients undergo major cardiac surgery, there is not a clear picture of the progressive changes that affect the myocardium throughout the course of the disease.

To shed light on the early alterations that subtend AF, we can draw information from post-operative AF (poAF), which is a form of AF that develops in a controlled setting. PoAF occurs in roughly a third of patients in the first week following cardiac surgery, and correlates with prolonged hospital stay, increased co-morbidities and mortality (Aranki et al., 1996; Mathew et al., 2004; Whitlock et al., 2014). The development of poAF is associated with increased risk of developing persistent AF later (Açil et al., 2007; Lee et al., 2014), suggesting that early myocardial alterations and/or risk factors may be common to both pathologies.

Given this, I set up a study aimed at identifying molecular changes associated with the development of poAF in a cohort of human atrial samples. The samples were collected at the time of surgery, in patients undergoing CABG that had no

history of AF. The patients were then monitored with a Holter during their post-operative recovery, and divided into two groups depending on whether or not they developed poAF or not (poAF vs. SR group, respectively). By analysing the differential expression of microRNAs and genes between the two groups, I sought to identify the molecular alterations that may predispose the atria to the development of poAF.

3.2. METHODS

3.2.1. Patients

To obtain a transcriptional profile, I used right atrial tissues from 48 patients enrolled in the Statin Therapy in Cardiac Surgery (STICS) trial (ClinicalTrials.gov Identifier: NCT01573143). The main aim of the trial, completed in September 2014, was to evaluate whether perioperative treatment with statins prevents poAF and reduces perioperative irreversible myocardial damage in patients undergoing elective cardiac surgery. 1922 patients in SR with no history of AF were randomised to receive either placebo or Rosuvastatin 20 mg once daily for up to 8 days before undergoing cardiac surgery, and for 5 days after surgery. The development of poAF was detected through continuous ECG monitoring from immediately after surgery until the end of post-operative day 5 (Zheng et al., 2016). Cardiac tissue from the right atrial appendage was collected during surgery.

24 patients who developed poAF and 24 matched (by sex, age and co-morbidities) patients who remained in SR after surgery were identified by the study statistician from the placebo-treated group. Each sample was profiled for gene expression using a microarray, and for microRNA expression using small RNA-Sequencing.

3.2.2. RNA extraction

Samples were homogenised using a polytron hand homogeniser, and RNA was extracted using the mirVana kit (ThermoFisher). DNase treatment was performed on column using the gDNA wipeout buffer from the QuantiTect RT kit (Qiagen). Quantification of RNA yield was done with Nanodrop, and quality assessed on agarose gel. A further quality assessment was performed with Bioanalyzer at the WTCHG before proceeding further. RNA samples were stored at -80°C before sending them for microarray and RNA-Seq.

3.2.3. mRNA profiling with microarray

24 paired samples were run on an Illumina HT-12 v4 BeadChip at the Wellcome Trust Centre for Human Genetics (WTCHG) facility (University of Oxford). Resulting probe intensities were normalised with Variance Stabilizing Normalisation (VSN) and processed for differential gene expression with the `limma` package for R. Sample size calculation was performed with the `ssize` package.

A second batch of 24 samples was submitted for microarray analysis and run on Affymetrix Human ST 2.1 arrays at the WTCHG facility. The change in microarray platform was necessary as Illumina chips were no longer produced by the company. Resulting probe intensities were normalised with the Robust Multi Array (RMA) algorithm and processed for differential gene expression with the `limma` package for R.

A combined analysis of the two batches (total n = 48) was carried out by adjusting the linear model for batch effects.

All the R scripts used to pre-process and analyse microarray data are reported in the Appendix.

3.2.4. miRNA profiling with small RNA-Sequencing

48 paired samples were sequenced with the Illumina MiSeq platform at the WTCHG facility (Oxford) in two batches. Reads were mapped to the GRCh37 reference genome, normalised with the `vsd` function (`DESeq` package) and processed for differential gene expression with the `EdgeR` package for R. The linear model was adjusted for batch effect.

Sample size calculation was performed with the **`RNAseqPS`** software through its web interface (https://cqs.mc.vanderbilt.edu/shiny/RNAseqPS/).

All the R scripts used to process and analyse RNA-Seq data are reported in the Appendix.

3.3. RESULTS

3.3.1. RNA profiling with microarray

3.3.1.1. First batch of samples

The distribution of probe intensities for the microarray show the same outline for all the samples, suggesting an overall comparably good quality (Figure 22).

An initial dimensionality reduction by means of Principal Component Analysis (PCA) was not able to cluster the samples into the two experimental groups (i.e., poAF and poSR), indicated as A and B since the study has not been unblinded yet (Figure 22).

Results from differential gene expression on mRNA highlighted 917 genes that are differentially expressed among the two groups. However, after correcting for multiple testing, no differentially expressed gene at an adjusted p-value (false discovery rate) below the accepted threshold of 0.05 are identified. The 20 genes with the lowest adjusted p-values are reported in Table 5; it should be noted that the lowest adjusted p-value overall is only ~0.881, which is very far from the established cut-off.

The full transcript set, and the top 30 genes were also visualised in a volcano plot and heatmap, respectively (Figure 23).

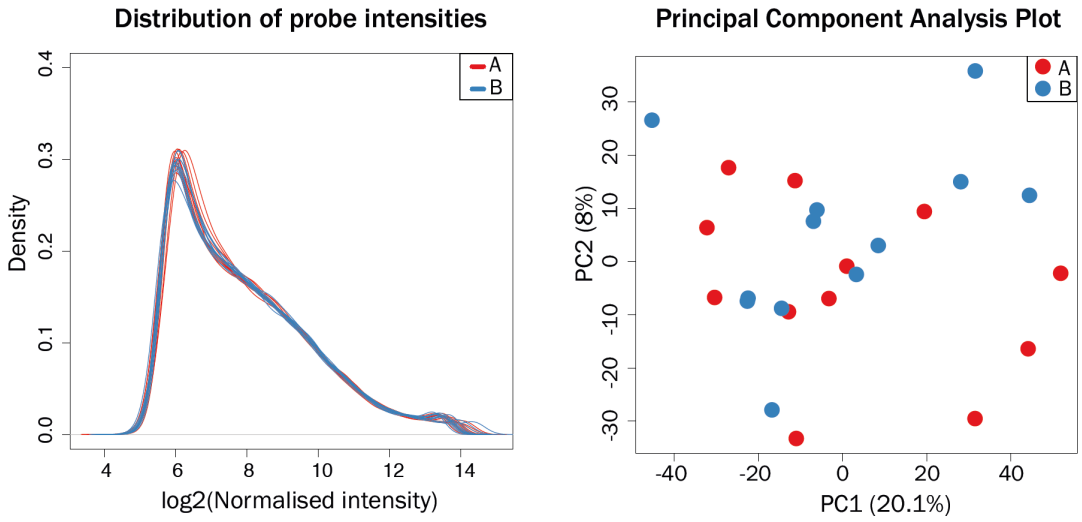


Figure 22: Quality control plots for Illumina microarray

Left panel: density plot of probe intensities. The x-axis shows the fluorescence intensity for all the probes on the microarray, normalised with the RMA function, and log-transformed. The y-axis shows the kernel density function of the intensities, computed with the `density()` function in R.

Right panel: Principal Component Analysis distribution for the microarray. Data points are transformed in such a way that the first Principal Component (PC1) expresses the largest possible variance for the data set, the second PC (PC2) expresses the second largest variance, and so on. PC1 and PC2 are reported on the x- and y-axis, respectively, and the associated explained variance is reported in brackets on the axis labels.

Samples belonging to experimental group A are reported in red, and samples belonging to group B in blue.

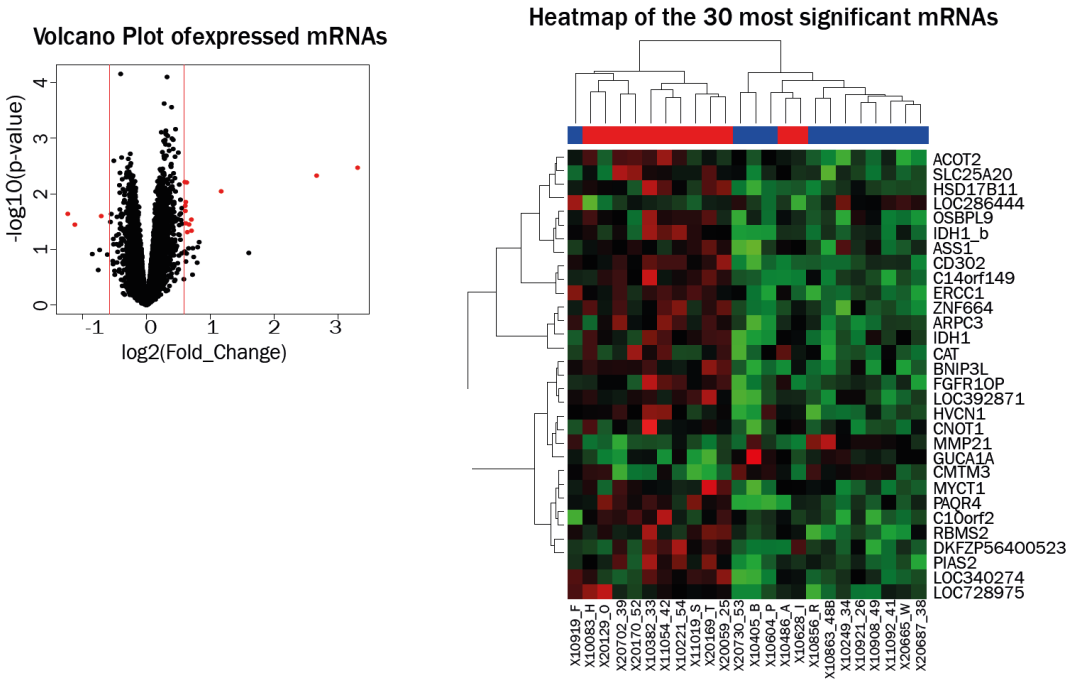


Figure 23: Differential gene expression plots for mRNAs

Left panel: Volcano plots of expressed mRNAs. Log-transformed fold changes are plotted on the x-axis. Negative, log-transformed p-values are plotted on the y-axis. Transcripts of interest will lie in the topleft and topright corners of the plot (i.e., will have high fold-change and low p-value). Red dots correspond to mRNAs that have an absolute fold-change higher than 1.5, and a p-value below 0.05. Note that the p-values reported have not been corrected for multiple testing. Right panel: Heatmap of the most significantly differentially expressed genes among the two experimental groups. Samples are clustered according to the expression of the top 30 genes (dendrogram at the top), and experimental groups are colour-coded in the bar below the dendrogram, where red is group A, and blue is group B.

Table 5: Top 20 differentially expressed mRNAs (by adjusted p-value)

Gene	log (fold-change)	Average expression	t-test	p-value	Adjusted p-value
MMP21	-0.4070977	5.9496990	-4.7162371	0.0000703	0.8809042
IDH1	0.3201293	10.1234786	4.6688346	0.0000797	0.8809042
ZNF664	0.2745734	7.7079818	4.2521604	0.0002396	0.9998960
PIAS2	0.3911765	6.8948188	4.1953826	0.0002782	0.9998960
ARPC3	0.3015388	10.2959349	3.8232823	0.0007354	0.9998960
C10orf2	0.2302809	6.7747505	3.7978221	0.0007856	0.9998960
CD302	0.3005111	9.0574842	3.7731138	0.0008375	0.9998960
RBMS2	0.3078796	6.7953999	3.7624468	0.0008609	0.9998960
MYCT1	0.3306213	5.7850435	3.7397792	0.0009128	0.9998960
ASS1	0.3974624	6.2890375	3.6973617	0.0010183	0.9998960
ACOT2	0.4166325	9.6060855	3.6768314	0.0010736	0.9998960
HVCN1	0.2900359	6.1709221	3.6741327	0.0010811	0.9998960
BNIP3L	0.2672241	10.8814267	3.6590283	0.0011239	0.9998960
OSBPL9	0.2826700	8.8271433	3.6351627	0.0011950	0.9998960
PAQR4	0.2965801	6.8240213	3.6030721	0.0012975	0.9998960
ERCC1	0.3475652	7.8992945	3.5924864	0.0013331	0.9998960
HSD17B11	0.4166622	9.4904935	3.5287113	0.0015690	0.9998960
FGFR1OP	0.2474907	6.3419219	3.4772732	0.0017885	0.9998960
LOC728975	0.4969402	6.6474285	3.4695044	0.0018241	0.9998960

3.3.1.2. Sample size calculation

Using these data, I ran a sample size calculation with the `ssize` package for R (Figure 24). The sample size needed to identify 1.25-fold change differences in at least 83% of the genes is 72 per group (Figure 25). Due to experimental constrains (i.e., insufficient RNA yield, poor RNA quality and pairwise matching) only roughly half of the RNA samples can be included in the final microarray and RNA-Seq. This means that in order to achieve the aforementioned statistical power for 1.25-fold change differences I would have to extract RNA from 72 * 2 * 2 = 292 samples in total. This would be unfeasible in the context of the current project due to cost and labour, and therefore I focused on larger fold-changes.

The sample size calculation indicate that 24 samples per group would be needed to measure significant differences of 1.5 fold-changes and above in at least 81% of the genes, while 36 samples per group would allow me to measure significant differences of 1.5 fold-changes and above in at least 91% of the genes (Figure 24). Based on this result, I decided to increase the sample size to 24 per group.

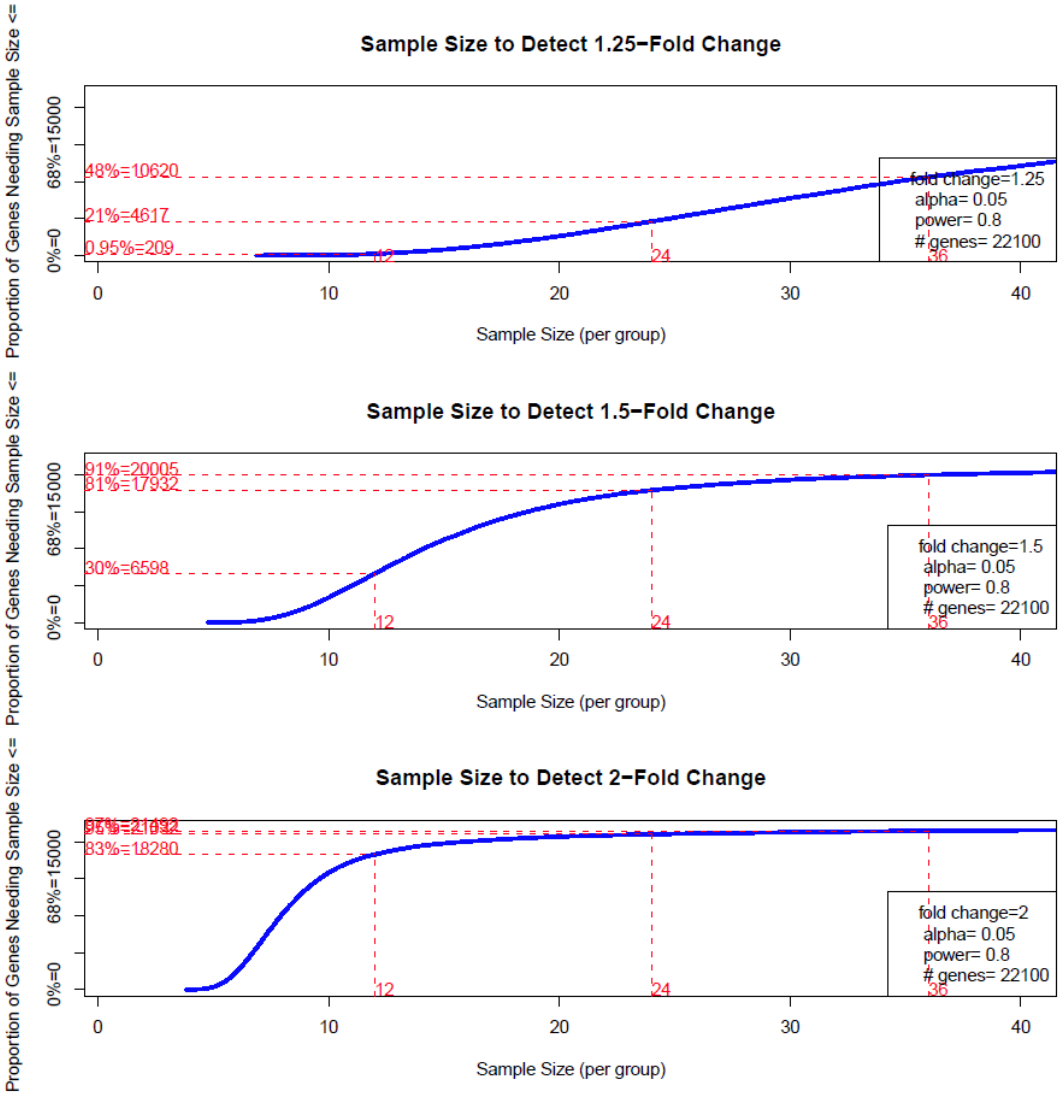


Figure 24: Sample size calculation for mRNA differential gene expression

Plots showing the sample size (x-axis) necessary to obtain a statistically sound analysis of a minimum percentage of mRNAs (y-axis). The plots assume alpha error = 0.05, power = 0.8, and 22100 genes detected above background. The top panel is simulates the sample size needed to measure at least a 1.25-fold changes, the middle panel 1.5-fold changes and the bottom panel 2-fold changes.

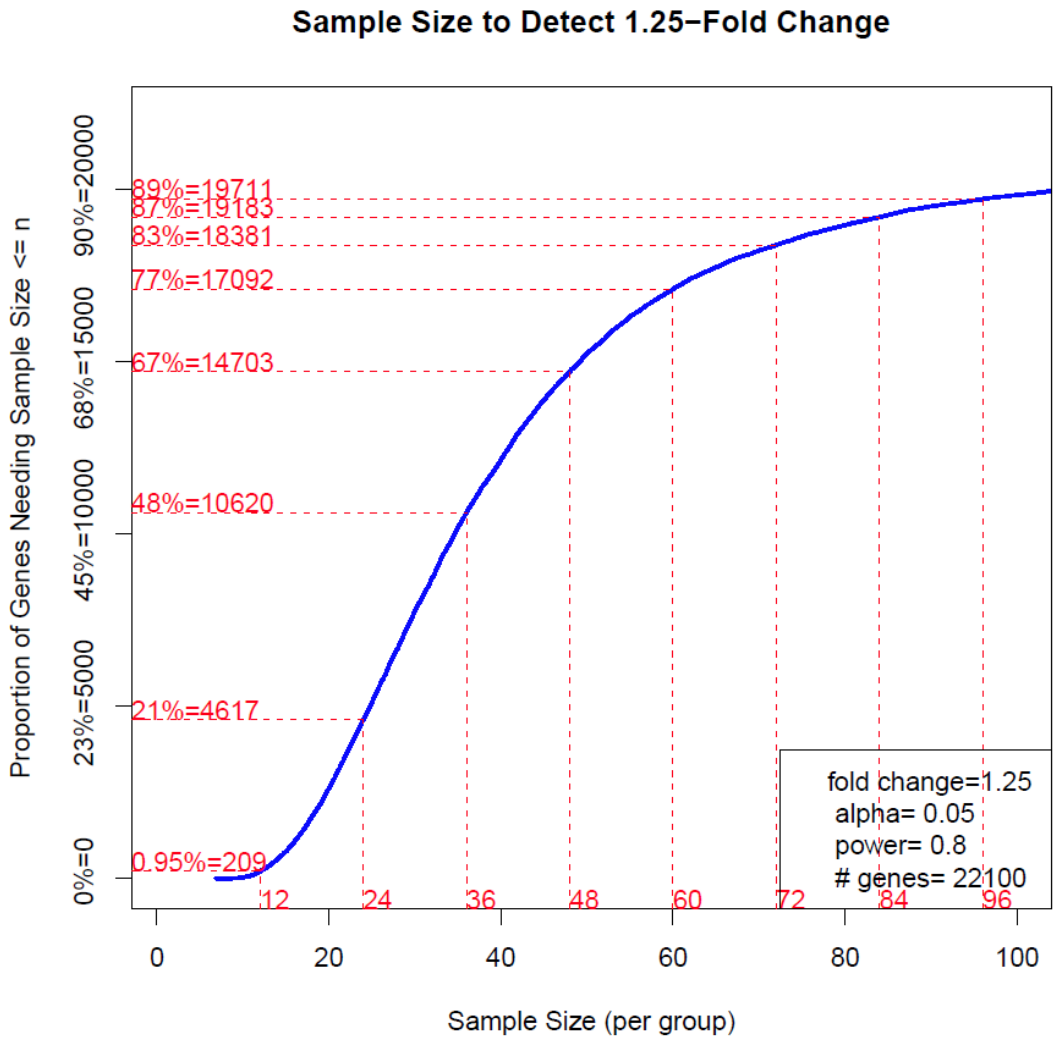


Figure 25: Sample size calculation to detect 1.25-Fold Changes

Plot showing the sample size (x-axis) necessary to measure a statistically significant difference of 1.25-fold changes and higher in a minimum percentage of mRNAs (y-axis) . The plots assume alpha error = 0.05, power = 0.8, and 22100 genes detected above background. The same plot is also reported in Figure 24.

3.3.1.3. Second batch of samples

Due to the Illumina microarray used for the first 24 samples being discontinued, the second batch of samples was run on an Affymetrix microarray. Given the difference in platforms, I analysed the two sample sets separately first, and then combined all the samples together for a full analysis.

Similarly to the Illumina microarray, the density distribution of probe intensity shows no abnormal sample (Figure 26, left panel). The unsupervised clustering by PCA does not separate the samples into the two experimental groups, or into any groups, indicating that differences in mRNA expression between the groups are not a major source of variability (Figure 26, right panel).

The top 20 differentially expressed mRNAs are reported in Table 6, sorted by adjusted p-values. As in the first batch of samples, no gene is below the required adjusted p-value threshold of 0.05, while 1497 have a p-value < 0.05 before correcting for multiple testing.

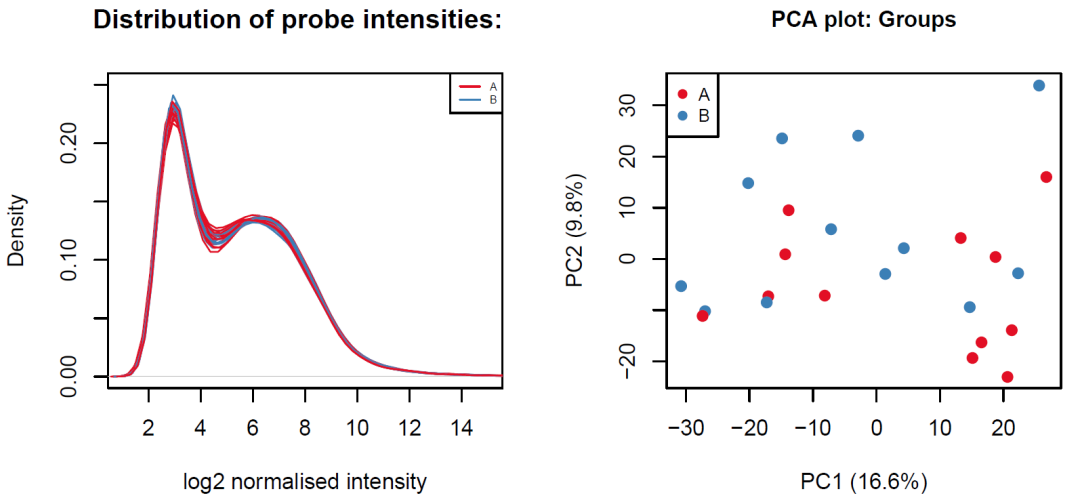


Figure 26: Quality control plots for Affymetrix microarray

Left panel: density plot of probe intensities. The x-axis shows the fluorescence intensity for all the probes on the microarray, normalised with the RMA function, and log-transformed. The y-axis shows the kernel density function of the intensities, computed with the `density()` function in R.

Right panel: Principal Component Analysis distribution for the microarray. Data points are transformed in such a way that the first Principal Component (PC1) expresses the largest possible variance for the data set, the second PC (PC2) expresses the second largest variance, and so on. PC1 and PC2 are reported on the x- and y-axis, respectively, and the associated explained variance is reported in brackets on the axis labels.

Samples belonging to experimental group A are reported in red, and samples belonging to group B in blue.

Table 6: Top 20 differentially expressed mRNAs (by adjusted p-value)

Gene	log (fold-change)	Average expression	t-test	p-value	Adjusted p-value
ARHGAP27	-0.5454375	10.3139944	-4.9343358	0.0000406	0.4474066
NMNAT3	-0.3128558	5.7099340	-4.7405541	0.0000676	0.4474066
BAMBI	0.4258362	8.6651650	4.6232726	0.0000920	0.4474066
LETMD1	-0.2314341	7.3698194	-4.6004465	0.0000977	0.4474066
CYP46A1	0.3463560	5.2274100	4.5768601	0.0001040	0.4474066
DNAJB11	0.2527740	7.4924448	4.4731676	0.0001366	0.4476229
LPAR4	0.5700273	4.8079863	4.3845446	0.0001724	0.4476229
NES	0.4850082	9.5350991	4.3278920	0.0002001	0.4476229
SPATC1	-0.4051368	3.7108978	-4.3230998	0.0002026	0.4476229
SMIM5	-0.5661747	3.6686333	-4.2353360	0.0002550	0.4476229
SLC25A28	-0.2557106	6.9214310	-4.2254829	0.0002617	0.4476229
MAPK4	0.5040793	8.1238957	4.2193268	0.0002659	0.4476229
CCDC93	-0.4124352	8.1265997	-4.2129255	0.0002704	0.4476229
MROH5	0.3407517	3.6093353	4.1218457	0.0003432	0.4979593
FAM58A	0.3995206	3.8036925	4.0342733	0.0004312	0.4979593
MSTN	-0.4296887	3.1854030	-4.0211945	0.0004462	0.4979593
ERLEC1	0.2008484	8.4254217	4.0052940	0.0004650	0.4979593
SNORD15B	-0.3411471	9.8418752	-4.0042958	0.0004662	0.4979593
CD58	0.2559057	6.5703835	3.9317986	0.0005629	0.4979593

3.3.1.4. Combined analysis

Unsupervised clustering on all the samples by means of PCA indicates that experimental group, sex and presence of diabetes do not determine any meaningful data segmentation (Figure 27 and Figure 28). The samples strongly cluster based on sample batch, which is the major source of variance (Figure 27). This is to be expected given that the two batches were run on two different microarray platforms, and it will be taken into account by adjusting the linear model used in the analysis for batch effects.

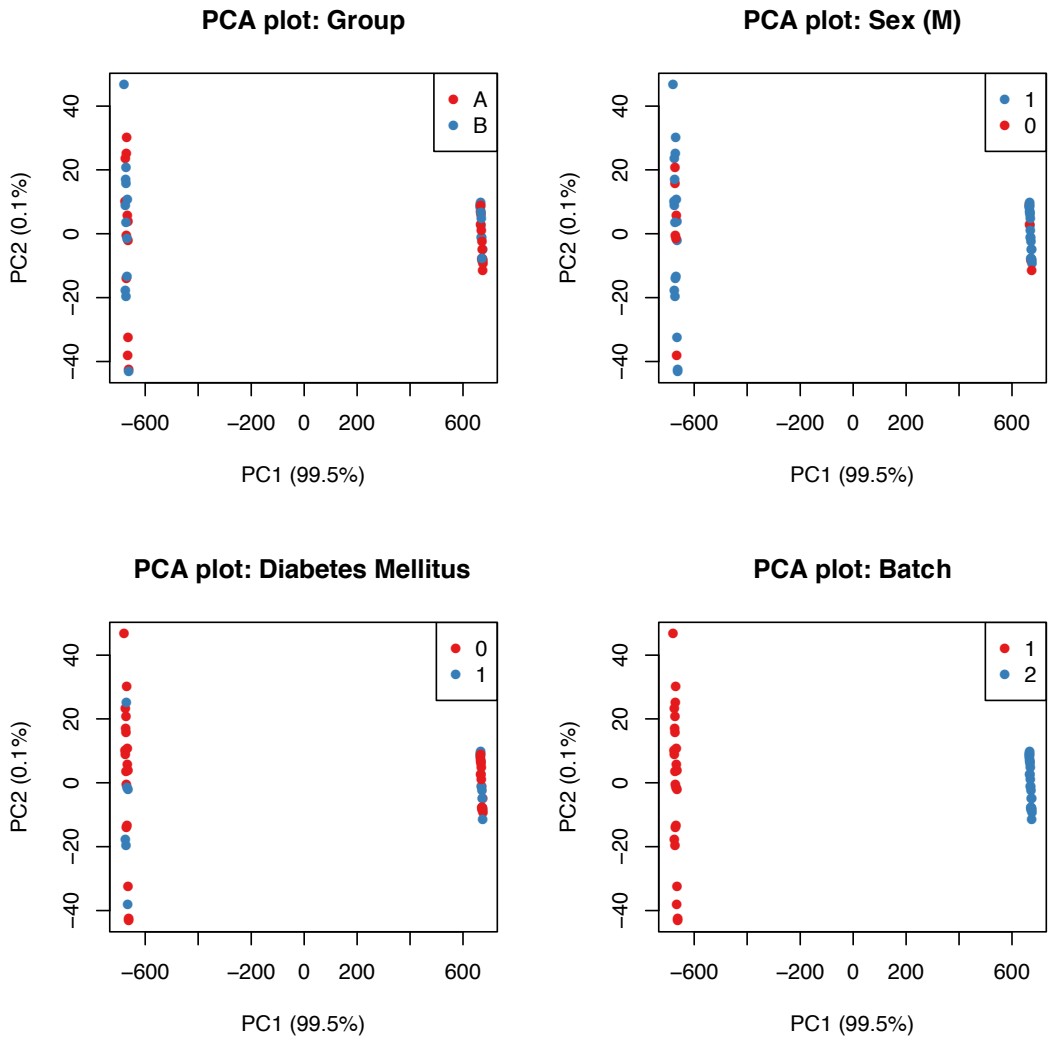


Figure 27: PCA plots of mRNA microarray data (PC1 and PC2)

Principal Component Analysis plots of the mRNA expression data.

The plots are coloured by group (topleft; red: group A, blue: group B), sex (topright; red: female, blue: male), presence of diabetes type I (bottomleft; red: no diabetes, blue: patient affected by diabetes), and batch (bottomright; red: first batch of samples, blue: second batch of samples). The samples clearly cluster according to batch number only.

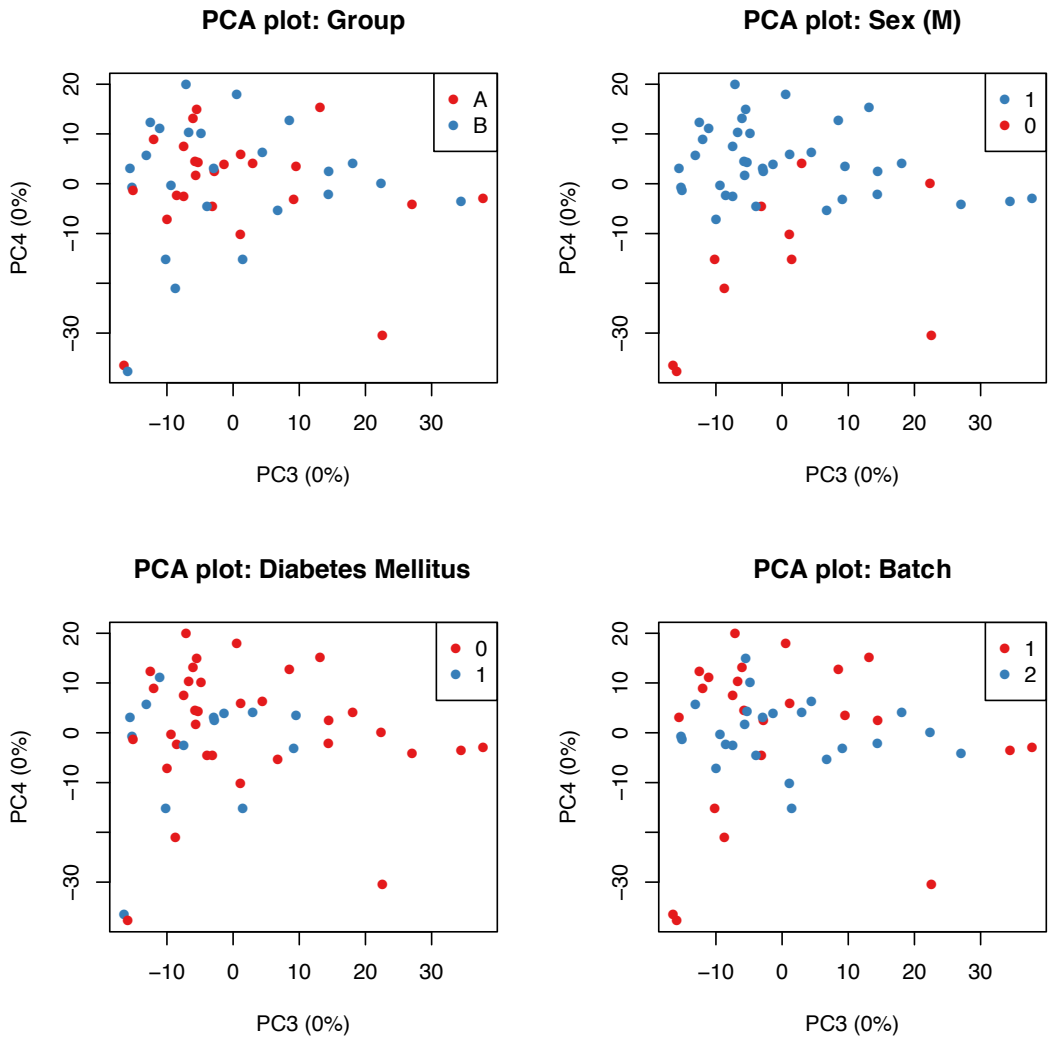


Figure 28: PCA plots of mRNA microarray data (PC3 and PC4)

Principal Component Analysis plots of the mRNA expression data.

The plots are coloured by group (topleft; red: group A, blue: group B), sex (topright; red: female, blue: male), presence of diabetes type I (bottomleft; red: no diabetes, blue: patient affected by diabetes), and batch (bottomright; red: first batch of samples, blue: second batch of samples). The samples do not cluster according along PC3 and PC4.

I combined together all the samples and applied the same analysis pipeline as before, and added batch-adjustment to the linear model. The linear model is now constructed in the form of $Y \sim G + B$, where Y is the response variable, G is the experimental group and B is the batch. The model now calculates the coefficients for the effect of G and B (each independent of the other), and their combined effect, allowing us to consider the effect of the experimental group on its own (hence adjusting for batch effect).

The top 20 differentially expressed transcripts sorted by adjusted p-value are reported in Table 7. As before, the adjusted p-values are suboptimal and none is below the 0.05 threshold.

The standard analysis for differential gene expression does not apply a fold-change cut off, and one can see how the absolute fold-change of the top 20 mRNAs in Table 7 range from 1.13 to 1.92. Considering that very small differences may not be clinically meaningful and would require extremely accurate detection methods, I also run the analysis giving higher weights to absolute fold-changes of 1.5 and higher using the `treat/topTreat` functions in the `limma` package. The top 20 differentially expressed mRNAs are reported in Table 8. The p-values of these mRNAs are higher as an additional constraint is added into the model.

Table 7: Top 20 differentially expressed mRNAs (by adjusted p-value)

Gene	log (fold-change)	Average expression	t-test	p-value	Adjusted p-value
ANP32D	0.3671013	3.8715799	4.6497160	0.0000252	0.4169830
ARHGAP27	-0.3502840	10.881070	-4.4852736	0.0000437	0.4169830
CP	0.4947466	11.499893	4.1594949	0.0001274	0.4859222
PAQR4	0.3066371	5.8597579	4.1028300	0.0001529	0.4859222
MIR183	-0.2187388	3.6451991	-4.1438701	0.0001339	0.4859222
SCTR	-0.3239853	4.5772724	-4.2297564	0.0001014	0.4859222
ASS1	0.9434366	27.780230	3.9714194	0.0002327	0.5546398
SERPINH1	0.2697858	8.2067036	4.0002631	0.0002123	0.5546398
CENPN	0.3897860	5.9629789	3.7901280	0.0004114	0.6807527
OSTM1	0.1765750	7.2963975	3.7328260	0.0004913	0.6807527
TRMT2B	-0.2208305	9.5483732	-3.7524641	0.0004624	0.6807527
TRIM3	-0.2695091	8.6609272	-3.7273081	0.0004998	0.6807527
ZNF783	-0.3304214	9.3511558	-3.7799962	0.0004245	0.6807527
GALNT16	-0.4218473	17.531304	-3.8420539	0.0003498	0.6807527
TAGLN	0.6497082	18.267398	3.6126314	0.0007103	0.6815775
IDH1	0.3848352	12.507184	3.6927559	0.0005559	0.6815775
PTH1R	0.3138076	9.4252575	3.6532225	0.0006276	0.6815775
ETV1	0.2636454	5.9817203	3.6671401	0.0006014	0.6815775
LGALSL	0.2578923	7.4486527	3.6403089	0.0006529	0.6815775

Table 8: Top 20 differentially expressed mRNAs of 1.5-fold change and higher (by adjusted p-value)

Gene	log (fold-change)	Average expression	t-test	p-value	Adjusted p-value
HLA-DRB5	2.2526929	6.07358670	2.35106976	0.01159367	0.9999999
HLA-DRB1	2.1628122	6.13817822	2.18999031	0.01682599	0.9999999
ASS1	0.9434365	27.7802306	1.48779999	0.07159081	0.9999999
HCG27	-0.9345490	20.3136298	-1.0696954	0.14499621	0.9999999
YME1L1	-1.6482981	79.5045518	-1.1029503	0.14960202	0.9999999
GSTT1	1.1675667	69.9548308	0.96642809	0.17176463	0.9999999
LOC100272216	-1.8918524	45.3595953	-0.9912750	0.19554181	0.9999999
TCF19	-0.7570080	24.1843057	-0.5033892	0.30855553	0.9999999
MOG	0.7814837	16.5981373	0.49846293	0.31059283	0.9999999
PPARG	0.7365944	15.4703831	0.48554965	0.31474988	0.9999999
HLA-A	-0.8248554	55.5550714	-0.4629111	0.32648503	0.9999999
MCCD1	-0.7474534	16.1605455	-0.3967319	0.34737434	0.9999999
TAGLN	0.6497082	18.2673981	0.33200114	0.37064892	0.9999999
GYG2	0.6863359	11.7539093	0.33206992	0.37065216	0.9999999
MICA	-0.8175720	22.3948564	-0.3586770	0.37626756	0.9999999
HSD11B1	0.6729267	12.9619493	0.30881503	0.37939323	0.9999999
NPPB	0.7316955	10.4348449	0.31347639	0.38022671	0.9999999
SELE	0.7038893	8.55509020	0.29401361	0.38579314	0.9999999
TRIM15	0.6695962	22.8709714	0.26982416	0.39425617	0.9999999

3.3.2. Small RNA profiling with RNA-Sequencing

The distribution of normalised counts for RNA-Sequencing shows that two samples are different from the others (Figure 29). The different count distribution indicates that something went wrong during the processing, most likely RNA degradation, and therefore the samples need to be removed from the analysis.

Given that the study was designed to be matched pair-wise between the groups, I also removed the corresponding matches to the two anomalous samples, bringing the total number of samples to 44 (22 per group). The quality control plots for this set of samples can be seen in Figure 30. Both the density plot and the boxplots indicate that the remaining samples are uniform and comparable between each other; no clear outlier can be identified in the pool.

Unsupervised clustering by means of PCA indicates that the samples do not naturally cluster according to experimental groups, sex or presence of diabetes. It is clear that the major source of variability (indicated as principal component 1, PC1) is determined by the sample batch, which accounts for roughly 25% of the variance (Figure 31). This will be taken into account during the differential gene expression analysis by adjusting the linear model for batch effects.

Principal components 2, 3 and 4 only account for 3-10% of the variance, and do not segment the samples in any meaningful way (Figure 32).

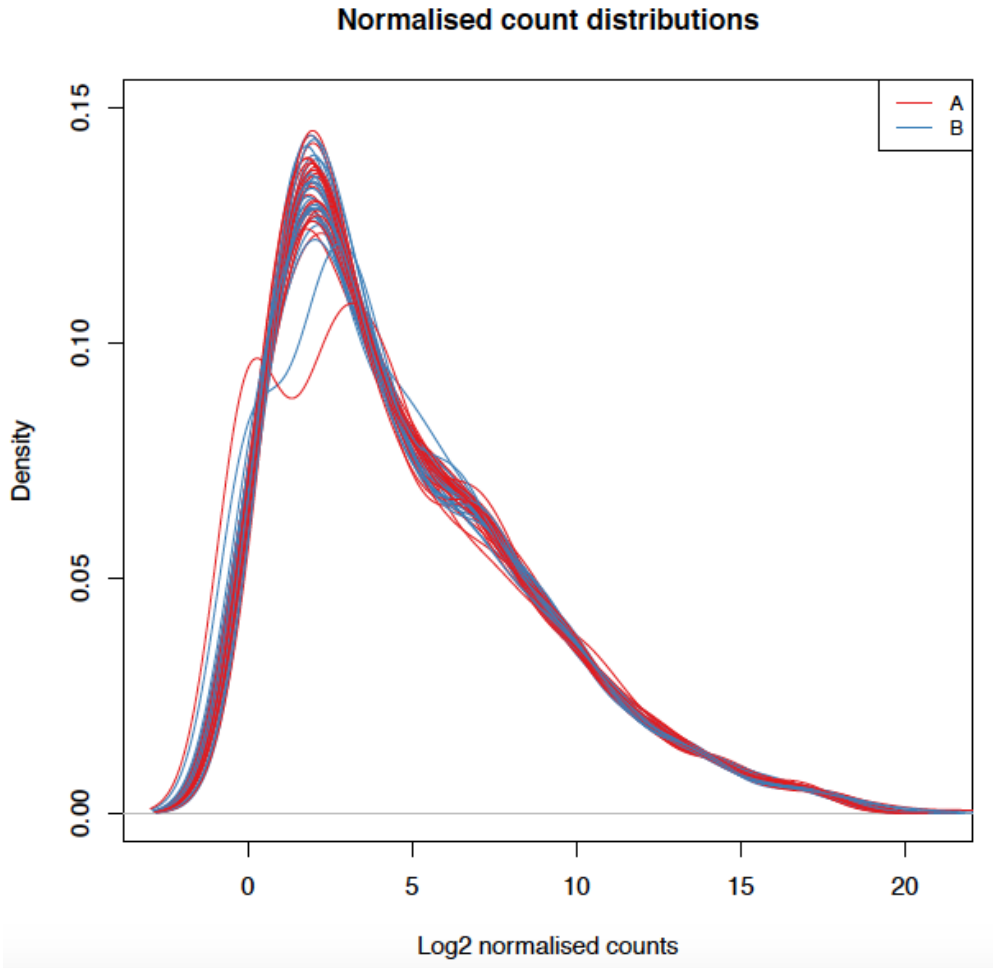


Figure 29: Normalised count distribution of samples from small RNA-Seq

The x-axis shows the counts for all the small RNAs identified in the study, and log-transformed. The y-axis shows the kernel density function of the intensities, computed with the `density()` function in R.

Two curves have a different distribution from the rest, indicating a problem with those two samples.

Samples belonging to experimental group A are reported in red, and samples belonging to group B in blue.

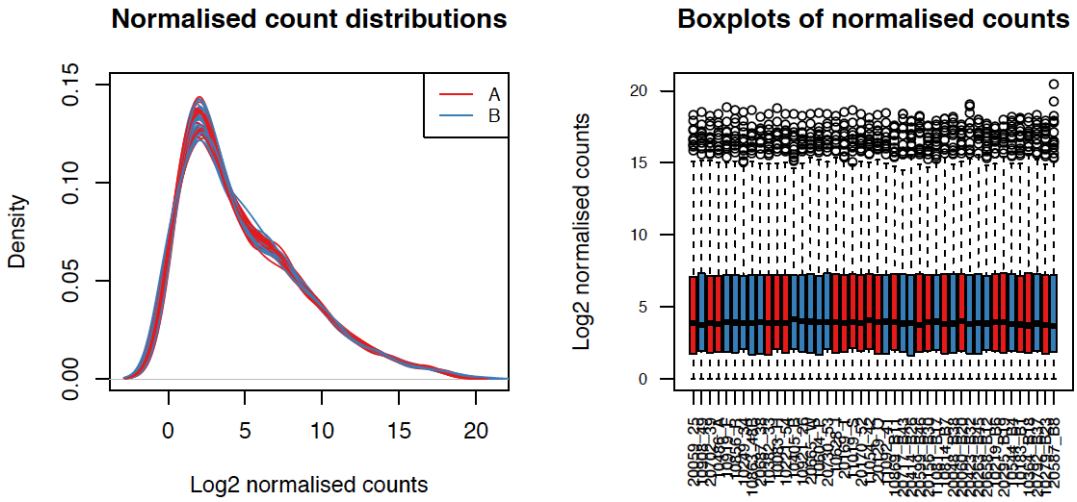


Figure 30: Quality control plots for small RNA-Seq

Left panel: density plot of count data from

Figure 29, after the removal of the two anomalous samples. The curves look now homogeneous.

Right panel: Boxplots of normalised counts (y-axis) per each sample (x-axis) after the removal of the anomalous samples. All the samples look homogeneous. Samples belonging to experimental group A are reported in red, and samples belonging to group B in blue.

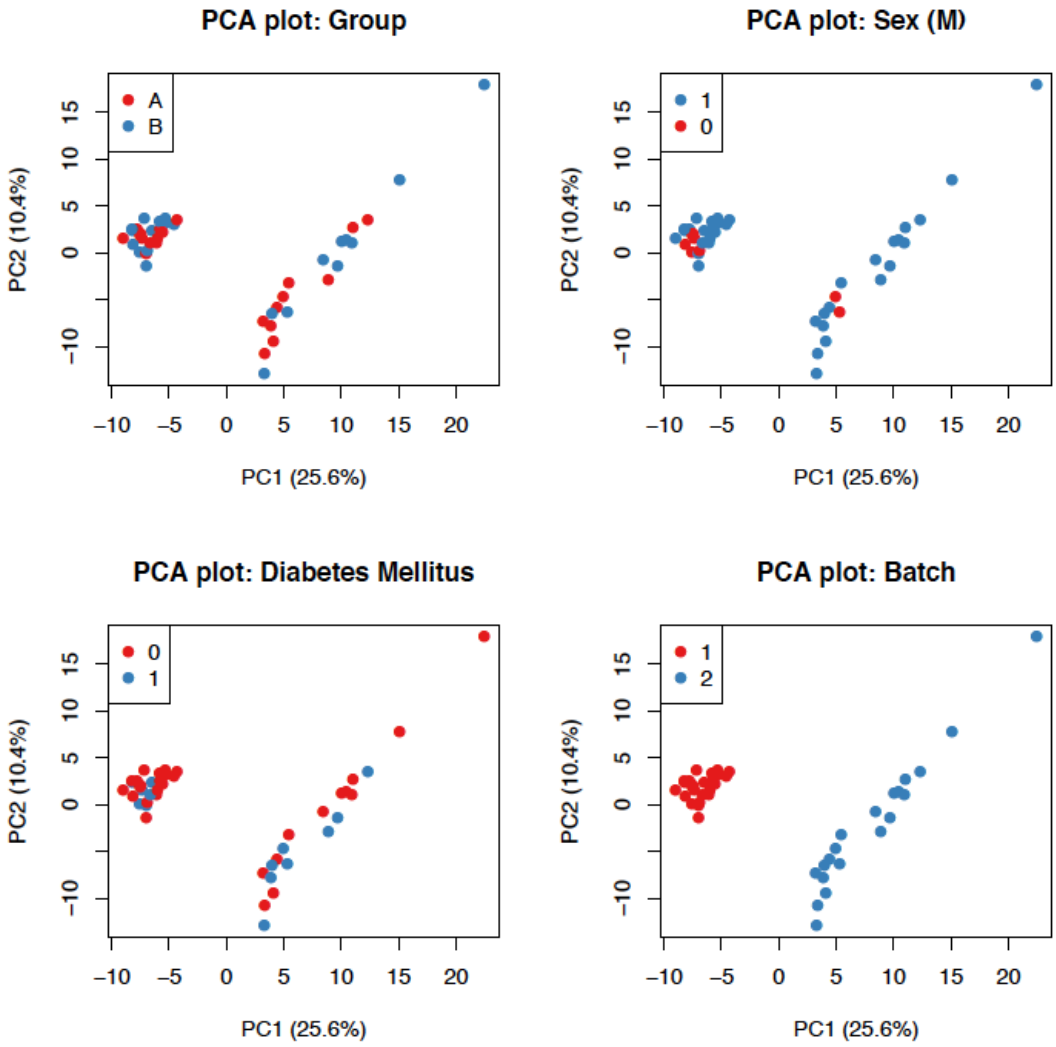


Figure 31: PCA plot of small RNA-Seq samples (PC1 and PC2)

Principal Component Analysis plots of the RNA-Seq data.

The plots are coloured by group (topleft; red: group A, blue: group B), sex (topright; red: female, blue: male), presence of diabetes type I (bottomleft; red: no diabetes, blue: patient affected by diabetes), and batch (bottomright; red: first batch of samples, blue: second batch of samples). The samples clearly cluster according only to batch number.

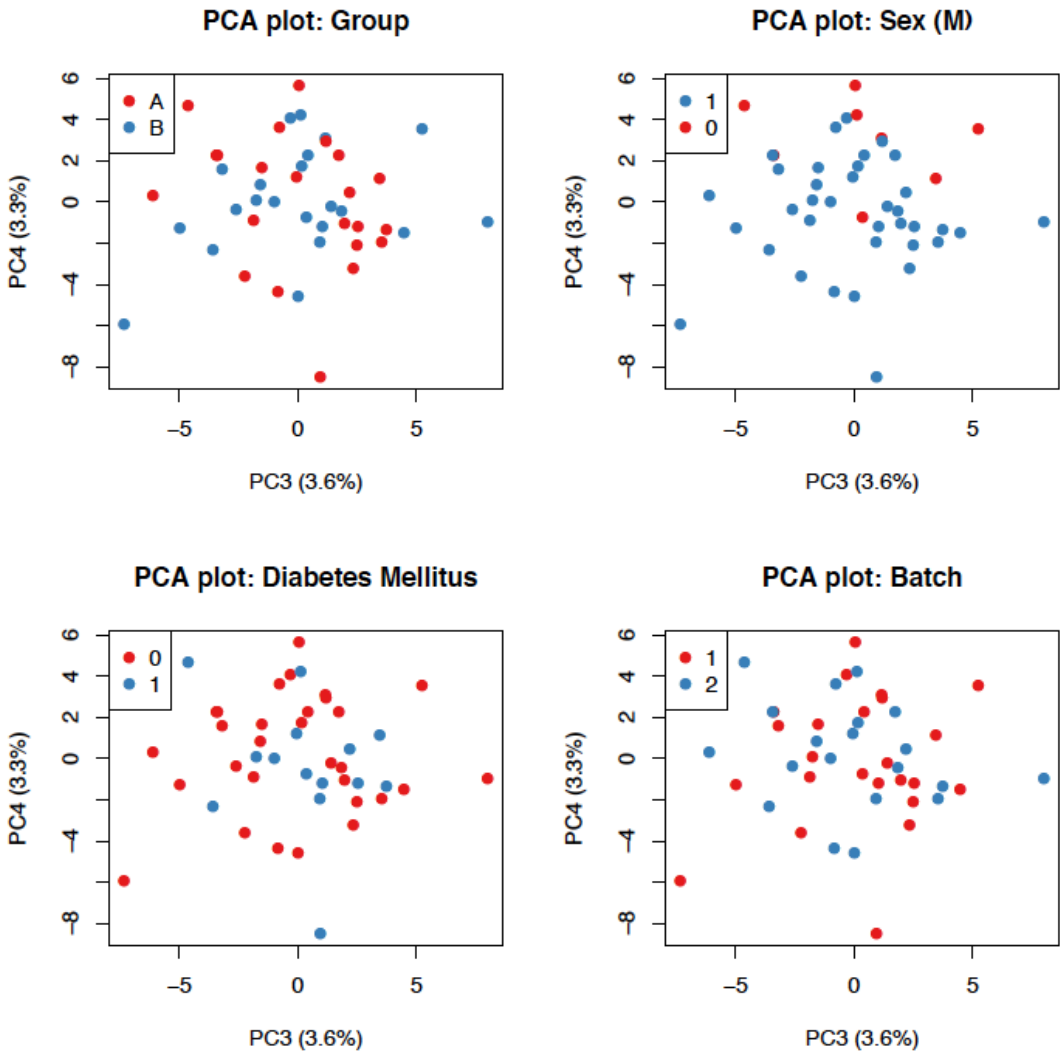


Figure 32: PCA plot of small RNA-Seq samples (PC3 and PC4)

Plots of Principal Components 3 and 4 for RNA-Seq samples.

The plots are coloured by group (topleft; red: group A, blue: group B), sex (topright; red: female, blue: male), presence of diabetes type I (bottomleft; red: no diabetes, blue: patient affected by diabetes), and batch (bottomright; red: first batch of samples, blue: second batch of samples). The samples do not form any visible cluster.

The differential gene expression analysis reveals that 30 miRNAs are differentially expressed at a p-value of 0.05 or lower. However, after correcting for multiple comparisons, none is below the adjusted p-value threshold of 0.05. The top 20 results are reported in Table 9.

I also tested for differentially expressed transcripts with an absolute fold change significantly higher than 1.25 (Table 10). As expected, all the p-values are above the significance threshold, and there is no statistically significant candidate to bring forward for further validation.

3.3.2.1. Sample size calculation

I ran a sample size calculation with the data obtained from the small RNA-Seq using the online tool RNAseqPS (Guo et al., 2014). A sample size of 55 per group would be needed to identify changes of 1.5 fold change or higher at a 0.8 power and 0.05 false discovery rate. A screenshot of the program output, detailing all the parameters used for the simulation, can be seen in Figure 33. Detecting a fold-change of 1.25 will require a sample size of 183, while a 2 fold-change will require 19 samples per group.

Table 9: Top 20 differentially expressed miRNAs (by p-value)

miRNA	log (fold-change)	log(CPM)	LR	p-value	Adjusted p-value
hsa-miR-152-5p	0.435084	4.061610	9.256273	0.0023469	0.518666
hsa-miR-381-3p	0.426682	8.499381	7.952482	0.0048021	0.518666
hsa-miR-10b-3p	-0.62671	3.048757	7.917574	0.0048956	0.518666
hsa-miR-148a-5p	0.414112	6.413791	7.564259	0.0059537	0.518666
hsa-miR-148a-3p	0.555941	15.84201	6.695932	0.0096633	0.518666
hsa-miR-193a-5p	0.239236	7.465997	6.236575	0.0125138	0.518666
hsa-miR-584-5p	0.331681	6.778976	5.949097	0.0147248	0.518666
hsa-miR-324-5p	-0.24133	5.582895	5.347028	0.0207576	0.518666
hsa-miR-143-3p	0.258063	17.22450	5.324340	0.0210295	0.518666
hsa-miR-15a-5p	-0.39864	5.069217	5.318274	0.0211029	0.518666
hsa-miR-139-5p	0.211595	10.87853	5.310248	0.0212003	0.518666
hsa-miR-151a-3p	0.409890	12.33548	5.263839	0.0217729	0.518666
hsa-miR-339-3p	0.170368	6.753298	5.218120	0.0223526	0.518666
hsa-miR-7151-5p	0.399845	2.801262	4.960788	0.0259283	0.518666
hsa-miR-139-3p	0.278835	6.518411	4.922407	0.0265105	0.518666
hsa-miR-190b	0.307778	5.580372	4.779085	0.0288074	0.518666
hsa-miR-1	0.225527	9.881436	4.697220	0.0302114	0.518666
hsa-miR-3909	0.351726	4.868978	4.620640	0.0315894	0.518666
hsa-miR-103a-2-5p	-0.28628	4.116096	4.599836	0.0319750	0.518666

log(CPM): average expression (counts per million); LR: Likelihood Ratio statistics

Table 10: Top 20 differentially expressed miRNAs above 1.25 fold change (by p-value)

miRNA	log (fold-change)	log(CPM)	p-value	Adjusted p-value
hsa-miR-10b-3p	-0.626713002	3.04875724	0.02564627	0.999990013
hsa-miR-148a-3p	0.55594104	15.8420155	0.046569342	0.999990013
hsa-miR-152-5p	0.435084374	4.06161095	0.077274512	0.999990013
hsa-miR-301a-3p	-0.492800037	3.28537163	0.089094638	0.999990013
hsa-miR-381-3p	0.426682065	8.49938146	0.09261045	0.999990013
hsa-miR-148a-5p	0.414112594	6.41379130	0.105338628	0.999990013
hsa-miR-151a-3p	0.40989057	12.3354899	0.128514881	0.999990013
hsa-miR-15a-5p	-0.398642279	5.06921746	0.137077315	0.999990013
hsa-miR-7151-5p	0.39984524	2.80126295	0.137182822	0.999990013
hsa-miR-6716-3p	0.432836252	3.14026565	0.137281653	0.999990013
hsa-miR-582-3p	0.397201778	4.52229730	0.15151581	0.999990013
hsa-miR-671-3p	0.386999747	4.48230719	0.158488613	0.999990013
hsa-miR-302a-3p	0.381721066	3.04560294	0.162813918	0.999990013
hsa-miR-7706	0.409687794	5.58669261	0.167011511	0.999990013
hsa-miR-708-3p	0.380993436	3.94910723	0.175183582	0.999990013
hsa-miR-654-5p	0.372050115	2.74275466	0.177972028	0.999990013
hsa-miR-1298-5p	-0.582883554	3.62040759	0.195022601	0.999990013
hsa-miR-3124-5p	0.356484658	3.16027693	0.195479832	0.999990013
hsa-miR-3909	0.351726191	4.86897802	0.196892905	0.999990013

log(CPM): average expression (counts per million)

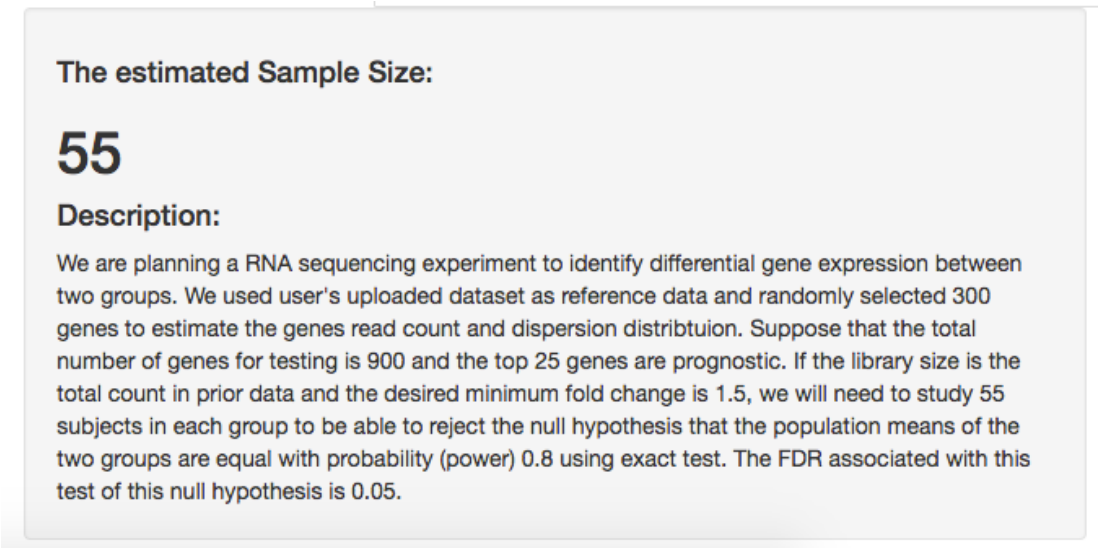


Figure 33: RNAseqPS output

Descriptive output of the web interface of RNAseqPS (https://cqs.mc.vanderbilt.edu/shiny/RnaSeqSampleSize/), indicating that a sample size of 55 per group would be needed to measure 1.5-fold changes difference in a pool of 900 miRNAs, with a FDR of 0.05 and power of 0.8.

3.4. DISCUSSION

The aim of this project was to identify the molecular signature of the atrial substrate that underpins the development of poAF by combining transcriptomic analysis of mRNAs and miRNAs in 48 paired human atrial samples. In this limited sample size, differential gene expression analysis did not return any significant hits. This is not entirely surprising as changes at this early stage (all patients were in SR at the time the atrial samples were taken) are expected to be subtle. Alternatively, one should be open to the possibility that such transcriptional markers may not be actually present in the myocardial samples.

For the microarray, the sample size calculation run on the first 24 samples (12 per group) indicated that 24 samples per group would be needed to detect a 1.5 fold change or higher in at least 81% of the genes. 12 samples per group would allow me to do so in only 30% of the genes. While the total sample size of this study was indeed 24, there was a change in microarray platform between the first and second batch of samples (i.e., the first 24 samples were run on an Illumina microarray, while the next 24 on an Affymetrix array). The change was due to the Illumina microarray production being discontinued by the manufacturer. This likely negatively impacted the overall statistical power of the study, by increasing the variability of the same transcript among all the samples. For this reason, a further set of 24 samples on an Affymetrix array should to be performed to obtain more robust results.

Looking at the differentially expressed mRNAs (although not above statistical significance) above 1.5 fold changes in poAF vs. SR, I noticed that there is no

overlap at all between the ones identified with the Illumina and the ones identified with the Affymetrix array. This observation suggests that the results of both arrays are indeed not statistically significant, as there is no replication from one batch to the other.

Similar conclusions can be drawn on the small RNA-Sequencing. No miRNA is differentially expressed at a significance level below an adjusted p-value of 0.05. No miRNA at the top of the differential expression list has a fold-change of 2 or higher. For this reason, the choice of sample size should be tailored to pick up differences of at least 1.5 fold-changes. Such sample size is predicted to be 55 per group. In order to detect 1.25 fold-change differences between the groups, one would need 183 samples per group. Both these scenarios would likely require the set up of a dedicated medium-large scale study.

Based on this prediction, the current study is underpowered to pick up differences of 1.5 fold-changes, and therefore no meaningful conclusions can be drawn from the results.

This project presents some significant technical challenges that may have an impact on the overall results. First of all, the quality and quantity of the RNA extracted from human samples are not always optimal. Further restrictions are imposed by the pair-wise matching: if one of the samples in the pair is not suitable, both may be discarded as re-pairing may insert significant bias in the results. In practice, only around 50% of the samples are used after RNA

extraction. Samples are rejected due to insufficient RNA yield (a minimum of 2 μg of total RNA is needed for the combined microarray and RNA-Seq) or poor quality (RNA degradation visible on agarose gel or presence of impurities). Even after rejecting such a high proportion of samples, some of them do not meet the strict criteria for optimal quality that are required from RNA-Sequencing guidelines (e.g., RIN score above 9).

If I were to bring this project forward I would:

i) Drop the miRNA investigation. The sample size calculations run on the data obtained so far are too high for a single-person run project, and would require a dedicated (large) budget.

ii) Run a second, and, budget allowing, a third batch of microarray samples on the Affymetrix platform. This would reduce the intra-platform variability that may make the results presented here less significant.

If no mRNAs would result differentially expressed between the two groups I would conclude that there are no alterations that can be successfully detected at the time of cardiac surgery, using the currently available microarray technology.

If, on the other hand, a further wave of screening would bring forward some differentially expressed mRNAs, I would validate the top 10-20 hits by qPCR in an appropriately sized independent cohort of

samples. I would then choose 1-5 of the ones validated in qPCR and investigate their functional role in poAF.

4. APPENDIX: R SCRIPTS

4.1. MICROARRAY ANALYSIS

4.1.1. Quality control of first batch of samples

```
## Illumina microarray QC
## My array from Dilair
## 2-3-2016
## Modified from P150367 Simon

library(vsn)

project.id <- "P150778"


## Read and parse raw data
source("./Illumina_GEX_preprocessing_v1.0.R")

raw.data <-
read.raw.data.files("SampleProbeProfile_Matilde.txt",
"SampleTableControl_Matilde.txt")
## in ProbeProfile samples are reported according to the
'Index' in TableControl which is NOT the order on the chip,
but ordered by experimental group and chip
## check signal data and control data in the same order

names(raw.data)

sig.data <- raw.data$sig.data
det.data <- raw.data$det.data
con.data <- raw.data$con.data

## samples in same order in all files (by group, rather than
chip)
colnames(sig.data)
colnames(sig.data) <- gsub(".AVG_Signal", "",
colnames(sig.data))
identical((as.vector(con.data$Sample.ID)), gsub("X", "",
colnames(sig.data))) # TRUE #the name of the columns in the
two files should match - means the samples are in the same
order

colnames(det.data) <- gsub(".Detection.Pval", "",
colnames(det.data))
```

```
identical(as.vector(con.data$Sample.ID), gsub("X", "",
colnames(det.data))) # TRUE


table(con.data$Sample.Group)
A  B
12 12


## check that everything is fine and there were no typos in
the identifyier supplied with the data

sample.ids <- as.vector(con.data$Sample.ID)
groups <- as.factor(con.data$Sample.Group)
chips <- as.factor(con.data$Sentrix.Barcode)


## define colour vectors for plots

# install.packages("RColorBrewer")
library(RColorBrewer)
#display.brewer.all()
col.pal <- c(brewer.pal(9, "Set1"), brewer.pal(12, "Set3"),
brewer.pal(3, "Pastel1"))
col.pal2 <- cbind(col.pal, c(rep("#E41A1C", 12),
rep("#377EB8", 12)))
colnames(col.pal2) <- c("different", "group")

palette(col.pal)


############################################################
## generate QC metric plot

source("./Illumina_GEX_preprocessing_v1.0.R")
library(ggplot2)
library(reshape2)

df2 <- qc.metrics(con.data, sample.ids, groups)

png(paste0(project.id,
"_QC_metrics.png"),width=1000,height=800)
     print(ggplot(data=df2, aes(x=Sample.ID, y=value,
fill=Group)) + geom_bar(stat = "identity") +
scale_fill_manual(values = col.pal) +
       facet_grid(variable ~ Sentrix.Barcode, scales="free")
+ theme(axis.text.x=element_text(angle = 90,vjust = 0.5,
hjust=0.5)) +
       ylab("Value") + xlab("Sample name"))
     dev.off()
```

```
############################################################
## data processing steps

## background subtraction

sig.data <- bkg.cor(sig.data, con.data)
head(sig.data)


## VSN normalisation
library(vsn)
fit <- vsn2(as.matrix(sig.data))
sig.v2 <- predict(fit, newdata = as.matrix(sig.data)) ##this
gives the same result as
meanSdPlot(sig.v)

## PREVIOUS VERSION
# sig.v <- vsn2(as.matrix(sig.data))
# sig.norm <- exprs(sig.v)
## sig.norm == sig.v2 TRUE


write.table(sig.v2, paste0(project.id,
"_VSN_normalised_data.txt"), sep="\t", quote=F, row.names=T)


############################################################
## Exploratory QC plots

## define plot functions
source("./Exploratory_QC_plot_functions_v1.3.R")

pdf(paste0(project.id, "_QC_plots_unfiltered_data.pdf"),
onefile=T)

make.density.plot(data=sig.v2, xlim.min=3.5, xlim.max = 15,
ylim.max = 0.6, title="VSN normalised data", col.vec=groups,
groups=groups, leg="T")

make.density.plot(data=sig.v2, xlim.min=3.5, xlim.max = 15,
ylim.max = 0.6, title="VSN normalised data", col.vec=col.pal,
groups=sample.ids, leg="T")

boxplot(as.data.frame(sig.v2), outline=F, names=sample.ids,
col=groups, las=2, cex.axis=0.8, cex.names=0.8,
main=paste0("Boxplots of probe intensities: ", "VSN normalised
data"), ylab="Log2 normalised intensity")


dists <- dist(t(sig.v2))
```

```
plot(hclust(dists, method="complete"),
main=paste0("Hierarchical clustering: ", "VSN normalised
data"), labels=sample.ids, cex=1)
plot(hclust(dists, method="complete"),
main=paste0("Hierarchical clustering: ", "VSN normalised
data"), labels=groups, cex=1)


  pca.norm <- prcomp(t(sig.v2))
  PC.norm <- pca.norm$x
  PC.norm <- data.frame(PC.norm)
  attach(PC.norm)

  var.cap <- summary(pca.norm)
  pc_var.cap <- round(var.cap$importance[2,]*100, 1)
  cum_var.cap <- round(var.cap$importance[3,]*100, 1)

  pc <- list("PC"=PC.norm, "PC1_var"=pc_var.cap[1],
"PC2_var"=pc_var.cap[2])


  ## identify samples by name
  plot(PC.norm$PC1, PC.norm$PC2, type="n", main=paste0("PCA
plot: ", "Samples"), cex.main=1, xlab=paste("PC1 (",pc
$PC1_var, sep="","%)"), ylab=paste("PC2 (",pc$PC2_var,
sep="","%)"))
  text(PC.norm$PC1, PC.norm$PC2, sample.ids, cex=0.8)

  plot(PC.norm$PC1, PC.norm$PC2, type="p", main=paste0("PCA
plot: ", "Groups"), cex.main=1, xlab=paste("PC1 (",pc$PC1_var,
sep="","%)"), ylab=paste("PC2 (",pc$PC2_var, sep="","%)"),
pch=19, col=groups)
  legend("topright", legend=unique(groups), pch=19,
col=unique(groups), cex=0.8)

  plot(PC.norm$PC1, PC.norm$PC2, type="p", main=paste0("PCA
plot: ", "Chip"), cex.main=1, xlab=paste("PC1 (",pc$PC1_var,
sep="","%)"), ylab=paste("PC2 (",pc$PC2_var, sep="","%)"),
pch=19, col=chips)
  legend("topright", legend=unique(chips), pch=19,
col=unique(chips), cex=0.8)


dev.off()


############################################################

## filter data (keeping those detected in at least one group)
and generate QC plots
## detection score <0.05 considered detected (or score >0.95)
```

```
keep <- rowSums(det.data < 0.05) >= 12  #cut-off of 10 gives:
22883    24
sig.v2.fil <- sig.v2[keep,]
dim(sig.v2.fil) # 22100    24

write.table(sig.v2.fil, paste0(project.id,
"_VSN_filtered_data.txt"), sep="\t", quote=F, row.names=T)

pdf(paste0(project.id, "_QC_plots_filtered_data.pdf"),
onefile=T)

make.density.plot(data=sig.v2.fil, xlim.min=3.5, xlim.max =
15, ylim.max = 0.4, title="VSN filtered data", col.vec=groups,
groups=groups, leg="T")

make.boxplot(sig.v2.fil, "VSN filtered data", sample.ids,
groups, 0.8)

dists.fil <- calc.dist(sig.v2.fil)
make.cluster.plot(dists.fil, "VSN filtered data", sample.ids,
1)
make.cluster.plot(dists.fil, "VSN filtered data", groups, 1)

pc.fil <- calc.pca(sig.v2.fil)

pca.ids(pc.fil$PC, pc.fil$PC1_var, pc.fil$PC2_var, "Samples",
sample.ids, 0.8)

pca.colours(pc.fil$PC, pc.fil$PC1_var, pc.fil$PC2_var,
"Groups", groups, groups, "T", "topright", 0.8)
pca.colours(pc.fil$PC, pc.fil$PC1_var, pc.fil$PC2_var,
"Chips", chips, chips, "T", "topright", 0.8)
dev.off()

# Similar pattern to unfiltered.

############################################################
## group relevant plots for report

png(paste0("./Plots_for_report/", project.id,
"_VSN_assessment.png"), height=480, width=480)
meanSdPlot(sig.norm, main="Mean-variance plot (normalised
data)", ylim=c(0,4))
dev.off()


## density plots

png(paste0("./Plots_for_report/", project.id,
"_density_plot.png"), height=480, width=960)
```

```
par(mfrow=c(1,2))
make.density.plot(sig.norm, 0.45, "VSN data", groups, groups)
make.density.plot(sig.norm.fil, 0.35, "VSN filtered data",
groups, groups)
dev.off()

## for filtered data only
## boxplots
png(paste0("./Plots_for_report/", project.id,
"_boxplots.png"), height=480, width=960)
make.boxplot(sig.norm.fil, "VSN filtered data", sample.ids,
groups, 1)
dev.off()

## clustering
png(paste0("./Plots_for_report/", project.id,
"_clustering.png"), height=480, width=600)
par(mfrow=c(1,1))
make.cluster.plot(dists.fil, "VSN filtered data", sample.ids,
0.8)
dev.off()

png(paste0("./Plots_for_report/", project.id, "_PCA.png"),
height=480, width=600)
pca.ids(pc.fil$PC, pc.fil$PC1_var, pc.fil$PC2_var, "Samples",
sample.ids, 0.7)
dev.off()

## colour by group
png(paste0("./Plots_for_report/", project.id,
"_PCA_groups.png"), height=480, width=600)
pca.colours(pc.fil$PC, pc.fil$PC1_var, pc.fil$PC2_var,
"Group", groups, groups, "topright", 0.8)
dev.off()

## colour by chip
png(paste0("./Plots_for_report/", project.id,
"_PCA_chip.png"), height=480, width=600)
pca.colours(pc.fil$PC, pc.fil$PC1_var, pc.fil$PC2_var, "Chip",
chips, chips, "topright", 0.8)
dev.off()

## colour by replicate
png(paste0("./Plots_for_report/", project.id, "_PCA_rep.png"),
height=480, width=600)
pca.colours(pc.fil$PC, pc.fil$PC1_var, pc.fil$PC2_var,
"Replicate", replicate, replicate, "topright", 0.8)
dev.off()
```

4.1.2. Illumina_GEX_preprocessing_v1.0.R

```
## Illumina microarray gene expression projects

## defines functions to:
## 1. read in raw data files
## 2. generate QC metrics plot
## 3. initial data preprocessing


##read and parse raw data
read.raw.data.files <- function(probeProfile, conTable)
{

      d <- read.delim(paste0("./DATA/", probeProfile),
header=T, sep="\t")
      print("Raw data matrix:")
      print(dim(d))
      print("Column names:")
      print(colnames(d))

      # assumes signal and detection P-VALUE (not detection
score) columns are present
      sig.data <- d[,grep("AVG_Signal", colnames(d))]
      det.data <- d[,grep("Detection.Pval", colnames(d))]

      print("Signal matrix: ")
      print(dim(sig.data))
      print("Detection matrix: ")
      print(dim(det.data))

      row.names(sig.data) <- d$PROBE_ID
      row.names(det.data) <- d$PROBE_ID

      ## read in control data file
      con.data <- read.table(paste0("./DATA/", conTable),
sep="\t", header=T)

      print("Control table: ")
      print(dim(con.data))

      return(list(sig.data=sig.data, det.data=det.data,
con.data=con.data))

}


############################################################
## QC metrics plot - use ggplot2 to summarize con.data
```

```
qc.metrics <- function(con.data, sample.ids, groups)
{

      df <-
con.data[,c("Index","Sample.ID","Sample.Group","Sentrix.Barcod
e","Sample.Section",

"BIOTIN","CY3_HYB","Signal.Average","HOUSEKEEPING","Detected.G
enes..0.05.")] #Create dataframe for ggplot2
      # df$name <- paste0(df$Sample.Section,"_",sample.ids)
#Add a column with sample names
      df$Sample.Group <- groups
      df$Index <- as.character(df$Index) # remove numeric info
from fields that should be used as ID variables
      df$Sample.ID <- as.character(df$Sample.ID)
      df$Sentrix.Barcode <- as.factor(df$Sentrix.Barcode)

      colnames(df) <- gsub("BIOTIN", "Biotin Signal",
colnames(df))
      colnames(df) <- gsub("CY3_HYB", "Hybridisation Signal",
colnames(df))
      colnames(df) <- gsub("Signal.Average", "Average Signal",
colnames(df))
      colnames(df) <- gsub("HOUSEKEEPING", "Housekeeping Gene
Signal", colnames(df))
      colnames(df) <- gsub("Detected.Genes..0.05.", "Number
Detected Genes", colnames(df))
      colnames(df) <- gsub("Sample.Group", "Group",
colnames(df))
      df2 <- melt(df) # remove numeric info from fields that
should be used as ID variables
      str(df2)
      return(df2)

}


## background correct data

bkg.cor <- function(sig.data, con.data)
{
      ## extract column containing negative signal measure -
average signal from negative probes for each array (must be
same order as sig.data)
      bkg <- con.data[,grep("NEGATIVE", colnames(con.data))]

      ## subtract average bkg signal for each array from each
intensity on that array
      for(i in 1:length(con.data[,1]))
      {
```

```
            sig.data[,i] <- sig.data[,i]-bkg[i]
      }
      return(sig.data)
}

## VSN for variance stabilisation and normalisation

vsn.norm <- function(sig.data, project.id)
{
      library(vsn)
      sig.v <- vsn2(as.matrix(sig.data))
      meanSdPlot(sig.v, ylim=c(0,4))

      sig.norm <- exprs(sig.v)
      write.table(sig.norm, paste0(project.id,
"_VSN_normalised_data.txt"), sep="\t", quote=F, row.names=T)
      return(sig.norm)
}
```

4.1.3. Exploratory_QC_plot_functions_v1.3.R

```
## define functions to make QC plots

##density plots

make.density.plot <- function(data, xlim.min, xlim.max,
ylim.max, title, col.vec, groups, leg)
{
      plot(density(data[,1]), xlim=c(xlim.min, xlim.max),
ylim=c(0,ylim.max), main=paste0("Distribution of probe
intensities: ", title), xlab="log2 normalised intensity",
col=col.vec[1])
      for(i in 2:dim(data)[2]){
      lines(density(data[,i]), col=col.vec[i])}
      ## add legend if multiple groups
      print(length(unique(col.vec)))
      if(leg == "T")
      {
            legend("topright", legend=unique(groups),
col=unique(col.vec), lty=1, cex=0.5, ncol = 2)
      }
}


make.boxplot <- function(data, title, labels, col.vec, mag)
{
      ## modify margins for legibility of sample names
      par(mar = c(8, 4, 3, 1)) ## default par(mar = c(5, 4, 4,
2)+ 0.1)
      boxplot(as.data.frame(data), outline=F, names=labels,
col=col.vec, las=2, cex.axis=mag, cex.names=mag,
main=paste0("Boxplots of probe intensities: ", title),
ylab="Log2 normalised intensity")
}


##dendogram plots
calc.dist <- function(data)
{
      d.tnorm <- dist(t(data)) #transpose matrix because
distance is computed by row
      return(d.tnorm)
}

make.cluster.plot <- function(dists, title, labels, mag)
{
      plot(hclust(dists, method="complete"),
main=paste0("Hierarchical clustering: ", title),
labels=labels, cex=mag)
```

```
}


## more efficient functions for PCA plots

calc.pca <- function(data)
{
      pca.norm <- prcomp(t(data))
      PC.norm <- pca.norm$x
      PC.norm <- data.frame(PC.norm)
      attach(PC.norm)

      var.cap <- summary(pca.norm)
      pc_var.cap <- round(var.cap$importance[2,]*100, 1)
      cum_var.cap <- round(var.cap$importance[3,]*100, 1)

      out <- list("PC"=PC.norm, "PC1_var"=pc_var.cap[1],
"PC2_var"=pc_var.cap[2])
      return(out)
}


pca.ids <- function(PC.norm, pc1, pc2, title, labels, mag)
{
      ## identify samples by name
      plot(PC.norm$PC1, PC.norm$PC2, type="n",
main=paste0("PCA plot: ", title), cex.main=1, xlab=paste("PC1
(",pc1, sep="","%)"), ylab=paste("PC2 (",pc2, sep="","%)"))
      text(PC.norm$PC1, PC.norm$PC2, labels, cex=mag)
}


pca.colours <- function(PC.norm, pc1, pc2, title, col.vec,
factor, leg, leg.pos, leg.mag)
{
      ## colour by factor
      plot(PC.norm$PC1, PC.norm$PC2, type="p",
main=paste0("PCA plot: ", title), cex.main=1, xlab=paste("PC1
(",pc1, sep="","%)"), ylab=paste("PC2 (",pc2, sep="","%)"),
pch=19, col=col.vec)
      if(leg == "T")
  {
    legend(paste0(leg.pos), legend=unique(factor), pch=19,
col=unique(col.vec), cex=leg.mag)
      }
}
```

4.1.4. Analysis of the first batch of samples

```
project.id <- "P150778"
today <- "20160403"


d.norm <- read.delim("P150778_VSN_filtered_data.txt",
header=T, sep="\t")
dim(d.norm) # 22100    24

head(d.norm)
colnames(d.norm)


####################################

# Make sorted table with average expression and annotation
d.norm.ave <- cbind(d.norm, rowMeans(d.norm))
col <- colnames(d.norm)
col <- c(col, "average")
colnames(d.norm.ave) <- col

d.norm.sort <- d.norm.ave[order(-d.norm.ave[,25]),]
head(d.norm.sort)

source("http://bioconductor.org/biocLite.R")
biocLite("illuminaHumanv4.db")
source("./Annotation/Annotate_HT12v4_topTableF.R")

annot <- read.delim("./Annotation/
HumanHT-12_V4_0_R1_15002873_B.txt", sep="\t", skip=8)
res.annot <- annotate.results(d.norm.sort, annot, out
=paste0(project.id,"_AveExpr_annotated.txt"))


####################################
# continue with analysis

s.info <- read.table("./DATA/SampleTableControl_Matilde.txt",
sep="\t", header=T)
s.info[,1:5]
# columns:  Index Sample.ID Sample.Group Sentrix.Barcode
Sample.Section

identical(as.vector(s.info$Sample.ID), gsub("X", "",
colnames(d.norm))) # TRUE

## compare groups for diff exp
library(limma)
```

```
group <- factor(s.info$Sample.Group, levels=c("A", "B"))

## calculate the array quality weights - arrays that are
different to the majority will be assigned lower weights and
the data from those arrays will have less influence on the
analysis. Idea is to be able to retain borderline outlier
samples without them having a negative impact on the ability
to find differentially expressed genes

arrayw <- arrayWeights(d.norm)
barplot(arrayw, xlab="Array", ylab="Weight", col="white",
las=2)
abline(h=1, lwd=1, lty=2)

design <- model.matrix(~group)
design
#  (Intercept) groupB


## fit the model to the VSN data, using the array weights
fit.w <- lmFit(d.norm, design, weights=arrayw)

fit.w2 <- eBayes(fit.w)

res.w <- topTable(fit.w2, coef="groupB", sort.by="P",
adjust.method="BH", number=length(d.norm[,1]))
res.wfc <- topTable(fit.w2, coef="groupB", sort.by="logFC",
adjust.method="BH", number=length(d.norm[,1]))
head(res.w)

length(which(res.w$adj.P.Val<0.05)) # 0
length(which(res.w$adj.P.Val<0.1)) # 0

write.table(res.w, paste0(project.id,
"_TopTable_weighted_data.txt"), sep="\t", quote=F,
row.names=T)

## without weights

fit <- lmFit(d.norm, design)
fit2 <- eBayes(fit)

res <- topTable(fit2, coef="groupB", sort.by="P",
adjust='fdr', number=length(d.norm[,1]))
res.fc <- topTable(fit2, coef="groupB", sort.by="logFC",
adjust.method="BH", number=length(d.norm[,1]))
length(which(res$adj.P.Val<0.05)) # 0 but Pvalues are better

write.table(res, paste0(project.id, "_TopTable.txt"),
sep="\t", quote=F, row.names=T)
```

```
## get a larger list of sig genes without the array weights
###########################################################

## overlap of the lists
# length(intersect(row.names(res.w)[1:304], row.names(res)
[1:250]))

## using all samples and no array weights gives the most sig
topTable
## annotate and write to file
source("http://bioconductor.org/biocLite.R")
biocLite("illuminaHumanv4.db")
source("./ma_150778/Annotation/Annotate_HT12v4_topTableF.R")

annot <- read.delim("/Users/user/Desktop/ma_150778/Annotation/
HumanHT-12_V4_0_R1_15002873_B.txt", sep="\t", skip=8)
res.annot <- annotate.results(res, annot, out
=paste0(project.id,"_TopTable_results_annotated.txt"))

###########################################################
## make boxplots of top genes
pdf("Boxplots_top10_genes_unweighted.pdf", onefile=T)
for(i in 1:10)
{
     plot.data <- d.norm[match(row.names(res)[i],
row.names(d.norm)),]
     plot.data <- cbind(t(plot.data), group)
     boxplot(get(colnames(plot.data)[1]) ~ group,
data=plot.data, names=levels(group),
main=paste0(colnames(plot.data)[1], ": ", res.annot
$Symbol[i]), sub="All samples", ylab="Log2 normalised
intensity")
}
dev.off()


################################
## SAMPLE SIZE CALCULATION

library(ssize)

st.dev <- apply(d.norm, 1, sd)
hist(st.dev, breaks=50, xlim=c(0, 1.5))
power <- 0.8
alpha <- 0.05
allsize2 <- ssize(sd=st.dev, delta=log2(2), sig.level = alpha,
power = power)
allsize1.5 <- ssize(sd=st.dev, delta=log2(1.5), sig.level =
alpha, power = power)
allsize1.25 <- ssize(sd=st.dev, delta=log2(1.25), sig.level =
alpha, power = power)
```

```
pdf("Sample_size1.25_20170416.pdf")
par(mfcol=c(1, 1))
ssize.plot(allsize1.25, xlim=c(1,100), lwd=2, col="blue",
marks=c(12, 24, 36, 48, 60, 72, 84, 96))
legend("bottomright", legend=strsplit( x=paste("fold
change=1.25",", alpha=", alpha, ", power=",power,", # genes=",
nobs(st.dev), sep=" " ), split=",")[[1]])
title(paste0("Sample Size to Detect 1.25-Fold Change"))
dev.off()

pdf("Sample_size_20170416.pdf")
par(mfcol=c(3, 1))
ssize.plot(allsize1.25, xlim=c(1,40), lwd=2, col="blue",
marks=c(12, 24, 36))
legend("bottomright", legend=strsplit( x=paste("fold
change=1.25",", alpha=", alpha, ", power=",power,", # genes=",
nobs(st.dev), sep=" " ), split=",")[[1]])
title(paste0("Sample Size to Detect 1.25-Fold Change"))

ssize.plot(allsize1.5, xlim=c(1,40), lwd=2, col="blue",
marks=c(12, 24, 36))
legend("bottomright", legend=strsplit( x=paste("fold
change=1.5",", alpha=", alpha, ", power=",power,", # genes=",
nobs(st.dev), sep=" " ), split=",")[[1]])
title(paste0("Sample Size to Detect 1.5-Fold Change"))

ssize.plot(allsize2, xlim=c(1,40), lwd=2, col="blue",
marks=c(12, 24, 36))
legend("bottomright", legend=strsplit( x=paste("fold
change=2",", alpha=", alpha, ", power=",power,", # genes=",
nobs(st.dev), sep=" " ), split=",")[[1]])
title(paste0("Sample Size to Detect 2-Fold Change"))
dev.off()

################################
# VOLCANO PLOT

res.fc <- read.delim("./
P150778_TopTable_results_annotated_20160310.txt", header=TRUE)

for (i in 1:22100) {
  ifelse (test=((abs(res.fc$logFC[i]) > log2(1.5)) & (res.fc
$P.Value[i] < 0.05)), res.fc$color[i] <- "red", res.fc
$color[i] <- "black")
}

pdf(paste0(project.id, "_volcano_", today, ".pdf"))
plot(res.fc$logFC, -log10(res.fc$P.Value), pch=19, col=res.fc
$color,
     xlab="log2(Fold_Change)", ylab="-log10(p-value)")
```

```
abline(v=c(log2(1.5), -log2(1.5)), col = "red")
dev.off()

##############################
# DRAW HEATMAP

# d.norm <- read.delim("P150778_VSN_filtered_data.txt",
header=T, sep="\t")
head(res.fc)
head(d.norm)

keep <- res.fc[1,]
count <- 0

for (i in 1:22100) {
  if ((abs(res.fc$logFC[i]) > log2(1.5)) | (res.fc$P.Value[i]
< 0.05)) {count <- count+1; keep[count,] <- res.fc[i,]}
}

keep$Symbol <- as.character(keep$Symbol)
keep$Symbol[5] <- "IDH1_b"
keep$Symbol[41] <- "GTDC1_b"
keep <- keep[-22,]
keep <- keep[-48,]
keep <- keep[-50,]
keep$Symbol[50]
pull <- match(keep$Probe_Id, rownames(d.norm))
for.hm <- d.norm[pull,]
for.hm50 <- for.hm[1:50,]

rownames(for.hm50) <- keep$Symbol[1:50]
rownames(for.hm50)
head(for.hm50)

install.packages("gplots")
library(grDevices)
library("gplots")
group <- as.data.frame(group)
group$col <- c(rep("red", 12), rep("blue", 12))

pdf(paste0(project.id, "_heatmap30_", today, ".pdf"))
heatmap(as.matrix(for.hm50[1:30,]), col=redgreen(200),
ColSideColors = group$col, main="Heatmap of 30 most
significant mRNAs")
dev.off()
```

4.1.5. Quality control and analysis of second batch of samples

```
## Affymetrix Human ST 1.2 array
## pipeline from here https://support.bioconductor.org/p/
59396/

project.id <- "P170168"

setwd("Desktop/Script/") # location of working directory
setwd("/Volumes/Matilde_blue/Seq_array_study/2017/array/")
#hard drive on mac

library(oligo)
library(hugene21sttranscriptcluster.db)
library(limma)


## Read CEL files with oligo package, and use RMA to normalise
them.
## 24 samples (n=12 poAF and n=20 poSR)
## extract raw probe-level data, perform pre-processing,
normalisation and summarise to probe-set level intensities
(using RMA algorithms)

setwd("/Volumes/Matilde_blue/Seq_array_study/2017/array/
CEL_files/")
rawData <- read.celfiles(list.celfiles("."))
setwd("/Volumes/Matilde_blue/Seq_array_study/2017/array/")

eset <- rma(rawData, target="core")

## annotate transcripts
annotateGene <- function ( db , what , missing ) {
  tab = toTable(db[intersect(featureNames(eset),
mappedkeys(db))])
  mt = match ( featureNames ( eset ) , tab$probe_id )
  ifelse ( is.na(mt), missing , tab[[what]][mt])
}

fData(eset)$symbol <-
annotateGene(hugene21sttranscriptclusterSYMBOL, "symbol",
missing = NA)
fData(eset)$genename <-
annotateGene(hugene21sttranscriptclusterGENENAME , "gene_name"
, missing = NA)
fData(eset)$ensembl <-
annotateGene(hugene21sttranscriptclusterENSEMBL,
"ensembl_id" , missing = NA )

length(fData(eset)$ensembl) # 53617
```

```
sum(is.na(fData(eset)$ensembl)) # 30952 (22665 transcripts
identified in ensembl)

d.norm <- as.data.frame(exprs(eset))
dim(d.norm) # 53617    24
annot.table <- data.frame(cbind(fData(eset)$symbol,
fData(eset)$genename, fData(eset)$ensembl, rownames(d.norm)),
stringsAsFactors = F)
colnames(annot.table) <- c("symbol", "genename", "ensembl",
"probeID")

discard <- which(is.na(fData(eset)$ensembl))
d.norm <- d.norm[ -discard, ]
dim(d.norm) # 22665    24
head(d.norm)

d.norm <- cbind (d.norm, annot.table$ensembl[-discard])
colnames(d.norm)[25] <- "ensembl"
d.norm <- aggregate(. ~ ensembl, d.norm, sum) # add together
duplicate IDs and remove duplicated rows
rownames(d.norm) <- d.norm$ensembl
d.norm <- d.norm[ , -1]
d.norm <- as.matrix(d.norm)

write.table(d.norm, paste0(project.id,
"_RMA_normalised_data_", Sys.Date(), ".txt"), sep="\t",
quote=F, row.names=T)
# rownames are read as 1st column, but there is no column
name, therefore the colnames are shifted by one column

# create variable for groups
info <- read.delim("info_table.txt", header = T, sep = "\t")
groups <- as.factor(info$group)

library(RColorBrewer)
#display.brewer.all()
col.pal <- c(brewer.pal(9, "Set1"), brewer.pal(12, "Set3"),
brewer.pal(3, "Pastel1"))
palette(col.pal)


## make QC plots
sample.ids <- colnames(d.norm)

source("./Exploratory_QC_plot_functions_v1.3.R")
library(vsn)

pdf(paste0(project.id, "_QC_plots_filtered_data_", Sys.Date(),
".pdf"), onefile=T)
```

```
meanSdPlot(d.norm, main="Mean-variance plot (RMA normalised
data)")

par(mfrow=c(2,1))
make.density.plot(d.norm, title = "RMA normalised data",
xlim.min = 1, xlim.max = 15, ylim.max = 0.25, col.vec =
col.pal, groups = colnames(d.norm), leg = "T")

make.density.plot(d.norm, title = "RMA normalised data",
xlim.min = 1, xlim.max = 15, ylim.max = 0.25, col.vec =
groups, groups = colnames(d.norm), leg = "N")
legend("topright", legend=unique(groups), col=unique(groups),
lty=1, cex=0.5)

par(mfrow=c(2,1))
make.boxplot(d.norm, "RMA normalised data", sample.ids,
col.pal, 0.8)
make.boxplot(d.norm, "RMA normalised data", sample.ids,
groups, 0.8)

par(mfrow=c(1,1))
dists <- calc.dist(d.norm)
make.cluster.plot(dists, "VSN normalised data", sample.ids, 1)


pc <- calc.pca(d.norm)
pca.ids(pc$PC, pc$PC1_var, pc$PC2_var, "Samples", sample.ids,
0.8)

par(mfrow=c(2,1))
pca.colours(pc$PC, pc$PC1_var, pc$PC2_var, "Samples", col.vec
= col.pal, factor = colnames(d.norm), "topright", 0.8, leg =
"N")
pca.colours(pc$PC, pc$PC1_var, pc$PC2_var, "Groups", col.vec =
groups, factor = groups, "topright", 0.8, leg = "T")

dev.off()


###################
# keep genes with reasonable intensity (~2) in at least 12
samples
keep <- as.integer(rowSums(d.norm > 2 )) >= 12
d.norm.filt <- d.norm[keep,]

arrayw <- arrayWeights(d.norm.filt)
barplot(arrayw, xlab="Array", ylab="Weight", col="white",
las=2)
abline(h=1, lwd=1, lty=2)

design <- model.matrix(~ groups)
```

```
design
fit.w <- lmFit(d.norm, design, weights=arrayw)
fit.w2 <- eBayes(fit.w)

res.w <- topTable(fit.w2, coef="groupsB", sort.by="P",
adjust.method="BH", number=length(d.norm.filt[,1]))
res.wfc <- topTable(fit.w2, coef="groupsB", sort.by="logFC",
adjust.method="BH", number=length(d.norm.filt[,1]))
head(res.w)

length(which(res.w$P.Val<0.05)) # 1497
length(which(res.w$adj.P.Val<0.1)) # 0

write.table(res.w, paste0(project.id,
"_TopTable_weighted_data_", Sys.Date(), ".txt"), sep="\t",
quote=F, row.names=T)

order <- match(rownames(res.w), annot.table$ensembl)
res.w <- cbind(res.w, annot.table[order, ])
write.table(res.w, paste0(project.id,
"_TopTable_weighted_data_annotated_", Sys.Date(), ".txt"),
sep="\t", quote=F, row.names=T)
```

4.1.6. Combined microarray analysis

```
## Modified from GSE2240 and P15778

project.id <- "combined"

library(oligo)
library(hugene21sttranscriptcluster.db)
library(limma)


## Read normalised and filtered data sets from the two
experiments
## 48 samples (n=24 poAF and n=24 poSR)
## Will use ENSEMBL IDs to merge the data sets

data2 <-
read.table("P170188_RMA_normalised_data_2017-07-10.txt",
header = T, sep = "\t", stringsAsFactors = F)
data2 <- as.matrix(data2)

data1 <- read.table("P150778_VSN_normalised_data.txt", header
= T, sep = "\t", stringsAsFactors = F)

source("http://bioconductor.org/biocLite.R")
biocLite("illuminaHumanv4.db")

x <- illuminaHumanv4ENSEMBL
# Get the entrez gene IDs that are mapped to an Ensembl ID
mapped_genes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_genes])

info <- sapply(xx, "[[", 1)
ensembl <- data.frame(info[match(rownames(data1),
names(info))], stringsAsFactors = F) # list of ensembl IDs for
data1, in right order
sum(is.na(ensembl)) # 13847 NAs
length(unique(ensembl)[, 1]) # 20524 unique transcripts

data1 <- cbind(data1, ensembl)
colnames(data1)[25] <- "ensembl"
data1 <- aggregate(. ~ ensembl, data1, sum)
rownames(data1) <- data1$ensembl
data1 <- data1[, -1]
data1 <- as.matrix(data1)
## data1 is now ready

## MATCH IDs and merge into single table
dim(data1)  # 20523    24
```

```
dim(data2)  # 21518    24

keep <- match(rownames(data1), rownames(data2))
keep <- keep[!is.na(keep)] # all indexes of rows in data2 that
have a match in data1, excluding NAs
data2sh <- data2[keep, ]

keep <- match(rownames(data2), rownames(data1))
keep <- keep[!is.na(keep)] # all indexes of rows in data2 that
have a match in data1, excluding NAs
data1sh <- data1[keep, ]

test <- rownames(data1sh) == rownames(data2sh)
sum(test) #19069 all rownames match!


## COMBINE all data in a single matrix
data.all <- cbind(data1sh, data2sh)
dim(data.all)  # 19069    48
write.table(data.all, paste0(project.id,
"_all_normalised_data_", Sys.Date(), ".txt"), sep="\t",
quote=F, row.names=T)

# data.all <-
read.table("combined_all_normalised_data_2017-07-11.txt", sep
= "\t", stringsAsFactors = FALSE)

info <- read.csv("combined_info_table.csv", header = T)
groups <- as.factor(info$group)
set <- as.factor(info$batch)

library(RColorBrewer)
#display.brewer.all()
col.pal <- c(brewer.pal(9, "Set1"), brewer.pal(12, "Set3"),
brewer.pal(3, "Pastel1"))
palette(col.pal)


## make QC plots
sample.ids <- colnames(data.all)

source("./Exploratory_QC_plot_functions_v1.3.R")
library(vsn)

pdf(paste0(project.id, "_QC_plots_filtered_data_", Sys.Date(),
".pdf"), onefile=T)

meanSdPlot(data.all, main="Mean-variance plot (RMA normalised
data)")

par(mfrow=c(2,1))
```

```
make.density.plot(data.all, title = "RMA normalised data",
xlim.min = 1, xlim.max = 20, ylim.max = 0.25, col.vec =
col.pal, groups = colnames(data.all), leg = "T")

make.density.plot(data.all, title = "RMA normalised data",
xlim.min = 1, xlim.max = 20, ylim.max = 0.25, col.vec =
groups, groups = colnames(data.all), leg = "N")
legend("topright", legend=unique(groups), col=unique(groups),
lty=1, cex=0.5)

par(mfrow=c(2,1))
make.boxplot(data.all, "RMA normalised data", sample.ids,
col.pal, 0.8)
make.boxplot(data.all, "RMA normalised data", sample.ids,
groups, 0.8)

par(mfrow=c(1,1))
dists <- calc.dist(data.all)
make.cluster.plot(dists, "VSN normalised data", sample.ids, 1)


pc <- calc.pca(data.all)
pca.ids(pc$PC, pc$PC1_var, pc$PC2_var, "Samples", sample.ids,
0.8)

par(mfrow=c(2,1))
pca.colours(pc$PC, pc$PC1_var, pc$PC2_var, "Samples", col.vec
= col.pal, factor = colnames(data.all), "topright", 0.8, leg =
"N")
pca.colours(pc$PC, pc$PC1_var, pc$PC2_var, "Groups", col.vec =
groups, factor = groups, "topright", 0.8, leg = "T")

dev.off()


###################
arrayw <- arrayWeights(data.all)
barplot(arrayw, xlab="Array", ylab="Weight", col="white",
las=2)
abline(h=1, lwd=1, lty=2)

design <- model.matrix(~ set+groups)
design
fit.w <- lmFit(data.all, design)
fit.w2 <- eBayes(fit.w)

res.w <- topTable(fit.w2, coef="groupsB", sort.by="P",
adjust.method="BH", number=length(data.all[,1]))
res.wfc <- topTable(fit.w2, coef="groupsB", sort.by="logFC",
adjust.method="BH", number=length(data.all[,1]))
head(res.w)
```

```
write.table(res.w, paste0(project.id,
"_TopTable_weighted_data_", Sys.Date(), ".txt"), sep="\t",
quote=F, row.names=T)

length(which(res.w$adj.P.Val<0.05)) # 0
length(which(res.w$adj.P.Val<0.1)) # 0

annot.table <- read.table("annotation_table.txt", header = T,
sep = "\t", stringsAsFactors = F)
order <- match(rownames(res.w), annot.table$ensembl)
res.w <- cbind(res.w, annot.table[order, ])
write.table(res.w, paste0(project.id,
"_TopTable_weighted_data_annotated_", Sys.Date(), ".txt"),
sep="\t", quote=F, row.names=T)

write.table(res.w[1:500, ], paste0(project.id,
"_TopTable_weighted_data_annotated_500_", Sys.Date(), ".txt"),
sep="\t", quote=F, row.names=T)

res.batch <- topTable(fit.w2, coef="set2", sort.by="P",
adjust.method="BH", number=length(data.all[,1]))

## treat and topTreat are recommended to give more weight to
genes above a desired FC
fit.fc <- treat(fit.w, lfc = 0.32)
res.fc <- topTreat(fit.fc, coef="groupsB", sort.by="P",
adjust.method="BH", number=length(data.all[,1]))
write.table(res.fc, paste0(project.id,
"_TopTable_weighted_data_FC032_", Sys.Date(), ".txt"),
sep="\t", quote=F, row.names=T)

annot.table <- read.table("annotation_table.txt", header = T,
sep = "\t", stringsAsFactors = F)
order <- match(rownames(res.fc), annot.table$ensembl)
res.fc <- cbind(res.fc, annot.table[order, ])
write.table(res.fc, paste0(project.id,
"_TopTable_weighted_data_FC032_annotated_", Sys.Date(),
".txt"), sep="\t", quote=F, row.names=T)

write.table(res.w[1:500, ], paste0(project.id,
"_TopTable_weighted_data_annotated_500_", Sys.Date(), ".txt"),
sep="\t", quote=F, row.names=T)
```

4.2. SMALL RNA-SEQUENCING

4.2.1. Quality control of RNA-Seq

```
## QC plots for RNA-Seq projects

## Pipeline from P150269, then used for own samples P150775
2015, own samples 2017, and combined 2017


project.id <- "combined"

library(DESeq)


## Read in pre-processed data (from seq_combined_analysis.R),
zero counts already excluded
d <- read.csv("combined_data_sum.csv", header=T,
stringsAsFactors = FALSE)
dim(d) # 920  49
head(d)
tail(d)
rownames(d) <- d[, 1]
d <- d[, -1]


s.info <- read.csv("./combined_info_table.csv", header=T)
s.info$seq_ID == colnames(d) # check order of samples is the
same
head(s.info)
conds <- factor(s.info$group)

## remove miRs that have fewer than 48 counts in 48 samples
## cutoff 48 normalises boxplots
zero.count <- which(apply(d, 1, sum)<48)
length(zero.count) # 231 (cutoff 24: 95)
head(d[zero.count,])

final.mir.counts <- d[-zero.count,]
dim(final.mir.counts) # cutoff 48: 689  48

keep <- row.names(final.mir.counts) ## keep just the non-zero
genes


cds <- newCountDataSet(final.mir.counts, conds)
cds <- estimateSizeFactors(cds)
```

```
cds <- estimateDispersions(cds, method="blind")
vsd <- getVarianceStabilizedData(cds)
dists <- dist( t( vsd ) )

base.counts <- counts(cds, normalized=TRUE)
log.base.counts <- log2(base.counts+1)
# log.base.counts <-
log.base.counts[which(row.names(log.base.counts) %in% keep ==
TRUE),] #this is to select only the counts of the non-zero
genes, not necessary here cause selected them with
final.mir.counts


## make QC plots

## colour vector by groups, pairs
library(RColorBrewer)
col.pal <- c(brewer.pal(9, "Set1"), brewer.pal(12, "Set3"),
brewer.pal(3, "Dark2"))
palette(col.pal)
col.vec <- as.factor(s.info$group)
col.pair <- as.factor(s.info$pair)
col.sex <- as.factor(s.info$male)
col.dm <- as.factor(s.info$DM)
col.batch <- as.factor(s.info$set)


pdf(paste0(project.id, "_QC_plots_cutoff48_", Sys.Date(),
".pdf"), onefile=T)

plot(density(log.base.counts[,1]), col = col.vec,
main="Normalised count distributions", xlab="Log2 normalised
counts", ylim = c(0, 0.15))
for(i in 2:ncol(log.base.counts)){
  lines(density(log.base.counts[,i]), col=col.vec[i])
}
legend("topright", legend=unique(conds), lty=1,
col=unique(col.vec), cex=0.8)
## samples 28 (X261_s2 20706_B31) and 30 (X263_s2 20434_B47)
have anomalous curves and need to be removed


boxplot(as.matrix(log.base.counts),las=2, names=s.info
$sample_ID, cex.names=0.7, ylab="Log2 normalised counts",
main="Boxplots of normalised counts", col=col.vec)


plot(hclust(dists), labels=s.info$sample_ID,
main="Hierarchical Clustering")
```

```
heatmap( as.matrix( dists ), symm=TRUE, labRow=s.info
$sample_ID, labCol=s.info$sample_ID, cexCol=0.7, cexRow=0.7,
main="Distance matrix heatmap" )

pca.n = prcomp(t(vsd))
scores.n  = as.data.frame(pca.n$x)
attach(scores.n)

var.capn = summary(pca.n)
pc_var.capn = round(var.capn$importance[2,]*100, 1)
cum_var.capn = round(var.capn$importance[3,]*100, 1)

plot(PC1, PC2, type="n", main="PCA plot", xlab=paste("PC1
(",pc_var.capn[1], sep="","%)"), ylab=paste("PC2
(",pc_var.capn[2], sep="","%)"))
text(PC1, PC2, s.info$sample_ID, cex=0.7)

plot(PC1, PC2, type="p", main="PCA plot: Group", pch=19,
col=col.vec, xlab=paste("PC1 (",pc_var.capn[1], sep="","%)"),
ylab=paste("PC2 (",pc_var.capn[2], sep="","%)"))
legend("topleft", legend=unique(conds), pch=19,
col=unique(col.vec))

plot(PC1, PC2, type="p", main="PCA plot: Pairs", pch=19,
col=col.pair, xlab=paste("PC1 (",pc_var.capn[1], sep="","%)"),
ylab=paste("PC2 (",pc_var.capn[2], sep="","%)"))
legend("topright", legend=unique(s.info$pair), pch=19,
col=unique(col.pair))

plot(PC1, PC2, type="p", main="PCA plot: Sex (M)", pch=19,
col=as.factor(s.info$male), xlab=paste("PC1 (",pc_var.capn[1],
sep="","%)"), ylab=paste("PC2 (",pc_var.capn[2], sep="","%)"))
legend("topleft", legend=unique(s.info$male), pch=19,
col=unique(col.sex))

plot(PC1, PC2, type="p", main="PCA plot: Diabetes Mellitus",
pch=19, col=as.factor(s.info$DM), xlab=paste("PC1
(",pc_var.capn[1], sep="","%)"), ylab=paste("PC2
(",pc_var.capn[2], sep="","%)"))
legend("topleft", legend=unique(s.info$DM), pch=19,
col=unique(col.dm))

plot(PC1, PC2, type="p", main="PCA plot: Batch", pch=19,
col=as.factor(s.info$set), xlab=paste("PC1 (",pc_var.capn[1],
sep="","%)"), ylab=paste("PC2 (",pc_var.capn[2], sep="","%)"))
legend("topleft", legend=unique(s.info$set), pch=19,
col=unique(col.dm))

dev.off()
```

4.2.2. Analysis of RNA-Seq

```
## Pipeline from P150269, then used for own samples P150775
2015, own samples 2017, and combined 2017

## Remove anomalous samples identified in seq_combined_QC.R
and redo plots
## samples 28 (20706_B31 pair 12) and 30 (20434_B47 pair 8)
## Also remove their pair match to be consistent with study
design
## samples 29 (X262_s2 10713_B10) and 33 (X266_s2 20777_B24)

project.id <- "combined"

library(DESeq)


## Read in pre-processed data (from seq_combined_analysis.R),
zero counts already excluded
d <- read.csv("combined_data_sum.csv", header=T,
stringsAsFactors = FALSE)
dim(d) # 920  49
head(d)
tail(d)
rownames(d) <- d[, 1]
d <- d[, -1]

s.info.all <- read.csv("./combined_info_table.csv", header=T)
s.info.all$seq_ID == colnames(d) # check order of samples is
the same
head(s.info.all)

## remove miRs that have fewer than 48 counts in 48 samples
## cutoff 48 normalises boxplots
zero.count <- which(apply(d, 1, sum)<48)
length(zero.count) # 231 (cutoff 24: 95)
head(d[zero.count,])

final.mir.counts <- d[-zero.count,]
dim(final.mir.counts) # cutoff 48: 689  48

keep <- row.names(final.mir.counts) ## keep just the non-zero
genes


## Remove anomalous samples 28 (X261_s2 20706_B31) and 30
(X263_s2 20434_B47)
colnames(final.mir.counts)[28] # X261_s2
colnames(final.mir.counts)[30] # X263_s2
final.mir.counts <- final.mir.counts[ , -c(28, 29, 30, 33)]
```

```
s.info <- s.info.all[-c(28, 29, 30, 33), ]
s.info$seq_ID == colnames(final.mir.counts) # check order of
samples is the same
conds <- factor(s.info$group)


cds <- newCountDataSet(final.mir.counts, conds)
cds <- estimateSizeFactors(cds)
cds <- estimateDispersions(cds, method="blind")
vsd <- getVarianceStabilizedData(cds)
dists <- dist( t( vsd ) )

base.counts <- counts(cds, normalized=TRUE)
log.base.counts <- log2(base.counts+1)
# log.base.counts <-
log.base.counts[which(row.names(log.base.counts) %in% keep ==
TRUE),] #this is to select only the counts of the non-zero
genes, not necessary here cause selected them with
final.mir.counts


## make QC plots

## colour vector by groups, pairs
library(RColorBrewer)
col.pal <- c(brewer.pal(9, "Set1"), brewer.pal(12, "Set3"),
brewer.pal(3, "Dark2"))
palette(col.pal)
col.vec <- as.factor(s.info$group)
col.pair <- as.factor(s.info$pair)
col.sex <- as.factor(s.info$male)
col.dm <- as.factor(s.info$DM)
col.batch <- as.factor(s.info$set)


pdf(paste0(project.id, "_QC_plots_cutoff48_removed_pair_",
Sys.Date(), ".pdf"), onefile=T)

plot(density(log.base.counts[,1]), col = col.vec,
main="Normalised count distributions", xlab="Log2 normalised
counts", ylim = c(0, 0.15))
for(i in 2:ncol(log.base.counts)){
  lines(density(log.base.counts[,i]), col=col.vec[i])
}
legend("topright", legend=unique(conds), lty=1,
col=unique(col.vec), cex=0.8)
```

```
boxplot(as.matrix(log.base.counts),las=2, names=s.info
$sample_ID, cex.axis=0.7, ylab="Log2 normalised counts",
main="Boxplots of normalised counts", col=col.vec)


plot(hclust(dists), labels=s.info$sample_ID,
main="Hierarchical Clustering")

heatmap( as.matrix( dists ), symm=TRUE, labRow=s.info
$sample_ID, labCol=s.info$sample_ID, cexCol=0.7, cexRow=0.7,
main="Distance matrix heatmap" )

pca.n = prcomp(t(vsd))
scores.n  = as.data.frame(pca.n$x)
attach(scores.n)

var.capn = summary(pca.n)
pc_var.capn = round(var.capn$importance[2,]*100, 1)
cum_var.capn = round(var.capn$importance[3,]*100, 1)

plot(PC1, PC2, type="n", main="PCA plot", xlab=paste("PC1
(",pc_var.capn[1], sep="","%)"), ylab=paste("PC2
(",pc_var.capn[2], sep="","%)"))
text(PC1, PC2, s.info$sample_ID, cex=0.7)

plot(PC1, PC2, type="p", main="PCA plot: Group", pch=19,
col=col.vec, xlab=paste("PC1 (",pc_var.capn[1], sep="","%)"),
ylab=paste("PC2 (",pc_var.capn[2], sep="","%)"))
legend("topleft", legend=unique(conds), pch=19,
col=unique(col.vec))

plot(PC1, PC2, type="p", main="PCA plot: Pairs", pch=19,
col=col.pair, xlab=paste("PC1 (",pc_var.capn[1], sep="","%)"),
ylab=paste("PC2 (",pc_var.capn[2], sep="","%)"))
legend("topright", legend=unique(s.info$pair), pch=19,
col=unique(col.pair))

plot(PC1, PC2, type="p", main="PCA plot: Sex (M)", pch=19,
col=as.factor(s.info$male), xlab=paste("PC1 (",pc_var.capn[1],
sep="","%)"), ylab=paste("PC2 (",pc_var.capn[2], sep="","%)"))
legend("topleft", legend=unique(s.info$male), pch=19,
col=unique(col.sex))

plot(PC1, PC2, type="p", main="PCA plot: Diabetes Mellitus",
pch=19, col=as.factor(s.info$DM), xlab=paste("PC1
(",pc_var.capn[1], sep="","%)"), ylab=paste("PC2
(",pc_var.capn[2], sep="","%)"))
legend("topleft", legend=unique(s.info$DM), pch=19,
col=unique(col.dm))
```

```
plot(PC1, PC2, type="p", main="PCA plot: Batch", pch=19,
col=as.factor(s.info$set), xlab=paste("PC1 (",pc_var.capn[1],
sep="","%)"), ylab=paste("PC2 (",pc_var.capn[2], sep="","%)"))
legend("topleft", legend=unique(s.info$set), pch=19,
col=unique(col.dm))

dev.off()
```

REFERENCES

Anderson, E.J., Efird, J.T., Davies, S.W., O'Neal, W.T., Darden, T.M., Thayne, K.A., Katunga, L.A., Kindell, L.C., Ferguson, T.B., Anderson, C.A*., et al.* (2014). Monoamine oxidase is a major determinant of redox balance in human atrial myocardium and is associated with postoperative atrial fibrillation. J Am Heart Assoc *3*, e000713.

Angelini, C., Tasca, E., Nascimbeni, A.C., and Fanin, M. (2014). Muscle fatigue, nNOS and muscle fiber atrophy in limb girdle muscular dystrophy. Acta Myol *33*, 119-126.

Antoniades, C., Demosthenous, M., Reilly, S., Margaritis, M., Zhang, M.H., Antonopoulos, A., Marinou, K., Nahar, K., Jayaram, R., Tousoulis, D*., et al.* (2012). Myocardial redox state predicts in-hospital clinical outcome after cardiac surgery effects of short-term pre-operative statin treatment. J Am Coll Cardiol *59*, 60-70.

Aranki, S.F., Shaw, D.P., Adams, D.H., Rizzo, R.J., Couper, G.S., VanderVliet, M., Collins, J.J., Cohn, L.H., and Burstin, H.R. (1996). Predictors of atrial fibrillation after coronary artery surgery. Current trends and impact on hospital resources. Circulation *94*, 390-397.

Açil, T., Cölkesen, Y., Türköz, R., Sezgin, A.T., Baltali, M., Gülcan, O., Demircan, S., Yildirir, A., Ozin, B., and Müderrisoğlu, H. (2007). Value of preoperative echocardiography in the prediction of postoperative atrial fibrillation following isolated coronary artery bypass grafting. Am J Cardiol *100*, 1383-1386.

Banach, M., Goch, A., Misztal, M., Rysz, J., Jaszewski, R., and Goch, J.H. (2007). Predictors of paroxysmal atrial fibrillation in patients undergoing aortic valve replacement. J Thorac Cardiovasc Surg *134*, 1569-1576.

Barresi, R., Di Blasi, C., Negri, T., Brugnoni, R., Vitali, A., Felisari, G., Salandi, A., Daniel, S., Cornelio, F., Morandi, L*., et al.* (2000). Disruption of heart sarcoglycan complex and severe cardiomyopathy caused by beta sarcoglycan mutations. J Med Genet *37*, 102-107.

Barth, A.S., Merk, S., Arnoldi, E., Zwermann, L., Kloos, P., Gebauer, M., Steinmeyer, K., Bleich, M., Kääb, S., Hinterseer, M*., et al.* (2005). Reprogramming of the human atrial transcriptome in permanent atrial fibrillation: expression of a ventricular-like genomic signature. Circ Res *96*, 1022-1029.

Beynon, R.P., and Ray, S.G. (2008). Cardiac involvement in muscular dystrophies. QJM *101*, 337-344.

Bia, B.L., Cassidy, P.J., Young, M.E., Rafael, J.A., Leighton, B., Davies, K.E., Radda, G.K., and Clarke, K. (1999). Decreased myocardial nNOS, increased iNOS and abnormal ECGs in mouse models of Duchenne muscular dystrophy. J Mol Cell Cardiol *31*, 1857-1862.

Björnberg, J., Albert, U., and Mellander, S. (1990). Resistance responses in proximal arterial vessels, arterioles and veins during reactive hyperaemia in skeletal muscle and their underlying regulatory mechanisms. Acta Physiol Scand *139*, 535-550.

Blake, D.J. (2002). Dystrobrevin dynamics in muscle-cell signalling: a possible target for therapeutic intervention in Duchenne muscular dystrophy? Neuromuscul Disord *12 Suppl 1*, S110-117.

Bostick, B., Yue, Y., Long, C., and Duan, D. (2008). Prevention of dystrophin-deficient cardiomyopathy in twenty-one-month-old carrier mice by mosaic dystrophin expression or complementary dystrophin/utrophin expression. Circulation research *102*, 121-130.

Brand, T., and Schneider, M.D. (1996). Transforming growth factor-beta signal transduction. Circ Res *78*, 173-179.

Brenman, J.E., Chao, D.S., Xia, H., Aldape, K., and Bredt, D.S. (1995). Nitric oxide synthase complexed with dystrophin and absent from skeletal muscle sarcolemma in Duchenne muscular dystrophy. Cell *82*, 743-752.

Brock, R.W., Tschakovsky, M.E., Shoemaker, J.K., Halliwill, J.R., Joyner, M.J., and Hughson, R.L. (1998). Effects of acetylcholine and nitric oxide on forearm blood flow at rest and after a single muscle contraction. J Appl Physiol (1985) *85*, 2249-2254.

Brundel, B.J., Ausma, J., van Gelder, I.C., Van der Want, J.J., van Gilst, W.H., Crijns, H.J., and Henning, R.H. (2002). Activation of proteolysis by calpains and structural changes in human paroxysmal and persistent atrial fibrillation. Cardiovasc Res *54*, 380-389.

Brundel, B.J., Van Gelder, I.C., Henning, R.H., Tieleman, R.G., Tuinenburg, A.E., Wietses, M., Grandjean, J.G., Van Gilst, W.H., and Crijns, H.J. (2001). Ion channel remodeling is related to intraoperative atrial effective refractory periods in patients with paroxysmal and persistent atrial fibrillation. Circulation *103*, 684-690.

Brussee, V., Tardif, F., and Tremblay, J.P. (1997). Muscle fibers of mdx mice are more vulnerable to exercise than those of normal mice. Neuromuscul Disord *7*, 487-492.

Bukowska, A., Lendeckel, U., Hirte, D., Wolke, C., Striggow, F., Röhnert, P., Huth, C., Klein, H.U., and Goette, A. (2006). Activation of the calcineurin signaling pathway induces atrial hypertrophy during atrial fibrillation. Cell Mol Life Sci *63*, 333-342.

Bulfield, G., Siller, W.G., Wight, P.A., and Moore, K.J. (1984). X chromosome-linked muscular dystrophy (mdx) in the mouse. Proc Natl Acad Sci U S A *81*, 1189-1192.

Bushby, K., Finkel, R., Birnkrant, D.J., Case, L.E., Clemens, P.R., Cripe, L., Kaul, A., Kinnett, K., McDonald, C., Pandya, S.*, et al.* (2010). Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. Lancet Neurol *9*, 77-93.

Butler, M.H., Douville, K., Murnane, A.A., Kramarcy, N.R., Cohen, J.B., Sealock, R., and Froehner, S.C. (1992). Association of the Mr 58,000 postsynaptic protein of electric tissue with Torpedo dystrophin and the Mr 87,000 postsynaptic protein. J Biol Chem *267*, 6213-6218.

Cacchiarelli, D., Incitti, T., Martone, J., Cesana, M., Cazzella, V., Santini, T., Sthandier, O., and Bozzoni, I. (2011). miR-31 modulates dystrophin expression: new implications for Duchenne muscular dystrophy therapy. EMBO Rep *12*, 136-141.

Cai, G.L., Chen, J., Hu, C.B., Yan, M.L., Xu, Q.H., and Yan, J. (2014). Value of plasma brain natriuretic peptide levels for predicting postoperative atrial fibrillation: a systemic review and meta-analysis. World J Surg *38*, 51-59.

Camerino, G.M., Cannone, M., Giustino, A., Massari, A.M., Capogrosso, R.F., Cozzoli, A., and De Luca, A. (2014). Gene expression in mdx mouse muscle in relation to age and

exercise: aberrant mechanical-metabolic coupling and implications for pre-clinical studies in Duchenne muscular dystrophy. Hum Mol Genet *23*, 5720-5732.

Campbell, K.P., and Kahl, S.D. (1989). Association of dystrophin and an integral membrane glycoprotein. Nature *338*, 259-262.

Carlsson, J., Miketic, S., Windeler, J., Cuneo, A., Haun, S., Micus, S., Walter, S., Tebbe, U., and Investigators, S. (2003). Randomized trial of rate-control versus rhythm-control in persistent atrial fibrillation: the Strategies of Treatment of Atrial Fibrillation (STAF) study. J Am Coll Cardiol *41*, 1690-1696.

Carnes, C.A., Janssen, P.M., Ruehr, M.L., Nakayama, H., Nakayama, T., Haase, H., Bauer, J.A., Chung, M.K., Fearon, I.M., Gillinov, A.M.*, et al.* (2007). Atrial glutathione content, calcium current, and contractility. J Biol Chem *282*, 28063-28073.

Carnicer, R., Crabtree, M.J., Sivakumaran, V., Casadei, B., and Kass, D.A. (2013). Nitric oxide synthases in heart failure. Antioxid Redox Signal *18*, 1078-1099.

Chamberlain, J.S., Metzger, J., Reyes, M., Townsend, D., and Faulkner, J.A. (2007). Dystrophin-deficient mdx mice display a reduced life span and are susceptible to spontaneous rhabdomyosarcoma. FASEB J *21*, 2195-2204.

Chandrasekharan, K., Yoon, J.H., Xu, Y., deVries, S., Camboni, M., Janssen, P.M., Varki, A., and Martin, P.T. (2010). A human-specific deletion in mouse Cmah increases disease severity in the mdx model of Duchenne muscular dystrophy. Science translational medicine *2*, 42ra54.

Chao, D.S., Silvagno, F., and Bredt, D.S. (1998). Muscular dystrophy in mdx mice despite lack of neuronal nitric oxide synthase. J Neurochem *71*, 784-789.

Chiang, D.Y., Zhang, M., Voigt, N., Alsina, K.M., Jakob, H., Martin, J.F., Dobrev, D., Wehrens, X.H., and Li, N. (2015). Identification of microRNA-mRNA dysregulations in paroxysmal atrial fibrillation. Int J Cardiol *184*, 190-197.

Chu, V., Otero, J.M., Lopez, O., Sullivan, M.F., Morgan, J.P., Amende, I., and Hampton, T.G. (2002). Electrocardiographic findings in mdx mice: a cardiac phenotype of Duchenne muscular dystrophy. Muscle Nerve *26*, 513-519.

Church, J.E., Gehrig, S.M., Chee, A., Naim, T., Trieu, J., McConell, G.K., and Lynch, G.S. (2011). Early functional muscle regeneration after myotoxic injury in mice is unaffected by nNOS absence. Am J Physiol Regul Integr Comp Physiol *301*, R1358-1366.

Collins, J.M., Despa, F., and Lee, R.C. (2007). Structural and functional recovery of electropermeabilized skeletal muscle in-vivo after treatment with surfactant poloxamer 188. Biochim Biophys Acta *1768*, 1238-1246.

Colussi, C., Berni, R., Rosati, J., Straino, S., Vitale, S., Spallotta, F., Baruffi, S., Bocchi, L., Delucchi, F., Rossi, S.*, et al.* (2010). The histone deacetylase inhibitor suberoylanilide hydroxamic acid reduces cardiac arrhythmias in dystrophic mice. Cardiovascular research *87*, 73-82.

Colussi, C., Mozzetta, C., Gurtner, A., Illi, B., Rosati, J., Straino, S., Ragone, G., Pescatori, M., Zaccagnini, G., Antonini, A.*, et al.* (2008). HDAC2 blockade by nitric oxide and histone deacetylase inhibitors reveals a common target in Duchenne muscular dystrophy treatment. Proc Natl Acad Sci U S A *105*, 19183-19187.

Cooley, N., Cowley, M.J., Lin, R.C., Marasco, S., Wong, C., Kaye, D.M., Dart, A.M., and Woodcock, E.A. (2012). Influence of atrial fibrillation on microRNA expression profiles in left and right atria from patients with valvular heart disease. Physiol Genomics *44*, 211-219.

Crabtree, G.R., and Olson, E.N. (2002). NFAT signaling: choreographing the social lives of cells. Cell *109 Suppl*, S67-79.

Crosbie, R.H., Straub, V., Yun, H.Y., Lee, J.C., Rafael, J.A., Chamberlain, J.S., Dawson, V.L., Dawson, T.M., and Campbell, K.P. (1998). mdx muscle pathology is independent of nNOS perturbation. Hum Mol Genet *7*, 823-829.

D'Angelo, M.G., Gandossini, S., Martinelli Boneschi, F., Sciorati, C., Bonato, S., Brighina, E., Comi, G.P., Turconi, A.C., Magri, F., Stefanoni, G.*, et al.* (2012). Nitric oxide donor and non steroidal anti inflammatory drugs as a therapy for muscular dystrophies: evidence from a safety study with pilot efficacy measures in adult dystrophic patients. Pharmacol Res *65*, 472-479.

Danelich, I.M., Lose, J.M., Wright, S.S., Asirvatham, S.J., Ballinger, B.A., Larson, D.W., and Lovely, J.K. (2014). Practical management of postoperative atrial fibrillation after noncardiac surgery. J Am Coll Surg *219*, 831-841.

de Bono, J.P., Warrick, N., Bendall, J.K., Channon, K.M., and Alp, N.J. (2007). Radiochemical HPLC detection of arginine metabolism: measurement of nitric oxide synthesis and arginase activity in vascular tissue. Nitric Oxide *16*, 1-9.

De Palma, C., and Clementi, E. (2012). Nitric oxide in myogenesis and therapeutic muscle repair. Mol Neurobiol *46*, 682-692.

Deshmukh, A., Barnard, J., Sun, H., Newton, D., Castel, L., Pettersson, G., Johnston, D., Roselli, E., Gillinov, A.M., McCurry, K.*, et al.* (2015). Left atrial transcriptional changes associated with atrial fibrillation susceptibility and persistence. Circ Arrhythm Electrophysiol *8*, 32-41.

Dinçer, P., Balci, B., Yuva, Y., Talim, B., Brockington, M., Dinçel, D., Torelli, S., Brown, S., Kale, G., Haliloglu, G.*, et al.* (2003). A novel form of recessive limb girdle muscular dystrophy with mental retardation and abnormal expression of alpha-dystroglycan. Neuromuscul Disord *13*, 771-778.

Divet, A., and Huchet-Cadiou, C. (2002). Sarcoplasmic reticulum function in slow- and fast-twitch skeletal muscles from mdx mice. Pflugers Arch *444*, 634-643.

Dzeshka, M.S., Lip, G.Y., Snezhitskiy, V., and Shantsila, E. (2015). Cardiac Fibrosis in Patients With Atrial Fibrillation: Mechanisms and Clinical Implications. J Am Coll Cardiol *66*, 943-959.

Fairclough, R.J., Wood, M.J., and Davies, K.E. (2013). Therapy for Duchenne muscular dystrophy: renewed optimism from genetic approaches. Nat Rev Genet *14*, 373-378.

Fauconnier, J., Thireau, J., Reiken, S., Cassan, C., Richard, S., Matecki, S., Marks, A.R., and Lacampagne, A. (2010). Leaky RyR2 trigger ventricular arrhythmias in Duchenne muscular dystrophy. Proc Natl Acad Sci U S A *107*, 1559-1564.

Finsterer, J., and Cripe, L. (2014). Treatment of dystrophin cardiomyopathies. Nat Rev Cardiol *11*, 168-179.

Finsterer, J., and Stöllberger, C. (2003). The heart in human dystrophinopathies. Cardiology *99*, 1-19.

Finsterer, J., and Stöllberger, C. (2008). Atrial fibrillation/flutter in myopathies. Int J Cardiol *128*, 304-310.

Florian, A., Rösch, S., Bietenbeck, M., Engelen, M., Stypmann, J., Waltenberger, J., Sechtem, U., and Yilmaz, A. (2016). Cardiac involvement in female Duchenne and Becker muscular dystrophy carriers in comparison to their first-degree male relatives: a comparative cardiovascular magnetic resonance study. Eur Heart J Cardiovasc Imaging *17*, 326-333.

García-Cardeña, G., Martasek, P., Masters, B.S., Skidd, P.M., Couet, J., Li, S., Lisanti, M.P., and Sessa, W.C. (1997). Dissecting the interaction between nitric oxide synthase (NOS) and caveolin. Functional significance of the nos caveolin binding domain in vivo. J Biol Chem *272*, 25437-25440.

Giliberto, F., Radic, C.P., Luce, L., Ferreiro, V., de Brasi, C., and Szijan, I. (2014). Symptomatic female carriers of Duchenne muscular dystrophy (DMD): genetic and clinical characterization. J Neurol Sci *336*, 36-41.

Gillard, E.F., Chamberlain, J.S., Murphy, E.G., Duff, C.L., Smith, B., Burghes, A.H., Thompson, M.W., Sutherland, J., Oss, I., and Bodrug, S.E. (1989). Molecular and phenotypic analysis of patients with deletions within the deletion-rich region of the Duchenne muscular dystrophy (DMD) gene. Am J Hum Genet *45*, 507-520.

Godfrey, C., Clement, E., Mein, R., Brockington, M., Smith, J., Talim, B., Straub, V., Robb, S., Quinlivan, R., Feng, L.*, et al.* (2007). Refining genotype phenotype correlations in muscular dystrophies with defective glycosylation of dystroglycan. Brain *130*, 2725-2735.

Goette, A., Arndt, M., Röcken, C., Staack, T., Bechtloff, R., Reinhold, D., Huth, C., Ansorge, S., Klein, H.U., and Lendeckel, U. (2002). Calpains and cytokines in fibrillating human atria. Am J Physiol Heart Circ Physiol *283*, H264-272.

Goette, A., Staack, T., Röcken, C., Arndt, M., Geller, J.C., Huth, C., Ansorge, S., Klein, H.U., and Lendeckel, U. (2000). Increased expression of extracellular signal-regulated kinase and angiotensin-converting enzyme in human atria during atrial fibrillation. J Am Coll Cardiol *35*, 1669-1677.

Gonzalez, D.R., Treuer, A.V., Lamirault, G., Mayo, V., Cao, Y., Dulce, R.A., and Hare, J.M. (2014). NADPH oxidase-2 inhibition restores contractility and intracellular calcium handling and reduces arrhythmogenicity in dystrophic cardiomyopathy. Am J Physiol Heart Circ Physiol *307*, H710-721.

Gordon, B.S., Lowe, D.A., and Kostek, M.C. (2014). Exercise increases utrophin protein expression in the mdx mouse model of Duchenne muscular dystrophy. Muscle Nerve *49*, 915-918.

Grady, R.M., Grange, R.W., Lau, K.S., Maimone, M.M., Nichol, M.C., Stull, J.T., and Sanes, J.R. (1999). Role for alpha-dystrobrevin in the pathogenesis of dystrophin-dependent muscular dystrophies. Nat Cell Biol *1*, 215-220.

Grady, R.M., Merlie, J.P., and Sanes, J.R. (1997a). Subtle neuromuscular defects in utrophin-deficient mice. J Cell Biol *136*, 871-882.

Grady, R.M., Teng, H., Nichol, M.C., Cunningham, J.C., Wilkinson, R.S., and Sanes, J.R. (1997b). Skeletal and cardiac myopathies in mice lacking utrophin and dystrophin: a model for Duchenne muscular dystrophy. Cell *90*, 729-738.

Granchelli, J.A., Pollina, C., and Hudecki, M.S. (2000). Pre-clinical screening of drugs using the mdx mouse. Neuromuscul Disord *10*, 235-239.

Grange, R.W., Isotani, E., Lau, K.S., Kamm, K.E., Huang, P.L., and Stull, J.T. (2001). Nitric oxide contributes to vascular smooth muscle relaxation in contracting fast-twitch muscles. Physiol Genomics *5*, 35-44.

Guo, Y., Zhao, S., Li, C.I., Sheng, Q., and Shyr, Y. (2014). RNAseqPS: A Web Tool for Estimating Sample Size and Power for RNAseq Experiment. Cancer Inform *13*, 1-5.

Hack, A.A., Cordier, L., Shoturma, D.I., Lam, M.Y., Sweeney, H.L., and McNally, E.M. (1999). Muscle degeneration without mechanical injury in sarcoglycan deficiency. Proc Natl Acad Sci U S A *96*, 10723-10728.

Hack, A.A., Ly, C.T., Jiang, F., Clendenin, C.J., Sigrist, K.S., Wollmann, R.L., and McNally, E.M. (1998). Gamma-sarcoglycan deficiency leads to muscle membrane defects and apoptosis independent of dystrophin. J Cell Biol *142*, 1279-1287.

Haffajee, J.A., Lee, Y., Alsheikh-Ali, A.A., Kuvin, J.T., Pandian, N.G., and Patel, A.R. (2011). Pre-operative left atrial mechanical function predicts risk of atrial fibrillation following cardiac surgery. JACC Cardiovasc Imaging *4*, 833-840.

Hara, H., Nolan, P.M., Scott, M.O., Bucan, M., Wakayama, Y., and Fischbeck, K.H. (2002). Running endurance abnormality in mdx mice. Muscle Nerve *25*, 207-211.

Hara, Y., Balci-Hayta, B., Yoshida-Moriguchi, T., Kanagawa, M., Beltrán-Valero de Bernabé, D., Gündeşli, H., Willer, T., Satz, J.S., Crawford, R.W., Burden, S.J.*, et al.* (2011). A dystroglycan mutation associated with limb-girdle muscular dystrophy. N Engl J Med *364*, 939-946.

Hayashi, Y.K., Ogawa, M., Tagawa, K., Noguchi, S., Ishihara, T., Nonaka, I., and Arahata, K. (2001). Selective deficiency of alpha-dystroglycan in Fukuyama-type congenital muscular dystrophy. Neurology *57*, 115-121.

Hickner, R.C., Fisher, J.S., Ehsani, A.A., and Kohrt, W.M. (1997). Role of nitric oxide in skeletal muscle blood flow at rest and during dynamic exercise in humans. Am J Physiol *273*, H405-410.

Hohnloser, S.H., Kuck, K.H., and Lilienthal, J. (2000). Rhythm or rate control in atrial fibrillation--Pharmacological Intervention in Atrial Fibrillation (PIAF): a randomised trial. Lancet *356*, 1789-1794.

Hosaka, Y., Yokota, T., Miyagoe-Suzuki, Y., Yuasa, K., Imamura, M., Matsuda, R., Ikemoto, T., Kameya, S., and Takeda, S. (2002). Alpha1-syntrophin-deficient skeletal muscle exhibits hypertrophy and aberrant formation of neuromuscular junctions during regeneration. J Cell Biol *158*, 1097-1107.

Hussain, S.N. (1998). Activity of nitric oxide synthase in the ventilatory muscle vasculature. Comp Biochem Physiol A Mol Integr Physiol *119*, 191-201.

Ibraghimov-Beskrovnaya, O., Milatovich, A., Ozcelik, T., Yang, B., Koepnick, K., Francke, U., and Campbell, K.P. (1993). Human dystroglycan: skeletal muscle cDNA, genomic

structure, origin of tissue specific isoforms and chromosomal localization. Hum Mol Genet *2*, 1651-1657.

Imbornoni, L., Price, E.T., Andrews, J., Meaney, F.J., Ciafaloni, E., and Cunniff, C. (2014). Diagnostic and clinical characteristics of early-manifesting females with Duchenne or Becker muscular dystrophy. Am J Med Genet A *164A*, 2769-2774.

Iskesen, I., Eserdag, M., Kurdal, A.T., Cerrahoglu, M., and Sirin, B.H. (2011). Preoperative NT-proBNP levels: a reliable parameter to estimate postoperative atrial fibrillation in coronary artery bypass patients. Thorac Cardiovasc Surg *59*, 213-216.

Ito, N., Ruegg, U.T., Kudo, A., Miyagoe-Suzuki, Y., and Takeda, S. (2013). Activation of calcium signaling through Trpv1 by nNOS and peroxynitrite as a key trigger of skeletal muscle hypertrophy. Nat Med *19*, 101-106.

Iwase, T., Takao, S., Akaike, M., Adachi, K., Sumitomo-Ueda, Y., Yagi, S., Niki, T., Kusunose, K., Tomita, N., Hirata, Y.*, et al.* (2010). Diagnostic utility of cardiac magnetic resonance for detection of cardiac involvement in female carriers of Duchenne muscular dystrophy. Heart Asia *2*, 52-55.

Johnson, E.K., Zhang, L., Adams, M.E., Phillips, A., Freitas, M.A., Froehner, S.C., Green-Church, K.B., and Montanaro, F. (2012). Proteomic analysis reveals new cardiac-specific dystrophin-associated proteins. PLoS One *7*, e43515.

Jung, C., Martins, A.S., Niggli, E., and Shirokova, N. (2008). Dystrophic cardiomyopathy: amplification of cellular damage by Ca2+ signalling and reactive oxygen species-generating pathways. Cardiovasc Res *77*, 766-773.

Kargacin, M.E., and Kargacin, G.J. (1996). The sarcoplasmic reticulum calcium pump is functionally altered in dystrophic muscle. Biochim Biophys Acta *1290*, 4-8.

Karpati, G., Carpenter, S., Morris, G.E., Davies, K.E., Guerin, C., and Holland, P. (1993). Localization and quantitation of the chromosome 6-encoded dystrophin-related protein in normal and pathological human muscle. J Neuropathol Exp Neurol *52*, 119-128.

Kaspar, R.W., Allen, H.D., and Montanaro, F. (2009). Current understanding and management of dilated cardiomyopathy in Duchenne and Becker muscular dystrophy. J Am Acad Nurse Pract *21*, 241-249.

Khairallah, M., Khairallah, R., Young, M.E., Dyck, J.R., Petrof, B.J., and Des Rosiers, C. (2007). Metabolic and signaling alterations in dystrophin-deficient hearts precede overt cardiomyopathy. Journal of molecular and cellular cardiology *43*, 119-129.

Kharlap, M.S., Timofeeva, A.V., Goryunova, L.E., Khaspekov, G.L., Dzemeshkevich, S.L., Ruskin, V.V., Akchurin, R.S., Golitsyn, S.P., and Beabealashvilli, R.S.h. (2006). Atrial appendage transcriptional profile in patients with atrial fibrillation with structural heart diseases. Ann N Y Acad Sci *1091*, 205-217.

Kim, N.H., Ahn, Y., Oh, S.K., Cho, J.K., Park, H.W., Kim, Y.S., Hong, M.H., Nam, K.I., Park, W.J., Jeong, M.H.*, et al.* (2005). Altered patterns of gene expression in response to chronic atrial fibrillation. Int Heart J *46*, 383-395.

Kim, Y.H., Lim, D.S., Lee, J.H., Shim, W.J., Ro, Y.M., Park, G.H., Becker, K.G., Cho-Chung, Y.S., and Kim, M.K. (2003). Gene expression profiling of oxidative stress on atrial fibrillation in humans. Exp Mol Med *35*, 336-349.

Kim, Y.M., Kattach, H., Ratnatunga, C., Pillai, R., Channon, K.M., and Casadei, B. (2008). Association of atrial nicotinamide adenine dinucleotide phosphate oxidase activity with the development of atrial fibrillation after cardiac surgery. J Am Coll Cardiol *51*, 68-74.

Kinoshita, T., Asai, T., Ishigaki, T., Suzuki, T., Kambara, A., and Matsubayashi, K. (2011). Preoperative heart rate variability predicts atrial fibrillation after coronary bypass grafting. Ann Thorac Surg *91*, 1176-1181.

Kirchmann, C., Kececioglu, D., Korinthenberg, R., and Dittrich, S. (2005). Echocardiographic and electrocardiographic findings of cardiomyopathy in Duchenne and Becker-Kiener muscular dystrophies. Pediatr Cardiol *26*, 66-72.

Koenig, M., Beggs, A.H., Moyer, M., Scherpf, S., Heindrich, K., Bettecken, T., Meng, G., Müller, C.R., Lindlöf, M., and Kaariainen, H. (1989). The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. Am J Hum Genet *45*, 498-506.

Koenig, X., Dysek, S., Kimbacher, S., Mike, A.K., Cervenka, R., Lukacs, P., Nagl, K., Dang, X.B., Todt, H., Bittner, R.E.*, et al.* (2011). Voltage-gated ion channel dysfunction precedes cardiomyopathy development in the dystrophic heart. PLoS One *6*, e20300.

Koenig, X., Kovar, M., Rubi, L., Mike, A.K., Lukacs, P., Gawali, V.S., Todt, H., Hilber, K., and Sandtner, W. (2013). Anti-addiction drug ibogaine inhibits voltage-gated ionic currents: a study to assess the drug's cardiac ion channel profile. Toxicol Appl Pharmacol *273*, 259-268.

Kramarcy, N.R., Vidal, A., Froehner, S.C., and Sealock, R. (1994). Association of utrophin and multiple dystrophin short forms with the mammalian M(r) 58,000 dystrophin-associated protein (syntrophin). J Biol Chem *269*, 2870-2876.

Kyrychenko, S., Poláková, E., Kang, C., Pocsai, K., Ullrich, N.D., Niggli, E., and Shirokova, N. (2013). Hierarchical accumulation of RyR post-translational modifications drives disease progression in dystrophic cardiomyopathy. Cardiovasc Res *97*, 666-675.

Lai, Y., Thomas, G.D., Yue, Y., Yang, H.T., Li, D., Long, C., Judge, L., Bostick, B., Chamberlain, J.S., Terjung, R.L.*, et al.* (2009). Dystrophins carrying spectrin-like repeats 16 and 17 anchor nNOS to the sarcolemma and enhance exercise performance in a mouse model of muscular dystrophy. J Clin Invest *119*, 624-635.

Lai, Y., Zhao, J., Yue, Y., and Duan, D. (2013a). alpha2 and alpha3 helices of dystrophin R16 and R17 frame a microdomain in the alpha1 helix of dystrophin R17 for neuronal NOS binding. Proc Natl Acad Sci U S A *110*, 525-530.

Lai, Y., Zhao, J., Yue, Y., and Duan, D. (2013b). α2 and α3 helices of dystrophin R16 and R17 frame a microdomain in the α1 helix of dystrophin R17 for neuronal NOS binding. Proc Natl Acad Sci U S A *110*, 525-530.

Lamirault, G., Gaborit, N., Le Meur, N., Chevalier, C., Lande, G., Demolombe, S., Escande, D., Nattel, S., Léger, J.J., and Steenman, M. (2006). Gene expression profile associated with chronic atrial fibrillation and underlying valvular heart disease in man. J Mol Cell Cardiol *40*, 173-184.

Lanza, G.A., Dello Russo, A., Giglio, V., De Luca, L., Messano, L., Santini, C., Ricci, E., Damiani, A., Fumagalli, G., De Martino, G.*, et al.* (2001). Impairment of cardiac autonomic function in patients with Duchenne muscular dystrophy: relationship to myocardial and respiratory function. Am Heart J *141*, 808-812.

Lee, R.C., River, L.P., Pan, F.S., Ji, L., and Wollmann, R.L. (1992). Surfactant-induced sealing of electropermeabilized skeletal muscle membranes in vivo. Proc Natl Acad Sci U S A *89*, 4524-4528.

Lee, S., Nam, H., Kim, J., Cho, H., Jang, Y., Lee, E., Choi, E., Jin, D.I., and Moon, H. (2012). Body Weight Changes of Laboratory Animals during Transportation. Asian-Australas J Anim Sci *25*, 286-290.

Lee, S.H., Kang, D.R., Uhm, J.S., Shim, J., Sung, J.H., Kim, J.Y., Pak, H.N., Lee, M.H., and Joung, B. (2014). New-onset atrial fibrillation predicts long-term newly developed atrial fibrillation after coronary artery bypass graft. Am Heart J *167*, 593-600.e591.

Li, D., Shinagawa, K., Pang, L., Leung, T.K., Cardin, S., Wang, Z., and Nattel, S. (2001). Effects of angiotensin-converting enzyme inhibition on the development of the atrial fibrillation substrate in dogs with ventricular tachypacing-induced congestive heart failure. Circulation *104*, 2608-2614.

Li, Z., Li, Y., Zhang, L., Zhang, X., Sullivan, R., Ai, X., Szeto, C., Cai, A., Liu, L., Xiao, W.*, et al.* (2017). Reduced Myocardial Reserve in Young X-Linked Muscular Dystrophy Mice Diagnosed by Two-Dimensional Strain Analysis Combined with Stress Echocardiography. J Am Soc Echocardiogr.

Lin, C.C., Lin, J.L., Lin, C.S., Tsai, M.C., Su, M.J., Lai, L.P., and Huang, S.K. (2004). Activation of the calcineurin-nuclear factor of activated T-cell signal transduction pathway in atrial fibrillation. Chest *126*, 1926-1932.

Lin, H., Yin, X., Lunetta, K.L., Dupuis, J., McManus, D.D., Lubitz, S.A., Magnani, J.W., Joehanes, R., Munson, P.J., Larson, M.G.*, et al.* (2014). Whole blood gene expression and atrial fibrillation: the Framingham Heart Study. PLoS One *9*, e96794.

Liu, H., Chen, G.X., Liang, M.Y., Qin, H., Rong, J., Yao, J.P., and Wu, Z.K. (2014). Atrial fibrillation alters the microRNA expression profiles of the left atria of patients with mitral stenosis. BMC Cardiovasc Disord *14*, 10.

Loufrani, L., Matrougui, K., Gorny, D., Duriez, M., Blanc, I., Lévy, B.I., and Henrion, D. (2001). Flow (shear stress)-induced endothelium-dependent dilation is altered in mice lacking the gene encoding for dystrophin. Circulation *103*, 864-870.

Love, D.R., Hill, D.F., Dickson, G., Spurr, N.K., Byth, B.C., Marsden, R.F., Walsh, F.S., Edwards, Y.H., and Davies, K.E. (1989). An autosomal transcript in skeletal muscle with homology to dystrophin. Nature *339*, 55-58.

Madhavan, R., and Jarrett, H.W. (1994). Calmodulin-activated phosphorylation of dystrophin. Biochemistry *33*, 5797-5804.

Madhavan, R., and Jarrett, H.W. (1999). Phosphorylation of dystrophin and alpha-syntrophin by Ca(2+)-calmodulin dependent protein kinase II. Biochim Biophys Acta *1434*, 260-274.

Mah, J.K., Korngut, L., Dykeman, J., Day, L., Pringsheim, T., and Jette, N. (2014). A systematic review and meta-analysis on the epidemiology of Duchenne and Becker muscular dystrophy. Neuromuscul Disord *24*, 482-491.

Markham, L., Barone, A., Kinnett, K., Spicer, R., Wong, B., Benson, D., and Cripe, L. (2006). Revising the cardiac phenotype of Duchenne muscular dystrophy. Neuromuscular Disorders *16*, 699-699.

Mathew, J.P., Fontes, M.L., Tudor, I.C., Ramsay, J., Duke, P., Mazer, C.D., Barash, P.G., Hsu, P.H., Mangano, D.T., Foundation, I.o.t.I.R.a.E.*, et al.* (2004). A multicenter risk index for atrial fibrillation after cardiac surgery. JAMA *291*, 1720-1729.

Matsuda, R., Nishikawa, A., and Tanaka, H. (1995). Visualization of dystrophic muscle fibers in mdx mouse by vital staining with Evans blue: evidence of apoptosis in dystrophin-deficient muscle. J Biochem *118*, 959-964.

Matsumura, K., Ervasti, J.M., Ohlendieck, K., Kahl, S.D., and Campbell, K.P. (1992). Association of dystrophin-related protein with dystrophin-associated proteins in mdx mouse muscle. Nature *360*, 588-591.

McGreevy, J.W., Hakim, C.H., McIntosh, M.A., and Duan, D. (2015). Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy. Dis Model Mech *8*, 195-213.

McManus, D.D., Tanriverdi, K., Lin, H., Esa, N., Kinno, M., Mandapati, D., Tam, S., Okike, O.N., Ellinor, P.T., Keaney, J.F.*, et al.* (2015). Plasma microRNAs are associated with atrial fibrillation and change after catheter ablation (the miRhythm study). Heart Rhythm *12*, 3-10.

Melduni, R.M., Suri, R.M., Seward, J.B., Bailey, K.R., Ammash, N.M., Oh, J.K., Schaff, H.V., and Gersh, B.J. (2011). Diastolic dysfunction in patients undergoing cardiac surgery: a pathophysiological mechanism underlying the initiation of new-onset post-operative atrial fibrillation. J Am Coll Cardiol *58*, 953-961.

Menke, A., and Jockusch, H. (1991). Decreased osmotic stability of dystrophin-less muscle cells from the mdx mouse. Nature *349*, 69-71.

Mercier, S., Toutain, A., Toussaint, A., Raynaud, M., de Barace, C., Marcorelles, P., Pasquier, L., Blayau, M., Espil, C., Parent, P.*, et al.* (2013). Genetic and clinical specificity of 26 symptomatic carriers for dystrophinopathies at pediatric age. Eur J Hum Genet *21*, 855-863.

Mihm, M.J., Yu, F., Carnes, C.A., Reiser, P.J., McCarthy, P.M., Van Wagoner, D.R., and Bauer, J.A. (2001). Impaired myofibrillar energetics and oxidative injury during human atrial fibrillation. Circulation *104*, 174-180.

Mizuno, Y., Nonaka, I., Hirai, S., and Ozawa, E. (1993). Reciprocal expression of dystrophin and utrophin in muscles of Duchenne muscular dystrophy patients, female DMD-carriers and control subjects. J Neurol Sci *119*, 43-52.

Moens, A.L., Yang, R., Watts, V.L., and Barouch, L.A. (2010). Beta 3-adrenoreceptor regulation of nitric oxide in the cardiovascular system. J Mol Cell Cardiol *48*, 1088-1095.

Monaco, A.P., Bertelson, C.J., Liechti-Gallati, S., Moser, H., and Kunkel, L.M. (1988). An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. Genomics *2*, 90-95.

Morine, K.J., Sleeper, M.M., Barton, E.R., and Sweeney, H.L. (2010). Overexpression of SERCA1a in the mdx diaphragm reduces susceptibility to contraction-induced damage. Hum Gene Ther *21*, 1735-1739.

Mosqueira, M., Zeiger, U., Förderer, M., Brinkmeier, H., and Fink, R.H. (2013). Cardiac and respiratory dysfunction in Duchenne muscular dystrophy and the role of second messengers. Med Res Rev *33*, 1174-1213.

Muntoni, F., Cau, M., Ganau, A., Congiu, R., Arvedi, G., Mateddu, A., Marrosu, M.G., Cianchetti, C., Realdi, G., and Cao, A. (1993). Brief report: deletion of the dystrophin muscle-promoter region associated with X-linked dilated cardiomyopathy. N Engl J Med *329*, 921-925.

Muntoni, F., Torelli, S., and Ferlini, A. (2003). Dystrophin and mutations: one gene, several proteins, multiple phenotypes. Lancet Neurol *2*, 731-740.

Mázala, D.A., Pratt, S.J., Chen, D., Molkentin, J.D., Lovering, R.M., and Chin, E.R. (2015). SERCA1 overexpression minimizes skeletal muscle damage in dystrophic mouse models. Am J Physiol Cell Physiol *308*, C699-709.

Nattel, S., Maguy, A., Le Bouter, S., and Yeh, Y.H. (2007). Arrhythmogenic ion-channel remodeling in the heart: heart failure, myocardial infarction, and atrial fibrillation. Physiol Rev *87*, 425-456.

Neef, S., Dybkova, N., Sossalla, S., Ort, K.R., Fluschnik, N., Neumann, K., Seipelt, R., Schöndube, F.A., Hasenfuss, G., and Maier, L.S. (2010). CaMKII-dependent diastolic SR Ca2+ leak and elevated diastolic Ca2+ levels in right atrial myocardium of patients with atrial fibrillation. Circ Res *106*, 1134-1144.

Neri, M., Valli, E., Alfano, G., Bovolenta, M., Spitali, P., Rapezzi, C., Muntoni, F., Banfi, S., Perini, G., Gualandi, F*., et al.* (2012). The absence of dystrophin brain isoform expression in healthy human heart ventricles explains the pathogenesis of 5' X-linked dilated cardiomyopathy. BMC Med Genet *13*, 20.

Nigro, G., Comi, L.I., Politano, L., and Bain, R.J. (1990). The incidence and evolution of cardiomyopathy in Duchenne muscular dystrophy. Int J Cardiol *26*, 271-277.

Nishi, H., Sakaguchi, T., Miyagawa, S., Yoshikawa, Y., Fukushima, S., Saito, S., Ueno, T., Kuratani, T., and Sawa, Y. (2013). Impact of microRNA expression in human atrial tissue in patients with atrial fibrillation undergoing cardiac surgery. PLoS One *8*, e73397.

Nolan, M.A., Jones, O.D., Pedersen, R.L., and Johnston, H.M. (2003). Cardiac assessment in childhood carriers of Duchenne and Becker muscular dystrophies. Neuromuscul Disord *13*, 129-132.

Obler, D., Wu, B.L., Lip, V., Estrella, E., Keck, S., Haggan, C., Semigran, M., and Smoot, L.B. (2010). Familial dilated cardiomyopathy secondary to dystrophin splice site mutation. J Card Fail *16*, 194-199.

Opolski, G., Torbicki, A., Kosior, D.A., Szulc, M., Wozakowska-Kaplon, B., Kolodziej, P., Achremczyk, P., and Study, I.o.t.P.H.t.T.C.A.F. (2004). Rate control vs rhythm control in patients with nonvalvular persistent atrial fibrillation: the results of the Polish How to Treat Chronic Atrial Fibrillation (HOT CAFE) Study. Chest *126*, 476-486.

Palladino, M., Gatto, I., Neri, V., Straino, S., Smith, R.C., Silver, M., Gaetani, E., Marcantoni, M., Giarretta, I., Stigliano, E*., et al.* (2013). Angiogenic impairment of the vascular endothelium: a novel mechanism and potential therapeutic target in muscular dystrophy. Arterioscler Thromb Vasc Biol *33*, 2867-2876.

Parvatiyar, M.S., Marshall, J.L., Nguyen, R.T., Jordan, M.C., Richardson, V.A., Roos, K.P., and Crosbie-Watson, R.H. (2015). Sarcospan Regulates Cardiac Isoproterenol Response and Prevents Duchenne Muscular Dystrophy-Associated Cardiomyopathy. J Am Heart Assoc *4*.

Pasternak, C., Wong, S., and Elson, E.L. (1995). Mechanical function of dystrophin in muscle cells. J Cell Biol *128*, 355-361.

Percival, J.M., Anderson, K.N., Gregorevic, P., Chamberlain, J.S., and Froehner, S.C. (2008). Functional deficits in nNOSmu-deficient skeletal muscle: myopathy in nNOS knockout mice. PLoS One *3*, e3387.

Perkins, K.J., and Davies, K.E. (2002). The role of utrophin in the potential therapy of Duchenne muscular dystrophy. Neuromuscul Disord *12 Suppl 1*, S78-89.

Perloff, J.K. (1984). Cardiac rhythm and conduction in Duchenne's muscular dystrophy: a prospective study of 20 patients. J Am Coll Cardiol *3*, 1263-1268.

Perloff, J.K., de Leon, A.C., and O'Doherty, D. (1966). The cardiomyopathy of progressive muscular dystrophy. Circulation *33*, 625-648.

Perloff, J.K., Roberts, W.C., de Leon, A.C., and O'Doherty, D. (1967). The distinctive electrocardiogram of Duchenne's progressive muscular dystrophy. An electrocardiographic-pathologic correlative study. Am J Med *42*, 179-188.

Petrof, B.J., Shrager, J.B., Stedman, H.H., Kelly, A.M., and Sweeney, H.L. (1993). Dystrophin protects the sarcolemma from stresses developed during muscle contraction. Proc Natl Acad Sci U S A *90*, 3710-3714.

Pons, F., Augier, N., Léger, J.O., Robert, A., Tomé, F.M., Fardeau, M., Voit, T., Nicholson, L.V., Mornet, D., and Léger, J.J. (1991). A homologue of dystrophin is expressed at the neuromuscular junctions of normal individuals and DMD patients, and of normal and mdx mice. Immunological evidence. FEBS Lett *282*, 161-165.

Pons, F., Robert, A., Marini, J.F., and Léger, J.J. (1994). Does utrophin expression in muscles of mdx mice during postnatal development functionally compensate for dystrophin deficiency? J Neurol Sci *122*, 162-170.

Rando, T.A. (2001). The dystrophin-glycoprotein complex, cellular signaling, and the regulation of cell survival in the muscular dystrophies. Muscle Nerve *24*, 1575-1594.

Reilly, S.N., Jayaram, R., Nahar, K., Antoniades, C., Verheule, S., Channon, K.M., Alp, N.J., Schotten, U., and Casadei, B. (2011). Atrial sources of reactive oxygen species vary with the duration and substrate of atrial fibrillation: implications for the antiarrhythmic effect of statins. Circulation *124*, 1107-1117.

Reilly, S.N., Liu, X., Carnicer, R., Recalde, A., Muszkiewicz, A., Jayaram, R., Carena, M.C., Wijesurendra, R., Stefanini, M., Surdo, N.C.*, et al.* (2016). Up-regulation of miR-31 in human atrial fibrillation begets the arrhythmia by depleting dystrophin and neuronal nitric oxide synthase. Science translational medicine *8*, 340ra374.

Roberts, R.G. (2001). Dystrophins and dystrobrevins. Genome Biol *2*, REVIEWS3006.

Ronan, J.A., Perloff, J.K., Bowen, P.J., and Mann, O. (1972). [The vectorcardiogram in Duchenne's progressive muscular dystrophy]. Am Heart J *84*, 588-596.

Ruan, Z., Sun, X., Sheng, H., and Zhu, L. (2015). Long non-coding RNA expression profile in atrial fibrillation. Int J Clin Exp Pathol *8*, 8402-8410.

Sadeghi, A., Doyle, A.D., and Johnson, B.D. (2002). Regulation of the cardiac L-type Ca2+ channel by the actin-binding proteins alpha-actinin and dystrophin. Am J Physiol Cell Physiol *282*, C1502-1511.

Sander, M., Chavoshan, B., Harris, S.A., Iannaccone, S.T., Stull, J.T., Thomas, G.D., and Victor, R.G. (2000). Functional muscle ischemia in neuronal nitric oxide synthase-deficient skeletal muscle of children with Duchenne muscular dystrophy. Proc Natl Acad Sci U S A *97*, 13818-13823.

Sanyal, S.K., and Johnson, W.W. (1982). Cardiac conduction abnormalities in children with Duchenne's progressive muscular dystrophy: electrocardiographic features and morphologic correlates. Circulation *66*, 853-863.

Sapp, J.L., Bobet, J., and Howlett, S.E. (1996). Contractile properties of myocardium are altered in dystrophin-deficient mdx mice. J Neurol Sci *142*, 17-24.

Sarma, S., Li, N., van Oort, R.J., Reynolds, C., Skapura, D.G., and Wehrens, X.H. (2010). Genetic inhibition of PKA phosphorylation of RyR2 prevents dystrophic cardiomyopathy. Proc Natl Acad Sci U S A *107*, 13165-13170.

Schneider, J.S., Shanmugam, M., Gonzalez, J.P., Lopez, H., Gordan, R., Fraidenraich, D., and Babu, G.J. (2013). Increased sarcolipin expression and decreased sarco(endo)plasmic reticulum Ca2+ uptake in skeletal muscles of mouse models of Duchenne muscular dystrophy. J Muscle Res Cell Motil *34*, 349-356.

Schotten, U., Verheule, S., Kirchhof, P., and Goette, A. (2011). Pathophysiological mechanisms of atrial fibrillation: a translational appraisal. Physiological reviews *91*, 265-325.

Shaul, P.W. (2002). Regulation of endothelial nitric oxide synthase: location, location, location. Annu Rev Physiol *64*, 749-774.

Shirokova, N., and Niggli, E. (2013). Cardiac phenotype of Duchenne Muscular Dystrophy: insights from cellular studies. J Mol Cell Cardiol *58*, 217-224.

Sicinski, P., Geng, Y., Ryder-Cook, A.S., Barnard, E.A., Darlison, M.G., and Barnard, P.J. (1989). The molecular basis of muscular dystrophy in the mdx mouse: a point mutation. Science *244*, 1578-1580.

Simon, J.N., Duglan, D., Casadei, B., and Carnicer, R. (2014). Nitric oxide synthase regulation of cardiac excitation-contraction coupling in health and disease. J Mol Cell Cardiol *73*, 80-91.

Slucka, C. (1968). The electrocardiogram in Duchenne progressive muscular dystrophy. Circulation *38*, 933-940.

Spurney, C.F., Knoblach, S., Pistilli, E.E., Nagaraju, K., Martin, G.R., and Hoffman, E.P. (2008). Dystrophin-deficient cardiomyopathy in mouse: expression of Nox4 and Lox are associated with fibrosis and altered functional parameters in the heart. Neuromuscul Disord *18*, 371-381.

Stamler, J.S., and Meissner, G. (2001). Physiology of nitric oxide in skeletal muscle. Physiol Rev *81*, 209-237.

Stuckey, D.J., Carr, C.A., Camelliti, P., Tyler, D.J., Davies, K.E., and Clarke, K. (2012). In vivo MRI characterization of progressive cardiac dysfunction in the mdx mouse model of muscular dystrophy. PloS one *7*, e28569.

Su, Y., Li, L., Zhao, S., Yue, Y., and Yang, S. (2018). The long noncoding RNA expression profiles of paroxysmal atrial fibrillation identified by microarray analysis. Gene *642*, 125-134.

Surgery, A.A.o.P.S.o.C.a.C. (2005). Cardiovascular health supervision for individuals affected by Duchenne or Becker muscular dystrophy. Pediatrics *116*, 1569-1573.

Suzuki, N., Motohashi, N., Uezumi, A., Fukada, S., Yoshimura, T., Itoyama, Y., Aoki, M., Miyagoe-Suzuki, Y., and Takeda, S. (2007). NO production results in suspension-induced muscle atrophy through dislocation of neuronal NOS. J Clin Invest *117*, 2468-2476.

Thomas, G.D. (2013). Functional muscle ischemia in Duchenne and Becker muscular dystrophy. Front Physiol *4*, 381.

Thomas, G.D., Sander, M., Lau, K.S., Huang, P.L., Stull, J.T., and Victor, R.G. (1998). Impaired metabolic modulation of alpha-adrenergic vasoconstriction in dystrophin-deficient skeletal muscle. Proc Natl Acad Sci U S A *95*, 15090-15095.

Thomas, G.D., Shaul, P.W., Yuhanna, I.S., Froehner, S.C., and Adams, M.E. (2003). Vasomodulation by skeletal muscle-derived nitric oxide requires alpha-syntrophin-mediated sarcolemmal localization of neuronal Nitric oxide synthase. Circ Res *92*, 554-560.

Tidball, J.G., and Wehling-Henricks, M. (2004). Expression of a NOS transgene in dystrophin-deficient muscle reduces muscle membrane damage without increasing the expression of membrane-associated cytoskeletal proteins. Mol Genet Metab *82*, 312-320.

Tinsley, J.M., Blake, D.J., Roche, A., Fairbrother, U., Riss, J., Byth, B.C., Knight, A.E., Kendrick-Jones, J., Suthers, G.K., and Love, D.R. (1992). Primary structure of dystrophin-related protein. Nature *360*, 591-593.

Townsend, D., Yasuda, S., Li, S., Chamberlain, J.S., and Metzger, J.M. (2008). Emergent dilated cardiomyopathy caused by targeted repair of dystrophic skeletal muscle. Mol Ther *16*, 832-835.

Tran, C.T., Schmidt, T.A., Christensen, J.B., and Kjeldsen, K. (2009). Atrial Na,K-ATPase increase and potassium dysregulation accentuate the risk of postoperative atrial fibrillation. Cardiology *114*, 1-7.

Tsai, C.T., Lai, L.P., Kuo, K.T., Hwang, J.J., Hsieh, C.S., Hsu, K.L., Tseng, C.D., Tseng, Y.Z., Chiang, F.T., and Lin, J.L. (2008). Angiotensin II activates signal transducer and activators of transcription 3 via Rac1 in atrial myocytes and fibroblasts: implication for the therapeutic effect of statin in atrial structural remodeling. Circulation *117*, 344-355.

Ueda, K., Valdivia, C., Medeiros-Domingo, A., Tester, D.J., Vatta, M., Farrugia, G., Ackerman, M.J., and Makielski, J.C. (2008). Syntrophin mutation associated with long QT syndrome through activation of the nNOS-SCN5A macromolecular complex. Proc Natl Acad Sci U S A *105*, 9355-9360.

Van Gelder, I.C., Hagens, V.E., Bosker, H.A., Kingma, J.H., Kamp, O., Kingma, T., Said, S.A., Darmanata, J.I., Timmermans, A.J., Tijssen, J.G.*, et al.* (2002). A comparison of rate

control and rhythm control in patients with recurrent persistent atrial fibrillation. N Engl J Med *347*, 1834-1840.

Van Wagoner, D.R., Pond, A.L., Lamorgese, M., Rossie, S.S., McCarthy, P.M., and Nerbonne, J.M. (1999). Atrial L-type Ca2+ currents and human atrial fibrillation. Circ Res *85*, 428-436.

Verheule, S., Sato, T., Everett, T.t., Engle, S.K., Otten, D., Rubart-von der Lohe, M., Nakajima, H.O., Nakajima, H., Field, L.J., and Olgin, J.E. (2004). Increased vulnerability to atrial fibrillation in transgenic mice with selective atrial fibrosis caused by overexpression of TGF-beta1. Circulation research *94*, 1458-1465.

Vest, J.A., Wehrens, X.H., Reiken, S.R., Lehnart, S.E., Dobrev, D., Chandra, P., Danilo, P., Ravens, U., Rosen, M.R., and Marks, A.R. (2005). Defective cardiac ryanodine receptor regulation during atrial fibrillation. Circulation *111*, 2025-2032.

Viggiano, E., Picillo, E., Cirillo, A., and Politano, L. (2013). Comparison of X-chromosome inactivation in Duchenne muscle/myocardium-manifesting carriers, non-manifesting carriers and related daughters. Clin Genet *84*, 265-270.

Viola, H.M., Adams, A.M., Davies, S.M., Fletcher, S., Filipovska, A., and Hool, L.C. (2014). Impaired functional communication between the L-type calcium channel and mitochondria contributes to metabolic inhibition in the mdx heart. Proc Natl Acad Sci U S A *111*, E2905-2914.

Viola, H.M., Davies, S.M., Filipovska, A., and Hool, L.C. (2013). L-type Ca(2+) channel contributes to alterations in mitochondrial calcium handling in the mdx ventricular myocyte. Am J Physiol Heart Circ Physiol *304*, H767-775.

Wagner, S., Knipp, S., Weber, C., Hein, S., Schinkel, S., Walther, A., Bekeredjian, R., Muller, O.J., and Friedrich, O. (2012). The heart in Duchenne muscular dystrophy: early detection of contractile performance alteration. J Cell Mol Med *16*, 3028-3036.

Wehling, M., Spencer, M.J., and Tidball, J.G. (2001). A nitric oxide synthase transgene ameliorates muscular dystrophy in mdx mice. J Cell Biol *155*, 123-131.

Whitlock, R., Healey, J.S., Connolly, S.J., Wang, J., Danter, M.R., Tu, J.V., Novick, R., Fremes, S., Teoh, K., Khera, V*., et al.* (2014). Predictors of early and late stroke following cardiac surgery. CMAJ *186*, 905-911.

Williams, I.A., and Allen, D.G. (2007a). Intracellular calcium handling in ventricular myocytes from mdx mice. Am J Physiol Heart Circ Physiol *292*, H846-855.

Williams, I.A., and Allen, D.G. (2007b). The role of reactive oxygen species in the hearts of dystrophin-deficient mdx mice. Am J Physiol Heart Circ Physiol *293*, H1969-1977.

Williamson, R.A., Henry, M.D., Daniels, K.J., Hrstka, R.F., Lee, J.C., Sunada, Y., Ibraghimov-Beskrovnaya, O., and Campbell, K.P. (1997). Dystroglycan is essential for early embryonic development: disruption of Reichert's membrane in Dag1-null mice. Hum Mol Genet *6*, 831-841.

Wu, G., Majewski, J., Ege, C., Kjaer, K., Weygand, M.J., and Lee, K.Y. (2004). Lipid corralling and poloxamer squeeze-out in membranes. Phys Rev Lett *93*, 028101.

Wyse, D.G., Waldo, A.L., DiMarco, J.P., Domanski, M.J., Rosenberg, Y., Schron, E.B., Kellen, J.C., Greene, H.L., Mickel, M.C., Dalquist, J.E., *et al.* (2002). A comparison of rate control and rhythm control in patients with atrial fibrillation. N Engl J Med *347*, 1825-1833.

Xiao, J., Liang, D., Zhang, Y., Liu, Y., Zhang, H., Li, L., Liang, X., Sun, Y., and Chen, Y.H. (2011). MicroRNA expression signature in atrial fibrillation with mitral stenosis. Physiol Genomics *43*, 655-664.

Xu, Y., Huang, R., Gu, J., and Jiang, W. (2016). Identification of long non-coding RNAs as novel biomarker and potential therapeutic target for atrial fibrillation in old adults. Oncotarget *7*, 10803-10811.

Yanagisawa, A., Miyagawa, M., Yotsukura, M., Tsuya, T., Shirato, C., Ishihara, T., Aoyagi, T., and Ishikawa, K. (1992). The prevalence and prognostic significance of arrhythmias in Duchenne type muscular dystrophy. Am Heart J *124*, 1244-1250.

Yasuda, S., Townsend, D., Michele, D.E., Favre, E.G., Day, S.M., and Metzger, J.M. (2005). Dystrophic heart failure blocked by membrane sealant poloxamer. Nature *436*, 1025-1029.

Yokota, T., Miyagoe-Suzuki, Y., Ikemoto, T., Matsuda, R., and Takeda, S. (2014). α1-Syntrophin-deficient mice exhibit impaired muscle force recovery after osmotic shock. Muscle Nerve *49*, 728-735.

Yotsukura, M., Sasaki, K., Kachi, E., Sasaki, A., Ishihara, T., and Ishikawa, K. (1995). Circadian rhythm and variability of heart rate in Duchenne-type progressive muscular dystrophy. Am J Cardiol *76*, 947-951.

Youn, J.Y., Zhang, J., Zhang, Y., Chen, H., Liu, D., Ping, P., Weiss, J.N., and Cai, H. (2013). Oxidative stress in atrial fibrillation: an emerging role of NADPH oxidase. J Mol Cell Cardiol *62*, 72-79.

Yue, Y., Wasala, N.B., Bostick, B., and Duan, D. (2016). 100-fold but not 50-fold dystrophin overexpression aggravates electrocardiographic defects in the mdx model of Duchenne muscular dystrophy. Mol Ther Methods Clin Dev *3*, 16045.

Zakkar, M., Ascione, R., James, A.F., Angelini, G.D., and Suleiman, M.S. (2015). Inflammation, oxidative stress and postoperative atrial fibrillation in cardiac surgery. Pharmacol Ther.

Zhang, W., ten Hove, M., Schneider, J.E., Stuckey, D.J., Sebag-Montefiore, L., Bia, B.L., Radda, G.K., Davies, K.E., Neubauer, S., and Clarke, K. (2008). Abnormal cardiac morphology, function and energy metabolism in the dystrophic mdx mouse: an MRI and MRS study. J Mol Cell Cardiol *45*, 754-760.

Zheng, Z., Jayaram, R., Jiang, L., Emberson, J., Zhao, Y., Li, Q., Du, J., Guarguagli, S., Hill, M., Chen, Z., *et al.* (2016). Perioperative Rosuvastatin in Cardiac Surgery. N Engl J Med *374*, 1744-1753.