Imperial College London

Imperial College London, Faculty of Medicine,

National Heart & Lung Institute.

Project Title:

Investigating the role of RNA-editing in the human hearts with dilated cardiomyopathy at single-cell type resolution.

Submitted in partial fulfilment of the requirements for an MPhil in Computational Biology

Name: Mohamed HASSANI

CID: 00170435

Supervisor: Dr Prashant Srivastava

Word count: 14219

Date: April 2023

STATEMENT OF ORIGINALITY

‘I certify that this thesis, and the research to which it refers, are the product of my own work, conducted during the current and part of the past year of the MPhil project in Genomic Medicine at Imperial College London. Any ideas, quotations, or software from the work of other people, published or otherwise, or from my own previous work are fully acknowledged in accordance with the standard referencing practices of the discipline.’

ACKNOWLEDGEMENTS

I am grateful to my supervisor Dr Prashant Srivastava for giving me his full support and directing me during this research project. He provided some of his invaluable time to answer many of my questions and guided me into exploring and extracting useful information from a large single nuclei dataset.

I am also grateful to my colleague and PhD student Jiahui Ji for providing the initial single nuclei bam files and its reference article, and guiding me into splitting those files in the early stages of my research.

I am also grateful to all my colleagues in the group for their friendship and scientific exchanges.

I am grateful to many for various reasons, but I shall withhold naming them.

Abstract

RNA editing dysregulation has been linked to various cardiac diseases. In this study we have explored the role of RNA editing in dilated cardiomyopathy (DCM) using single-cell transcriptomics to profile RNA editing events in different cell types of the human heart.

Cardiomyopathy is a disease that affects the heart muscle and can lead to heart failure or other complications. One type of Cardiomyopathy is Dilated Cardiomyopathy (DCM). In DCM the heart muscle becomes stretched and thin, causing the heart to become enlarged. Thereafter, the weakened heart muscle is unable to pump blood effectively, thus leading to heart failure. This can be due to inherited genetic factors and / or epigenetic (gene regulation) factors.

Our data analysis revealed a significant correlation between RNA editing dysregulation, histone protein post-translational modifications (PTMs), and chromatin reorganisation in DCM heart muscle cells. Consequently, we have identified gene transcription regulation impacting nucleic acid metabolism pathways, as well as downstream impairment of cardiac muscle motility. Furthermore, our results also demonstrate a strong correlation between RNA editing dysregulation and non-histone protein PTMs, highlighting their impact on the cytoskeleton (including microfilaments, intermediate filaments, and microtubules), which, in turn, affects organelle dynamics and cell motility.

Significantly, we have found cardiomyocytes to be the most RNA-edited cells in the heart, with ADAR enzymes being the main RNA editors. This discovery underscores a significant correlation between RNA editing dysregulation, nucleic acid metabolism, ubiquitination, and cellular motility.

Exploring aberrant RNA editing may provide a deeper understanding of the underlying mechanisms that alter gene expression and RNA regulation in DCM. This insight holds promise as an early biomarker, enabling the detection of DCM prior to the manifestation of clinical symptoms. Moreover, by reprogramming RNA editing enzymes, it may be possible to restore healthy metabolism and motility.

Table of Content

TITLE PAGE 1
STATEMENT OF ORIGINALITY 2
ACKNOWLEDGEMENTS 3
ABSTRACT 4
TABLE OF CONTENTS 5
LIST OF FIGURES AND TABLES 6
ABBREVIATIONS 8
1. INTRODUCTION 9
1.1. Background 9
1.2. RNA editing importance 11
1.3. Role of RNA editing in Cardiac diseases 11
1.4. Hypothesis and aims 12
1.5. Timeliness 12
2. METHODS 14
2.1. Single-cell data processing and classification to cell types 14
2.2. Predicting RNA editing from single cell types 14
2.2.1. Predicting RNA edited sites 14
2.2.2. Filtering process to get quality RNA edited sites 15
2.2.3. Removal of SNPs from all files for each cell type 15
2.2.4. Differential RNA editing 16
2.3. RNA Structural Metrics (MFE and others) 17
2.4. Consequences of RNA Editing 20
2.5. Functional enrichments analysis 20
3. RESULTS 21
3.1. Single-cell data processing and classification to cell types 21
3.2. Predicting RNA editing from single cell types 22
3.2.1. Predicting RNA edited sites 22
3.2.2. Filtering process to get quality RNA edited sites 22
3.2.3. Removal of SNPs from all files for each cell type 24
3.2.4. RNA editing total count per cell type 25
3.2.4.1. RNA editing total count per cell type in healthy state 28
3.2.4.2. RNA editing total count per cell type in all health states 29
3.3. mRNA structure and health post RNA editing 45
3.3.1. mRNA structural metrics for unedited transcripts 45
3.3.2. mRNA structural metrics for edited transcripts 46
3.3.3. Healthy mRNA structure post RNA editing 47
3.3.4. Healthy and unhealthy mRNA structure post RNA editing 51
3.4. Consequences of RNA Editing 59
3.5. Differential RNA editing 60
3.6. Functional enrichments analysis 61
3.6.1. ORA KEGG pathways in Cardiomyocytes 62
3.6.2. Biological processes in Cardiomyocytes 63
3.6.3. Molecular functions in Cardiomyocytes 71
3.7. RNA structural pathology 74
4. DISCUSSION 75
5. CONCLUSION 78
6. REFERENCES 80

LIST OF FIGURES AND TABLES

Figure 1: Secondary Structure change after one A-to-I modification.
Figure 2: F2R MFE Secondary Structure
Figure 3: F2R Centroid Secondary Structure
Figure 4: Summary of RNA edit total count per cell type.
Figure 5: Summary of A-to-I and C-to-U total counts for highly edited healthy heart cells.
Figure 6: Summary of A-to-I and C-to-U total counts for relatively low edited healthy heart cells.
Figure 7: Cardiomyocytes RNA edits per health state.
Figure 8: Fibroblast RNA edits per health state.
Figure 9: Endothelium RNA edits per health state.
Figure 10: Macrophages RNA edits per health state.
Figure 11: Pericytes RNA edits per health state.
Figure 12: Endocardium RNA edits per health state.
Figure 13: Smooth muscle RNA edits per health state.
Figure 14: T NK cells RNA edits per health state.
Figure 15: Adipocytes RNA edits per health state.
Figure 16: Epicardium RNA edits per health state.
Figure 17: Lymphatic RNA edits per health state.
Figure 18: Neurons RNA edits per health state.
Figure 19: Mast cells RNA edits per health state.
Figure 20: Monocytes RNA edits per health state.
Figure 21: Unedited transcripts structural metrics.
Figure 22: Edited transcripts structural metrics.
Figure 23: MFE & Healthy and highly RNA Edited cells
Figure 24 MFE & Healthy and Less frequently RNA edited cells.
Figure 25: MFE & Cardiomyocytes RNA editing in all health states.
Figure 26: MFE & Fibroblasts RNA editing in all health states.
Figure 27: MFE & Endothelium RNA editing in all health states.
Figure 28: MFE & Macrophages RNA editing in all health states.
Figure 29: MFE & Pericytes RNA editing in all health states.
Figure 30: MFE & Endocardium RNA editing in all health states.
Figure 31: MFE & Smooth muscle RNA editing in all health states.
Figure 32: MFE & T NK cells RNA editing in all health states.
Figure 33: MFE & Adipocytes RNA editing in all health states.
Figure 34: MFE & Epicardium RNA editing in all health states.
Figure 35: MFE & Lymphatic RNA editing in all health states.
Figure 36: MFE & Neurons RNA editing in all health states.
Figure 37: MFE & Mast cells RNA editing in all health states.
Figure 38: MFE & Monocytes RNA editing in all health states.

Figure 39: Cardiomyocytes ORA KEGG pathways
Figure 40: Cardiomyocytes GO Biological Processes
Figure 41: Cardiomyocytes GO Molecular Function.
Figure 42: Study workflow.
Figure 43: RNA editing dysregulation and Dilated Cardiomyopathy

Table 1: Predicting RNA edited positions example.
Table 2: Predicted RNA edited positions as BED file.
Table 3: Predicted RNA edited positions summary.
Table 4: Summary of RNA edited positions with only 2 nucleotides in AllSubs.
Table 5: Summary of RNA edited positions removed with Coverage-q30 < 10
Table 6: Summary of RNA edited positions removed with BaseCount < 5.
Table 7: Summary of RNA edited positions with removed SNPs.
Table 8: Cardiomyocytes ORA KEGG pathways.
Table 9: Cardiomyocytes GO Biological Processes
Table 10: Cardiomyocytes GO Molecular Function.
Table 11: Overlap Genes from Chen et al study and our Cardiomyocytes Enrichment Analysis in DCM

ABREVIATIONS

A-to-I : Adenosine modified to become Inosine.

A-to-G : The same as A-to-I, but with the hint that the transcript is from the positive DNA strand.

T-to-C : The same as A-to-I, but with the hint that the transcript is from the negative DNA strand.

C-to-U: Cytosine modified to become Uracil.

C-to-T: The same as C-to-U, but the transcript comes from the positive DNA strand.

G-to-A: The same as C-to-U, but the transcript comes from the negative DNA strand.

ADAR: Adenosine deaminases acting on RNA.

AID / APOBEC: Activation-induced cytidine deaminase / apolipoprotein B mRNA editing enzyme.

RNA: Ribonucleic acid

DNA: Deoxyribonucleic acid

mRNA: messenger RNA or protein-coding RNA.

miRNA: micro-RNA.

SNP: Single nucleotide polymorphism. A germline substitution of a single nucleotide at a specific position but must be present with large numbers in a population.

scRNA: Single Cell RNA.

MFE: Minimum Free energy.

MEA: Maximum Expected Accuracy.

PF: Partition function.

EBI: European Bioinformatics Institute.

UCSC: University of California Santa Cruz (Genomics Institute).

FDR: False Discovery Rate.

GO: Gene ontology.

GO BP: GO biological processes.

GO MF: GO molecular function.

GO CC: GO cellular component.

KEGG: Kyoto Encyclopaedia of Genes and Genomes

Disgenet: A gene-disease associations database.

GLAD4U: Gene List Automatically Derived For You. A database of gene-disease associations

OMIM: Online Mendelian Inheritance in Man. A database of gene-disease associations

DCM: Dilated cardiomyopathy.

DB: Diabetes

1. Introduction:

1. 1. Background

RNA editing is a type of RNA modification and is a dynamically regulated process. RNA editing represents an important mechanism to expand and diversify functions of the protein repertoire [1, 2]. RNA editing alters mRNA sequence without changing the DNA template. RNA editing is a natural process that could occur during transcription and translation in healthy and unhealthy cells.

There are two RNA editing types: (i) A-to-I RNA-Editing, where the Adenosine Deaminase (ADARs) enzymes modify Adenosine (A) into Inosine (I). Inosine has a preferential binding to Cytosine [3], and it is processed by the cell as if it were Guanine (G). Hence, computationally, I (Inosine) is represented as G (Guanine), and (ii) C-to-U RNA-Editing, where the AID/APOBEC family enzymes modify Cytosine (C) into Uracil (U). At times, we may also use T (Thymine) instead of U (Uracil) when documenting RNA changes like T-to-C instead of U-to-C. As this would give us the added information of which DNA strand the RNA was transcribed from.

RNA Editing modifies the RNA's nucleotide composition and, thereby, RNA's secondary structure (Fig. 1). Unlike, the double stranded DNA, messenger RNA is single stranded. However, its structure is dynamic and can fold onto itself. Thereby, forming double strands in some parts, while some other parts may remain single stranded (Fig. 1). Indeed, Adenosine (A) has a double hydrogen bond with Uracil (U). When modified to Guanine by the ADAR enzymes, the newly formed GU wobble pair [4, 5] has also a double bond but weaker than AU double bonding. On the other hand, Guanine (G) has a triple hydrogen bond with Cytosine (C), and when the latter is modified to Uracil (U), the newly formed GU wobble pair will have a double and weaker bond. Therefore, at a nucleotide level the structure is weakened, but the overall structure is mainly reinforced and at times it can be weakened as well. Please see the results section for more details. Interestingly, the ADAR family enzymes act on double-stranded RNAs, while AID/APOBEC enzymes act on single-stranded RNAs. Furthermore, both types of enzymes are found to be active in the nucleus as well as the cytoplasm.

Please see the following figure (Fig 1) where one nucleotide A was changed to G and its impact on the transcript's secondary structure.

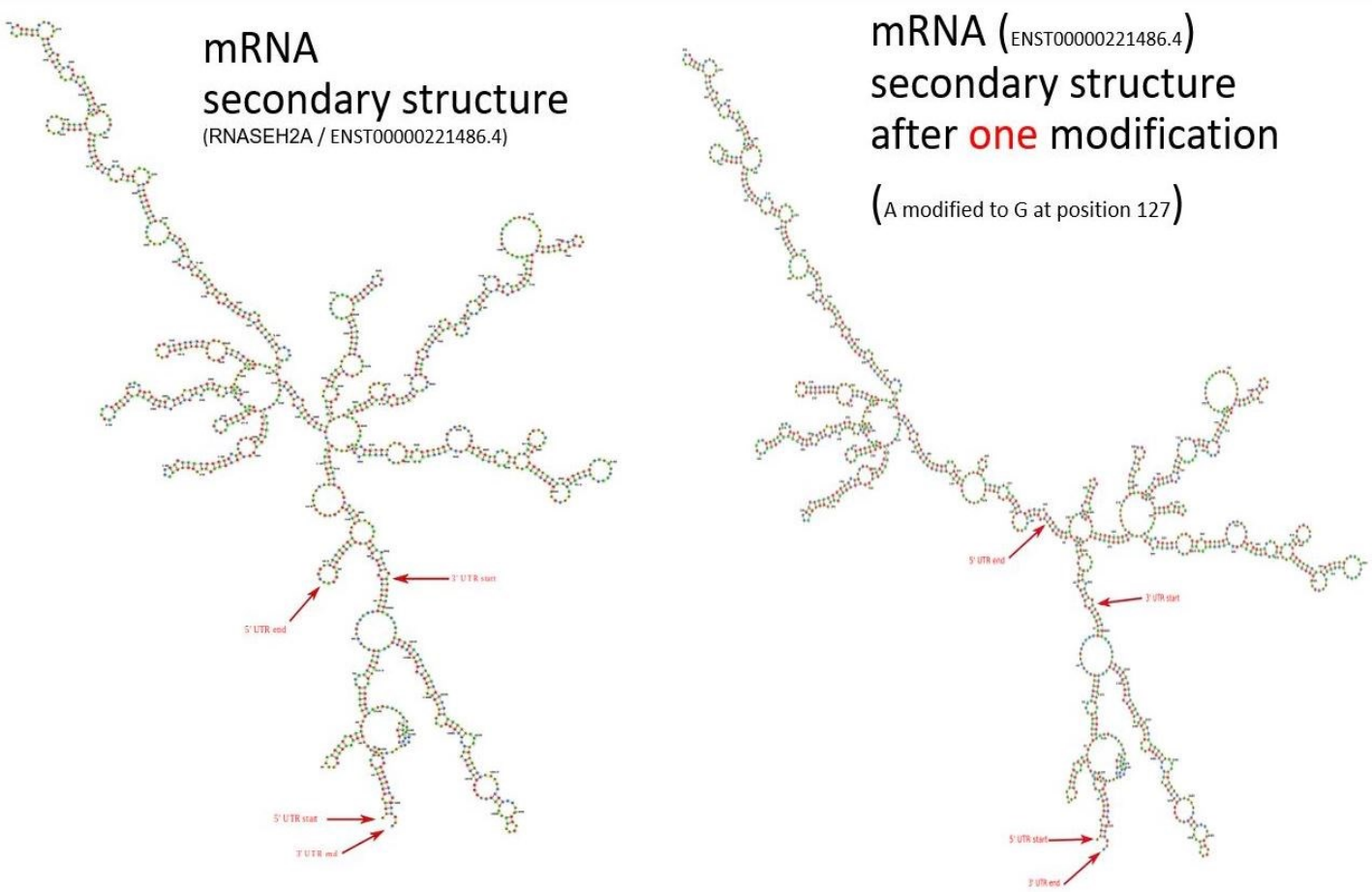


Figure 1 Secondary Structure change after one A-to-I modification. This modification changed the minimum free energy from -449 kcal/mol to -448.3 kcal/mol.

These modifications have several diverse outcomes, such as the degradation of the modified mRNA, the reduction, or the increase in the translation speed. Moreover, modified mRNA codons could result in a synonymous or non-synonymous translation. In other words, in the synonymous translation case, the RNA edited codon and its non-edited version translate to an identical amino acid. In contrast, in the non-synonymous case, the edited codon is translated into a different amino acid from its non-edited counterpart. Importantly, synonymous changes could (i) alter RNA stability and could control translation, and (ii) in the case of micro-RNA (miRNA), their binding site on their respective target mRNAs may change or may not be accessible because of RNA editing [2]. Although in this thesis, our focus is only on mRNAs, RNA Editing occurs in a wide range of RNAs; long non-coding RNAs (lncRNA), small nucleolar RNAs (snoRNA), small RNAs (sRNA), transfer RNAs (tRNAs), ribosomal RNAS (rRNAs), to name but a few.

1. 2. RNA editing Importance:

RNAs can be divided into protein coding RNAs, also called messenger RNAs (mRNAs), which are involved in protein synthesis, and non-coding RNAs, which regulate gene expression. RNAs interact with proteins and DNA [6]. These interactions allow them to have a variety of functions and play a major role in the carrying of ATP energy, metabolism, cell signalling and other pathways [7]. RNA editing is a natural and important regulatory process, with the brain being the most highly RNA edited organ in the human body, followed by the heart. Interestingly, while the heart has the highest RNA edited exosomes, the brain holds the highest intronic RNA editing incidents (Huntley et al., 2016). However, aberrant RNA editing has been implicated in various human diseases. For example, in cardiovascular diseases, RNA editing dysregulation has been associated with conditions such as atherosclerosis and heart failure [8]. In metabolic disorders, RNA editing has been linked to conditions like obesity and diabetes. In immune disorders, it has been shown to play a role in autoimmune diseases and chronic inflammation [9]. In cancer, RNA editing is involved in cancer progression and drug resistance [10] [11]. Finally, in neurological disorders, such as epilepsy, schizophrenia, and autism, RNA editing has been identified as a potential contributor [12] [13] [14]

1. 3. Role of RNA editing in cardiac diseases:

Dilated cardiomyopathy (DCM), is a heart disease caused by the stretching, thickening or stiffening of its wall muscles [15]. This makes it harder for the heart to pump blood around the body. Thereby, those affected by it, are at higher risk of heart attack, which is one of the leading causes of death worldwide [16, 17].

In cardiovascular diseases, Stellos *et al*. demonstrated ADAR-1-mediated changes in the mRNA expression and stability of cathepsin S (a protease) in endothelial cells, with a major consequence on the angiogenesis and atherosclerosis processes [18] ADAR-1 is essential in maintaining cardiac function [19]. This has been demonstrated using mice with cardiomyocyte-specific *Adar-1* gene knock-out [20]. Kokot *et al.* demonstrated that ADAR-2 mediates A-to-I RNA-editing in the human heart and regulates key pathways in heart failure [21]. Another recent study showed that ADAR-2 is up-regulated in exercise training and protects the heart from MI and doxorubicin-induced cardiotoxicity in rats [22]. Several questions remain, including the

mechanisms of action of ADAR-1 and -2 within the diseased heart and their relevance to human disease.

In the lab of my supervisor, Dr Prashant Srivastava, while investigating the role of RNA editing in ischemic heart disease associated with Diabetes Mellitus. The preliminary results from RNA sequencing of left ventricle biopsies from cardiac surgery patients as part of the ARCADIA study (Association of Non-coding RNAs with Coronary Artery Disease and Type 2 Diabetes, REC 13/LO/1687) have shown an A-to-I RNA editing signature in the diabetic heart that affects mRNA transcripts of genes contributing to mitochondrial dysfunction. This is exciting as mitochondrial dysfunction is strongly linked to type 2 Diabetes Mellitus associated cardiac complications [23]. Hence, understanding the complexity underlying the biological mechanisms that may lead to cardiac complications and providing targeted treatments may decrease those risk factors in the future.

1.4. Hypothesis and Aims

Our initial results lead to the novel hypothesis that changes in RNA editing levels drive cardiac complications leading to heart failure, for example, dilated cardiomyopathy. To investigate this hypothesis further, our project aims to achieve three key objectives: (i) develop a framework for profiling RNA editing events using single-cell transcriptomic datasets, which can help us identify specific cell types involved in RNA editing dysregulation, (ii) generate a comprehensive profile of RNA editing events and their impact on RNA structure in the human heart, and (iii) characterise RNA editing events that are altered in non-failing hearts and patients with DCM. By achieving these objectives, we hope to gain a deeper understanding of the molecular mechanisms underlying RNA editing dysregulation in the heart, which could ultimately lead to the development of more effective treatments for cardiac complications.

1.5. Timeliness

It is important to note that the transcriptome is a dynamic and flexible area, both in terms of its content and structure, unlike its counterparts, the genome or the

proteome. This makes the transcriptome a promising target for drug development, as any unwanted effects may be easily reversible. Understanding the dysregulation of RNA editing and its impact on health could lead to improved drug development, successful precision medicine, and affordable personalised treatment. RNA editing is a rapidly growing field of research. While there is a wealth of scientific literature on RNA editing and its relationship to pathology based on nucleotide content, there is a lack of documentation on RNA editing's structural pathology [24].

2. Methods

2. 1. Single-cell data processing and classification of cell types

The single-cell sequencing data, I used, was published by Koenig et al. [25]. They sequenced 35 samples from 11 donors with Dilated Cardiomyopathy (DCM) and 24 without DCM. Their age ranged between 8 and 75. Out of 35 samples sequenced, 19 were females, and 16 were males. All samples' tissues came from the heart's Transmural Left Ventricular Apex and they were collected after heart transplant or Assist Device installation. GPL24676 Illumina NovaSeq 6000 platform was used to process them all. The raw sequenced files (fastq) were downloaded and processed by Ji Jiahui (PhD student in Dr Srivastava's lab). Briefly, the fastq files were mapped to human genome version hg38 using CellRanger version 6.0.2 [26]. The single-cell data were processed and mapped to respective cell types.

All data processing in this thesis was performed using Imperial College HPC (High-Performance Computing) under the conda environment. Sinto version 0.9.0, a single-cell analysis software, takes classified cell barcodes as input and splits the post-sorted genome bam files into their respective cell types [27]. In essence, it reads the bam file using a list of cell barcodes per cell type and produces several bam files, one for each cell type found. Next, the resulting bam files were sorted and indexed using samtools [28].

2. 2. Predicting RNA editing from single cell types

2. 2. 1. Predicting RNA edited sites

I used Reditools version 2.0 [29] to estimate all modifications per chromosome per cell type. Reditools was called with the following options:

-S: only sites with edits will be included in the output file.

-s 2: first strand oriented.

-q 20: Reads whose mapping quality is below 20 are discarded.

-bq 25: Bases whose quality is below 25 were not included in the output.

Once the results were collated for each cell type, they were arranged in a table, please see the table below as an example of the output from Reditools (Table 1):

Table 1:

Region	Position	Reference	Strand	Coverage-q30	MeanQ	BaseCount[A,C,G,T]	AllSubs	Frequency
chr4	362804	G	1	22	35.91	[0, 1, 21, 0]	GC	0.05
chr7	55766687	A	1	18	34.33	[0, 0, 1, 17]	AT AG	1.00

Table 1: Predicting RNA edited positions example.

2. 2. 2. Filtering process to get quality RNA edited sites

Prediction of RNA edited sites is error-prone. Therefore, I set to filter the output produced by Reditools with the following criteria:

a) I removed all file rows which contained more than two nucleotides in AllSubs (See Table 1)

b) I removed all file rows with less than10 the total base count, i.e. Coverage-q30 < 10 for each row (See Table 1)

c) I remove all file rows with less than five counts per base of interest. (See Table 1)

d) I collated all unique RNA Edited positions per cell type. One file for each cell type.

2. 2. 3. Removal of SNPs (Single-Nucleotide Polymorphism) from all files per cell types

The samples profiled in the scRNA sequencing dataset were not genotyped. So, we didn't know if the RNA/DNA mismatch predicted by Reditools were real or because of SNPs. Therefore, in addition to the above filtering criteria, I also removed the SNPs from the predicted sites by using the following steps:

a) I made the previously filtered files, containing all the unique RNA Edited positions per cell type, into BED files (a file format) as in the following table example:

chr1	10518	10519
chr1	10531	10532
chr1	10532	10533

Table 2: Predicted RNA edited positions as BED file.

a) I used Linux Intersect utility and the BED files version of NCBI SNPs database, to get all the SNPs for each cell type.

b) I removed all these SNPs from all files from each cell type.

c) I filtered out all that is not A-to-I or C-to-U content from all files.

In other words, I only kept that which on the positive strand had AG or CT in AllSubs (See Table 1), and on the negative strand, that which had TC or GA in AllSubs. All other entries were discarded.

d) I removed RNA incidents that occurred less than 10 times in occurrence for each position in all the sample data.

e) As each sample has an associated health state and a total number of cell types associated with it, I collated all the data for each health state and each cell type, labelling each case appropriately.

2. 2. 4. Differential RNA editing

To perform differential RNA editing analysis, I used Wilcoxon Test (Mann-Whitley U test) to get the p-value for each position per cell type for the following groups:

1. Donors with no DCM (Dilated Cardiomyopathy) vs Donors with DCM
2. Donors with no DB (Diabetes) vs Donors with DB

3. Donors with no DCM and no DB vs Donors with no DCM and DB
4. Donors with DCM and no DB vs Donors with DCM and DB

Wilcoxon Test is a non-parametric alternative to the paired T-test, and Mann-Whitley U test is its implementation in a Python environment. As the sample data, I am handling, is not large and is not normally distributed, Wilcoxon Test seems to be the most suitable to use in our case.

2. 3. RNA Structural Metrics (MFE and others)

To estimate the consequence of RNA editing events, I used the Vienna RNA package software [30] to estimate the RNA structural metrics. The Vienna Package offers several metrics to measure / quantify secondary structures. At its core is the RNA fold software that enables us to calculate the Minimum Free Energy (MFE), and other metrics, of RNA structures. RNAs can fold into many different structures, but the one I focused on is the secondary structure. RNA Secondary Structure is the RNA structure formed by pairing Guanine (G) with Cytosine (C) and Adenine (A) with Uracil (U).

It is important to note that there are many possible RNA Secondary Structures, but there is only one optimal RNA structure. The sum of all RNA possible structures constitutes and is referred to as the Boltzmann ensemble [30]. From a thermodynamics point of view, all systems tend to move towards their lowest energy state or optimal structure, but this may be affected by environmental factors. Hence a non-optimal structure may be adopted by an mRNA structure.

To provide a sense of data and files, in the following figures and tables, I used a mRNA transcript example, from our dataset, and have described its metrics below.

Gene name: F2R

Ensembl Transcript Stable ID: ENST00000505600

Transcript Sequence:

>ENST00000505600

AUGGGGCCGCGGCGGCUGCUGCUGGUGGCCGCCUGCUUCAGUCUGUGCGG
CCCGCUGUUGUCUGCCCGCACCCGGGCCCGCAGGCCAGUCACGUUUAGGUG
GGGCGUGCCACCCCCUUCGCGGGCCCAGCCGAUGCCCCUUUGGACUCGAUC
UUGGAGGGUGCAGCCCGCCUGCCACGGGGUGUUGGAUAUGGAGGAGGAUG
GAGCGGAAGCCCCCUGGGGGAGCCUGCAGUCCUGCGUUGCACUUGUCAUUG
UGUUUCUCCCAGGACCACCCCCAAAAAGAAAAGCUCUCACGUUGCUCCACCA
AUAA

Minimum Free energy (MFE):

MFE = -136.03 kcal/ml.

MFE Secondary Structure in dot-bracket notation:

The dots are for unpaired bases, '(' for paired 5' base and ')' is for paired 3' base.

...((((((((((((((((...))))))))))).))....(((((((.((((((..((.((.....)))).))))))))))))).....(((..(((.(((.....))))))..))))))))..((((((((((..(((.....((.((((...)))).))..))).))))))))))...(((((.((..((((((.....(((((((....((..((((((... .((((.....)))).....)))))))))))))).....)))).))..)).)))))).....

MFE Secondary Structure as a plot (Fig. 2):

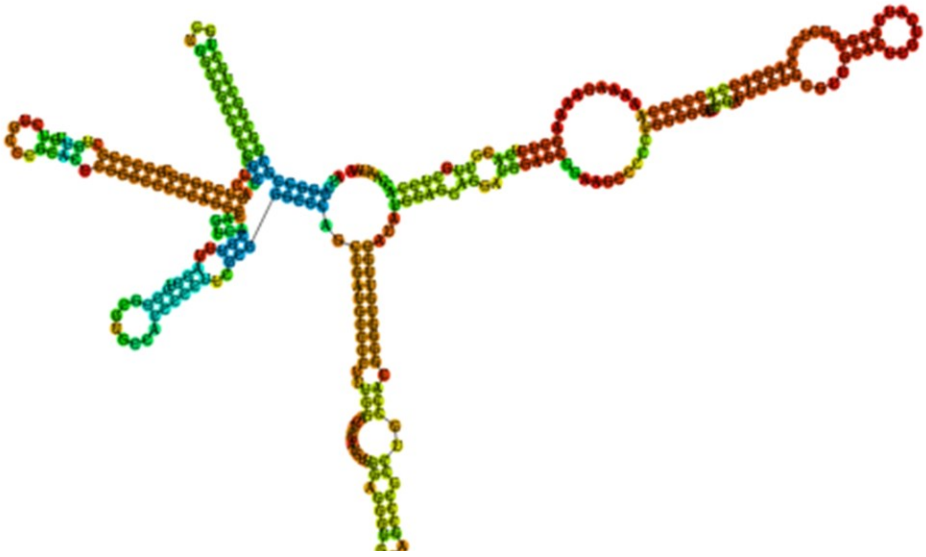


Figure 2
F2R MFE Secondary Structure

Frequency of the MFE structure:

The frequency of the MFE structure in the ensemble is 0.05 %.

The ensemble diversity:

It is the average base pair distance between all structures in the thermodynamic ensemble. The ensemble diversity is 72.68.

Centroid secondary structure:

It is the structure with a minimum total base-pair distance to all structures in the thermodynamic ensemble.

Centroid secondary structure in a dot-bracket notation:

.((((.(((((((((((.....)))))))))..((...(((((((.((((((.....(.........)...)))))))))))))).)).(((......)))..................)))
).)))).((((((((((..(((.....(((.(((((...))))).))..))).))))))))))...(((((.((..((((((.....((((((....((..((((((....((
((........))))......))))))))))))))).....)))))..))..)).)))))).....

Centroid secondary structure as a plot (Fig. 3):

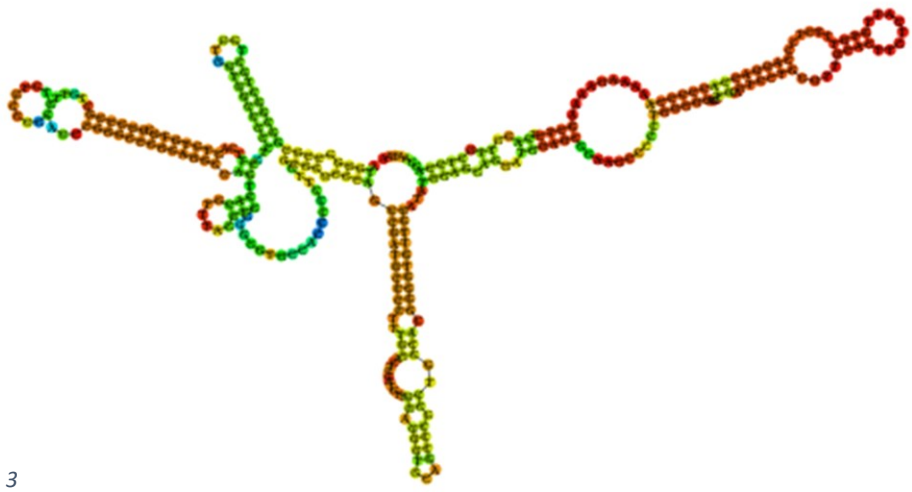


Figure 3
F2R Centroid Secondary Structure

Centroid Minimum Free Energy:

The Centroid's minimum free energy is -119.60 kcal/mol. The centroid distance is 47.48.

Partition Function (PF):

The partition function is a rough measure for the well-definedness of the MFE structure. PF = -138.07

Maximum Expected Accuracy (MEA)

MEA = 238.85

MEA energy is -123.10 kcal/mol

2. 4. Consequence of RNA editing

I installed the Ensembl VEP (Variant Effect Predictor) [31] on my local workstation, and then I formatted my data files in the Ensembl default format to serve as input. That helped me get the consequences of RNA-Editing of all the unique RNA Edited chromosome positions.

2. 5. Functional enrichments analysis

I only used the Dilated Cardiomyopathy data versus non-Dilated Cardiomyopathy data in my enrichment analysis as it is directly linked to Dilated Cardiomyopathy, and I only selected positions that had P values that were less than 0.05.

First, I used Ensembl VEP and Biomart, available online from Ensembl.org [31], to annotate RNA edited sites with their respective genes and transcripts. This allowed me to select only protein-coding transcripts. And, later on I checked those positions against the gencode human reference database [32] and again selected only those positions that belonged to protein coding genes.

For each cell type, I used its background gene list (i.e. all genes expressed in its respective cell type) and its corresponding significant gene list in my findings to do the

enrichment analysis. I used the WEB-based Gene SeT AnaLysis Toolkit (WebGestalt) [33] and selected the following options:

1. Method of Interest: Over-Representation Analysis (ORA)

    1. a. Geneontology

    1. b. Biological Process

2. Select Gene ID Type: Gene Symbol

    2. a. I uploaded my findings gene list

3. Select the ID type of reference set: Gene Symbol

    3. a. I uploaded my background gene list

4. In advanced parameters, I choose “FDR” as the significant level and have set the Maximum number of genes for my category accordingly.

I repeated the above four steps with “1. b.” option first as “Cellular Component” and then as “Molecular Function”.

I repeated all the above four steps with “1. a.” as “Pathway” with the option “Reactome” and then with the option “KEGG”.

Lastly, I repeated all the above four steps with “1. a.” as “Disease” with the option “Disgenet”, “GLAD4U”, and then “OMIM.”

GO and pathway terms were considered significant if FDR was less than 0.05.

3. Results

3. 1. Single-cell data processing and classification of cell types

The original sample data, I used, comes from 35 donors. Each donor’s heart had 15 cell types. Therefore, I expected a total number of 525 data files. However, few data files were missing; 2 files in Adipocytes, 10 files in B-Cells, and 14 files in Epicardium cells. They were found missing after splitting the bam files for each donor into their respective cell type. In total, we had 499 data files. I repeated the splitting process to be certain, and each time I got the same results. This is certainly due to how the tissues

were prepared for Single Nuclei sequencing. This is bound to happen at a practical level as well as the introduction and capture of ambient RNAs and other impurities. The quality check report for our BAM files, before splitting, was quite high. Please see Table 1 in the supplement files for more details.

3.2. Predicting RNA editing from single cell types

3.2.1 Predicting RNA edited sites

I used Reditools software options only to capture sites with RNA edits. I have discarded reads with mapping quality less than 20, and bases below 25 in quality. Please see the supplement file 'RNA_Edits-per-Donor-per-Cell_Type.tar.gz' for all my findings. Briefly, I have summarised my findings for each cell type below in (Table 3):

Cell_Type	A-to-I or C-to-U Edits count	Total counts including non A-to-I and C-to-U Edits
Adipocytes	315	1513
B_Cells	5	23
Cardiomyocytes	119723	402581
Endocardium	9307	35050
Endothelium	23531	86261
Epicardium	136	773
Fibroblasts	44410	154729
Lymphatic	245	1169
Macrophages	18720	65062
Mast_Cells	87	446
Monocytes	24	99
Neurons	128	644
Pericytes	10649	41993
Smooth_Muscle	681	3378
T_NK_Cells	373	1805

Table 3: Predicted RNA edited positions summary.

3. 2 .2. Filtering process to get quality RNA edited sites

1. I removed all rows which contained more than two nucleotides in AllSubs (Table 4):

Cell Type	Total rows removed containing more than 2 nucleotides in AllSubs	Overall total
Adipocytes	64942	6163720
B_Cells	3438	382951
Cardiomyocytes	5981182	175965035
Endocardium	602070	40853091
Endothelium	944518	63315841
Epicardium	26986	2832560
Fibroblasts	1983678	107778615
Lymphatic	59937	6400386
Macrophages	837865	60874906
Mast_Cells	25321	3389583
Monocytes	12599	1236727
Neurons	26800	3561350
Pericytes	532433	36132278
Smooth_Muscle	90743	8969896
T_NK_Cells	65275	6685482

Table 4: Summary of RNA edited positions with only 2 nucleotides in AllSubs.

2. I removed all rows with less than ten total base count, i.e. Coverage-q30 < 10 for each row (Table 5):

Cell Type	Total rows removed containing less than 10 in total coverage	Overall total
Adipocytes	3714693	6098778
B_Cells	250106	379513
Cardiomyocytes	57456342	162619319
Endocardium	20777459	38407927
Endothelium	30750556	57151069
Epicardium	1877442	2805574
Fibroblasts	45369707	94292764
Lymphatic	4188036	6010152
Macrophages	33088858	58056705

Mast_Cells	2296712	3290647
Monocytes	901588	1218655
Neurons	2392279	3361270
Pericytes	19025979	34579619
Smooth_Muscle	5773362	8613967
T_NK_Cells	4412965	6501547

Table 5: Summary of RNA edited positions removed with Coverage-q30 < 10

3. I removed all rows with less than five counts per base of interest (Table 6).

Cell Type	Total rows removed with less than 5 base counts	Overall total
Adipocytes	1836112	2384118
B_Cells	98767	129432
Cardiomyocytes	87943043	105163010
Endocardium	14183904	17630501
Endothelium	21129971	26400546
Epicardium	721433	928153
Fibroblasts	39744475	48923090
Lymphatic	1392775	1840549
Macrophages	18932926	23383049
Mast_Cells	735098	986806
Monocytes	234673	308157
Neurons	712915	969523
Pericytes	12459221	15579593
Smooth_Muscle	2107162	2780307
T_NK_Cells	1528787	1999012

Table 6: Summary of RNA edited positions removed with BaseCount < 5.

3.2.3 Removal of SNPs from all files per cell types

I removed all SNPs from the summary files I got after the previous filtering process, and about 50% of genomic positions were found to be SNPs. Please see the following table (Table 7):

Cell Type	Total Chr Positions including SNPs	Total Chr Positions excluding SNPs	no SNPs / SNPs %
Adipocytes	462309	258048	55.8%
B_Cells	29034	14736	50.8%
Cardiomyocytes	9182046	3776664	41.1%
Endocardium	2434813	982409	40.3%
Endothelium	3350076	1418525	42.3%
Epicardium	184487	99444	53.9%
Fibroblasts	5908402	2461200	41.7%
Lymphatic	386625	210848	54.5%
Macrophages	2922086	1212052	41.5%
Mast_Cells	218207	119721	54.9%
Monocytes	66861	35210	52.7%
Neurons	230016	127710	55.5%
Pericytes	1977250	808602	40.9%
Smooth_Muscle	535389	304613	56.9%
T_NK_Cells	386800	217525	56.2%

Table 7: Summary of RNA edited positions with removed SNPs.

3. 2. 4. RNA editing Total count per cell type:

I found that by using the RNA edit total count, the cardiac cell types could be classified into two groups:

1. **Highly RNA Edited cells** i.e., cell types with over 1000 RNA Edit counts, and they are in the following order:

1. Cardiomyocytes: A cardiac muscle cells with high mitochondrial content.
2. Fibroblasts: Common cells of connective tissue.
3. Endothelial cells: cells that line the interior single layer of blood and lymph vessels.
4. Macrophages: Innate immune system cells that destroy pathogens.
5. Pericytes: multi-functional mural cells

6. Endocardium cells: Cells that line the innermost layer of the heart cavities.

2. **Relatively Low RNA Edited cells** i.e., cell types with about 1000 or less RNA Edit counts and they are in the following order:

1. Smooth Muscle cells
2. T NK Cells: a subset of T cells which have properties of natural killer cells.
3. Adipocytes: fat cells used for storing energy.
4. Epicardium cells: cells that line the upper layer of the heart wall.
5. Lymphatic cells.
6. Neuron cells which have high mitochondrial content.
7. Mast Cells: cells that are part of the immune system.
8. Monocytes: a type of white blood cells

Please note that B cells are not represented here. It would be safe to infer from our findings that B Cells are the least RNA Edited heart cells. Indeed, we found less Epicardium cells in the original bam files than B Cells, yet RNA Editing in Epicardium is well represented. Furthermore, it would be reasonable to infer from our findings that the number of RNA Edited transcripts are higher in Epicardium cells than in Adipocytes, as 14 Epicardium cells are missing against only 2 Adipocyte cells.

Please see the following figure for more details (Fig. 4):

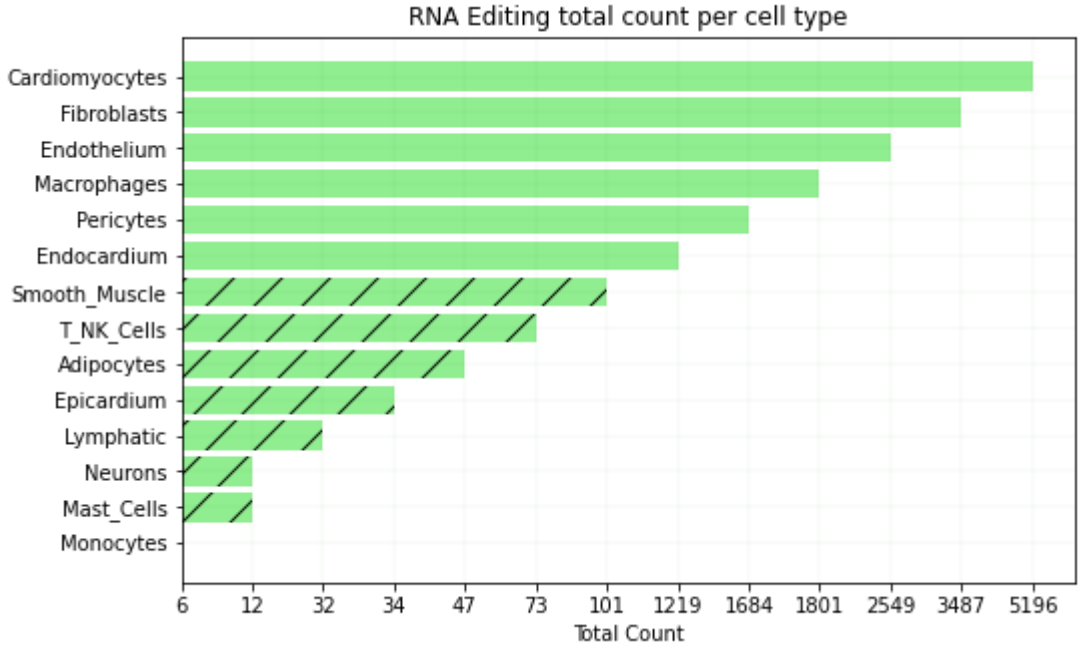


Figure 4:
Summary ofl RNA edit total count per cell type. The highly edited cells, in green bars with a total RNA edit count between 1219 and 5196. While relatively low edited cells, in hashed green bars with a total RNA edit count between 6 and 101.

3. 2. 4. 1. RNA editing (A-to-I and C-to-U) total counts per cell type in healthy state only –

In this study, I label a sample as “Healthy” only when they are not having Dilated Cardiomyopathy (DCM) or Diabetes. This is a digression from the original article [25] where they labelled a sample as “Healthy” when it was not from a donor with Dilated Cardiomyopathy (DCM) only.

Our results suggest that A-to-I is the dominant RNA editor which is also found in published scientific papers [34]. Moreover, C-to-U RNA editing looks completely absent in Monocytes and Neurons in our findings. However, in Adipocytes and Macrophages, C-to-U takes the lead as an RNA editor.

Please see the following two figures (Fig. 5 and 6):

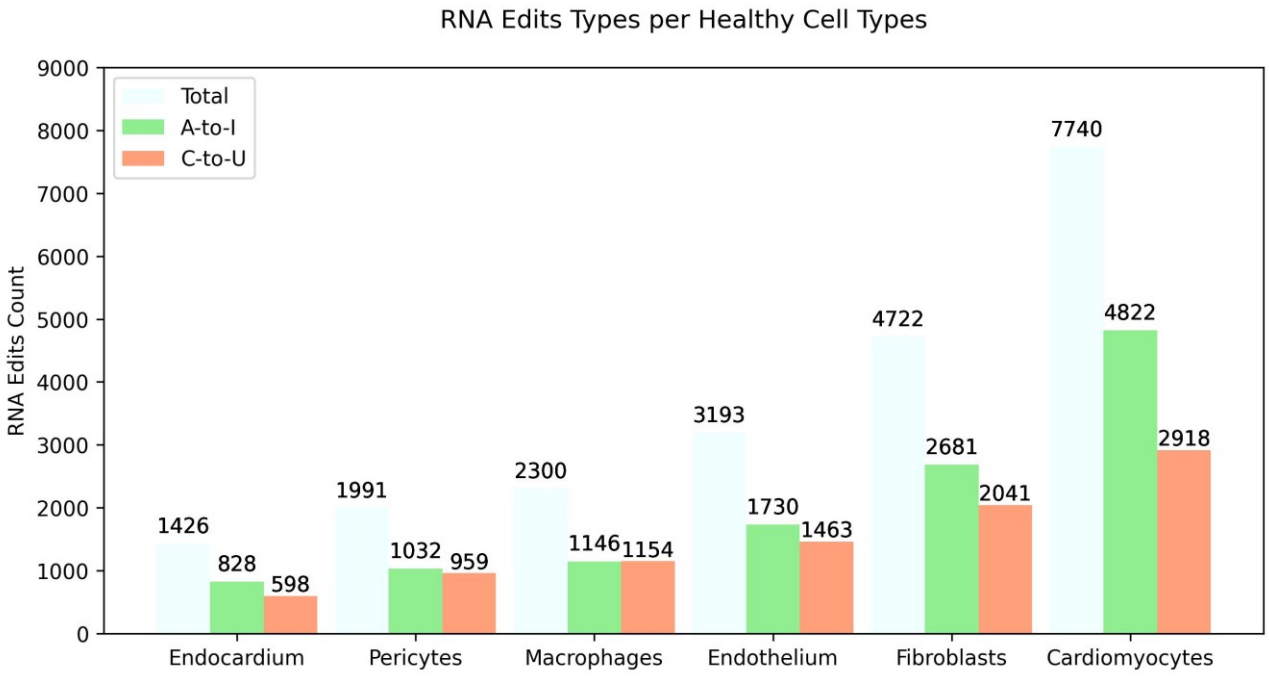


Figure 5
Summary of A-to-I and C-to-U total counts for highly edited healthy heart cells.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.

All these highly RNA edited cells have both type A-to-I and C-to-U edits. There is more of A-to-I edits that C-to-U. For instance, Cardiomyocytes exhibits 4822 RNA total edits against 2918 C-to-U total edits. There is one exception though: Macrophages with more C-to-U edits than A-to-I.

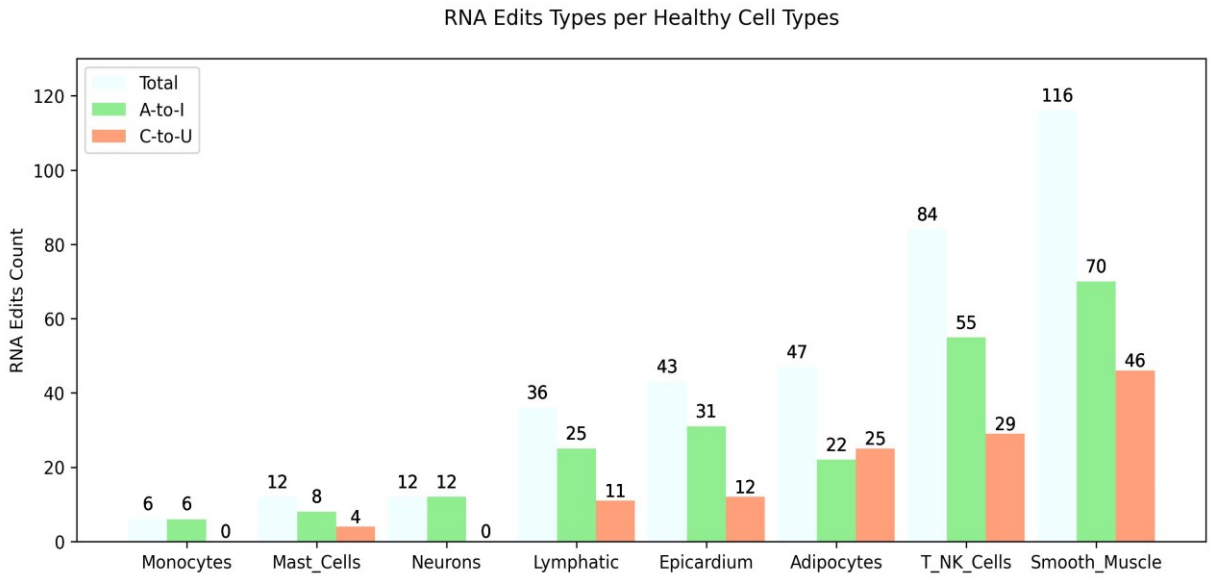


Figure 6
Summary of A-to-I and C-to-U total counts for relatively low edited healthy heart cells A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.

In all these relatively low edited cell types, there is a presence of A-to-I editing and it is the dominant one. As to C-to-U editing, it is also present, but it seems completely absent in Neurons and Monocytes.

3. 2. 4. 2. RNA Editing (A-to-I and C-to-U) total counts per cell type in all health states

In the following figures, I would refer to a healthy state as noDCM noDB i.e., no Dilated Cardiomyopathy and no Diabetes. All the other states are compromised and unhealthy. There is a clear dysregulation of A-to-I editing in patients with Dilated Cardiomyopathy. Additionally, Diabetes adds to the complexity of this dysregulation.

Please see the following 14 figures (Fig 7 to 20) for more details. In the figures, the green bars represent for A-to-I editing, the light red bars for C-to-U editing and the azure-coloured bars are for the sum of A-to-I and C-to-U edit counts. The number at the top of each bar is for the total count of RNA Edits. For example, in Cardiomyocytes "DCM DB" (Dilated Cardiomyopathy and Diabetes cells) case, there are 2206 C-to-U RNA Edits, 2805 A-to-I RNA Edits; and a total of 5011.

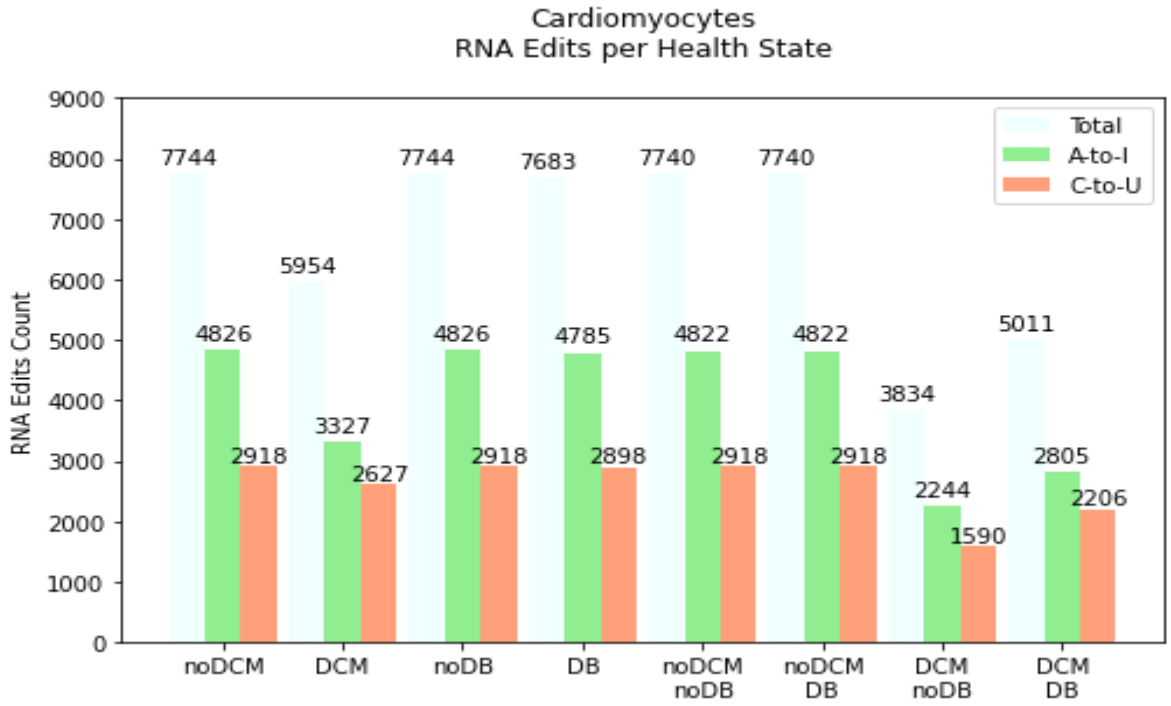


Figure 7 Cardiomyocytes RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

In Cardiomyocytes, there is a total of 7740 RNA edits for healthy individuals (noDCM noDB) with 4822 A-to-I edits and 2918 C-to-U total edits. In contrast these numbers are about halved in unhealthy cases such as individuals with no Diabetes, but with Dilated Cardiomyopathy (DCM noDB).

There is no recorded change, in my findings, in total RNA edit counts between healthy individuals (noDCM noDB) and Diabetic individuals with no Dilated Cardiomyopathy (noDCM DB).

We can imply from this that there is a major dysregulation of Cardiomyocytes' A-to-I and C-to-U editing in patients with Dilated Cardiomyopathy.

Our analysis of RNA editing in different cell types revealed interesting patterns. In fibroblasts (Figure 7), we observed a decrease in A-to-I and C-to-U editing in unhealthy cases (DCM noDB) compared to healthy cases (noDCM noDB). In endothelial cells and macrophages (Figures 8 and 9), there was also a change in RNA editing between healthy and unhealthy states. In macrophages, there was a clear distinction between healthy and unhealthy states, with a more pronounced decrease in A-to-I and C-to-U editing in the unhealthy case. Pericytes (Figure 10) and endocardium cells (Figure 11) also showed a decrease in RNA editing in unhealthy cases, particularly in C-to-U editing. Smooth muscle cells (Figure 12) exhibited a decrease in A-to-I editing in healthy individuals (DCM noDB) compared to unhealthy ones, and C-to-U editing remained stable except for DCM DB cases where it was affected. T-NK cells (Figure 13) showed no significant difference in A-to-I editing between healthy and unhealthy cases, but there was a noticeable decrease in C-to-U editing, which was even more pronounced in cases of DCM with diabetes. Adipocytes (Figure 14) exhibited a noticeable decrease in A-to-I and C-to-U editing in unhealthy cases, with a complete absence of C-to-U editing in complex cases of DCM and DB. Epicardium (Figure 15) also showed a decrease in A-to-I and C-to-U editing in unhealthy cases, with a serious drop in C-to-U editing levels. Lymph (Figure 16) showed no change in RNA editing levels. The case of patients with diabetes only having a drop in C-to-U editing needs further investigation. Neurons (Figure 17) were slightly affected by A-to-I editing in unhealthy DCM cases but showed no C-to-U editing. In mast cells (Figure 18), there was no change in C-to-U editing in all health states, but a clear drop in A-to-I editing in DCM cases. Monocytes (Figure 19) showed a striking absence of C-to-U editing, and A-to-U editing remained stable throughout all health states.

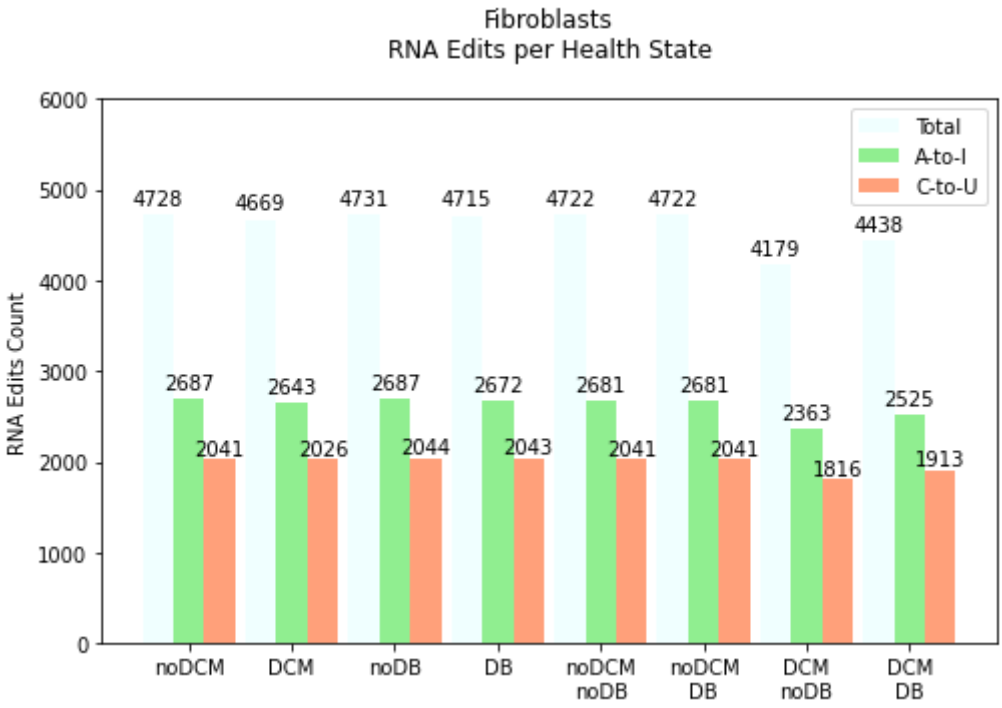


Figure 8 Fibroblasts RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

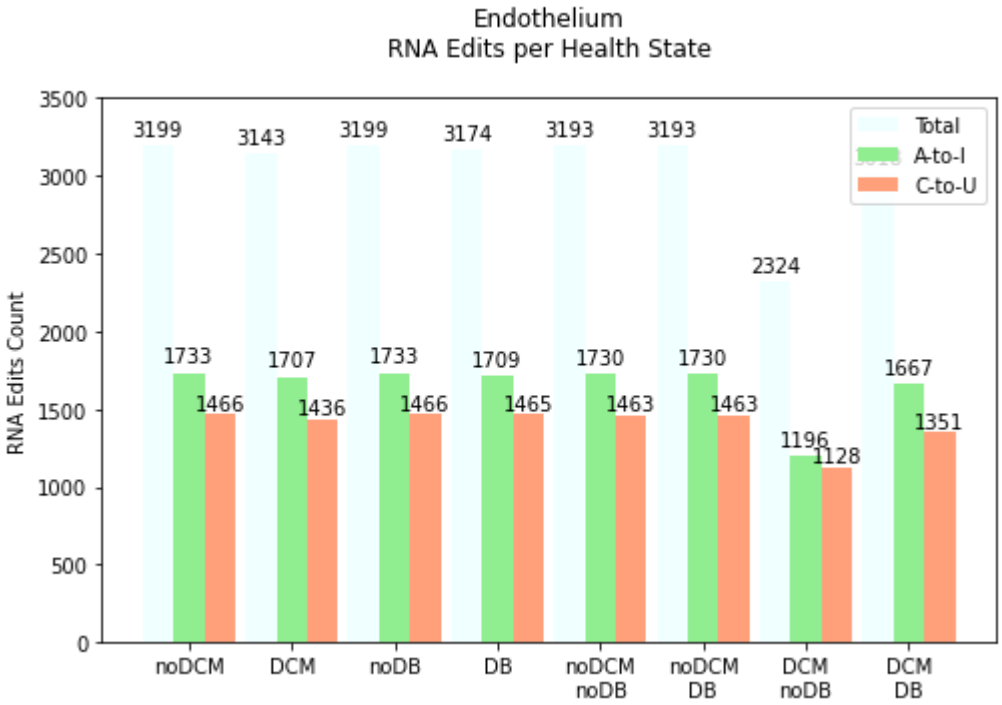


Figure 9 Endothelium RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

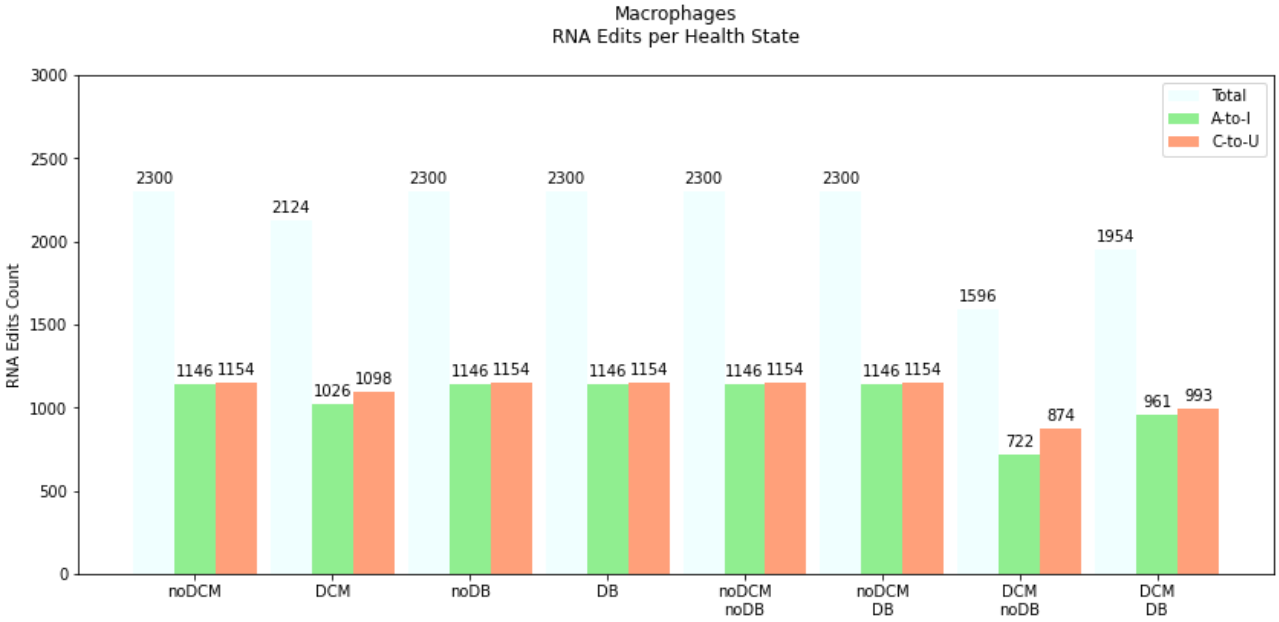


Figure 10 Macrophages RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

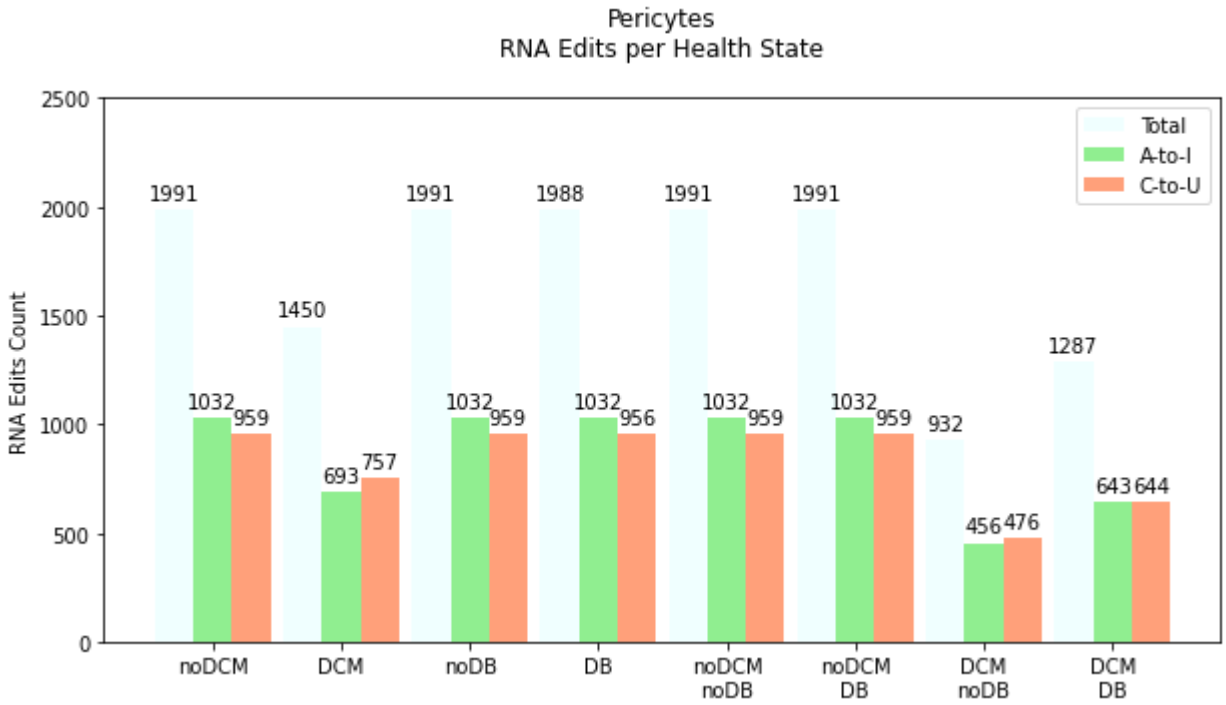


Figure 11 Pericytes RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

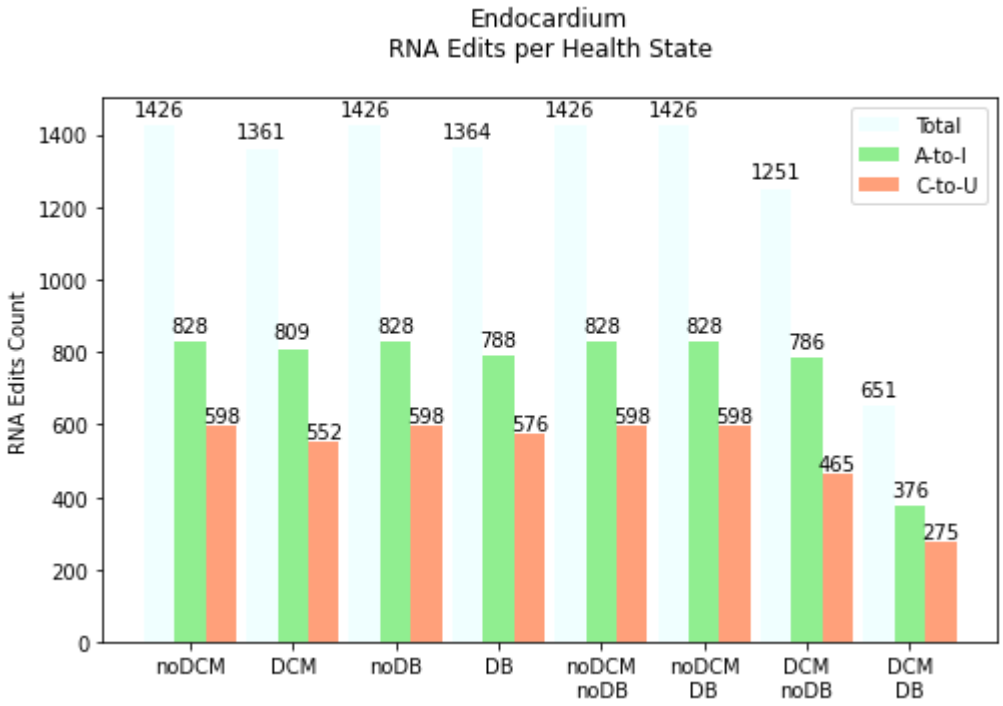


Figure 12 Endocardium RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

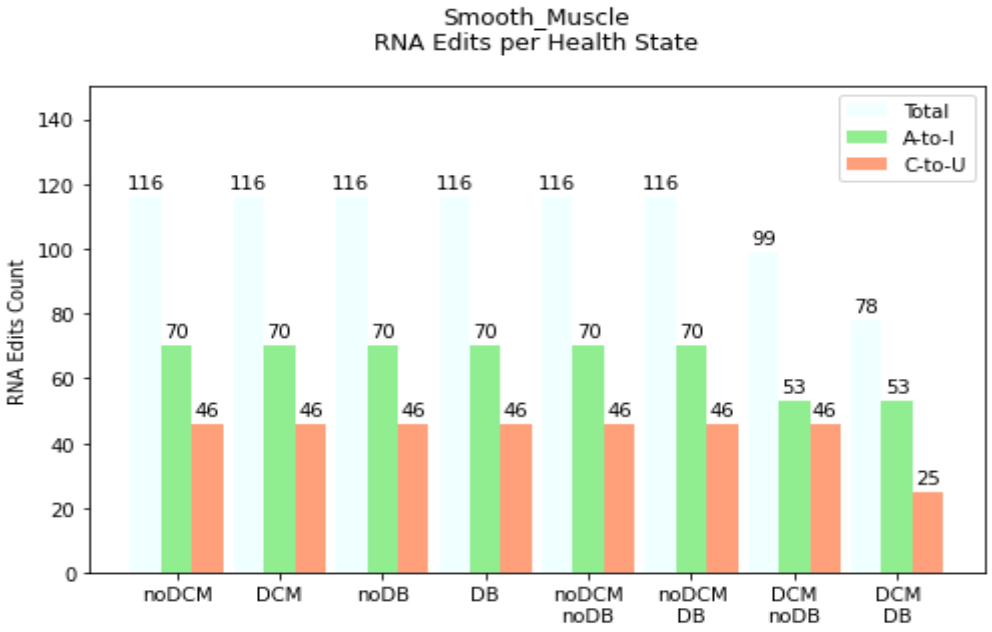


Figure 13 Smooth muscle RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

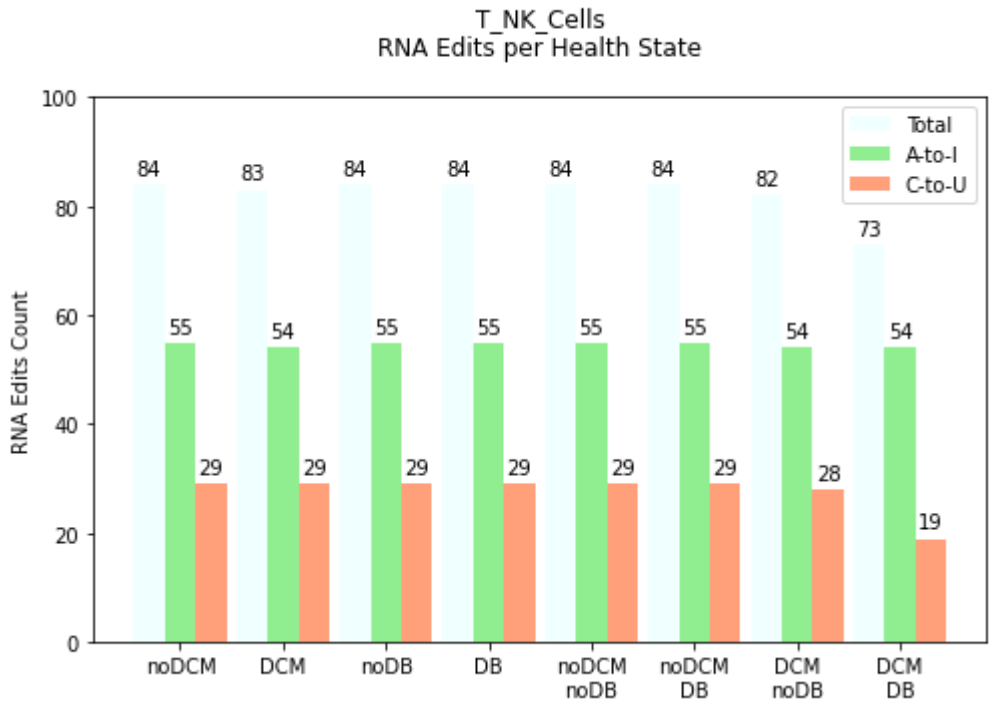


Figure 14 T NK cells RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

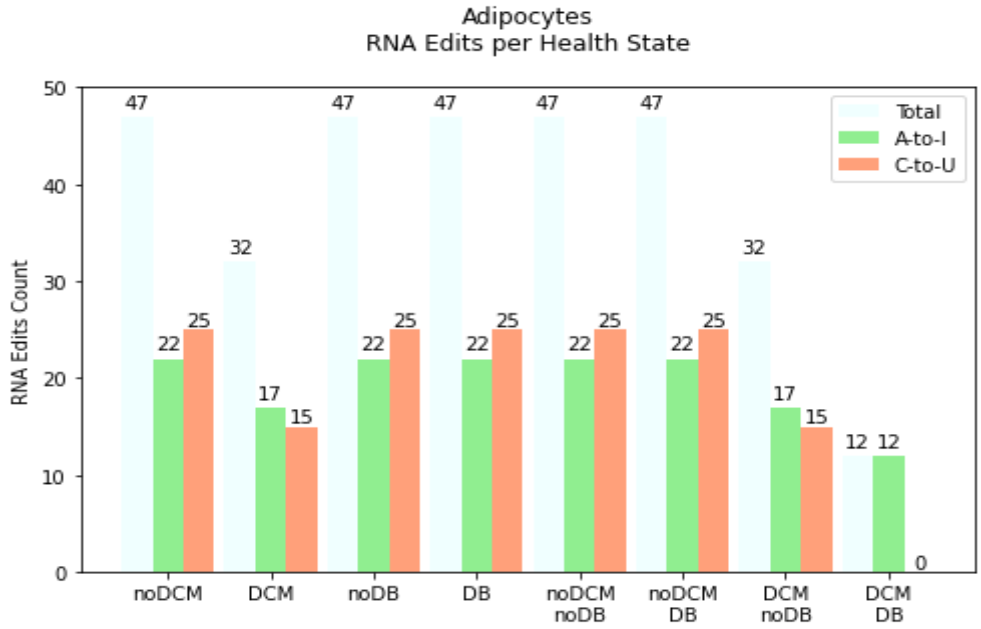


Figure 15 Adipocytes RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

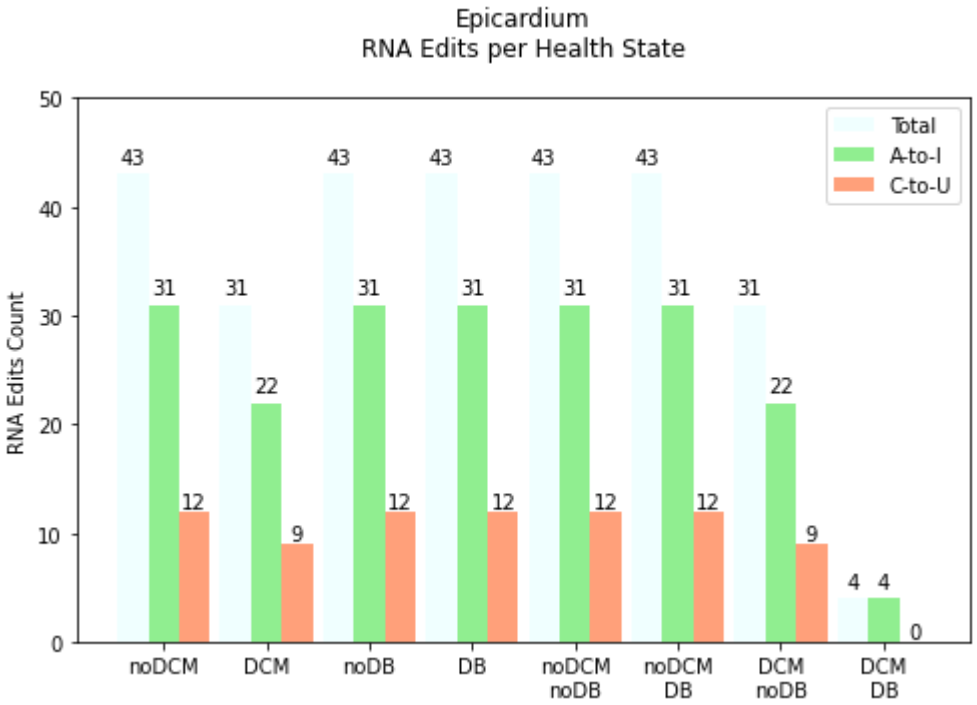


Figure 16 Epicardium RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

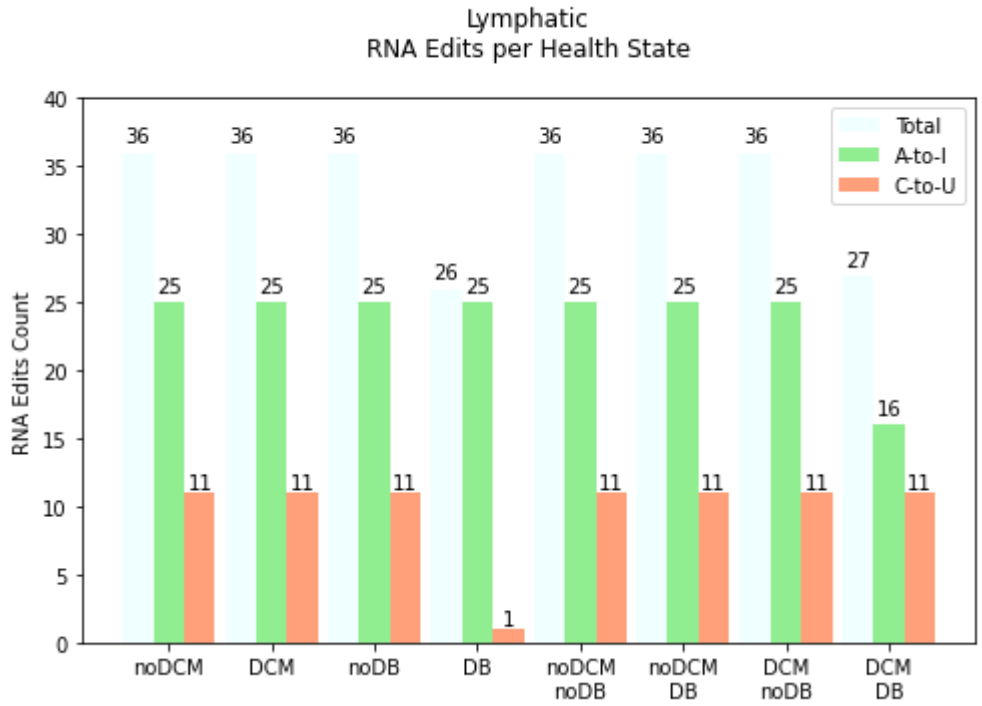


Figure 17 Lymphatic RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

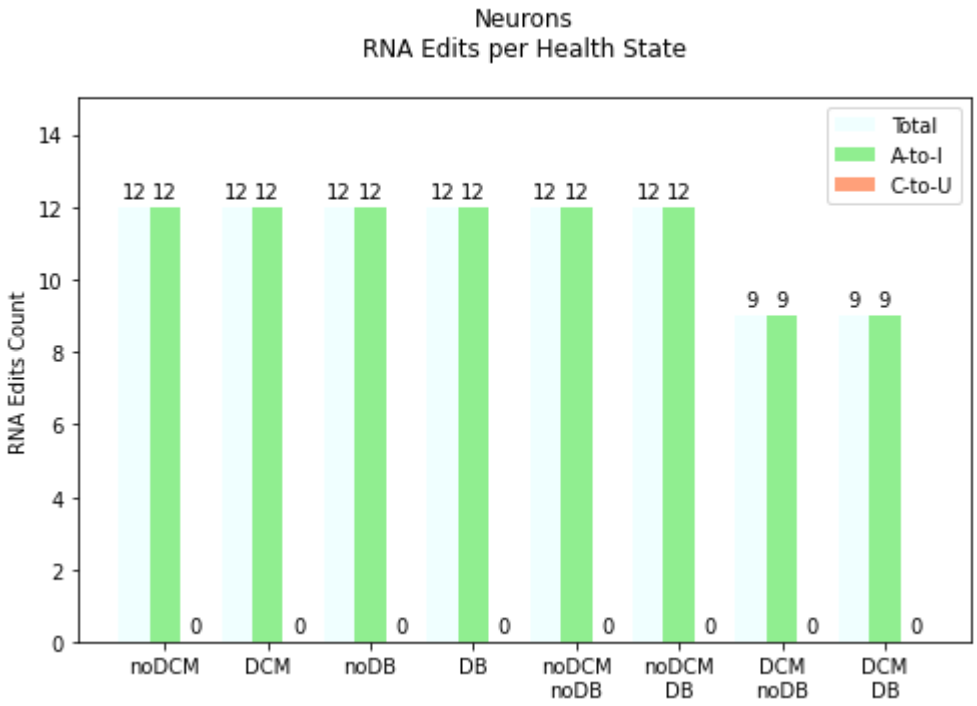


Figure 18 Neurons RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

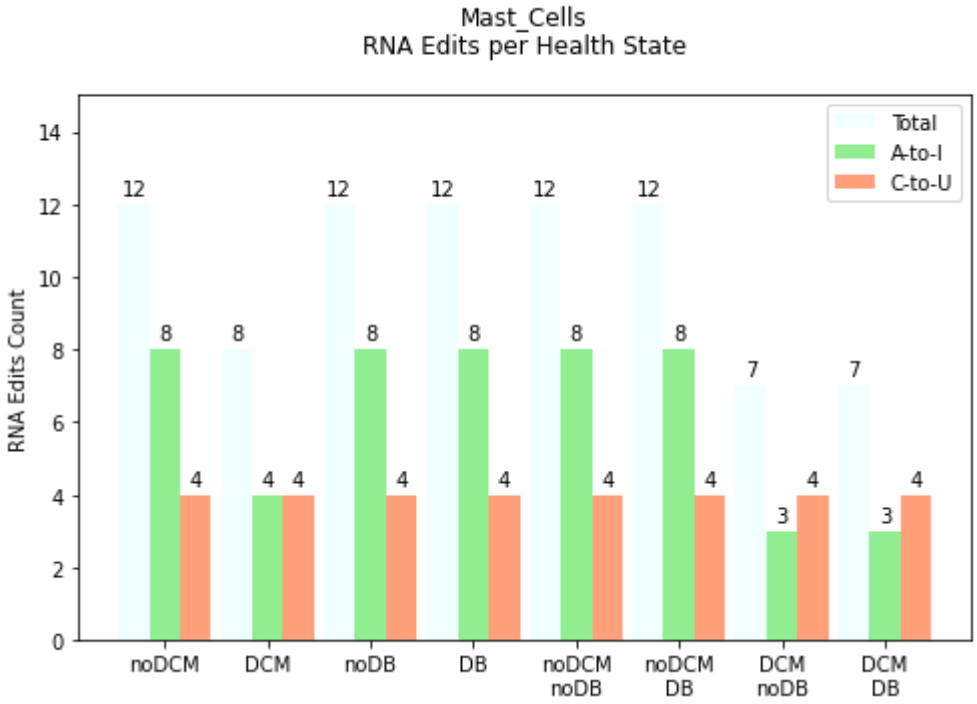


Figure 19 Mast cells RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

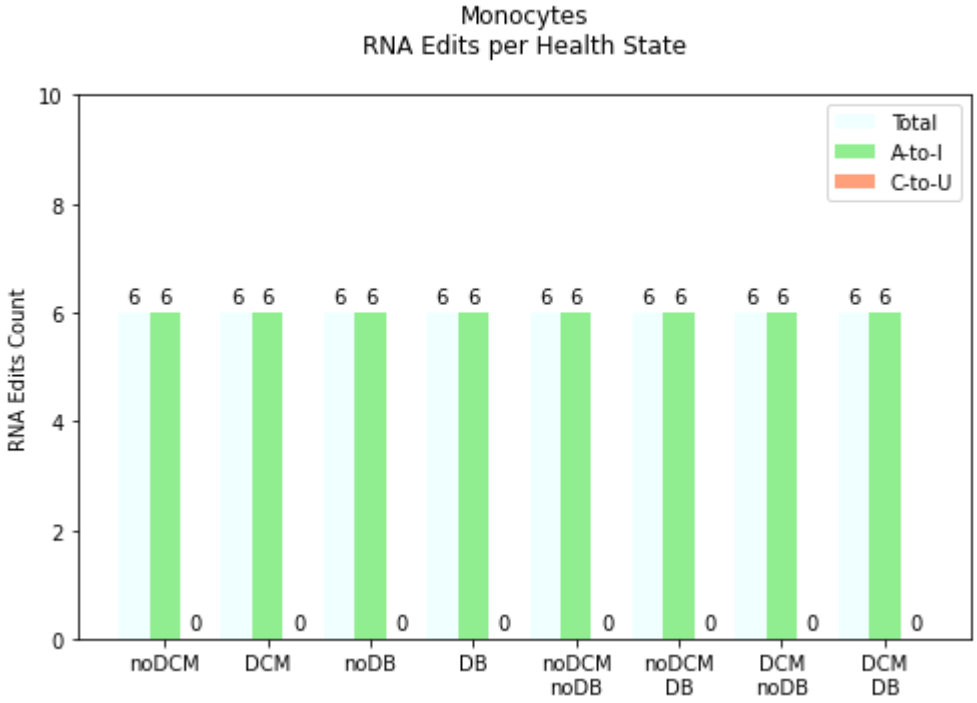


Figure 20 Monocytes RNA edits per health state.
A-to-I RNA editing total counts in green bars. On top of the bar is the total count.
C-to-U RNA editings total counts in red bars. On top of the bar is the total count.
Total of all edit counts, in azure bars, irrespective of it being A-to-I or C-to-U. On top of the bar is the total count.
DB = Diabetes; DCM= Dilated Cardiomyopathy.

3. 3. mRNA structure and health post RNA editing

When calculating the structural metrics, I only selected messenger RNAs. These are transcripts with a quality RNA Editing position whose Ensembl VEP predicted their modification consequence and their cDNA position.

3. 3. 1. mRNA structural metrics for unedited transcripts:

I used the Vienna package [30] to calculate 8 structural metrics for 15411 ‘unedited’ transcripts, and I compared them using a scatter matrix plot. I found that there was no apparent additional information that can be added to the minimum free energy (MFE) values. Please see figure 21.

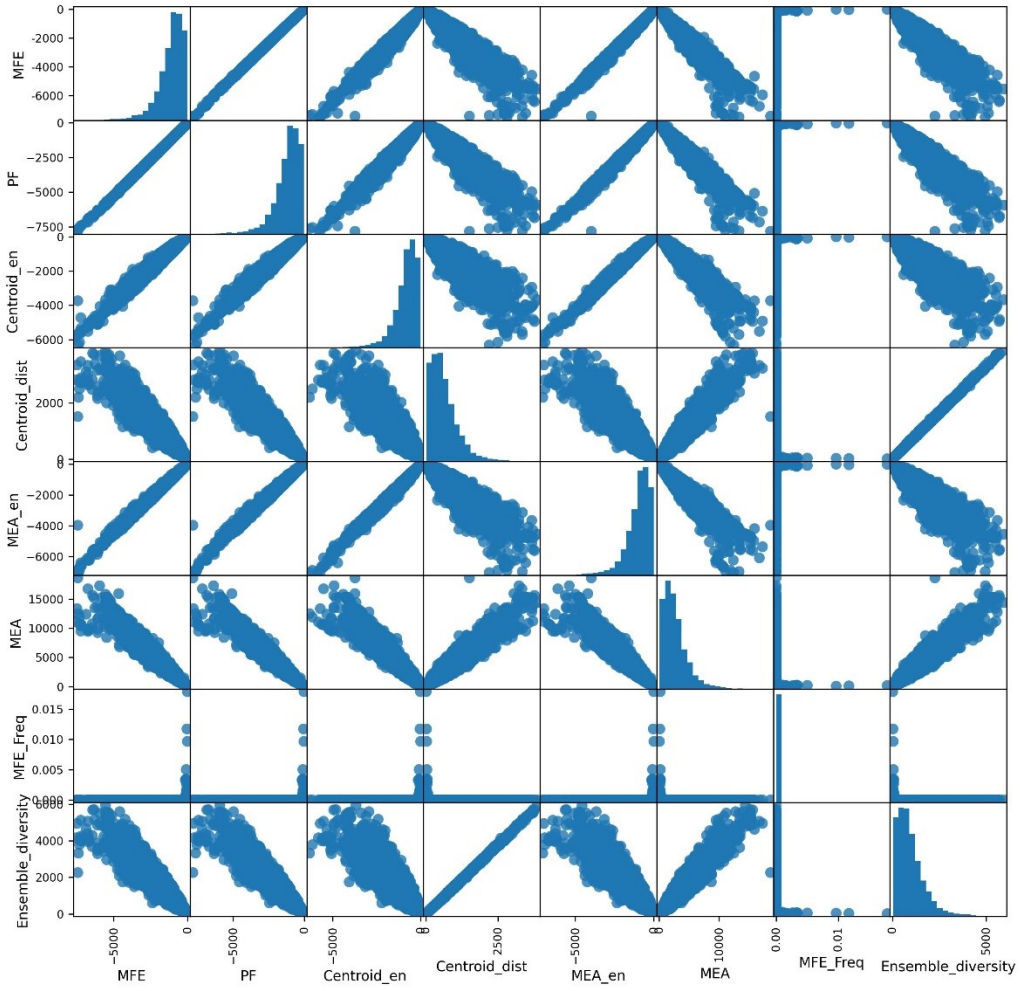


Figure 21 Unedited transcripts structural metrics

3. 3. 2. RNA Structural Metrics for edited transcripts:

I took all the RNA Edited transcripts from all cells, removed all duplications, and removed any transcripts that did not have an entry from a health state. Afterwards, I compared the 8 metrics using a scatter matrix plot for each health state. The results did not show any significant addition to the MFE values. Please see figure 22 for DCM cases. All other figures for all other health states will be made part of the supplementary files.

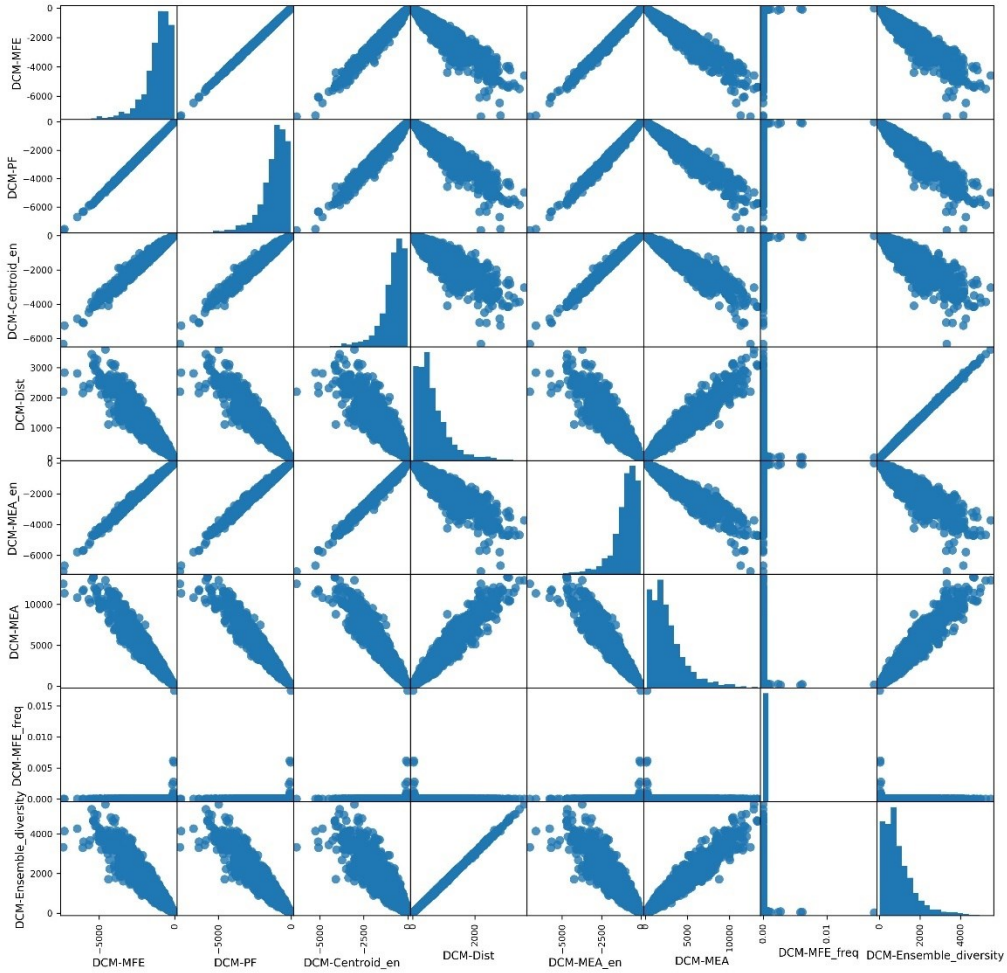


Figure 22 DCM transcripts structural metrics

Based on the above findings, I only used MFE metric to assess the structural stability of an mRNA.

3. 3. 3. Healthy mRNA structures post RNA editing:

RNA Editing is a normal regulatory process, and it is important to know if an RNA editing event will increase or decrease the stability of RNA transcripts. Hence, I set a task to calculate the Minimum Free Energy (MFE) for healthy cells, that is, cells from Donors with no Dilated Cardiomyopathy or Diabetes. I repeated the same task but for cells from unhealthy donors. It is important to observe that the lower is the MFE, the more structurally stable is the transcript.

In the following two figures (Fig. 23 and 24), each cell has two bars:

1. 'no editing' bar: this refers to the raw transcript before editing. The number in the middle of the bar is the total number of RNA-Edited transcripts for a given cell type. e.g. 5196 is the total number of mRNA transcripts that I found to be RNA-Edited in Cardiomyocytes.

2. 'Cell type name with it associated ΔMFE' bar: this refers to the changes to the MFE of an mRNA structure post-RNA-Editing. For example, in Cardiomyocytes, I found 788 mRNA transcripts with no change to their MFE, i.e. There was no change to the stability of their MFE structure. 1431 changed, but they had a higher MFE value. Hence their ΔMFE is positive, which is a destabilisation of the mRNA MFE structure. 2977 changed their MFE structure, i.e. they gained a lower MFE value after Editing. Hence their ΔMFE is negative, which means their MFE structure has been reinforced. This bar reflects the MFE changes in healthy individuals and will be used as a reference for unhealthy states in the following section.

For now, we can conclude that:

1. All heart cells are RNA Edited

2. And, from a structural point of view, all these RNA Edited transcripts must fall into one of three categories:

a) Some structures are reinforced, highlighted in a green colour in the following figures (Fig. 23 and 24).

b) Some destabilised; highlighted in a light red colour in the following figures (Fig. 23 and 24).

c) Some remain unchanged, highlighted in an azure colour in the following figures (Fig. 23 and 24).

These three types must be present in any healthy heart cell, as per our findings.

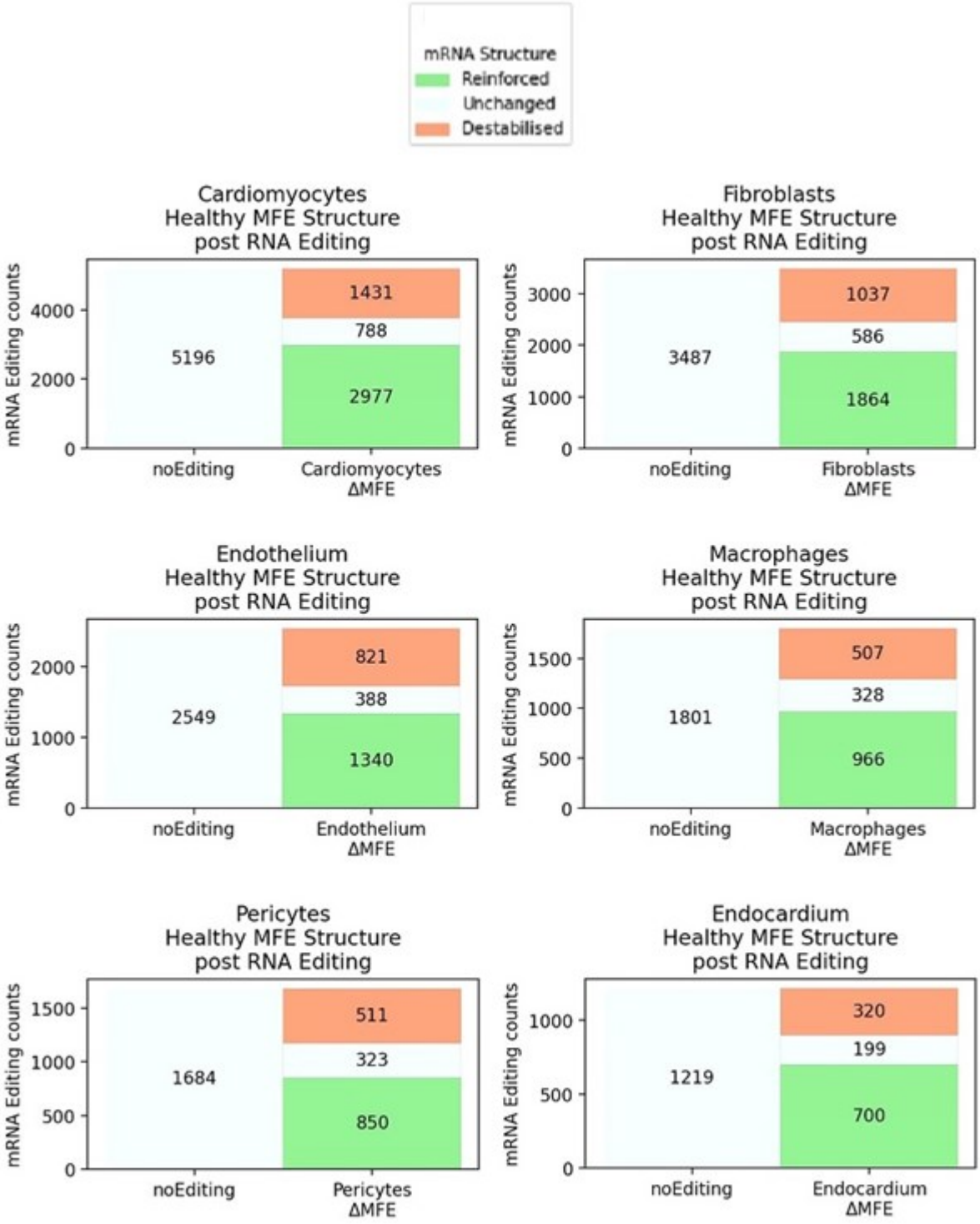


Figure 23 MFE & Healthy and highly RNA Edited cells

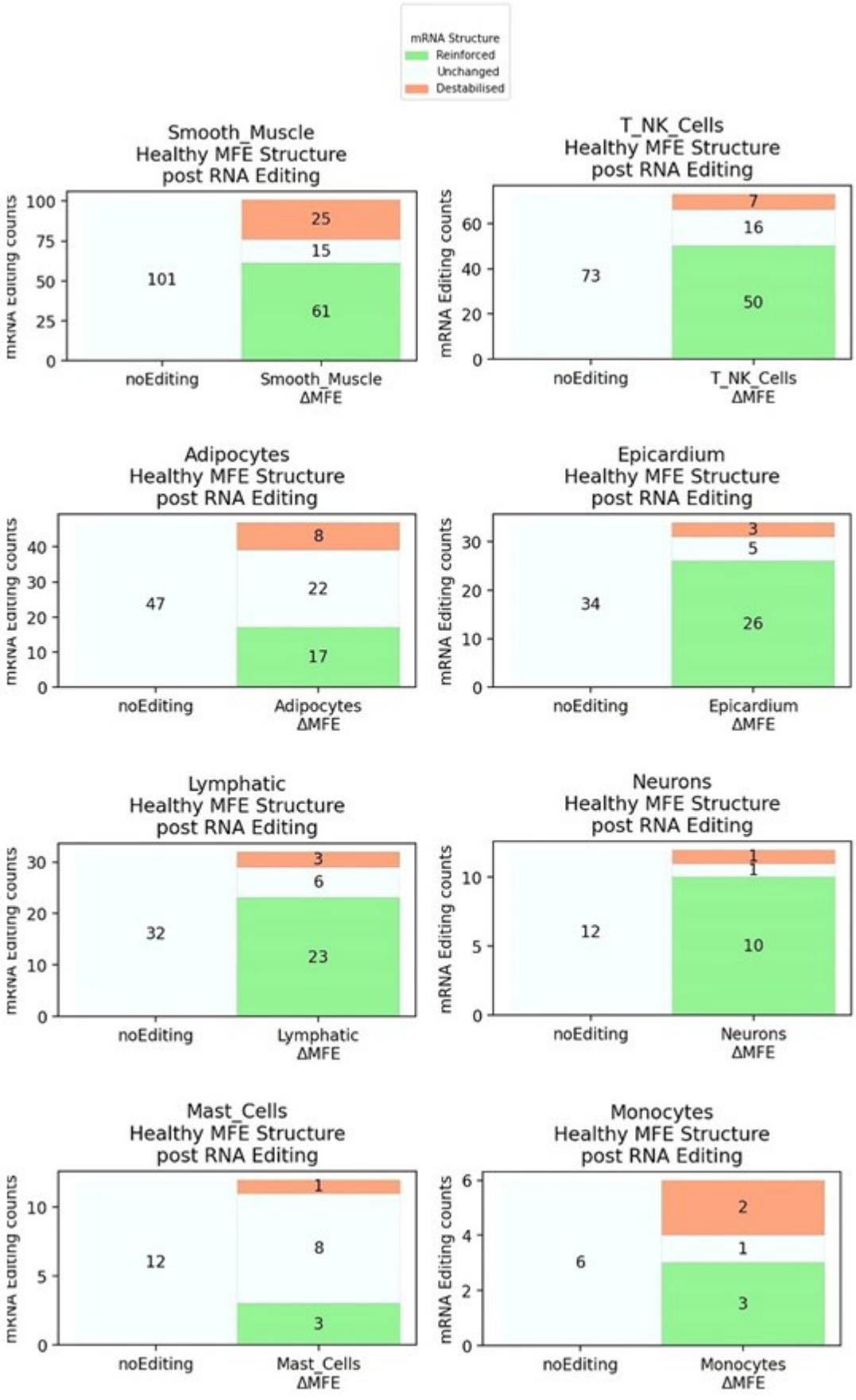


Figure 24 MFE & Healthy and Less frequently RNA edited cells.

3. 3. 4. Healthy and unhealthy mRNA structure post RNA editing:

In the following figures (Fig. 25 to 38), “noDCM noDB” refers to the healthy state. All others are unhealthy. The “no editing” is only for reference for the total number of transcripts I found to be RNA edited for each cell type. These findings show clearly that there are disruptions to the mRNA structure in all cells in unhealthy states. The only cell at odds with this and clearly stands out is “Monocyte cells”, which remain unchanged in all health states.

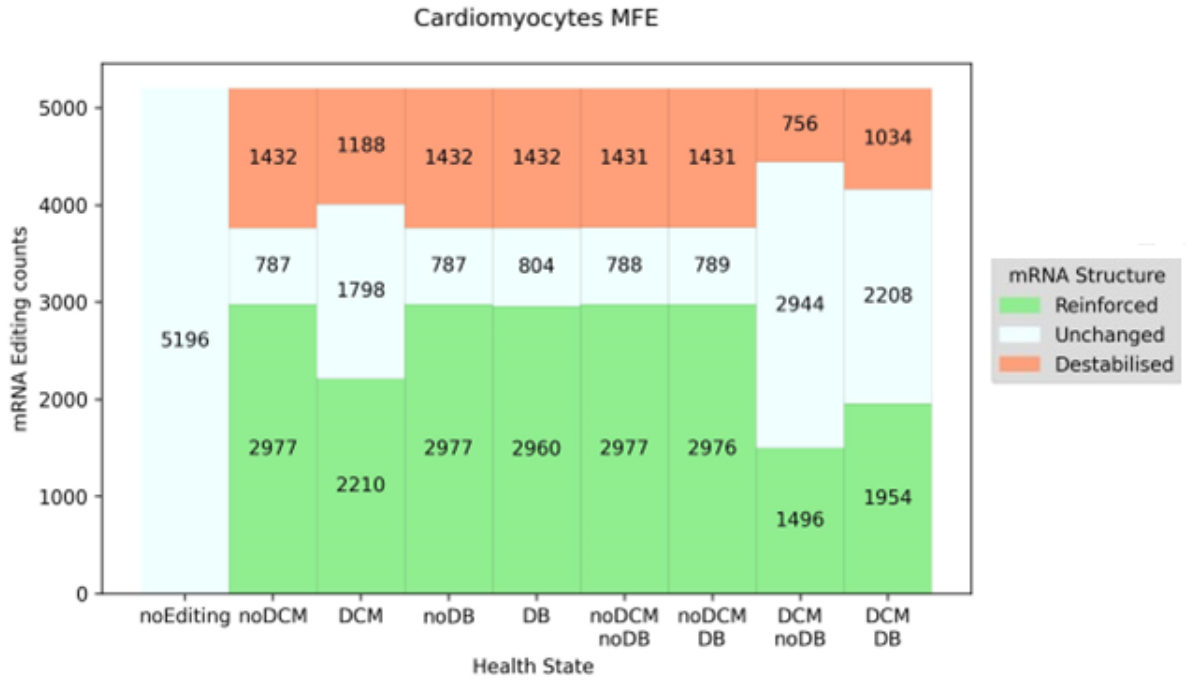


Figure 25 MFE & Cardiomyocytes RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

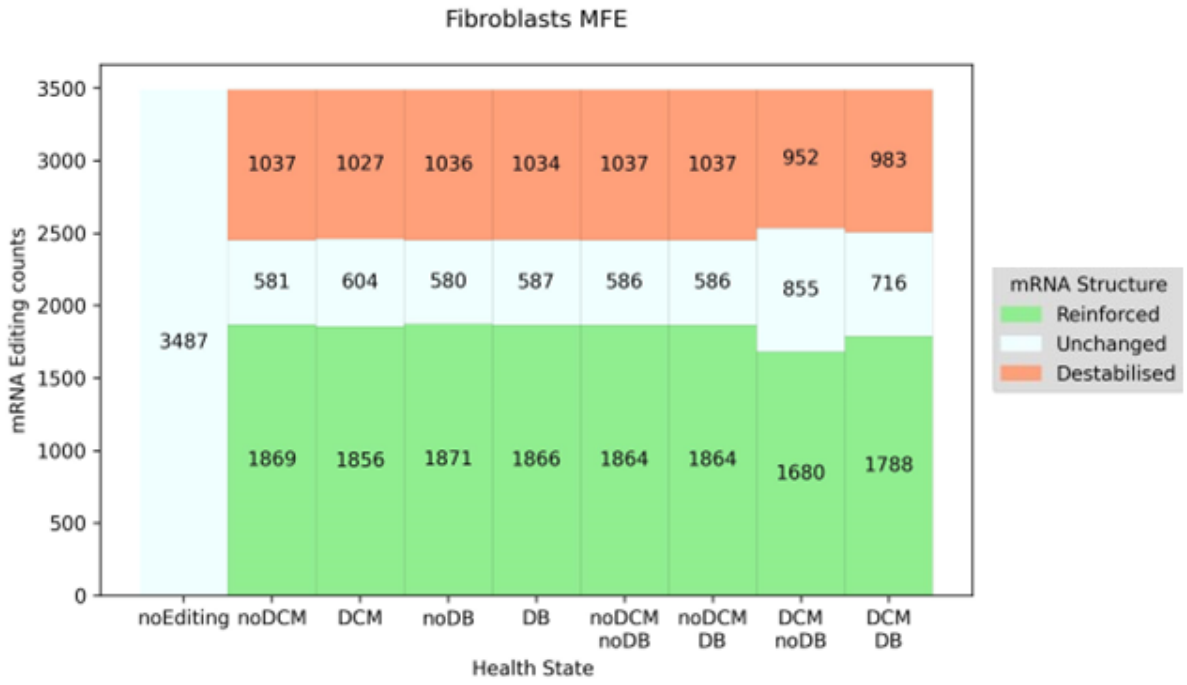


Figure 26 MFE & Fibroblasts RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

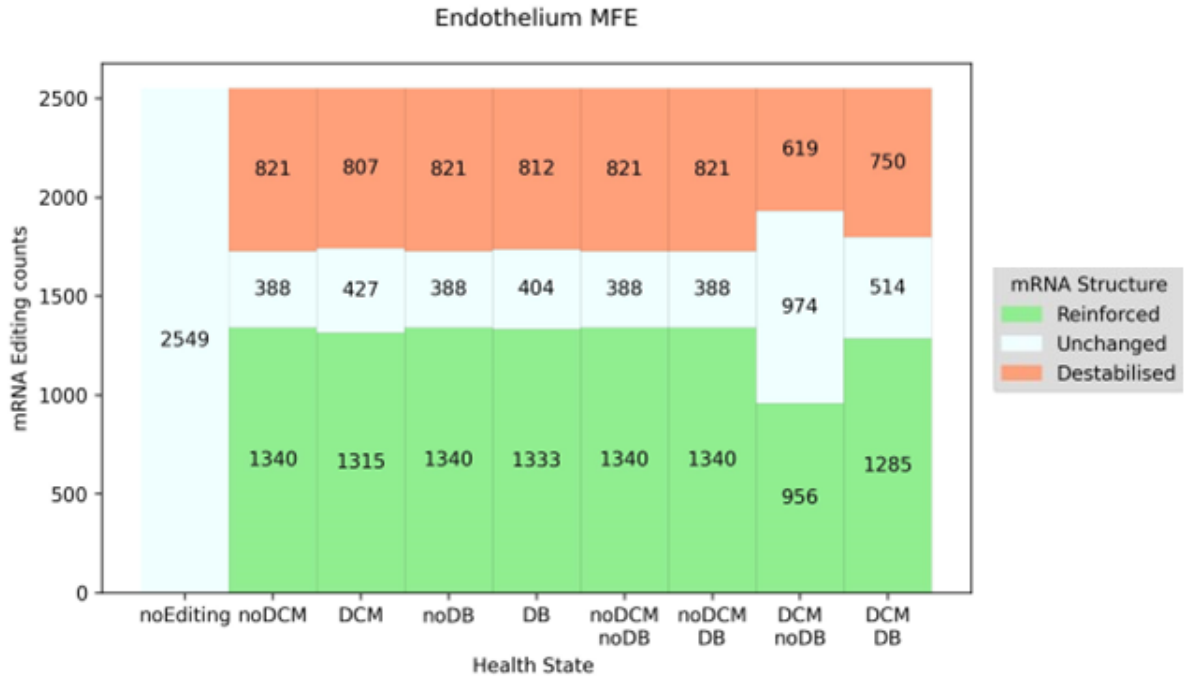


Figure 27 MFE & Endothelium RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

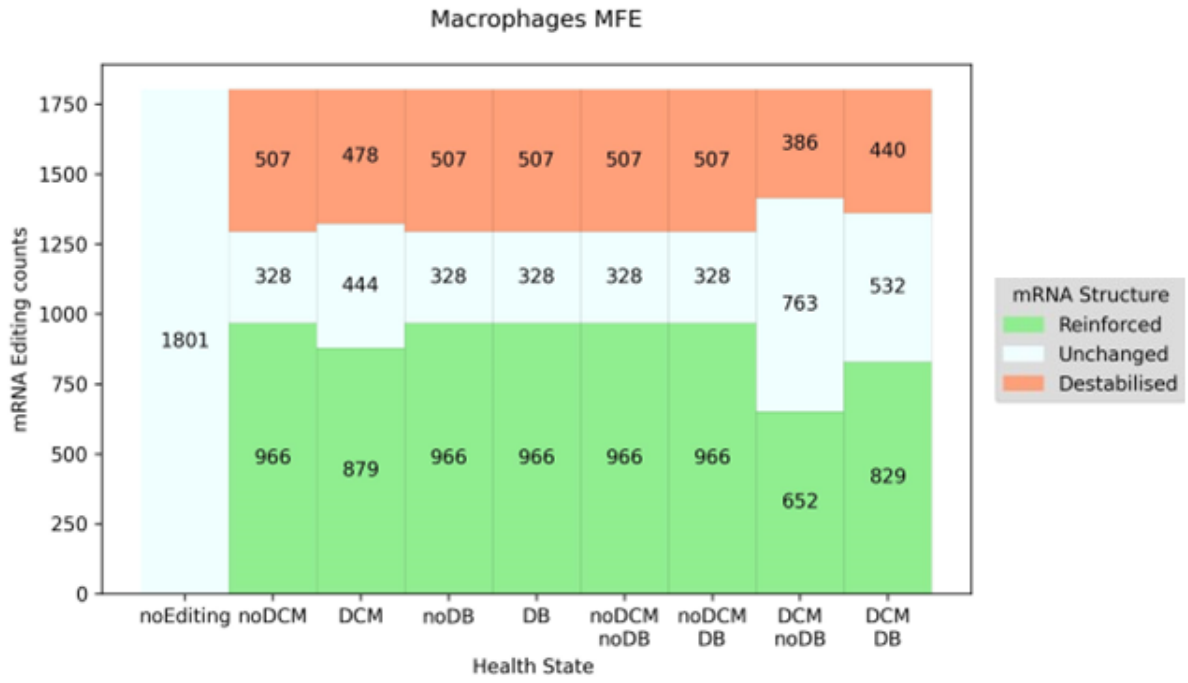


Figure 28 MFE & Macrophages RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

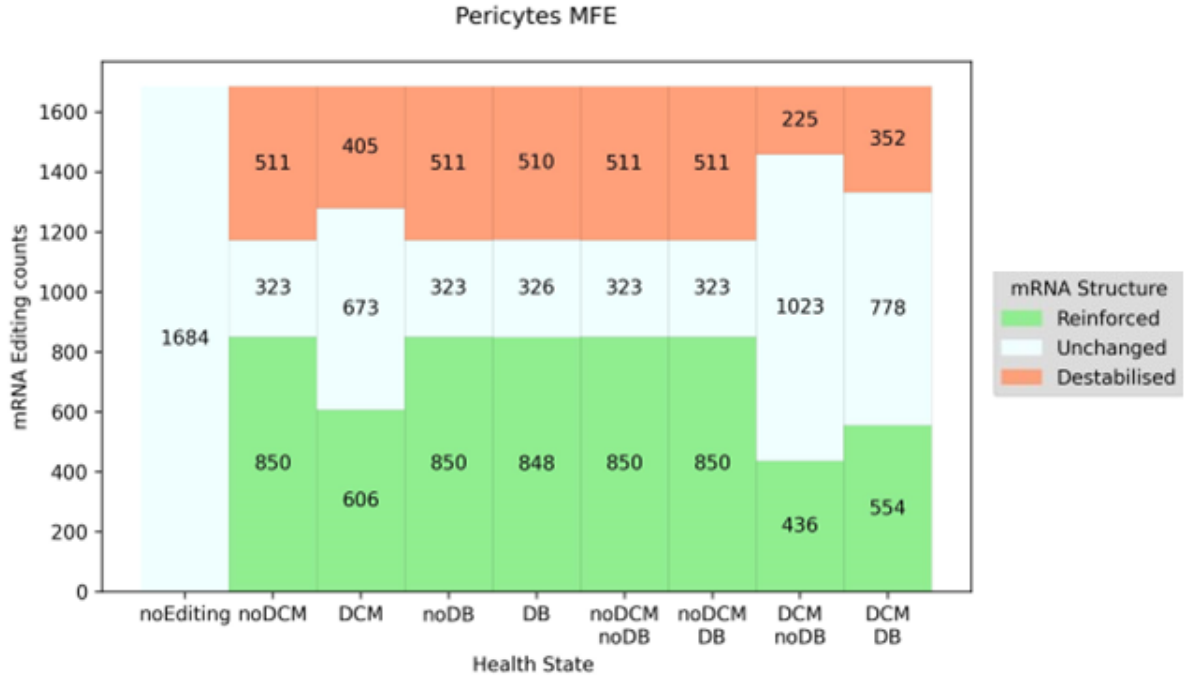


Figure 29 MFE & Pericytes RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

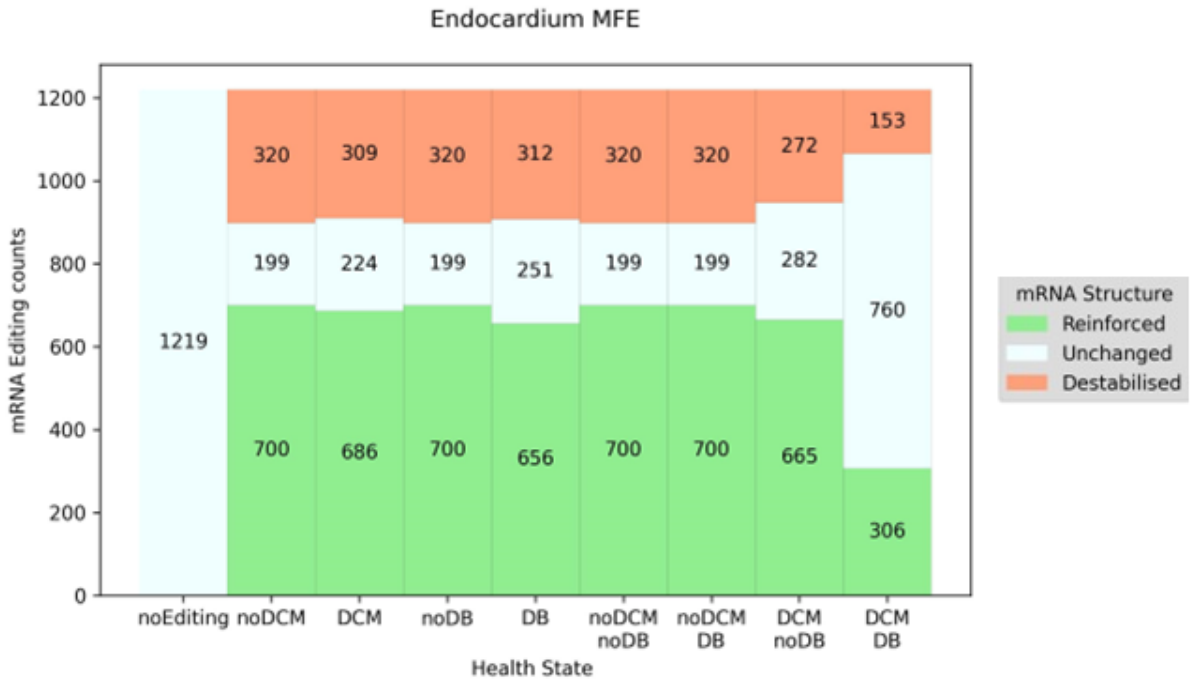


Figure 30 MFE & Endocardium RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

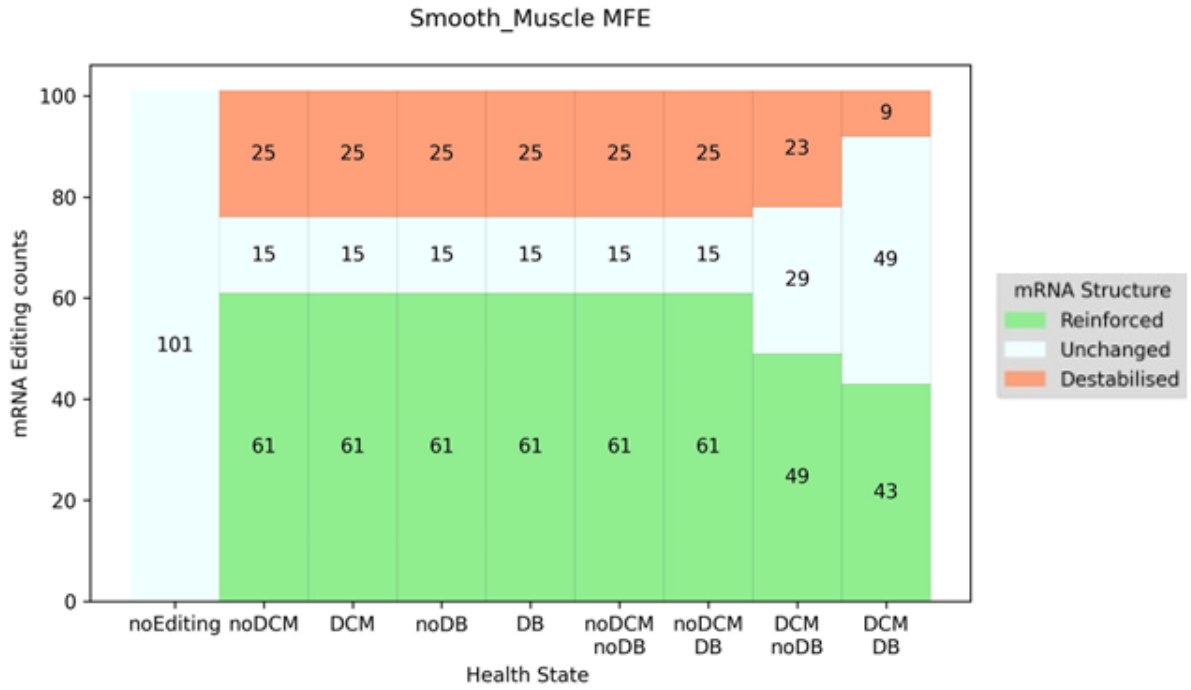


Figure 31 MFE & Smooth muscle RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

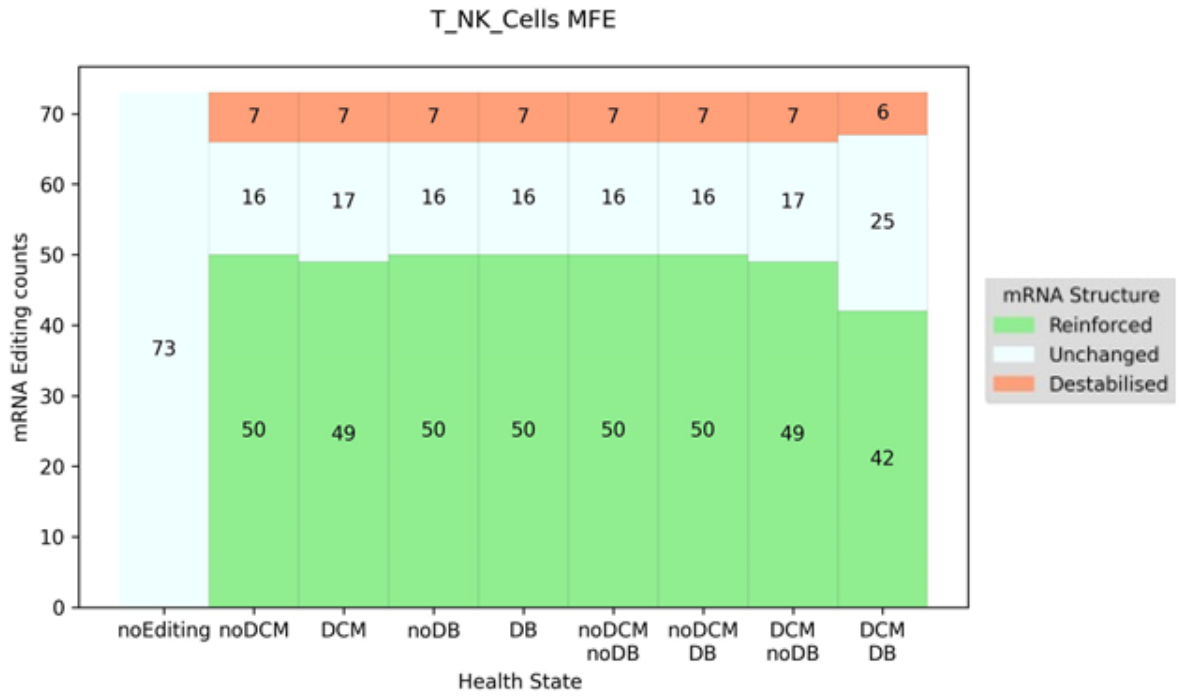


Figure 32 MFE & T NK cells RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

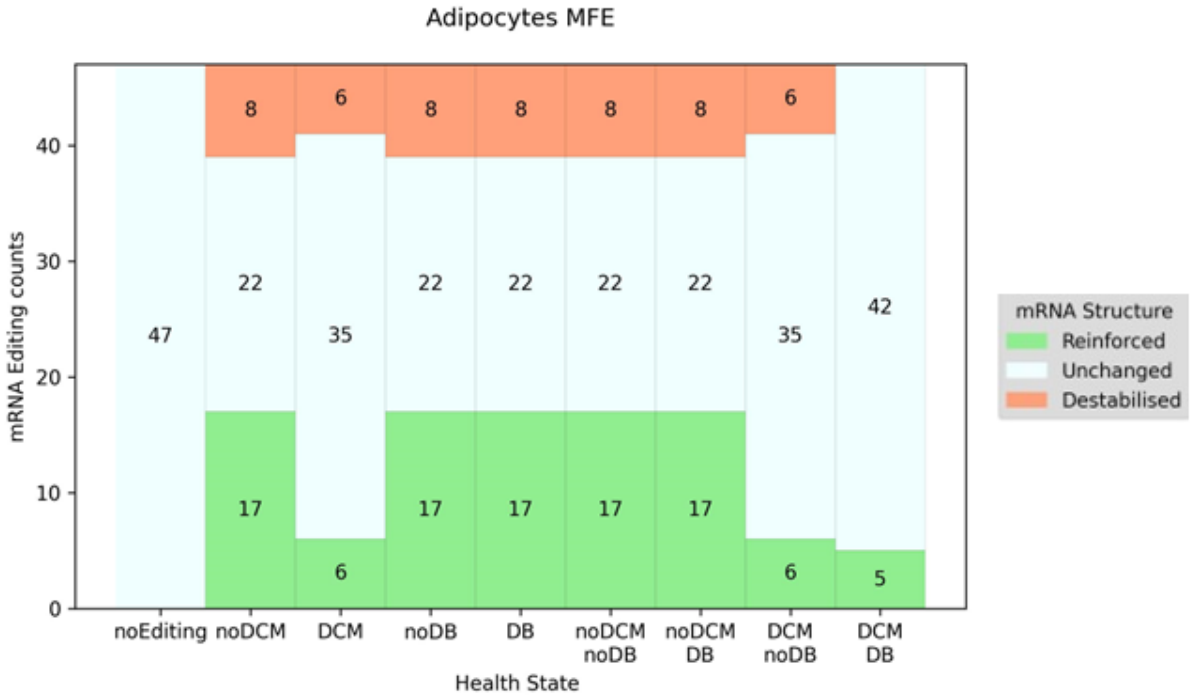


Figure 33 MFE & Adipocytes RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

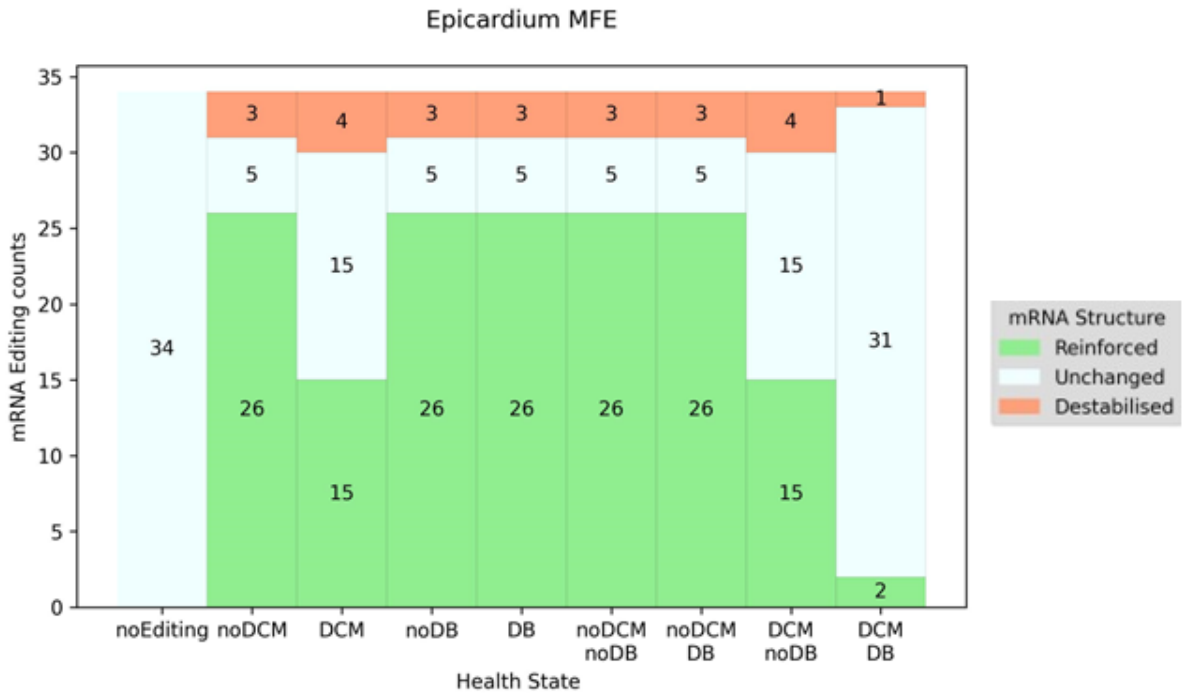


Figure 34 MFE & Epicardium RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

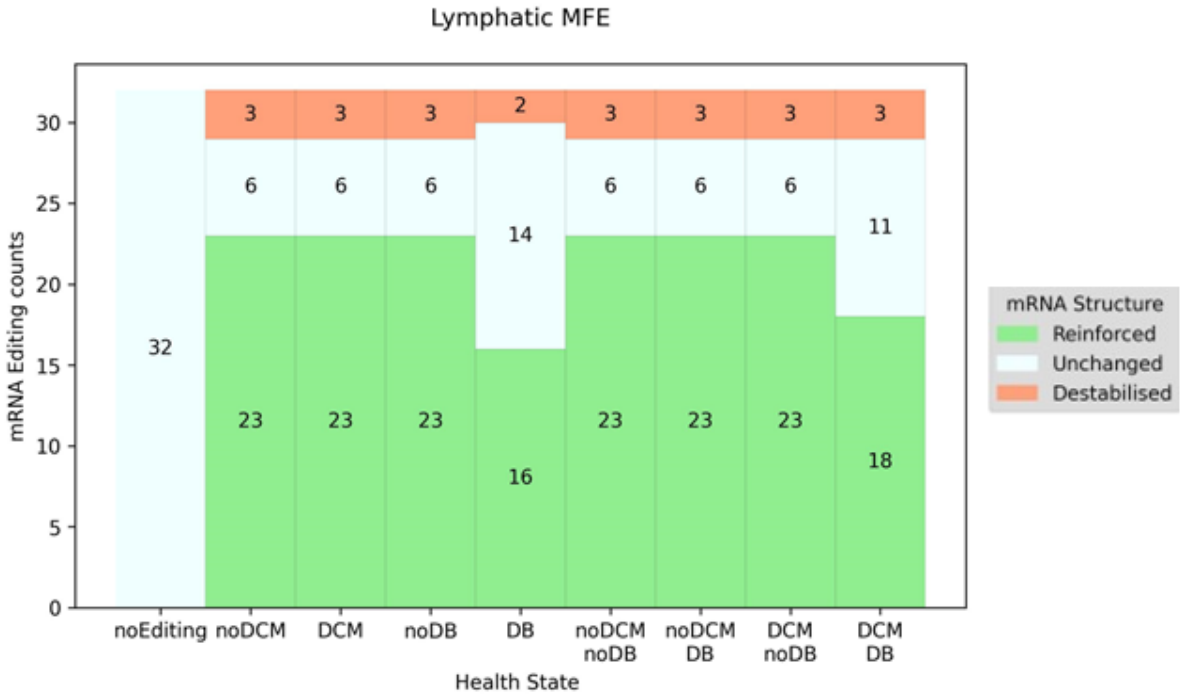


Figure 35 MFE & Lymphatic RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

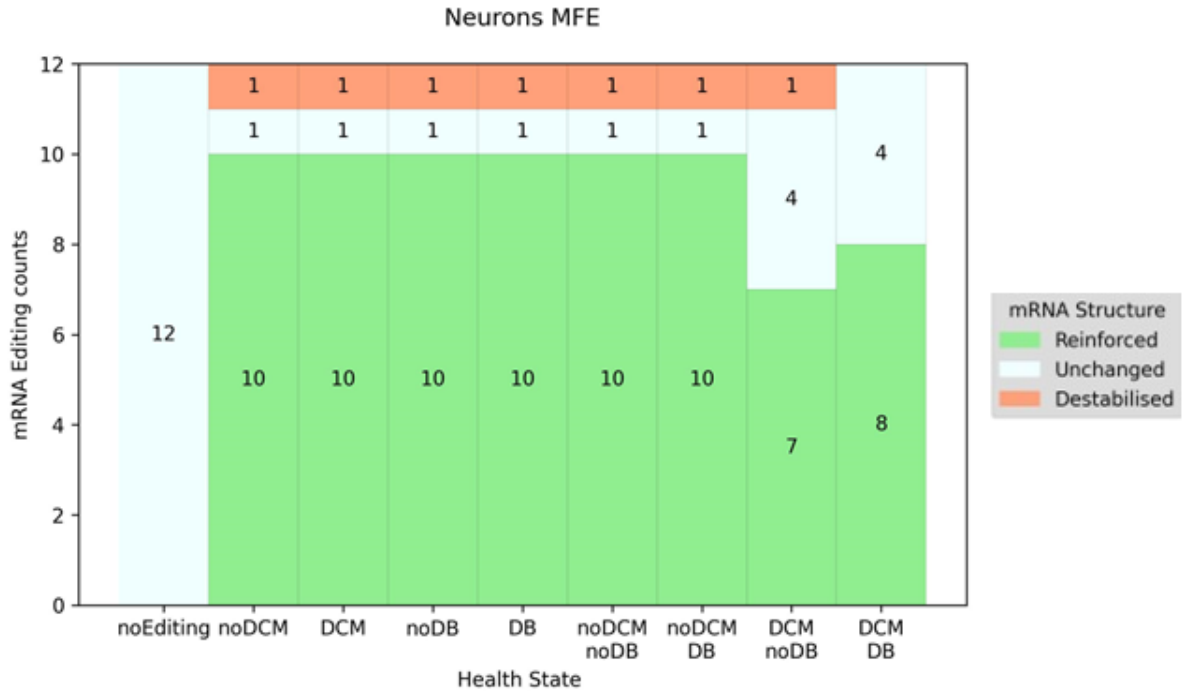


Figure 36 MFE & Neurons RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

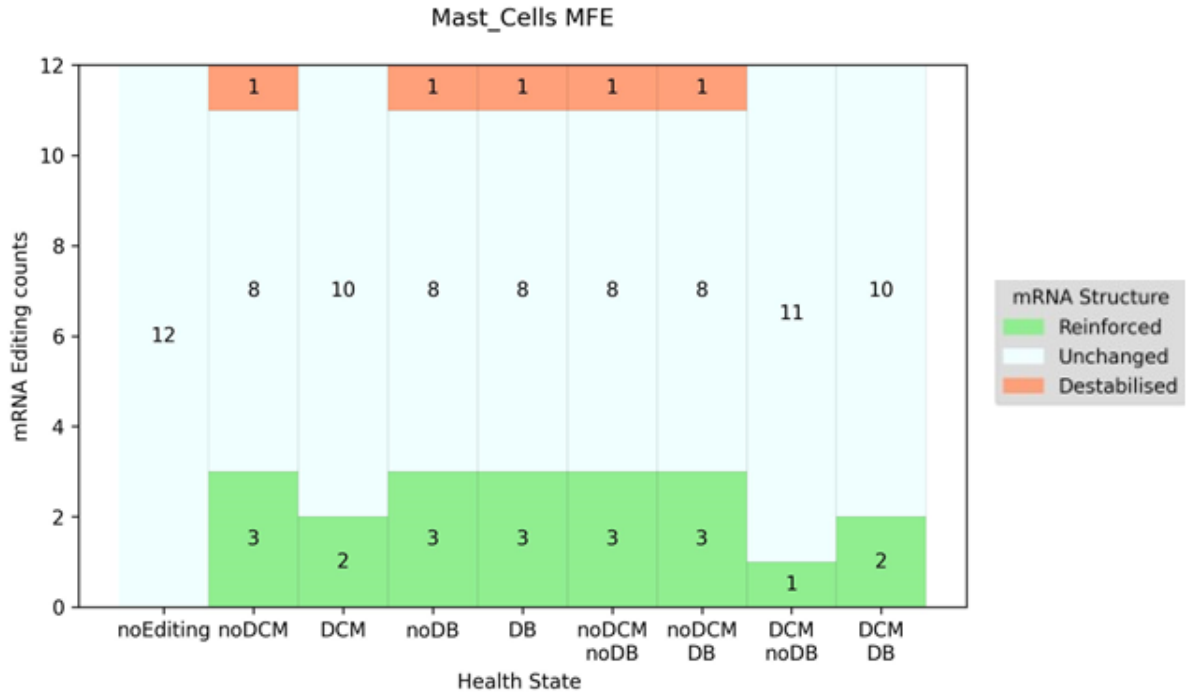


Figure 37 MFE & Mast cells RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

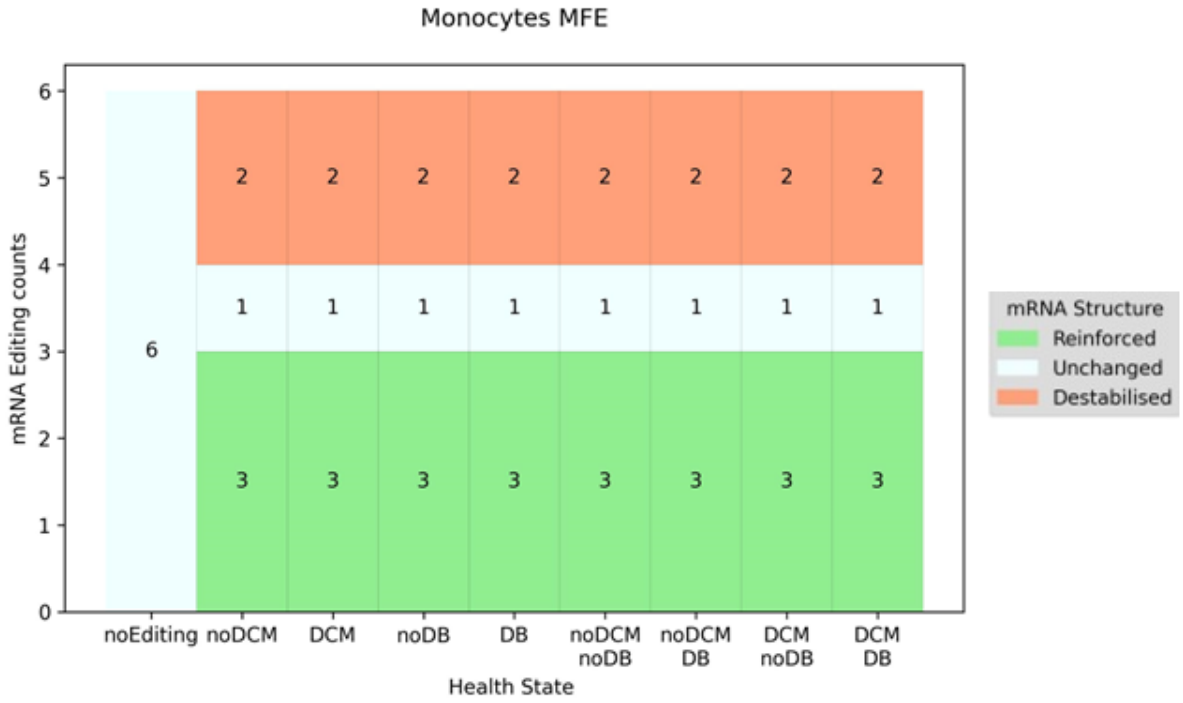


Figure 38 MFE & Monocytes RNA editing in all health states.

DB = Diabetes; DCM= Dilated Cardiomyopathy.

3. 5. Consequence/s of RNA Editing

I used the Ensembl Variant Effect Predictor [31] to categories all the RNA Editing hits. The ones relevant to our study can be summarised as follows:

1) A variation to the mRNA 3' UTR (Untranslated Region) sequence, which may have an impact on microRNAs targeting.
2) A variation to the mRNA 5' UTR sequence, which may have an impact the start of the translation by the tRNA and rRNA, and on microRNAs targeting.
3) A variation to the coding sequence that may or may not result in a different amino acid after translation.
4) Incomplete terminal codon variant, where at least one base of the final codon of an incompletely annotated transcript is changed.
5) Missense variant: A sequence variant, that changes one or more bases, resulting in a different amino acid sequence but where the length is preserved.
6) NMD transcript variant: A variant in a transcript that is the target of nonsense-mediated mRNA decay.
7) Splice region variant: A sequence variant in which a change has occurred within the region of the splice site, either within 1-3 bases of the exon or 3-8 bases of the intron.
8) Start lost: A codon variant that changes at least one base of the canonical start codon.
9) Stop gained: A sequence variant whereby at least one base of a codon is changed, resulting in a premature stop codon, leading to a shortened polypeptide.
10) Stop lost: A sequence variant where at least one base of the terminator (stop) codon is changed, resulting in an elongated transcript.
11) Stop retained variant: A sequence variant where at least one base in the terminator codon is changed, but the terminator remains.
12) Synonymous variant: A sequence variant where there is no resulting change to the encoded amino acid. Hence, this type of variant is of interest to those studying mRNA structural impact on disease.
13) Mature miRNA variant: A transcript variant located within the sequence of the mature miRNA. Hence, it may change its target RNA.
14) Non-coding transcript exon variant: A sequence variant that changes non-coding exon sequence in a non-coding transcript.

Additionally, RNA Editing can lead to the reinforcement of the mRNA structure, or it may cause its destabilisation. At times, it may not have an impact on the overall structure as shown previously in section 3.3.

3.5 Differential RNA editing

I could not possibly anticipate the results I would be getting. Hence, I divided the 35 samples set into 4 groups that kept enough diversity and variety: DCM vs noDCM, DB vs noDB, noDCM no DB vs noDCM DB, and DCM noDB vs DCM DB. Then I run Wilcoxon's pairwise test on the 4 groups and calculated the P value for each unique RNA edited chromosomal position in the 4 groups. All the results are made part of the supplement files.

As DCM vs noDCM was core to our research in Dilated Cardiomyopathy, we only chose that group and only chose chromosomal positions that had a P value that were less than 0.05 when doing the functional enrichment analysis.

3. 6. Functional enrichments analysis

The enrichment was done with the data belonging to the category DCM vs noDCM only.

Next, to estimate functional categories, I only selected chromosome positions with a P_value less than 0.05 and only those reported by Reditools software [29] with identical strand and modification across all 35 data samples. And I only chose A-to-I and C-to-U type modifications.

I cross referenced all those chromosome positions with the reference genome (gencode v43 basic annotation) from Ensembl and European Bioinformatics Institute [32]

I left duplicate chromosome positions as they had an updated Ensembl name as opposed to the generic gene symbol name. For they reported a slight change to the start and / or end of the protein coding gene. Later, these duplicates were removed automatically by the sites I used for enrichment.

I only selected gene names that are protein coding, and I calculated the Fold Change (DCM / noDCM) for each position. Hence, I was able to select only dysregulated genes and discarded those entries with Fold Changes equal to one. I have added all the input enrichment files to the supplement files as “Input Enrichment files.”

I only considered dysregulated genes when doing the enrichment using www.webgestalt.org and I only considered enrichment results with False Discovery Rate (FDR) less than 0.05 as significant. The cell types and their associated results are made part of the supplement files. Please see “Enrichment-results.zip” in the supplement files for more details. I further used Expressanalyst [35] for comparison and to get more possible insights.

I am using Cardiomyocytes’ list of over 2201 genes, for it is directly linked to Dilated Cardiomyopathy, to illustrate our enrichment results.

Cardiomyocytes’ enrichment results are summarised below in the following figures (Fig 37 and 38) and tables (Table 8 and 9).

3. 6. 1 ORA KEGG pathways in Cardiomyocytes

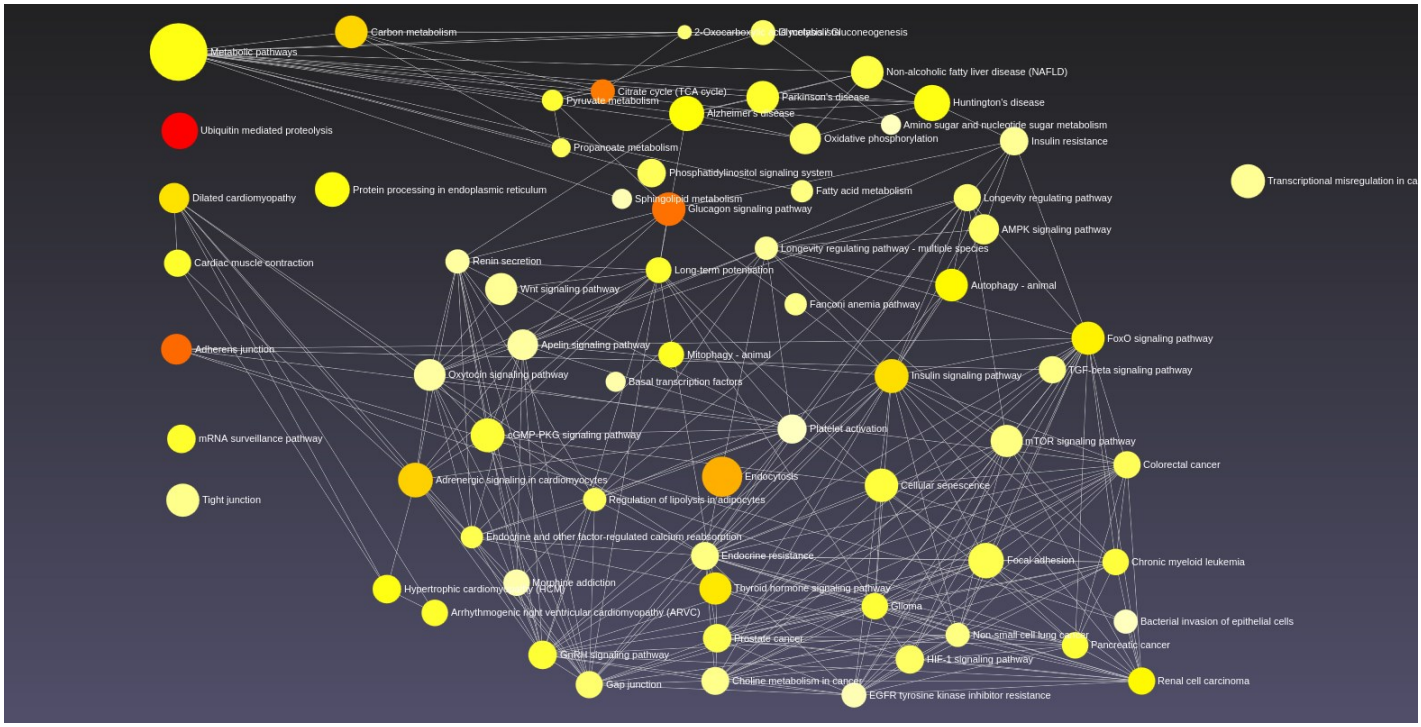


Figure 39 Cardiomyocytes ORA KEGG pathways

geneSet	description	pValue	FDR
hsa04120	Ubiquitin mediated proteolysis	7.7416E-07	0.000255473
hsa04520	Adherens junction	2.67055E-06	0.000374136
hsa04922	Glucagon signaling pathway	4.27732E-06	0.000374136
hsa00020	Citrate cycle (TCA cycle)	4.53498E-06	0.000374136
hsa04714	Thermogenesis	7.5788E-05	0.00500201
hsa04144	Endocytosis	0.000145806	0.008019324
hsa01200	Carbon metabolism	0.000215018	0.009710993
hsa04261	Adrenergic signaling in cardiomyocytes	0.000235418	0.009710993
hsa04910	Insulin signaling pathway	0.000292898	0.010739609
hsa05414	Dilated cardiomyopathy (DCM)	0.000464018	0.0153126
hsa04919	Thyroid hormone signaling pathway	0.000804346	0.024130374
hsa04140	Autophagy	0.001085763	0.029858482
hsa04068	FoxO signaling pathway	0.001817685	0.046141239

Table 8 Cardiomyocytes ORA KEGG pathways.

3. 6. 2 Biological processes in Cardiomyocytes

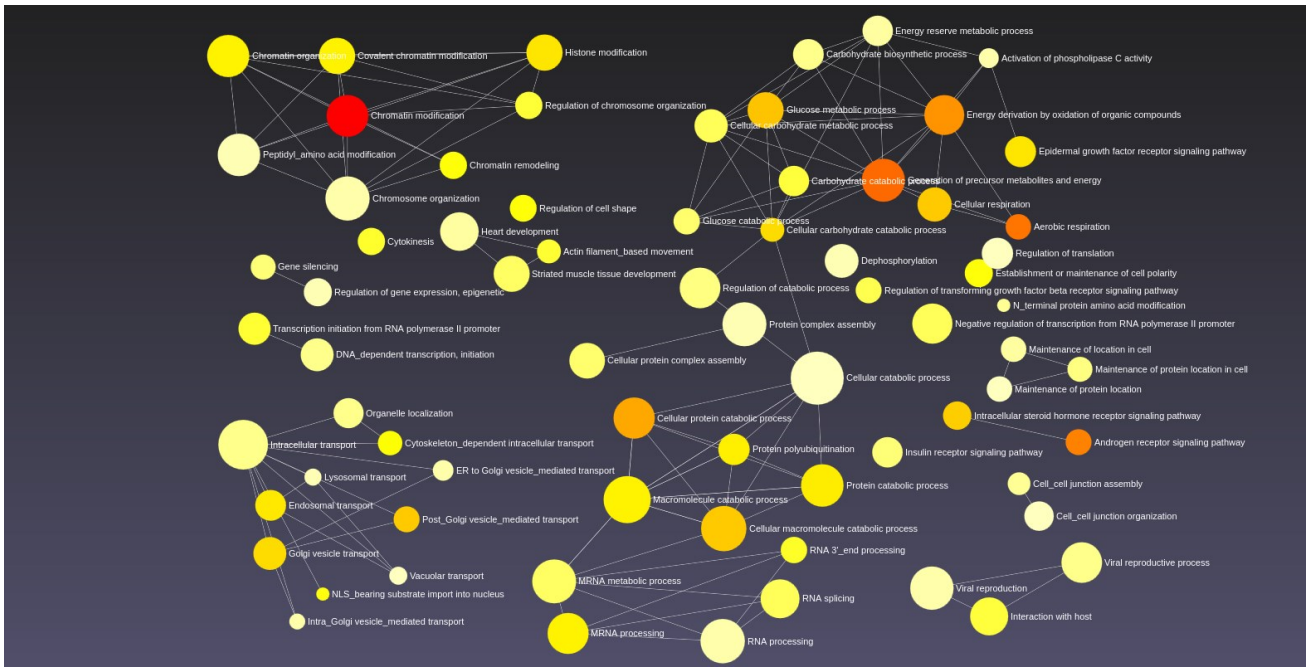


Figure 40 Cardiomyocytes GO Biological Processes

geneSet	description	pValue	FDR
GO:0006397	mRNA processing	0.00	0.00
GO:0006396	RNA processing	0.00	0.00
GO:0016071	mRNA metabolic process	0.00	0.00
GO:0044265	cellular macromolecule catabolic process	0.00	0.00
GO:0008380	RNA splicing	0.00	0.00
GO:0000377	RNA splicing	2.08	0.00
GO:0000398	mRNA splicing	2.08	0.00
GO:0051276	chromosome organization	0.00	0.00
GO:0000375	RNA splicing	2.06	0.00
GO:0009057	macromolecule catabolic process	0.00	0.00
GO:0031329	regulation of cellular catabolic process	0.00	0.00
GO:0016570	histone modification	0.00	0.00
GO:0016569	covalent chromatin modification	0.00	0.00
GO:0070727	cellular macromolecule localization	0.00	0.00
GO:0006325	chromatin organization	0.00	0.00
GO:0034613	cellular protein localization	0.00	0.00
GO:1901575	organic substance catabolic process	0.00	0.00
GO:0015980	energy derivation by oxidation of organic compounds	0.00	0.00
GO:0044257	cellular protein catabolic process	0.00	0.00
GO:0046907	intracellular transport	0.00	0.00
GO:0009894	regulation of catabolic process	0.00	0.00
GO:0006091	generation of precursor metabolites and energy	0.00	0.00
GO:0018205	peptidyl-lysine modification	0.00	0.00
GO:0009060	aerobic respiration	0.00	0.00
GO:0051603	proteolysis involved in cellular protein catabolic process	0.00	0.00
GO:0045333	cellular respiration	0.00	0.00
GO:1903311	regulation of mRNA metabolic process	0.00	0.00
GO:0034622	cellular protein-containing complex assembly	0.00	0.00
GO:0070647	protein modification by small protein conjugation or removal	0.00	0.00
GO:0006886	intracellular protein transport	0.00	0.00
GO:0043632	modification-dependent macromolecule catabolic process	0.00	0.00
GO:0010629	negative regulation of gene expression	0.00	0.00
GO:0009117	nucleotide metabolic process	0.00	0.00
GO:0055086	nucleobase-containing small molecule metabolic process	0.00	0.00
GO:0033043	regulation of organelle organization	0.00	0.00
GO:0006753	nucleoside phosphate metabolic process	0.00	0.00
GO:0006511	ubiquitin-dependent protein catabolic process	0.00	0.00
GO:0050684	regulation of mRNA processing	0.00	0.00
GO:0045934	negative regulation of nucleobase-containing compound metabolic process	0.00	0.00
GO:0043543	protein acylation	0.00	0.00
GO:0009161	ribonucleoside monophosphate metabolic process	0.00	0.00
GO:0065003	protein-containing complex assembly	0.00	0.00
GO:0019693	ribose phosphate metabolic process	0.00	0.00
GO:0019941	modification-dependent protein catabolic process	0.00	0.00

GO:0009167	purine ribonucleoside monophosphate metabolic process	0.00	0.00
GO:0009126	purine nucleoside monophosphate metabolic process	0.00	0.00
GO:0010256	endomembrane system organization	0.00	0.00
GO:0009123	nucleoside monophosphate metabolic process	0.00	0.00
GO:0006259	DNA metabolic process	0.00	0.00
GO:0009259	ribonucleotide metabolic process	0.00	0.00
GO:0006281	DNA repair	0.00	0.00
GO:0016197	endosomal transport	0.00	0.00
GO:0051253	negative regulation of RNA metabolic process	0.00	0.00
GO:0046034	ATP metabolic process	0.00	0.00
GO:0030163	protein catabolic process	0.00	0.00
GO:0015031	protein transport	0.00	0.00
GO:0009150	purine ribonucleotide metabolic process	0.00	0.00
GO:0048024	regulation of mRNA splicing	2.56	0.00
GO:0006473	protein acetylation	0.00	0.00
GO:1901565	organonitrogen compound catabolic process	0.00	0.00
GO:0009205	purine ribonucleoside triphosphate metabolic process	0.00	0.00
GO:0019637	organophosphate metabolic process	0.00	0.00
GO:0009199	ribonucleoside triphosphate metabolic process	0.00	0.00
GO:0009144	purine nucleoside triphosphate metabolic process	0.00	0.00
GO:0006163	purine nucleotide metabolic process	0.00	0.00
GO:0007049	cell cycle	0.00	0.00
GO:0072521	purine-containing compound metabolic process	0.00	0.00
GO:2000113	negative regulation of cellular macromolecule biosynthetic process	0.00	0.00
GO:0015833	peptide transport	0.00	0.00
GO:0034655	nucleobase-containing compound catabolic process	0.00	0.00
GO:0042147	retrograde transport	2.55	0.00
GO:0006914	autophagy	0.00	0.00
GO:0061919	process utilizing autophagic mechanism	0.00	0.00
GO:0009141	nucleoside triphosphate metabolic process	0.00	0.00
GO:0010558	negative regulation of macromolecule biosynthetic process	0.00	0.00
GO:0031327	negative regulation of cellular biosynthetic process	0.00	0.00
GO:0043484	regulation of RNA splicing	0.00	0.00
GO:0033365	protein localization to organelle	0.00	0.00
GO:0016482	cytosolic transport	0.00	0.00
GO:0018394	peptidyl-lysine acetylation	0.00	0.00
GO:0006099	tricarboxylic acid cycle	0.00	0.00
GO:0045184	establishment of protein localization	0.00	0.00
GO:0016579	protein deubiquitination	0.00	0.00
GO:0070646	protein modification by small protein removal	0.00	0.00
GO:0006974	cellular response to DNA damage stimulus	0.00	0.00
GO:0042886	amide transport	0.00	0.00
GO:0006101	citrate metabolic process	0.00	0.00
GO:0022402	cell cycle process	0.00	0.00
GO:0048738	cardiac muscle tissue development	0.00	0.00

GO:0009890	negative regulation of biosynthetic process	0.00	0.00
GO:0072350	tricarboxylic acid metabolic process	0.00	0.00
GO:0007030	Golgi organization	0.00	0.00
GO:0044087	regulation of cellular component biogenesis	0.00	0.00
GO:0018393	internal peptidyl-lysine acetylation	0.00	0.00
GO:0044270	cellular nitrogen compound catabolic process	0.00	0.00
GO:0006475	internal protein amino acid acetylation	0.00	0.00
GO:0016573	histone acetylation	0.00	0.00
GO:1901566	organonitrogen compound biosynthetic process	0.00	0.00
GO:1903507	negative regulation of nucleic acid-templated transcription	0.00	0.00
GO:0000381	regulation of alternative mRNA splicing	2.68	0.00
GO:0018193	peptidyl-amino acid modification	0.00	0.00
GO:1902679	negative regulation of RNA biosynthetic process	0.00	0.00
GO:0044281	small molecule metabolic process	0.00	0.00
GO:0046700	heterocycle catabolic process	0.00	0.00
GO:0033554	cellular response to stress	0.00	0.00
GO:0045892	negative regulation of transcription	1.33	0.00
GO:0000380	alternative mRNA splicing	2.50	0.00
GO:1901135	carbohydrate derivative metabolic process	0.00	0.00
GO:0019439	aromatic compound catabolic process	0.00	0.00
GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process	0.00	0.00
GO:0006006	glucose metabolic process	0.00	0.00
GO:0006338	chromatin remodeling	0.00	0.00
GO:0030521	androgen receptor signaling pathway	0.00	0.00
GO:0030029	actin filament-based process	0.00	0.00
GO:0055114	oxidation-reduction process	0.00	0.00
GO:0016236	macroautophagy	0.00	0.00
GO:0071826	ribonucleoprotein complex subunit organization	0.00	0.00
GO:0032446	protein modification by small protein conjugation	0.00	0.00
GO:0019318	hexose metabolic process	0.00	0.00
GO:0019362	pyridine nucleotide metabolic process	0.00	0.00
GO:0046496	nicotinamide nucleotide metabolic process	0.00	0.00
GO:0010506	regulation of autophagy	0.00	0.00
GO:0072524	pyridine-containing compound metabolic process	0.00	0.00
GO:0006479	protein methylation	0.00	0.00
GO:0008213	protein alkylation	0.00	0.00
GO:0010498	proteasomal protein catabolic process	0.00	0.00
GO:0040029	regulation of gene expression	1.70	0.00
GO:0016032	viral process	0.00	0.00
GO:0007507	heart development	0.00	0.00
GO:0010608	posttranscriptional regulation of gene expression	0.00	0.00
GO:0030522	intracellular receptor signaling pathway	0.00	0.00
GO:0005996	monosaccharide metabolic process	0.00	0.00
GO:0005975	carbohydrate metabolic process	0.00	0.00
GO:0006733	oxidoreduction coenzyme metabolic process	0.00	0.00
GO:0045935	positive regulation of nucleobase-containing compound metabolic process	0.00	0.00
GO:1902115	regulation of organelle assembly	0.00	0.00
GO:1901361	organic cyclic compound catabolic process	0.00	0.00

GO:0031331	positive regulation of cellular catabolic process	0.00	0.00
GO:0043414	macromolecule methylation	0.00	0.00
GO:0051726	regulation of cell cycle	0.00	0.00
GO:0048193	Golgi vesicle transport	0.00	0.00
GO:0000122	negative regulation of transcription by RNA polymerase II	0.00	0.00
GO:0030048	actin filament-based movement	0.00	0.00
GO:0022618	ribonucleoprotein complex assembly	0.00	0.00
GO:0071108	protein K48-linked deubiquitination	0.00	0.00
GO:0006482	protein demethylation	0.00	0.00
GO:0008214	protein dealkylation	0.00	0.00
GO:0043254	regulation of protein complex assembly	0.00	0.00
GO:0007010	cytoskeleton organization	0.00	0.01
GO:0046031	ADP metabolic process	0.00	0.01
GO:0009185	ribonucleoside diphosphate metabolic process	0.00	0.01
GO:0016567	protein ubiquitination	0.00	0.01
GO:0009896	positive regulation of catabolic process	0.00	0.01
GO:0051254	positive regulation of RNA metabolic process	0.00	0.01
GO:0009108	coenzyme biosynthetic process	0.00	0.01
GO:0006732	coenzyme metabolic process	0.00	0.01
GO:0014706	striated muscle tissue development	0.00	0.01
GO:0072594	establishment of protein localization to organelle	0.00	0.01
GO:0061912	selective autophagy	0.00	0.01
GO:0090407	organophosphate biosynthetic process	0.00	0.01
GO:0061013	regulation of mRNA catabolic process	0.00	0.01
GO:1903827	regulation of cellular protein localization	0.00	0.01
GO:0016577	histone demethylation	0.00	0.01
GO:0060537	muscle tissue development	0.00	0.01
GO:0007063	regulation of sister chromatid cohesion	0.00	0.01
GO:0000278	mitotic cell cycle	0.00	0.01
GO:0000209	protein polyubiquitination	0.00	0.01
GO:0009135	purine nucleoside diphosphate metabolic process	0.00	0.01
GO:0009179	purine ribonucleoside diphosphate metabolic process	0.00	0.01
GO:0031330	negative regulation of cellular catabolic process	0.00	0.01
GO:0042866	pyruvate biosynthetic process	0.00	0.01
GO:0006119	oxidative phosphorylation	0.00	0.01
GO:0006096	glycolytic process	0.00	0.01
GO:0044403	symbiont process	0.00	0.01
GO:0060048	cardiac muscle contraction	0.00	0.01
GO:0009166	nucleotide catabolic process	0.00	0.01
GO:0030518	intracellular steroid hormone receptor signaling pathway	0.00	0.01
GO:0006757	ATP generation from ADP	0.00	0.01
GO:0016052	carbohydrate catabolic process	0.00	0.01
GO:0006090	pyruvate metabolic process	0.00	0.01
GO:0009156	ribonucleoside monophosphate biosynthetic process	0.00	0.01
GO:0018022	peptidyl-lysine methylation	0.00	0.01
GO:0046939	nucleotide phosphorylation	0.00	0.01
GO:0043631	RNA polyadenylation	0.00	0.01
GO:0016050	vesicle organization	0.00	0.01

GO:0032870	cellular response to hormone stimulus	0.00	0.01
GO:0010564	regulation of cell cycle process	0.00	0.01
GO:0006892	post-Golgi vesicle-mediated transport	0.00	0.02
GO:0071495	cellular response to endogenous stimulus	0.00	0.02
GO:0051188	cofactor biosynthetic process	0.00	0.02
GO:0030239	myofibril assembly	0.00	0.02
GO:0070988	demethylation	0.00	0.02
GO:0022411	cellular component disassembly	0.00	0.02
GO:0006754	ATP biosynthetic process	0.00	0.02
GO:0009895	negative regulation of catabolic process	0.00	0.02
GO:0009132	nucleoside diphosphate metabolic process	0.00	0.02
GO:0009127	purine nucleoside monophosphate biosynthetic process	0.00	0.02
GO:0009168	purine ribonucleoside monophosphate biosynthetic process	0.00	0.02
GO:0006302	double-strand break repair	0.00	0.02
GO:0086069	bundle of His cell to Purkinje myocyte communication	0.00	0.02
GO:0006012	galactose metabolic process	0.00	0.02
GO:0055013	cardiac muscle cell development	0.00	0.02
GO:0010927	cellular component assembly involved in morphogenesis	0.00	0.02
GO:0019359	nicotinamide nucleotide biosynthetic process	0.00	0.02
GO:0019363	pyridine nucleotide biosynthetic process	0.00	0.02
GO:0055006	cardiac cell development	0.00	0.02
GO:0032801	receptor catabolic process	0.00	0.02
GO:0031032	actomyosin structure organization	0.00	0.02
GO:0086065	cell communication involved in cardiac conduction	0.00	0.02
GO:0016311	dephosphorylation	0.00	0.02
GO:0006165	nucleoside diphosphate phosphorylation	0.00	0.02
GO:0022904	respiratory electron transport chain	0.00	0.02
GO:0032989	cellular component morphogenesis	0.00	0.02
GO:0022613	ribonucleoprotein complex biogenesis	0.00	0.02
GO:0016458	gene silencing	0.00	0.02
GO:0006476	protein deacetylation	0.00	0.02
GO:0000731	DNA synthesis involved in DNA repair	0.00	0.02
GO:0006998	nuclear envelope organization	0.00	0.02
GO:0016571	histone methylation	0.00	0.02
GO:0055007	cardiac muscle cell differentiation	0.00	0.02
GO:0051186	cofactor metabolic process	0.00	0.02
GO:0006104	succinyl-CoA metabolic process	0.00	0.02
GO:0071921	cohesin loading	0.00	0.02
GO:1905634	regulation of protein localization to chromatin	0.00	0.02
GO:0006417	regulation of translation	0.00	0.02
GO:0006378	mRNA polyadenylation	0.00	0.02
GO:0070076	histone lysine demethylation	0.00	0.02
GO:1901292	nucleoside phosphate catabolic process	0.00	0.02
GO:0043401	steroid hormone mediated signaling pathway	0.00	0.02
GO:0002027	regulation of heart rate	0.00	0.02
GO:0072359	circulatory system development	0.00	0.02
GO:0035194	posttranscriptional gene silencing by RNA	0.00	0.02
GO:0009142	nucleoside triphosphate biosynthetic process	0.00	0.02

GO:0043954	cellular component maintenance	0.00	0.02
GO:0009124	nucleoside monophosphate biosynthetic process	0.00	0.03
GO:0006402	mRNA catabolic process	0.00	0.03
GO:0033044	regulation of chromosome organization	0.00	0.03
GO:0055002	striated muscle cell development	0.00	0.03
GO:1901293	nucleoside phosphate biosynthetic process	0.00	0.03
GO:0046434	organophosphate catabolic process	0.00	0.03
GO:0046726	positive regulation by virus of viral protein levels in host cell	0.00	0.03
GO:0072367	regulation of lipid transport by regulation of transcription from RNA polymerase II promoter	0.00	0.03
GO:2000587	negative regulation of platelet-derived growth factor receptor-beta signaling pathway	0.00	0.03
GO:0060341	regulation of cellular localization	0.00	0.03
GO:0016441	posttranscriptional gene silencing	0.00	0.03
GO:0072384	organelle transport along microtubule	0.00	0.03
GO:0072525	pyridine-containing compound biosynthetic process	0.00	0.03
GO:0034968	histone lysine methylation	0.00	0.03
GO:0099518	vesicle cytoskeletal trafficking	0.00	0.03
GO:0070936	protein K48-linked ubiquitination	0.00	0.03
GO:0006941	striated muscle contraction	0.00	0.03
GO:0043487	regulation of RNA stability	0.00	0.03
GO:0035520	monoubiquitinated protein deubiquitination	0.00	0.03
GO:0009755	hormone-mediated signaling pathway	0.00	0.03
GO:0009165	nucleotide biosynthetic process	0.00	0.03
GO:0055003	cardiac myofibril assembly	0.00	0.03
GO:0030511	positive regulation of transforming growth factor beta receptor signaling pathway	0.00	0.03
GO:1903846	positive regulation of cellular response to transforming growth factor beta stimulus	0.00	0.03
GO:0034086	maintenance of sister chromatid cohesion	0.00	0.03
GO:0034088	maintenance of mitotic sister chromatid cohesion	0.00	0.03
GO:0006401	RNA catabolic process	0.00	0.03
GO:0034248	regulation of cellular amide metabolic process	0.00	0.03
GO:0006942	regulation of striated muscle contraction	0.00	0.04
GO:0035195	gene silencing by miRNA	0.00	0.04
GO:0009145	purine nucleoside triphosphate biosynthetic process	0.00	0.04
GO:0006352	DNA-templated transcription	1.56	0.00
GO:0043488	regulation of mRNA stability	0.00	0.04
GO:0007064	mitotic sister chromatid cohesion	0.00	0.04
GO:0043044	ATP-dependent chromatin remodeling	0.00	0.04
GO:0070252	actin-mediated cell contraction	0.00	0.04
GO:0035601	protein deacylation	0.00	0.04
GO:0007017	microtubule-based process	0.00	0.04
GO:0042752	regulation of circadian rhythm	0.00	0.04
GO:0009719	response to endogenous stimulus	0.00	0.04
GO:0071539	protein localization to centrosome	0.00	0.04
GO:0043984	histone H4-K16 acetylation	0.00	0.04
GO:1903312	negative regulation of mRNA metabolic process	0.00	0.04

GO:1903313	positive regulation of mRNA metabolic process	0.00	0.04
GO:0009301	snRNA transcription	0.00	0.04
GO:0003012	muscle system process	0.00	0.04
GO:1903508	positive regulation of nucleic acid-templated transcription	0.00	0.04
GO:0006103	2-oxoglutarate metabolic process	0.00	0.05
GO:0030705	cytoskeleton-dependent intracellular transport	0.00	0.05
GO:0003231	cardiac ventricle development	0.00	0.05
GO:0031047	gene silencing by RNA	0.00	0.05
GO:1902680	positive regulation of RNA biosynthetic process	0.00	0.05
GO:0098732	macromolecule deacylation	0.00	0.05
GO:0061014	positive regulation of mRNA catabolic process	0.00	0.05
GO:0043603	cellular amide metabolic process	0.00	0.05
GO:0048008	platelet-derived growth factor receptor signaling pathway	0.00	0.05
GO:0030036	actin cytoskeleton organization	0.00	0.05
GO:0007005	mitochondrion organization	0.00	0.05
GO:0051640	organelle localization	0.00	0.05

Table 9 Cardiomyocytes GO Biological Processes

3. 6. 3 Molecular functions in Cardiomyocytes

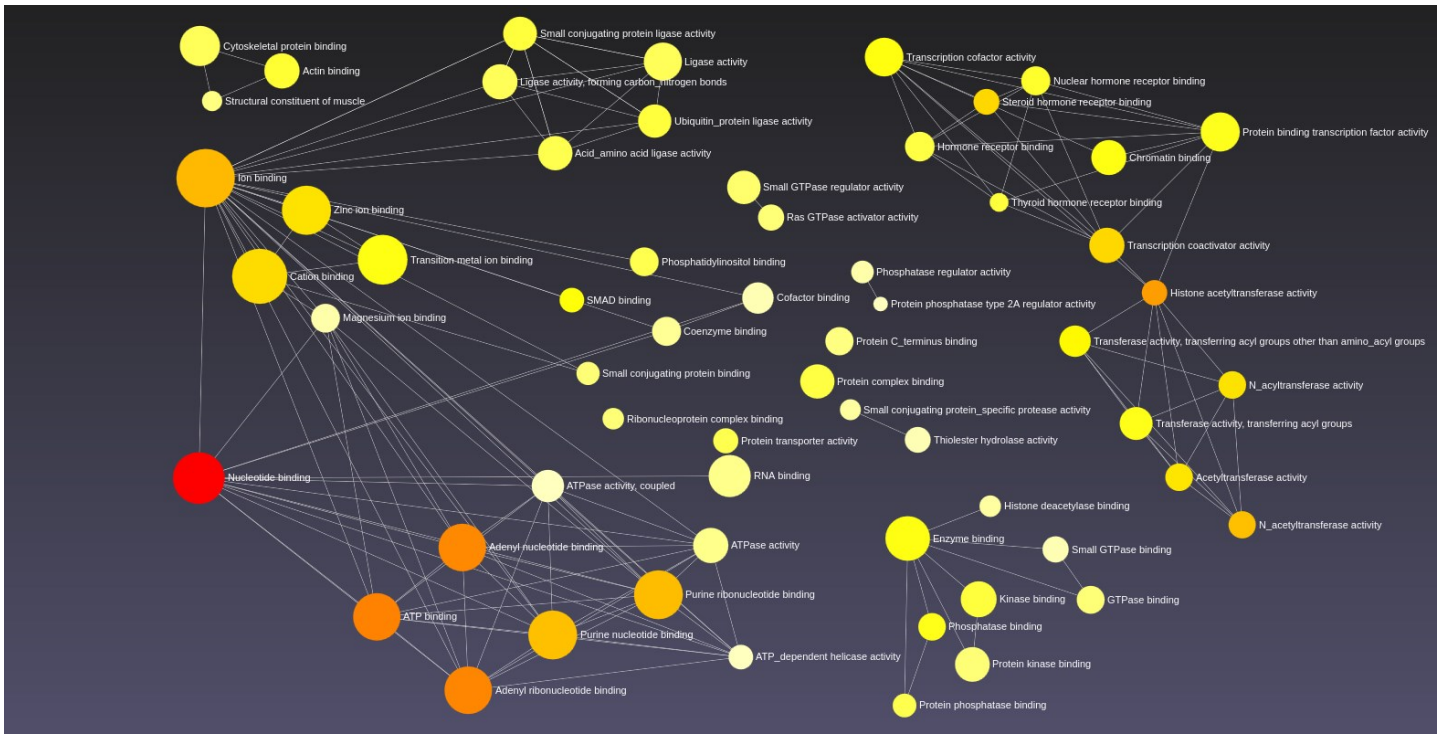


Figure 41 Cardiomyocytes GO Molecular Function

geneSet	description	pValue	FDR
GO:0003723	RNA binding	0.00	0.00
GO:0032559	adenyl ribonucleotide binding	0.00	0.00
GO:0030554	adenyl nucleotide binding	0.00	0.00
GO:0005524	ATP binding	0.00	0.00
GO:0003712	transcription coregulator activity	0.00	0.00
GO:0140110	transcription regulator activity	0.00	0.00
GO:0003713	transcription coactivator activity	0.00	0.00
GO:0032553	ribonucleotide binding	0.00	0.00
GO:0032555	purine ribonucleotide binding	0.00	0.00
GO:0017076	purine nucleotide binding	0.00	0.00
GO:0035639	purine ribonucleoside triphosphate binding	0.00	0.00
GO:0003682	chromatin binding	0.00	0.00
GO:0008144	drug binding	0.00	0.00
GO:0044877	protein-containing complex binding	0.00	0.00
GO:0008134	transcription factor binding	0.00	0.00
GO:0008080	N-acetyltransferase activity	0.00	0.00
GO:0034212	peptide N-acetyltransferase activity	0.00	0.00
GO:0035257	nuclear hormone receptor binding	0.00	0.00
GO:0019783	ubiquitin-like protein-specific protease activity	0.00	0.00
GO:0042393	histone binding	0.00	0.00
GO:0004843	thiol-dependent ubiquitin-specific protease activity	0.00	0.00
GO:0016772	transferase activity	1.42	0.00
GO:0036459	thiol-dependent ubiquitinyl hydrolase activity	0.00	0.00
GO:0101005	ubiquitinyl hydrolase activity	0.00	0.00
GO:0061733	peptide-lysine-N-acetyltransferase activity	0.00	0.00
GO:0019787	ubiquitin-like protein transferase activity	0.00	0.00
GO:0008092	cytoskeletal protein binding	0.00	0.00
GO:0140030	modification-dependent protein binding	0.00	0.00
GO:0016410	N-acyltransferase activity	0.00	0.00
GO:0004402	histone acetyltransferase activity	0.00	0.00
GO:0045296	cadherin binding	0.00	0.00
GO:0051427	hormone receptor binding	0.00	0.00
GO:0046332	SMAD binding	0.00	0.00
GO:0016407	acetyltransferase activity	0.00	0.00
GO:0019902	phosphatase binding	0.00	0.00
GO:0016746	transferase activity	1.78	0.00
GO:0017137	Rab GTPase binding	0.00	0.00
GO:0016747	transferase activity	1.83	0.00
GO:0004842	ubiquitin-protein transferase activity	0.00	0.00
GO:0019904	protein domain specific binding	0.00	0.00
GO:0008270	zinc ion binding	0.00	0.00
GO:0003700	DNA-binding transcription factor activity	0.00	0.00
GO:0050681	androgen receptor binding	0.00	0.00
GO:0008234	cysteine-type peptidase activity	0.00	0.00
GO:0019900	kinase binding	0.00	0.00
GO:0035258	steroid hormone receptor binding	0.00	0.00
GO:0000981	DNA-binding transcription factor activity	1.25	0.00
GO:0050839	cell adhesion molecule binding	0.00	0.01
GO:0035064	methylated histone binding	0.00	0.01
GO:0140034	methylation-dependent protein binding	0.00	0.01

GO:0043995	histone acetyltransferase activity (H4-K5 specific)	0.00	0.01
GO:0043996	histone acetyltransferase activity (H4-K8 specific)	0.00	0.01
GO:0046972	histone acetyltransferase activity (H4-K16 specific)	0.00	0.01
GO:0004674	protein serine/threonine kinase activity	0.00	0.01
GO:0004386	helicase activity	0.00	0.01
GO:0044389	ubiquitin-like protein ligase binding	0.00	0.01
GO:0003779	actin binding	0.00	0.01
GO:0032452	histone demethylase activity	0.00	0.01
GO:0003729	mRNA binding	0.00	0.01
GO:0031625	ubiquitin protein ligase binding	0.00	0.01
GO:0016773	phosphotransferase activity	1.34	0.00
GO:0046914	transition metal ion binding	0.00	0.01
GO:0051020	GTPase binding	0.00	0.01
GO:0016301	kinase activity	0.00	0.01
GO:0019903	protein phosphatase binding	0.00	0.01
GO:0003730	mRNA 3'-UTR binding	0.00	0.01
GO:0032451	demethylase activity	0.00	0.02
GO:0036002	pre-mRNA binding	0.00	0.02
GO:0019901	protein kinase binding	0.00	0.02
GO:0008047	enzyme activator activity	0.00	0.02
GO:0051018	protein kinase A binding	0.00	0.02
GO:0030374	nuclear receptor transcription coactivator activity	0.00	0.02
GO:0097157	pre-mRNA intronic binding	0.00	0.02
GO:0016922	nuclear receptor binding	0.00	0.03
GO:0034237	protein kinase A regulatory subunit binding	0.00	0.03
GO:0031267	small GTPase binding	0.00	0.03
GO:0017016	Ras GTPase binding	0.00	0.03
GO:0061659	ubiquitin-like protein ligase activity	0.00	0.03
GO:0008276	protein methyltransferase activity	0.00	0.04
GO:0061631	ubiquitin conjugating enzyme activity	0.00	0.04
GO:0016779	nucleotidyltransferase activity	0.00	0.04
GO:0031490	chromatin DNA binding	0.00	0.05
GO:0003714	transcription corepressor activity	0.00	0.05
GO:0070016	armadillo repeat domain binding	0.00	0.05
GO:0061980	regulatory RNA binding	0.00	0.05
GO:0016887	ATPase activity	0.00	0.05
GO:0140097	catalytic activity	1.63	0.00

Table 9 Cardiomyocytes GO Molecular Function

3.7 RNA Structural Pathology:

From the enrichment section, I selected 1064 protein coding genes with significant P values (less than 0.05). Using VEP (Variant Effect Predictor), I found their predicted transcripts and I selected all the 4904-protein coding only transcripts. However, there were various consequences to these RNA edited transcripts. So, out of these transcripts, I only selected those transcripts that belonged to “synonymous variants” types. There were 927 synonymous variant transcripts from 201 genes. These transcripts (mRNAs) when RNA edited would not give a different protein after translation if compared to their non-RNA edited counterpart. Hence the term synonymous variant. Furthermore, I only selected genes that had synonymous variant transcripts in my findings, and I did their functional enrichments. These genes are ARHGAP35, DMD, DYSF, FGD6, FOCAD, FZD3, HOMER1, MYOM2, PKN2, SNX5.And selecting only FDR values less than 0.05, I found they were involved in Cytoskeletal protein binding and Actin binding. In other words, RNA editing through its synonymous variants has a direct impact on the contractility of the heart muscles. Thus, RNA editing in our findings imply that RNA structural changes may be linked to DCM.

I have added a synonymous variants enrichment file with more details to the supplement files.

4. Discussion

In this study, we used the workflow in figure 42 to split the raw data files into 15 various cell types. And through a rigorous filtering process, we extracted quality results of RNA edited sites. We proceeded first do an exploration of our findings and then do the functional enrichments. We also highlight the case of synonymous variants and their intrinsic relationship to structural pathology.

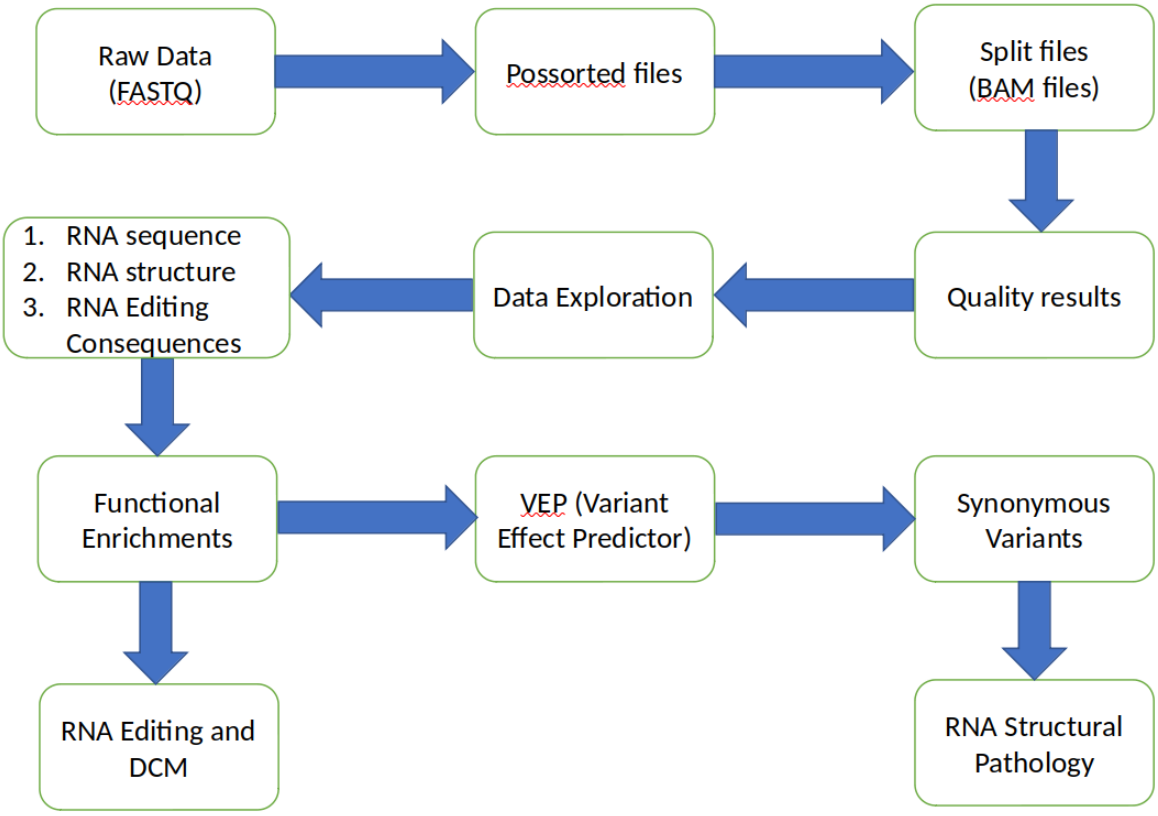


Figure 42- Study workflow

We have used single-cell transcriptomics to profile RNA editing events in different cell types of the human heart, such as fibroblasts, cardiomyocytes, endothelial cells, macrophages, pericytes, endocardium cells, smooth muscle cells, T-NK cells, adipocytes, epicardium, lymph, neurons, mast cells, and monocytes. We have shown that there are significant differences in RNA editing levels between healthy and DCM hearts and that changes in RNA editing may be driving cardiac complications in DCM. We have also shown that there are structural changes to the protein coding transcripts in DCM for various cell types. Furthermore, in our findings in the enrichment part, we can clearly see that aberrant RNA editing had an impact on

the silencing of many genes, protein degradation, and it had a major impact on the cardiac ion channels, the cytoskeleton (actin microtubules), and the metabolism pathways. This led to major changes in the contractility of the heart muscles which caused DCM. These findings are corroborated by several published scientific papers. Indeed ubiquitination (ORA.KEGG- figure.39), is the most abundant post translational modification to cellular proteins in the human body [36], and it has been implicated in Dilated Cardiomyopathy [37, 38]. Chromatin modifications (GO. BP – figure 40) with their impact on gene expression are implicated in DCM [39, 40]. Ion binding (GO.MF – figure 41) dysregulation were implicated in DCM [41, 42]. Nucleotide binding (GO.MF) [43, 44], Cytoskeletal (GO.MF) [45, 46], and last but not least the Metabolism (ORA.KEGG- figure.39) [47, 48] have all been involved in Dilated Cardiomyopathy.

Recent research studies [20, 49] have linked ADAR enzymes dysfunction to cardiovascular diseases especially cardiomyopathy. However, one recent study in Nature Medicine by Chen et al [50], investigated the microtubule detyrosination in human heart failure using mass spectrometry and single-myocyte mechanical assays. It linked the remodelling of the Cardiomyocyte's cytoskeleton in Cardiomyopathy with human heart failure. We should note that they have not used other techniques to identify other modifications. We highlight 9 genes found in this study and in our Cardiomyocytes enrichment analysis in DCM, and they are: ACTC1, MAP4, MYH6, MYH7, MYH9, SVIL, THBS4, TUBB6 and VIM. Please see table 11 below for more details. Importantly, the 9 genes as analysed by Reactome bioinformatics platform (reactome.org) shows they are mainly involvement in the structural constituent of cytoskeleton and structural molecule activity.

Description	Gene list
RNA binding	MAP4, MYH9, VIM
cell adhesion molecule binding	THBS4
protein-containing complex binding	MYH6, MYH7, MYH9, SVIL, THBS4, VIM
protein domain specific binding	VIM
adenyl ribonucleotide binding	ACTC1, MYH6, MYH7, MYH9
adenyl nucleotide binding	ACTC1, MYH6, MYH7, MYH9
ribonucleotide binding	ACTC1, MYH6, MYH7, MYH9, TUBB6
purine ribonucleotide binding	ACTC1, MYH6, MYH7, MYH9, TUBB6
purine nucleotide binding	ACTC1, MYH6, MYH7, MYH9, TUBB6
purine ribonucleoside triphosphate binding	ACTC1, MYH6, MYH7, MYH9, TUBB6
kinase binding	MYH6
protein kinase binding	MYH6
ATPase activity	ACTC1, MYH6, MYH7, MYH9
ATP binding	ACTC1, MYH6, MYH7, MYH9
actin binding	MYH6, MYH7, MYH9, SVIL
cadherin binding	MYH9
cytoskeletal protein binding	ACTC1, MAP4, MYH6, MYH7, MYH9, SVIL

Table 11: Overlap Genes from Chen et al study and our Cardiomyocytes Enrichment Analysis in DCM

The Cytoskeleton is a complex network of protein fibres distributed throughout the cytoplasm and it has 3 main components:

1. Microfilaments: They are actin protein-based filaments. Actin proteins together with myosin proteins form an actomyosin complex. This actomyosin complex allows muscle contraction and relaxation. In our enrichment analysis, ACTC1 (actin alpha cardiac muscle 1), MYH6 (myosin heavy chain 6), MYH7 (myosin heavy chain 7), MYH9 (myosin heavy chain 9) are RNA edited and significantly involved in Dilated Cardiomyopathy (DCM).

SVIL (supervillin) and THBS4 (thrombospondin 4), both associated with actin filaments and plasma membranes, are also significantly involved in Dilated Cardiomyopathy (DCM) in our analysis.

2. Intermediate fibres: They are structural filaments that strengthen the shape of the cell and prevent damage to it during mechanical movements. They are made of different proteins such as vimentin (VIM), desmin, keratin and lamin. We found in our analysis the VIM gene, in Cardiomyocytes, to be significantly RNA edited in DCM.

3. Microtubules: They are formed from subunits of the protein tubulin. These subunits assemble to form tubulin polymers, which then form hollow microtubules. These microtubules are involved in shaping the cell, in the movement of organelles, and in forming the spindle fibres that are involved in the movement of chromosomes during mitosis and meiosis. We found in our analysis the TUBB6 (Tubulin Beta 6 Class V) gene and MAP4 (Microtubule-Associated Protein 4) gene, in Cardiomyocytes, to be significantly RNA edited in DCM. It is worth noting that MAP4, is the most prominent MAP in the heart, and is instrumental in regulating the stability, transport, and interactome of the microtubules [51].

5. Conclusion:

In summary what has emerged from this research is the relationship between RNA editing and Cardiomyocytes' motility. This cell motility is regulated through transcriptional modifications as well as post-translational modifications as illustrated in figure 43 below, and both modifications seem to be regulated by RNA editing.

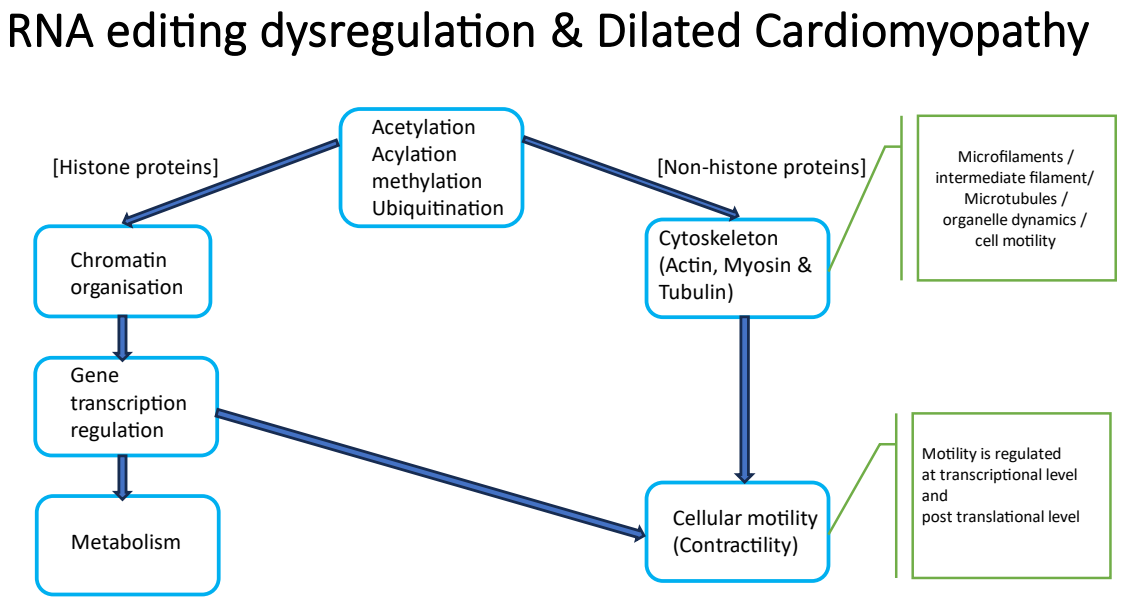


Figure 43- RNA editing dysregulation and Dilated Cardiomyopathy

This research is based on data taken from 35 human heart samples. There was diversity in our samples, but the number of samples remains small. A larger sample in the hundreds or thousands may yield higher quality results with higher confidence. However, it would take a longer time to process. Moreover, our analysis is based on Single Nuclei processed data. Therefore, we lack in this research the cytoplasmic data. Combining the results of this research with Single Cell data would give us a clearer view of RNA editing dysregulation in the transcriptome and its impact in DCM.

Understanding the intricate mechanisms of RNA editing dysregulation in DCM holds promise for drug development, precision medicine, and personalised treatment approaches tailored to individuals at risk or already diagnosed with DCM. By elucidating the molecular underpinnings of DCM, clinicians can potentially identify early biomarkers and develop targeted therapies aimed at restoring cytoskeletal integrity and improving cardiac muscle motility, ultimately enhancing patient outcomes and quality of life.

6. References

1. Tan, B.Z., et al., *Dynamic regulation of RNA editing of ion channels and receptors in the mammalian nervous system.* Mol Brain, 2009. **2**: p. 13.
2. Park, E., et al., *Genetic variation and microRNA targeting of A-to-I RNA editing fine tune human tissue transcriptomes.* Genome Biol, 2021. **22**(1): p. 77.
3. Roberts, J.T., et al., *ADAR Mediated RNA Editing Modulates MicroRNA Targeting in Human Breast Cancer.* Processes (Basel), 2018. **6**(5).
4. Varani, G. and W.H. McClain, *The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems.* EMBO Rep, 2000. **1**(1): p. 18-23.
5. Ruszkowska, A., et al., *Structural Insights Into the 5'UG/3'GU Wobble Tandem in Complex With Ba(2+) Cation.* Front Mol Biosci, 2021. **8**: p. 762786.
6. Cozzolino, F., et al., *Protein-DNA/RNA Interactions: An Overview of Investigation Methods in the -Omics Era.* J Proteome Res, 2021. **20**(6): p. 3018-3030.
7. Jarmoskaite, I., A.E. Helmers, and R. Russell, *Measurement of ATP utilization in RNA unwinding and RNA chaperone activities by DEAD-box helicase proteins.* Methods Enzymol, 2022. **673**: p. 53-76.
8. Kokot, K.E., European Heart Journal, Volume 43, Issue Supplement_2, October 2022, ehac544.2962.
9. Li, Q., et al., *RNA editing underlies genetic risk of common inflammatory diseases.* Nature, 2022. **608**(7923): p. 569-577.
10. Zhou, X., et al., *Genomic Landscape and Potential Regulation of RNA Editing in Drug Resistance.* Adv Sci (Weinh), 2023: p. e2207357.
11. Chan, T.W., et al., *RNA editing in cancer impacts mRNA abundance in immune response pathways.* Genome Biol, 2020. **21**(1): p. 268.
12. Nussbacher, J.K., et al., *Disruption of RNA Metabolism in Neurological Diseases and Emerging Therapeutic Interventions.* Neuron, 2019. **102**(2): p. 294-320.
13. Shafik, A.M., E.G. Allen, and P. Jin, *Epitranscriptomic dynamics in brain development and disease.* Mol Psychiatry, 2022. **27**(9): p. 3633-3646.
14. Silberberg, G., et al., *Deregulation of the A-to-I RNA editing mechanism in psychiatric disorders.* Hum Mol Genet, 2012. **21**(2): p. 311-21.
15. NHS. Available from: https://www.nhs.uk/conditions/cardiomyopathy/.
16. Townsend, N., et al., *Epidemiology of cardiovascular disease in Europe.* Nat Rev Cardiol, 2022. **19**(2): p. 133-143.
17. Ohira, T. and H. Iso, *Cardiovascular disease epidemiology in Asia: an overview.* Circ J, 2013. **77**(7): p. 1646-52.
18. Stellos, K., et al., *Adenosine-to-inosine RNA editing controls cathepsin S expression in atherosclerosis by enabling HuR-mediated post-transcriptional regulation.* Nat Med, 2016. **22**(10): p. 1140-1150.
19. Garcia-Gonzalez, C., et al., *ADAR1 Prevents Autoinflammatory Processes in the Heart Mediated by IRF7.* Circ Res, 2022. **131**(7): p. 580-597.
20. El Azzouzi, H., et al., *Cardiomyocyte Specific Deletion of ADAR1 Causes Severe Cardiac Dysfunction and Increased Lethality.* Front Cardiovasc Med, 2020. **7**: p. 30.
21. Kokot, K.E., et al., *Reduction of A-to-I RNA editing in the failing human heart regulates formation of circular RNAs.* Basic Res Cardiol, 2022. **117**(1): p. 32.
22. Wu, X., et al., *ADAR2 increases in exercised heart and protects against myocardial infarction and doxorubicin-induced cardiotoxicity.* Mol Ther, 2022. **30**(1): p. 400-414.
23. Kwak, S.H., et al., *Mitochondrial metabolism and diabetes.* J Diabetes Investig, 2010. **1**(5): p. 161-9.

24. Gaither, J.B.S., et al., *Synonymous variants that disrupt messenger RNA structure are significantly constrained in the human population.* Gigascience, 2021. **10**(4).
25. Koenig, A.L., et al., *Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure.* Nat Cardiovasc Res, 2022. **1**(3): p. 263-280.
26. 10X-Genomics. Available from: 10xgenomics.com.
27. TimOast. Available from: https://timoast.github.io/sinto/.
28. Samtools. Available from: http://www.htslib.org.
29. Picardi, E., et al., *Using REDItools to Detect RNA Editing Events in NGS Datasets.* Curr Protoc Bioinformatics, 2015. **49**: p. 12 12 1-12 12 15.
30. Lorenz, R., et al., *ViennaRNA Package 2.0.* Algorithms Mol Biol, 2011. **6**: p. 26.
31. Cunningham, F., et al., *Ensembl 2022.* Nucleic Acids Res, 2022. **50**(D1): p. D988-D995.
32. EBI. *EBI*. Available from: http://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_human/release_43/.
33. ZhangLab. *WebGestalt*. Available from: https://www.webgestalt.org/.
34. Eisenberg, E. and E.Y. Levanon, *A-to-I RNA editing - immune protector and transcriptome diversifier.* Nat Rev Genet, 2018. **19**(8): p. 473-490.
35. Expressanalyst. Available from: https://www.expressanalyst.ca.
36. Damgaard, R.B., *The ubiquitin system: from cell signalling to disease biology and new therapeutic opportunities.* Cell Death Differ, 2021. **28**(2): p. 423-426.
37. Spanig, S., et al., *The Ubiquitin Proteasome System in Ischemic and Dilated Cardiomyopathy.* Int J Mol Sci, 2019. **20**(24).
38. Luo, H., J. Wong, and B. Wong, *Protein degradation systems in viral myocarditis leading to dilated cardiomyopathy.* Cardiovasc Res, 2010. **85**(2): p. 347-56.
39. Chapski, D.J., et al., *Early adaptive chromatin remodeling events precede pathologic phenotypes and are reinforced in the failing heart.* J Mol Cell Cardiol, 2021. **160**: p. 73-86.
40. Tabish, A.M., et al., *Association of intronic DNA methylation and hydroxymethylation alterations in the epigenetic etiology of dilated cardiomyopathy.* Am J Physiol Heart Circ Physiol, 2019. **317**(1): p. H168-H180.
41. Molina-Navarro, M.M., et al., *Differential gene expression of cardiac ion channels in human dilated cardiomyopathy.* PLoS One, 2013. **8**(12): p. e79792.
42. El-Battrawy, I., et al., *Ion Channel Dysfunctions in Dilated Cardiomyopathy in Limb-Girdle Muscular Dystrophy.* Circ Genom Precis Med, 2018. **11**(3): p. e001893.
43. Yamada, T. and S. Nomura, *Recent Findings Related to Cardiomyopathy and Genetics.* Int J Mol Sci, 2021. **22**(22).
44. Luo, B., et al., *Association of nucleotide-binding oligomerization domain-like receptor 3 inflammasome and adverse clinical outcomes in patients with idiopathic dilated cardiomyopathy.* Clin Chem Lab Med, 2013. **51**(7): p. 1521-8.
45. Le Dour, C., et al., *Actin-microtubule cytoskeletal interplay mediated by MRTF-A/SRF signaling promotes dilated cardiomyopathy caused by LMNA mutations.* Nat Commun, 2022. **13**(1): p. 7886.
46. Towbin, J.A. and N.E. Bowles, *Dilated cardiomyopathy: a tale of cytoskeletal proteins and beyond.* J Cardiovasc Electrophysiol, 2006. **17**(8): p. 919-26.
47. Ampong, I., *Metabolic and Metabolomics Insights into Dilated Cardiomyopathy.* Ann Nutr Metab, 2022. **78**(3): p. 147-155.
48. Haas, J., et al., *Energy Metabolites as Biomarkers in Ischemic and Dilated Cardiomyopathy.* Int J Mol Sci, 2021. **22**(4).
49. Chen, J., et al., *Adenosine deaminase acting on RNA 1 (ADAR1) as crucial regulators in cardiovascular diseases: structures, pathogenesis, and potential therapeutic approach.* Front Pharmacol, 2023. **14**: p. 1194884.
50. Chen, C.Y., et al., *Suppression of detyrosinated microtubules improves cardiomyocyte function in human heart failure.* Nat Med, 2018. **24**(8): p. 1225-1233.

51. Caporizzo, M.A. and B.L. Prosser, *The microtubule cytoskeleton in cardiac mechanics and heart failure.* Nat Rev Cardiol, 2022. **19**(6): p. 364-378.