MSB Research Project 2023

Discovering regulatory switches in gene co-expression networks of dilated cardiomyopathy (DCM)

Ramón Reszat[a], Patrick Mertens[a], Luca Farolina[a], Alaric de Biolley[a]

[a]*Maastricht University, Faculty of Science and Engineering, Paul-Henri Spaaklaan 1, 6229 EN Maastricht, The Netherlands*

Abstract

In this paper, we use the SWIMMER data mining tool to discover switch genes in a gene co-expression network of dilated cardiomyopathy (DCM). The method aims to select for nodes that are characterized by their position at the periphery of clusters identified within the network. Those topological characteristics allows them to have a regulatory role on this different subgraphs of highly co-expressed genes. The resulting list of genes is linked to relevant pathways and regulatory elements using information from the WikiPathways and ENCODE databases. In addition to that, from exploratory analysis, we discover clusters in the transcriptome of cardiomyocytes that hint to a potential viral cause of DCM. Outliers in the values for *CEL*, *SELE*, *ADIPOQ*, *OSR1*, *PRKAG2-AS1* and *MYH6* are chosen due to a suspected association with viral causes. This resulting samples are labeled as Viral Cardiomyopathies (VCM). The discovered switch genes seem to be consistent with the phenotype for DCM. The affected pathways include T cell activation, selective expression of chemokines, sulfanation biotransformation, sphingolipid metabolism and striated muscle contraction. In the VCM cases, we observe an increased amount of signalling pathways while inflammatory pathways are less present compared to the the DCM phenotype. Further research is needed to clarify i.e. the role of sphinganine and sphingosine phosphorylation enabled by *SGPP2* in dilated cardiomyopathy and the roles of *SULT1A2* and *SULT1B1* in the sulfanation biotransformation pathway and their respective up- and down-regulation in dilated cardiomyopathy. Overall, this method represents a promising approach to hypothesis generation from a network of a complex disease.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2018).

Keywords: switch genes, co-expression networks, dilated cardiomyopathy, viral myocarditis

1. Introduction

Dilated cardiomyopathy (DCM) is a heart muscle disorder characterized by an enlargement of the left ventricle and reduced contractility, which is indicated by a left ventricular ejection fraction of less than 40% [1, 2, 3]. The complications that arise because of DCM increase the chance of developing heart failure (HF) [2, 3]. Therefore, determining the presence, as well as being able to discriminate between different sub-types of DCM has the potential to improve its treatment. At the moment, the most common method of determining DCM is through imaging techniques such as echocardiography or cardiac magnetic resonance [2, 3, 4, 5]. In this project, we dive deeper into genetic mechanisms focusing on gene expression collected from patients who were affected by DCM and healthy non-failing (NF) heart donors. Our goal is to identify regulatory mechanism-related genes comparing the this two phenotypes.

Therefore, we have decided to employ SWIMMER [6], which is a network-based algorithm that extracts switch genes within the co-expression network constructed from the left ventricular transcriptome. The main hypothesis behind this computational tool is that switch genes are essential for the transition from a healthy (NF) to a sick state (DCM). In fact, it is widely established that DCM has a genetic aspect; especially in the case of familial DCM [2, 3, 4]. However, DCM causes are not easily identifiable, and a lot of DCM cases are labelled with an idiopathic cause [2, 3]. Even so, some literature proves that further analyses of idiopathic DCM case could lead to a specific cause [2, 3, 7]. Albeit this possibility, researching DCM for possible genetic biomarkers or drug targets can be difficult because a wide variety of causes can lead to DCM [2, 8, 7]. Namely, DCM can be caused by viruses [9], bacteria, fungi, exposure to toxins, autoimmune disorders, and metabolic or endocrine dysfunction [2, 8, 7]. In fact, viruses can be confounding variables, thus influencing the validity of the research [10], due to their reproduction mechanisms, taking over the protein machinery of the host and rapidly creating more viruses [11, 12, 13]. This leads to a drastically different gene expression profile [11, 12]. The difference in gene expression between an infected cell and a healthy cell can be significant because it has been reported that some viruses can cause a 2-fold change of 10% of the genes [11, 12]. Therefore, we aim to create a subset of DCM based on a clustering of genes that have been associated with viral infections. The selected genes are: CEL, SELE, ADIPOQ, OSR1, PRKAG2-AS1 [9], and MYH6 [14, 15]. These genes were chosen after exploratory research of the data has shown that these were most optimal for clustering.

2. Methodology

For this research project, we decided to explore elements of gene regulation that can be extracted from transcriptomics data using a network analysis approach. Our goal is to find switch genes that play a connector role within the gene correlation network and show a negative association with the expression of genes in different functional hubs. The code for our analysis is adapted from the R implementation of the SWIMMER [16] data mining tool. The underlying MAGnet dataset [17] contains mRNA-seq expression profiles in counts per million (CPM), metadata with relevant phenotype information such as aetiology, age and gender for the study participants, as well as exon lengths for each gene that are used to normalize counts and compare relative expression levels. As the focus of our analysis is on dilated cardiomyopathy (DCM), we initially use a case-control study design of DCM cases versus non-failing heart donors with 166 samples in each group. After defining the design and contrast matrix accordingly, we run a differential expression analysis using the 'limma': mixed linear models for microarrays and mRNA-seq [18] package in R. We fit a linear mixed model to the gene expression data from the MAGnet dataset. Our model is corrected for gender and age, as these constitute confounding factors for the expression of the genes considered in the analysis.

$$Gene\ Expression = \begin{cases} \beta_1 + \beta_0 & \text{if case is DCM} \\ \beta_0 & \text{if case is NF} \end{cases} + \begin{cases} \beta_3 + \beta_2 & \text{if person is male} \\ \beta_2 & \text{if person is female} \end{cases} + \beta_4\ \text{Age} \tag{1}$$

The effects on the gene expression considered by the model are shown in equation (1). The differentially expressed genes are extracted from the first coefficient β_1 describing the contrast between dilated cardiomyopathy (DCM) and the non-failing (NF) heart. The coefficients β_3 and β_4 are correction terms that describe the mean differences between gender and age-related effects on gene expression. Afterwards, we apply a threshold of 0.05 for the adjusted p values and 1.5 log-fold change to derive a list of significantly up and down-regulated genes between these conditions. Moreover, as previously stated in the introduction, viruses have an effect on gene expression. Specific genes cluster the samples that are potentially affected by viruses. Therefore, samples that have a very high expression in the following genes: CEL, SELE, ADIPOQ, OSR1, PRKAG2-AS1 [9], and MYH6 [14, 15] are selected to be in a separate group labelled as viral cardiomyopathy (VCM). The selection of the group is made with hierarchical average clustering [19]. Furthermore, the samples that are selected will be excluded from the DCM group and thus, leading to the following experimental groups: NF, DCM, and VCM. In the following parts we are going to present the results obtained comparing NF vs DCM and NF vs VCM.

$$\rho_{R(x_i),R(x_j)} = \frac{cov(R(x_i), R(x_j))}{\sigma_{R(x_i)}\sigma_{R(x_j)}} \tag{2}$$

The next step is to calculate the Spearman correlation between the expression profiles of each gene (2). Using the correlation coefficient from ranked gene expression values has several advantages. First of all, it will remove noise

from the genomic dataset equivalently to perform rank normalization [20]. Additionally, the chosen correlation metric can capture monotonical changes that are not linear as we would expect with switching behaviour. The resulting values can be used to weight the edges in a gene co-expression network [21] by converting the correlation to a distance metric. Edges are included based on a hard threshold if their value lies above the 90th-percentile of edge weights. Moreover, we can use the average correlation coefficient of each node to subdivide important hub genes into three main categories [22]. Those can be visualized in the density correlation plot that is created for diagnostic purposes. The middle peak contains *Party-Hubs*, central nodes in the network with a positive average correlation and a high node degree. They are usually considered to be centres of modules with a common function. A second set of nodes with low connectivity and few high correlations are referred to as *Date-Hubs*. Existing network analysis approaches such as WGCNA [23] have been developed in order to find hubs with positive correlations between genes. However, in this paper, we will focus on a third type of topological group called *Fight-Clubs* [24]. Genes with this role have a high negative average correlation to their direct neighbours. In the final co-expression network gene clusters are detected using the k-means method. The optimal number of initial centroids [25] was chosen using the elbow method. After performing the clustering, we can calculate the metrics that are used for detecting switch genes.

$$K_{\pi}^{i} = 1 - \left(\frac{k_i^{in}}{k_i} \right)^2 \tag{3}$$

The *clusterophobic coefficient* K_{π}^{i} (3) counts the connection a gene has inside of its cluster compared to the overall connections within the network k_i. This metric is used to find peripheral genes with a connector function.

$$z_i = \frac{k_i^{in} - k_{C_i}^{in}}{\sigma_{C_i}^{in}} \tag{4}$$

The *within-module degree* z_i is a normalized measure of how many connections are within a cluster versus connections that are leaving the cluster. The SWIMMER algorithm selects genes that have a high clusterophobic coefficient K_{π}^{i} and a low within module degree z_i. Additionally, each gene has a high negative average correlation which corresponds to the fight club property. Those metrics are compared to each other in the heat cartography map following the Guimerà-Amaral [26, 27] approach of defining regions in complex networks. Thus switch genes can be found in the R4 region of the map. Our hypothesis is that those genes may have regulatory roles in the DCM versus the non-failing phenotype. At this point, we have filtered out a list of switch gene that can be loaded into Cytoscape [28] for further analysis using CyTargetLinker [29]. This allows us to identify transcription factors and pathways related to the hypothesized switch genes and map the transcriptomics data [30] to the pathways extracted from the WikiPathways [31] database. Therefore, we can generate hypotheses and discuss the role and function of switch genes in developing dilated cardiomyopathy.

3. Results

As it is mentioned in the method section, other than just exploring regulatory mechanisms, one of the results was to be able to select samples that are likely associated with viral causes within the DCM group. In order to avoid including possible confounding factors we have relabelled these samples in order to distinguish VCM and DCM. In the following section the two contrast DCM vs NF and VCM vs NF are described. From the first contrast of the differential gene expression analysis a total of 342 significantly altered genes are identified. Among those, 237 were up-regulated and 105 down-regulated. Those are the genes that are used to create a correlation-weighted network and to identify modules within it.

3.1. Network analysis of Dilated Cardiomyopathy

The network is a coarse-grained view of the regulatory effects observed in DCM. If every correlation is considered, we would have a fully connected network of every up- and down-regulated gene. Therefore, we need to filter out edges to consider for further analysis.

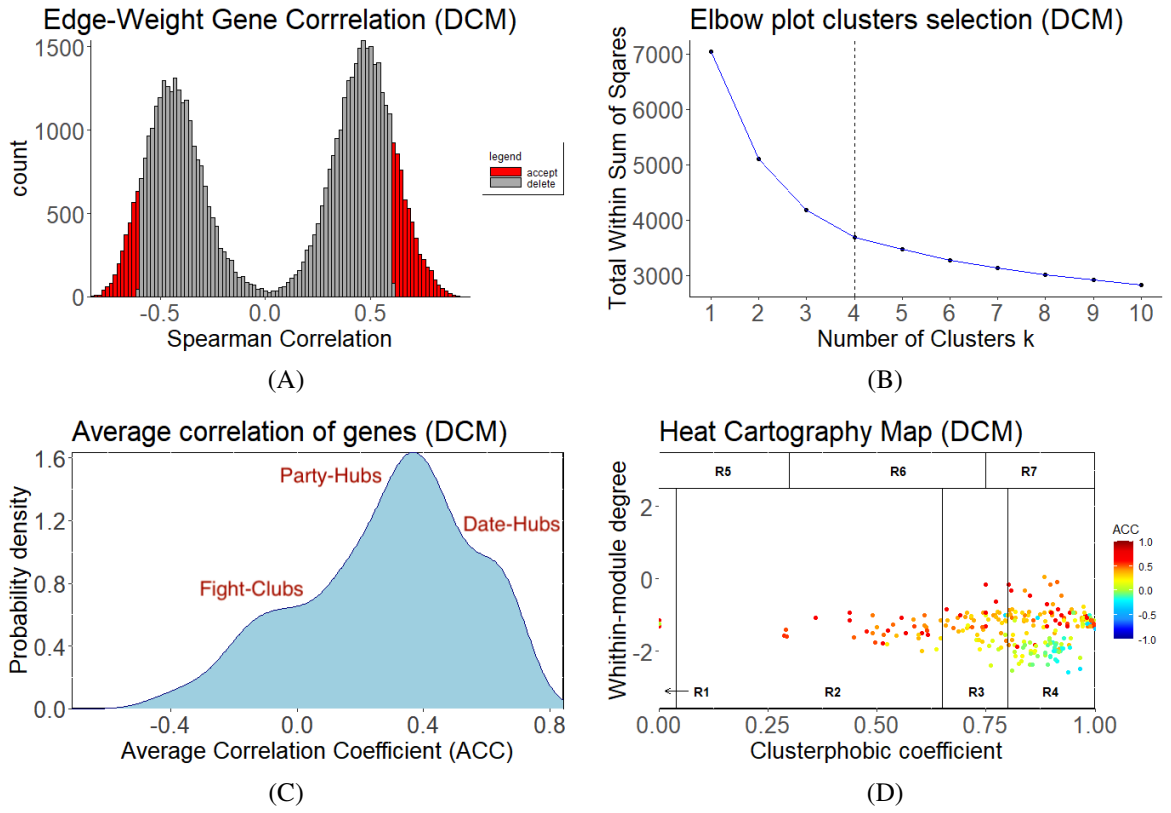


Fig. 1. Identification of switch genes in a gene correlation network of samples with suspected viral cardiomyopathy (A) histogram of edge-weights for the gene co-expression network with a hard threshold of $|\rho| \geq 0.6$ (B) Elbow-method to determine the optimal number of clusters in the network ($k = 4$) (C) density-correlation plot showing the density of *Fight-Clubs*, *Party-Hubs* and *Date-Hubs* in the network ($|\rho| \geq 0.6$, $k = 4$) (D) network cartography of differentially expression genes with regions according to the Amaral-Guimerà approach

Fig.1 presents the results from the SWIMMER algorithm. A panel of four plots is used to tune the parameters for this analysis. Plotting the edge weights in a histogram yields a symmetric distribution that can be observed in Fig. 1A. This plot is used to determine the threshold that is applied to the correlation on the edges before the network is constructed. In this case we follow an hard thresholding approach cutting out edges having a correlation coefficient lower than $|\rho| \geq 0.6$ in order to build the network with only the top 10% of the most strongly weighted edges. The positive and negative correlations between genes generate two distinct distributions centered around in -0.5 and 0.5 respectively. Considering negative correlations is crucial, especially if we aim to study regulatory mechanisms. The resulting network is divided into clusters using *NbClust* and the optimal number of clusters is chosen accordingly. The tool performs a grid search over commonly used clustering techniques such as k-means, Ward's and complete clustering to minimize the intracluster difference as the total within sum of squares according to multiple distance measures including Euclidean, Maximum and Canberra distance. For the optimal method and metric the elbow plot is shown Fig. 1B. In the average correlation density plot in Fig. 1C the presence of fight-clubs becomes visible as a small peak at a negative average correlation coefficients (ACC). The next step in the workflow is to extract the switch genes by considering the clusterophobic coefficient and the within module degree of each node. Using the metrics described in (3) and (4) it is possible to create the heat cartography map shown in Fig. 1D. It indicates the presence of switch genes falling into the R4 region, which corresponds to a high clusterophobic coefficient and low within module degree. Those genes are annotated with data from the Ensemble database and exported to be loaded as a node table in Cytoscape. From the contrast between dilated cardiomyopathy and healthy heart donors, we could extract 45 switch genes that fall into this category.

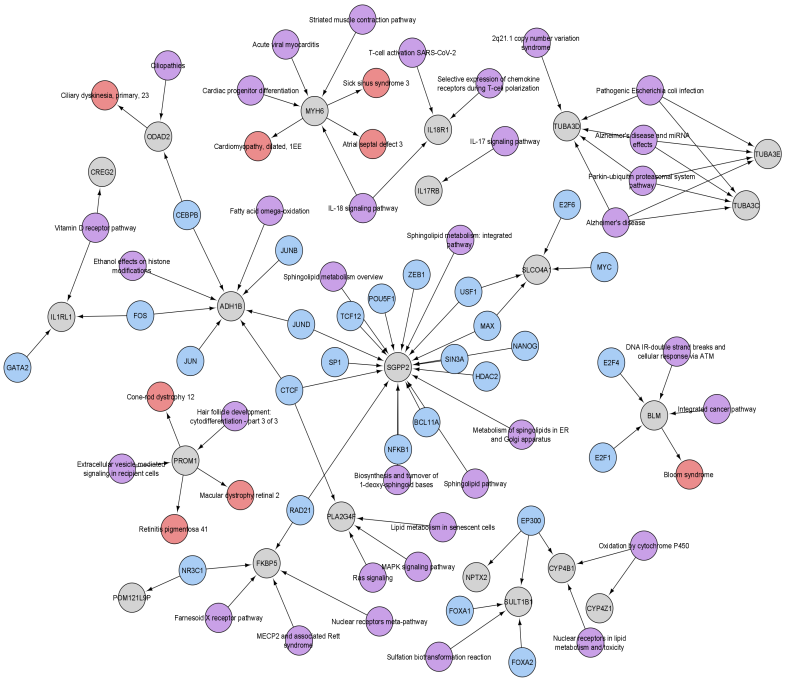


Fig. 2. List of switch genes was used to generate a Cytoskape network for DCM. Gray circles are used to depict the switch genes. The network is expanded using CyTargetLinker to include WikiPathways (purple circles), diseases (light corals circles) and ENCODE elements (blue circles).

Finally, we obtain a network in Cytoskape using CyTargetLinker to integrate data from different databases and interpreting the biological meaning. Among the switch genes, there is a subgroup targeted and linked by transcription factors (*SGPP2*, *SLCO4A1*, *ADH1B*, *PLA2G4F*, *FKBP5*). In Fig 2 this is shown toghether with other smaller subset of genes as for example (*NPTX2*, *SULT1B1*, *CYP4B1*) that are linked toghether by the transcription factor *EP300*. The *TUBA3D*, *TUB3E* and *TUBA3C* genes are expressing members of the alpha-tubulin family [32]. If instead the wikipathways(violet nodes) are taken into consideration it is possible to notice that some of those genes are linked to a possible immune response, such as *IL17RB* or *IL18R1*. In fact we can observe in the network WikiPathway entries such as T cell activation or selective expression of chemokines. Among the biological pathways we have further explored are the striated muscle contraction pathway (Appendix A), sphingolipid metabolism (Appendix B), sulfanation bio-transformation reactions (Appendix C), lipid metabolism in senescent cells (Appendix D) and the farnesoid X receptor pathway (Appendix E). The corresponding pathways with the changes in expression between NF and DCM are attached in the appendix.

3.2. Suspected Viral Causes of Dilated Cardiomyopathy

As we have observed some of the WikiPathways entries are related to viral infection, we decided to furtheranalyze these samples to see if it was possible to filter out viral causes. In order to do so the VCM cluster is established by mapping all the DCM cases against the following genes: CEL, SELE, ADIPOQ, OSR1, PRKAG2-AS1 [9], and MYH6 [14, 15]. This mapping can be seen in Fig. 3A+B. In the mapping, it can be seen that the majority of the samples in the rows show relatively small gene expression levels, and only a few samples with a notably high expression indicated by red. The outliers with higher expression values are selected for each gene (Fig. 3B) and a subgroup of susspected viral causes is created.

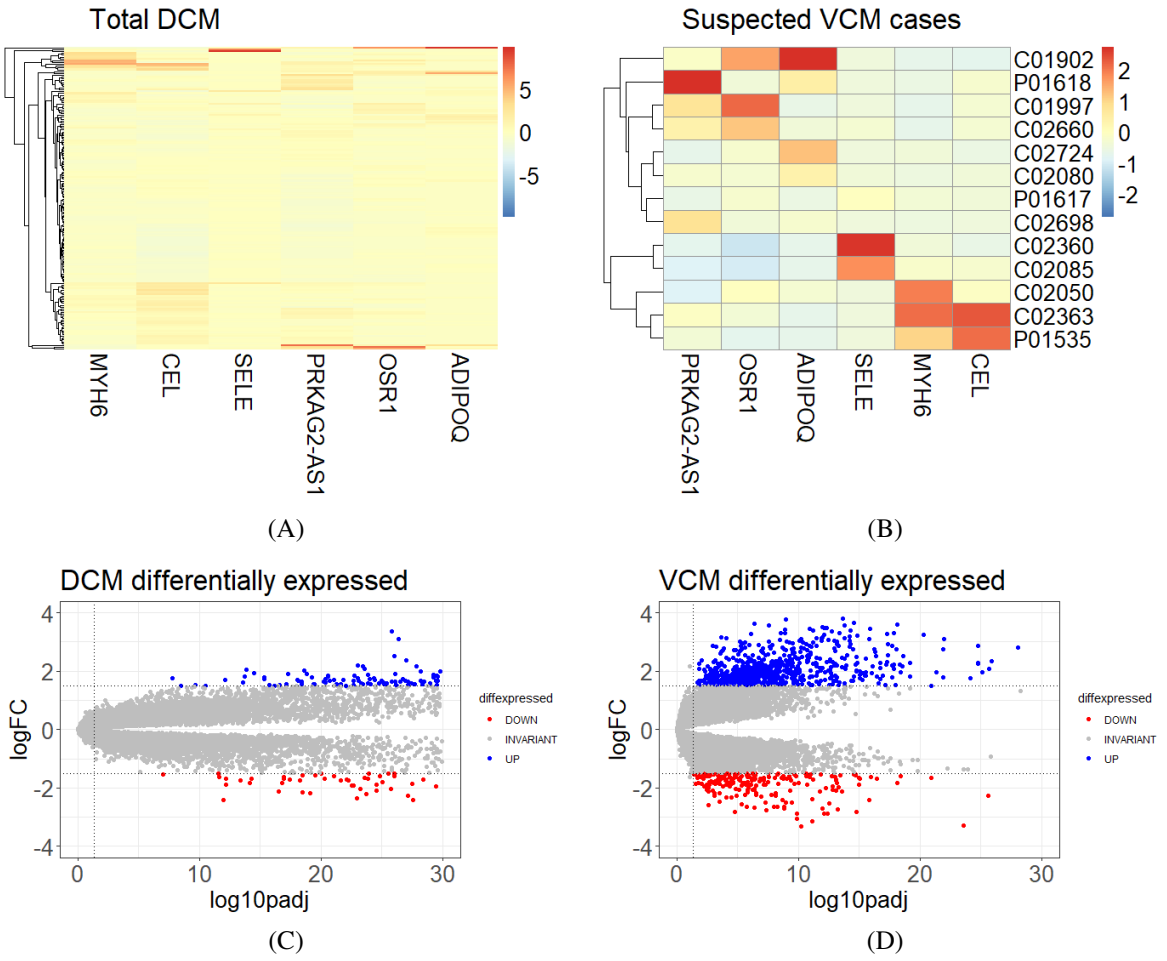
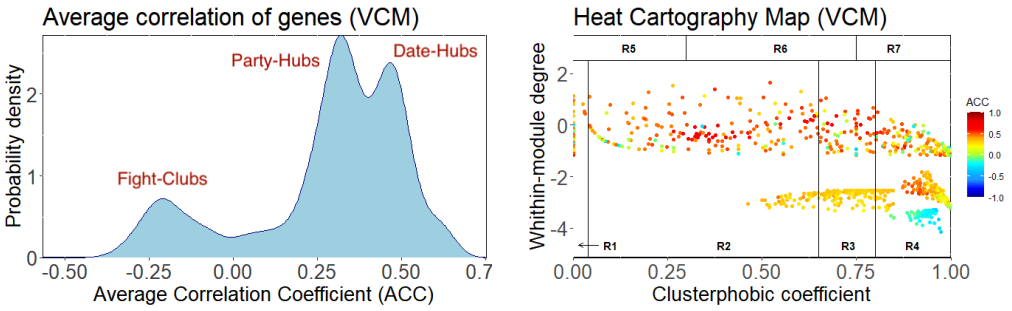


Fig. 3. Heatmap of relative expression levels in FPKM for (A) all samples labelled with the DCM aetiology, (B) selection of samples with suspected viral cause (C)volcano plot showing differential expressed gene according to pval = 0.05 and foldchange = 2 in all DCM patience and in (D) VCM patients

VCM vs NF turned out to be a very insightful contrast. Differential expression analysis resulted in more differentially expressed genes, which we were expecting to see. It is especially interesting to see the difference between the two contrasts already after this first filtering step as by maintaining the same thresholds for the adjusted p values and fold changes we are selecting a significantly greater number of genes. Notably many of those genes have considerably higher fold changes if compared to DCM vs NF (vulcanoplot). For the following steps required for the extraction of switch genes, to compare the two contrasts, we have used the same parameters to set the correlation threshold and established the same number of clusters. The average correlation density plot we have obtained shows a clear presence of party and date hub and even more significantly the presence of fight clubs. This separation based on average correlation is sharper with respect to DCM. More over the heat cartography map show very clearly a cluster of genes in the R4 region with a negative average correlation coefficient meaning switch genes are very well separated with respect to the other groups.



(a) (b)

Fig. 4. Identification of switch genes among the differentially expressed genes in VCM (a) density-correlation plot showing the density of *Fight-Clubs*, *Party-Hubs* and *Date-Hubs* in the network ($|\rho| \geq 0.5$, $k = 4$) (b) network cartography according to the Amaral-Guimerà approach

Similarly to what was done with DCM samples we have explored a network using cytoskape and cytarget linker to try to understand the similarity and difference with DCM patients. Many of the switch genes are present in both the groups. Indeed important nodes such as SGPP2 MYH6 are present in both contrasts. On the other side In VCM there is a notable increase in many kind of signalling pathways which is a coherent finding with the inspection that was done by observing the plots produced by SWIMMER. Nodes that we found interesting in this network include Hepatitis C and hepatocellular carciroma Wnt, MAPK, NRF2, PI3K-Akt-mTOR as we are going to further elaborate in the discussion section.

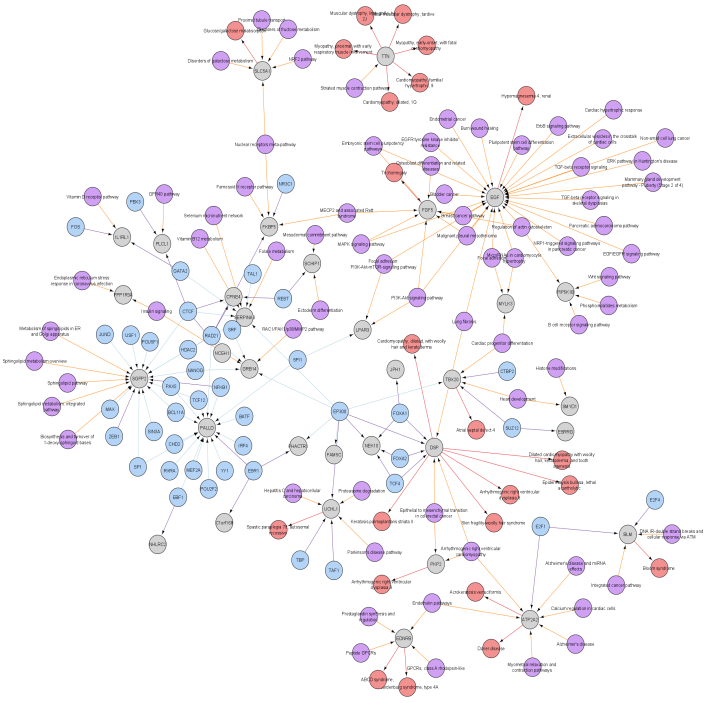


Fig. 5. List of switch genes was used to generate a Cytoscape network for VCM. Gray circles are used to depict the switch genes. The network was then expanded using CyTargetLinker to include WikiPathways (purple circles), diseases (light corals circles) and ENCODE elements (blue circles).

4. Discussion

First, we examine the striated muscle contraction pathway where we have observed down-regulation of the *MYH6* gene in dilated cardiomyopathy (Appendix A). This gene codes for the cardiac alpha-myosin heavy-chain protein [33], which functions as a component of the type II myosin protein and is involved in cardiac muscle contraction [34]. We have used this gene to select VCM patients, as it seem to be especially up-regulated in some viral cases [14, 15]. Hence, we have noticed that a pathway associated with this gene is acute viral myocarditis. In fact, Fig 3B shows that *MYH6* is toghether with *CEL* is useful to select for example patients C02363, P01535 together with *CEL* and more broadly seems to have a similar expression profile in some patients together with *SELE, OSR1* and *ADPOQ*. However it is important to mention that this gene is detected as a switch also in the DCM contrast. Further research might be needed to really understand the role of this gene in viral cases.

Furthermore, we discovered *SGPP2* among the switch genes for both the DCM and VCM samples. The gene, which is displayed in Fig. 2 is associated with the transcription factors *CTCF, JUND, MAX, RAD21* and *USF1* that are common to some of the switch genes, which places it in an interesting position within the DCM network. The WikiPathways entry show that *SGPP2* is involved in sphingolipid metabolism Appendix B) [35, 36]. These are a group of lipids that are produced in the ER [37] from the condensation of serine and palmityl-CoA [38] and are characterized by their eighteen-carbon amino-alcohol backbones [39]. Sphingosine and sphinganine are simple sphingolipids which are metabolized to more complex products. The large family of sphingolipids play important roles in membrane biology and supply a variety of bioactive metabolites that control cellular activity as a result of modifications to their fundamental structure [40]. In our results section, we found SGPP2 to be down-regulated. It catalyze the reaction from sphingosine to sphingosine-1-phosphate. While these metabolites are linked to a variety of cardiovascular diseases [41], such as hypertension, coronary artery disease, heart failure, arrhythmia, and stroke in addition to being an essential component of the cell membrane [42], their role in dilated cardiomyopathy is not yet fully discovered [43].

One of the most interesting results related to VCM in Fig. 5, because most of the differential switch genes are connected to signalling pathways. An example of a few reported pathways are: WnT [44, 45], MAPK [11, 46], NRF2 [47, 48], PI3K-Akt-mTOR [11, 49], TGF-beta [50], and EGFR [51, 52]. These pathways are interesting because of their importance for cellular mechanisms such as cell proliferation and apoptosis. In addition, these pathways have been reported to be affected during viral infection [11, 44, 46, 47, 48, 49, 50, 51]. Other noteworthy nodes are the "phosphoinositides metabolism", "regulation of actin cytoskeleton", and "Hepatitis C". These nodes are of interest because of their reported connection to viral infection [2, 11, 49, 53]. The phospoinositides and actin cytoskeleton play an important role in the live cycle of a virus, and because of their importance, these processes are heavily influenced by viruses [11, 49, 53]. Further, the presence of "Hepatitis C" as a WikiPathways entry in the VCM network is very promising, because our findings are coherent with the known association of Hepatitis C with DCM [2].

In conclusion, the switch genes related to DCM and VCM are strong candidates for further biological study. However, the current research is limited by the sample size of VCM in the minimal biological effect size that is significantly detectable, and by the fact that the distinction between VCM and DCM is based on the clustering of samples by manually selected genes. Determining genes that are affected by viruses would improve the detection of viral DCM. Overall, future research is needed to identify different causes of DCM in transcriptomics data.

5. Conclusion

We have explored a new and innovative approach to analyse a transcriptomics dataset that builds on top of the standard bioinformatics pipeline. The significance and fold changes are used as a first starting point to further study other important mathematical attributes that can be calculated by creating co-expression networks such as interesting topological properties or connectedness. This enables us to tackle interesting biological questions, such as whether crucial regulators can be observed in the transition between an healthy to a cardiomyopathic phenotype. Dilated cardiomyopathy cannot be characterized by a single set of genes. In order to understand the underlying mechanisms a

wide variety of integrated approaches are needed. Switch genes are a relatively new concept in network biology that is worth investigating. In particular, they represent an interesting concept both from a computational and biological point of view. In fact, this kind of algorithm can allow us to move from a data science. The availability of complex omics data on cardiovascular diseases necessitates a wide range of expertise, which the network language permits in a truly multidisciplinary fashion. This represents a great opportunity for further applications of systems biology to complex diseases.

Acknowledgements

The Declaration of Helsinki, the Ethical Principles for Medical Research Involving Human Subjects, issued by the World Medical Association, was upheld during our research. We always received informed consent before starting any scientific endeavor. The life, health, privacy, and dignity of the human subject associated with the data are likewise strictly protected by our policies. Additionally, the research was approved by the responsible ethical committee.

References

[1] Mahmaljy H, Yelamanchili VS, Singhal M. Dilated cardiomyopathy. In: StatPearls [Internet]. StatPearls Publishing; 2021. .
[2] Schultheiss HP, Fairweather D, Caforio ALP, Escher F, Hershberger RE, Liphshultz SE, et al. Dilated cardiomyopathy. Nature Reviews Disease Primers. 2019;5(32):1-19.
[3] McNally EM, Mestroni L. Dilated cardiomyopathy: genetic determinants and mechanisms. Circulation research. 2017;121(7):731-48.
[4] Mathew T, Williams L, Navaratnam G, Rana B, Wheeler R, Collins K, et al. Diagnosis and assessment of dilated cardiomyopathy: a guideline protocol from the British Society of Echocardiography. Echo research and practice. 2017;4(2):G1.
[5] Khalil H, Alzahrani T. Cardiomyopathy imaging. In: StatPearls [Internet]. StatPearls Publishing; 2022. .
[6] Paci P, Colombo T, Fiscon G, Gurtner A, Pavesi G, Farina L. SWIM: a computational tool to unveiling crucial nodes in complex biological networks. Scientific reports. 2017;7(1):1-16.
[7] Rubis P. The diagnostic work up of genetic and inflammatory dilated cardiomyopathy. EJ Cardiol Pract. 2015;13:19.
[8] Japp AG, Gulati A, Cook SA, Cowie MR, Prasad SK. The diagnosis and evaluation of dilated cardiomyopathy. Journal of the American college of cardiology. 2016;67(25):2996-3010.
[9] Zhang J, Xu M, Chen T, Zhou Y. Bioinformatics Analysis of Common Differential Genes of Viral Myocarditis and Dilated Cardiomyopathy: Screening for Potential Pharmacological Compounds. Journal of Cardiovascular Development and Disease. 2022;9(10):353.
[10] Skelly A, Dettori J, Brodt E. Assessing bias: the importance of considering confounding. Evid Based Spine Care J. 2012; 3 (01): 9–12;.
[11] Flint SJ, Racaniello VR, Rall GF, Hatziioannou T, Skalka AM. Principles of Virology, Volume 1: Molecular Biology, 5th Edition. John Wiley & Sons; 2020.
[12] Walsh D, Mohr I. Viral subversion of the host protein synthesis machinery. Nature Reviews Microbiology. 2011;9(12):860-75.
[13] Tong S, Revill P. Overview of hepatitis B viral replication and genetic variability. Journal of hepatology. 2016;64(1):S4-S16.
[14] Ludewig B, Krebs P, Metters H, Tatzel J, Tureci O, Sahin U. Molecular characterization of virus-induced autoantibody responses. The Journal of experimental medicine. 2004;200(5):637-46.
[15] Nishtala K, Phong TQ, Steil L, Sauter M, Salazar MG, Kandolf R, et al. Virus-induced dilated cardiomyopathy is characterized by increased levels of fibrotic extracellular matrix proteins and reduced amounts of energy-producing enzymes. Proteomics. 2011;11(22):4310-20.
[16] Paci P, Fiscon G. SWIMmeR: an R-based software to unveiling crucial nodes in complex biological networks. Bioinformatics. 2022;38(2):586-8.
[17] Lin H, Dolmatova EV, Morley MP, Lunetta KL, McManus DD, Magnani JW, et al. Gene expression and genetic variation in human atria. Heart Rhythm. 2014;11(2):266-71.
[18] Smyth G. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology. 2004;3(1).
[19] Herwanda R, Murdiono WE, Koesriharti K. Aplikasi Nitrogen dan Pupuk Daun terhadap Pertumbuhan dan Hasil Tanaman Bawang Merah (Allium cepa L. var. ascalonicum). Brawijaya University; 2017.
[20] Qiu X, Wu H, Hu R. The impact of quantile and rank normalization procedures on the testing power of gene differential expression analysis. BMC bioinformatics. 2013;14(1):1-10.
[21] Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Statistical applications in genetics and molecular biology. 2005;4(1).
[22] Hao D, Li C. The dichotomy in degree correlation of biological networks. PloS one. 2011;6(12):e28322.
[23] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics. 2008;9(1):1-13.
[24] Palumbo MC, Zenoni S, Fasoli M, Massonnet M, Farina L, Castiglione F, et al. Integrated network analysis identifies fight-club nodes as a class of hubs encompassing key putative switch genes that induce major transcriptome reprogramming during grapevine development. The Plant Cell. 2014;26(12):4617-35.
[25] Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. Journal of Statistical Software. 2014;61(6):1-36. Available from: https://www.jstatsoft.org/v61/i06/.

[26] Guimera R. L. a. N. Amaral. 2005a. Cartography of complex networks: modules and universal roles. Journal of Statistical Mechanics: Theory and Experiment PO. 2001.
[27] Guimera R, Nunes Amaral LA. Functional cartography of complex metabolic networks. nature. 2005;433(7028):895-900.
[28] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research. 2003;13(11):2498-504.
[29] Kutmon M, Ehrhart F, Willighagen EL, Evelo CT, Coort SL. CyTargetLinker app update: A flexible solution for network extension in Cytoscape. F1000Research. 2018;7.
[30] Gustavsen, A J, Pai, Shraddha, Isserlin, Ruth, et al. RCy3: Network Biology using Cytoscape from within R. F1000Research. 2019.
[31] Kelder T, Pico AR, Hanspers K, Van Iersel MP, Evelo C, Conklin BR. Mining biological pathways using WikiPathways web services. PloS one. 2009;4(7):e6447.
[32] Khodiyar VK, Maltais LJ, Sneddon KM, Smith JR, Shimoyama M, Cabral F, et al. A revised nomenclature for the human and rodent α-tubulin gene family. Genomics. 2007;90(2):285-9.
[33] Anfinson M, Fitts RH, Lough JW, James JM, Simpson PM, Handler SS, et al. Significance of α-Myosin Heavy Chain (MYH6) Variants in Hypoplastic Left Heart Syndrome and Related Cardiovascular Diseases. Journal of Cardiovascular Development and Disease. 2022;9(5):144.
[34] Barrick SK, Greenberg MJ. Cardiac myosin contraction and mechanotransduction in health and disease. Journal of Biological Chemistry. 2021;297(5).
[35] Merrill AH, Stokes TH, Momin A, Park H, Portz BJ, Kelly S, et al. Sphingolipidomics: a valuable tool for understanding the roles of sphingolipids in biology and disease. Journal of Lipid Research. 2009;50:S97-S102.
[36] Merrill Jr AH, Wang MD, Park M, Sullards MC. (Glyco) sphingolipidology: an amazing challenge and opportunity for systems biology. Trends in biochemical sciences. 2007;32(10):457-68.
[37] Tidhar R, Futerman AH. The complexity of sphingolipid biosynthesis in the endoplasmic reticulum. Biochimica Et Biophysica Acta (BBA)-Molecular Cell Research. 2013;1833(11):2511-8.
[38] Hanada K. Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids. 2003;1632(1-3):16-30.
[39] Gault CR, Obeid LM, Hannun YA. An overview of sphingolipid metabolism: from synthesis to breakdown. Sphingolipids as signaling and regulatory molecules. 2010:1-23.
[40] Hannun YA, Obeid LM. Principles of bioactive lipid signalling: lessons from sphingolipids. Nature reviews Molecular cell biology. 2008;9(2):139-50.
[41] Borodzicz-Jażdżyk S, Jażdżyk P, Łysik W, Cudnoch-Jdrzejewska A, Czarzasta K. Sphingolipid metabolism and signaling in cardiovascular diseases. Frontiers in cardiovascular medicine. 2022;9.
[42] Taguchi Y, Allende ML, Mizukami H, Cook EK, Gavrilova O, Tuymetova G, et al. Sphingosine-1-phosphate phosphatase 2 regulates pancreatic islet β-cell endoplasmic reticulum stress and proliferation. Journal of Biological Chemistry. 2016;291(23):12029-38.
[43] Baranowski M, Górski J. Heart sphingolipids in health and disease. Sphingolipids and Metabolic Disease. 2011:41-56.
[44] Ljungberg JK, Kling JC, Tran TT, Blumenthal A. Functions of the WNT signaling network in shaping host responses to infection. Frontiers in Immunology. 2019;10:2521.
[45] Schaale K, Brandenburg J, Kispert A, Leitges M, Ehlers S, Reiling N. Wnt6 is expressed in granulomatous lesions of Mycobacterium tuberculosis–infected mice and is involved in macrophage differentiation and proliferation. The Journal of Immunology. 2013;191(10):5182-95.
[46] Kumar R, Khandelwal N, Thachamvally R, Tripathi BN, Barua S, Kashyap SK, et al. Role of MAPK/MNK1 signaling in virus replication. Virus research. 2018;253:48-61.
[47] Ramezani A, Nahad MP, Faghihloo E. The role of Nrf2 transcription factor in viral infection. Journal of cellular biochemistry. 2018;119(8):6366-82.
[48] Ahmed SMU, Luo L, Namani A, Wang XJ, Tang X. Nrf2 signaling pathway: Pivotal roles in inflammation. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease. 2017;1863(2):585-97.
[49] Beziau A, Brand D, Piver E. The role of phosphatidylinositol phosphate kinases during viral infection. Viruses. 2020;12(10):1124.
[50] Mirzaei H, Faghihloo E. Viruses as key modulators of the TGF-β pathway; a double-edged sword involved in cancer. Reviews in medical virology. 2018;28(2):e1967.
[51] Qiu W, Chang L, He Y, Zhu L. The role of epidermal growth factor receptor signaling pathway during bovine herpesvirus 1 productive infection in cell culture. Viruses. 2020;12(9):927.
[52] Oda K, Matsuoka Y, Funahashi A, Kitano H. A comprehensive pathway map of epidermal growth factor receptor signaling. Molecular systems biology. 2005;1(1):2005-0010.
[53] Taylor MP, Koyuncu OO, Enquist LW. Subversion of the actin cytoskeleton during viral infection. Nature Reviews Microbiology. 2011;9(6):427-39.

Appendix A. MYH6 in striated muscle contraction pathway

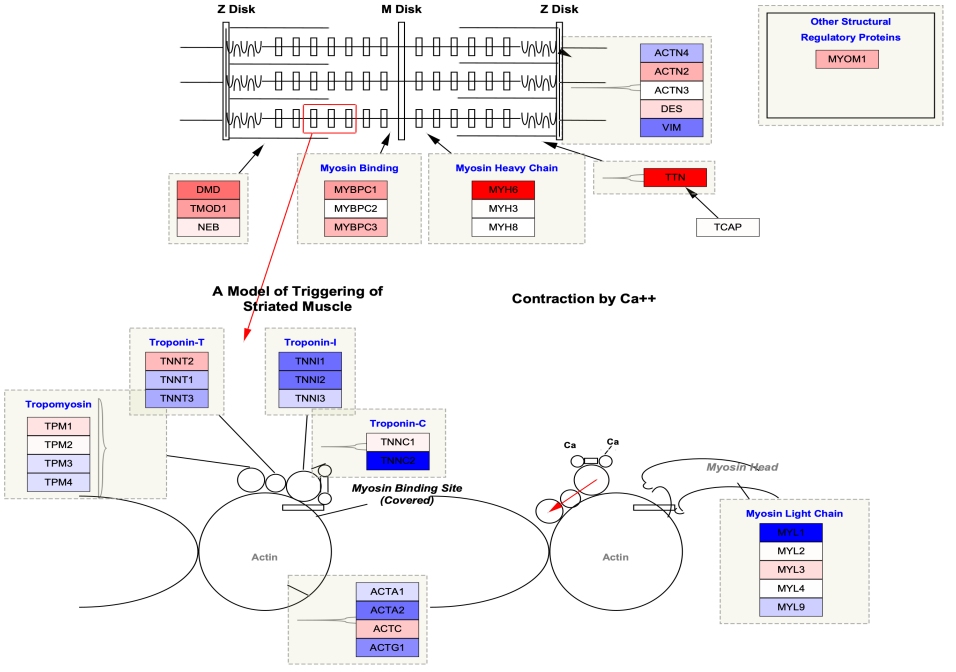


Fig. A.6. Striated muscle contraction pathway imported from WikiPathways (id=WP383) with log-fold changes between DCM and NF mapped to the nodes, blue corresponding to up regulated and red to down regulated

Appendix B. SGPP2 in sphingolipid metabolism

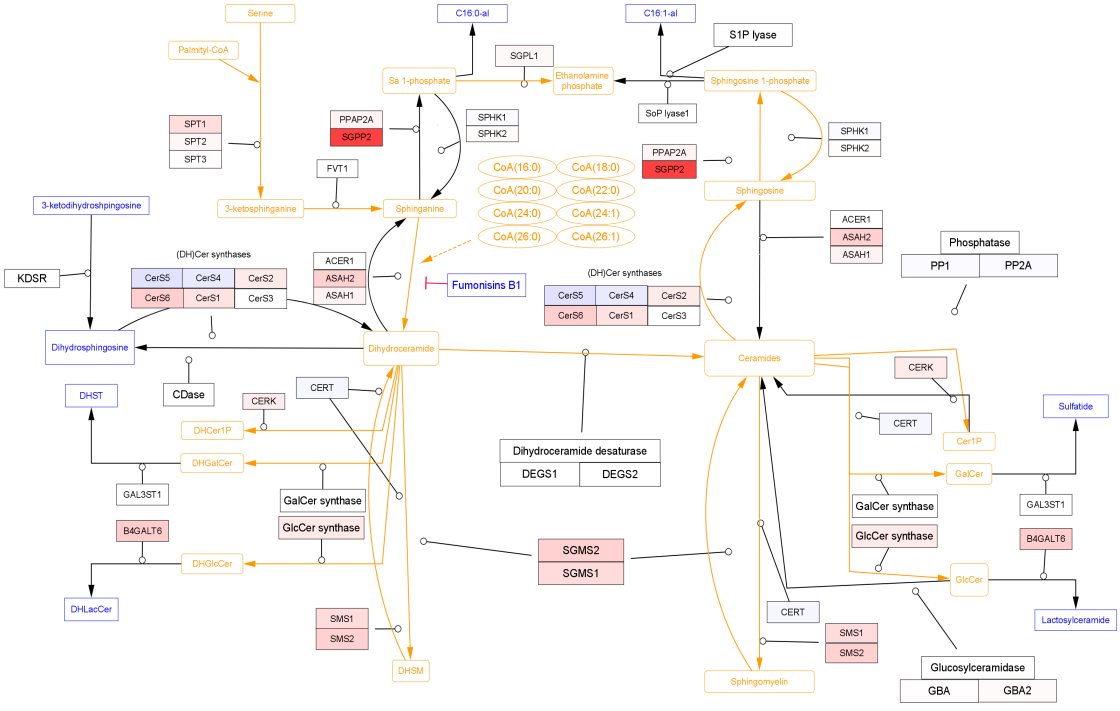


Fig. B.7. Sphingolipid metabolism pathway imported from WikiPathways (id=WP1422) with log-fold changes between DCM and NF mapped to the nodes, blue corresponding to up regulated and red to down regulated

Appendix C. SULT1B1 for sulfanation biotransformation

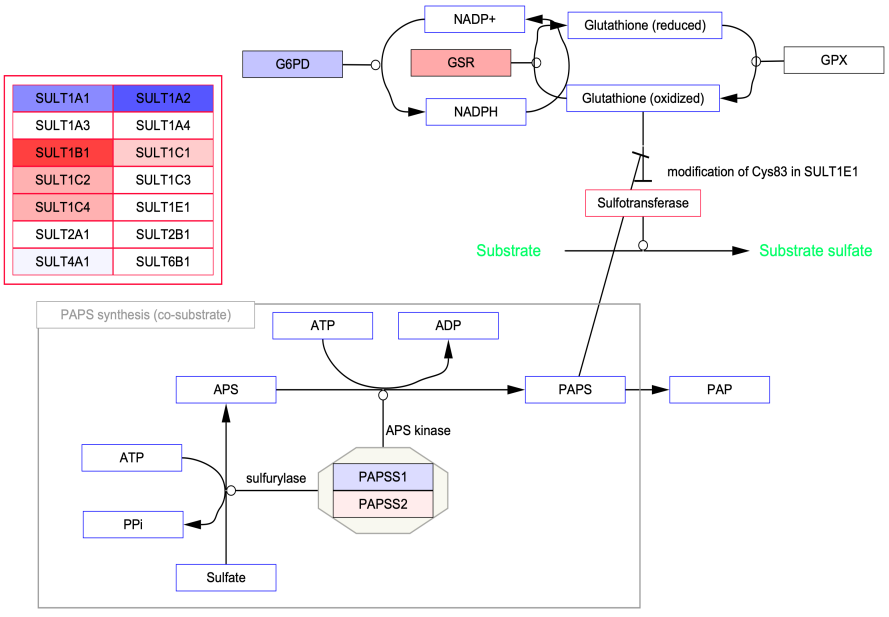


Fig. C.8. Sulfanation biotransformation reaction imported from WikiPathways (id=WP692) with log-fold changes between DCM and NF mapped to the nodes, blue corresponding to up regulated and red to down regulated

Appendix D. PLA2G4F for lipid metabolism in senescent cells

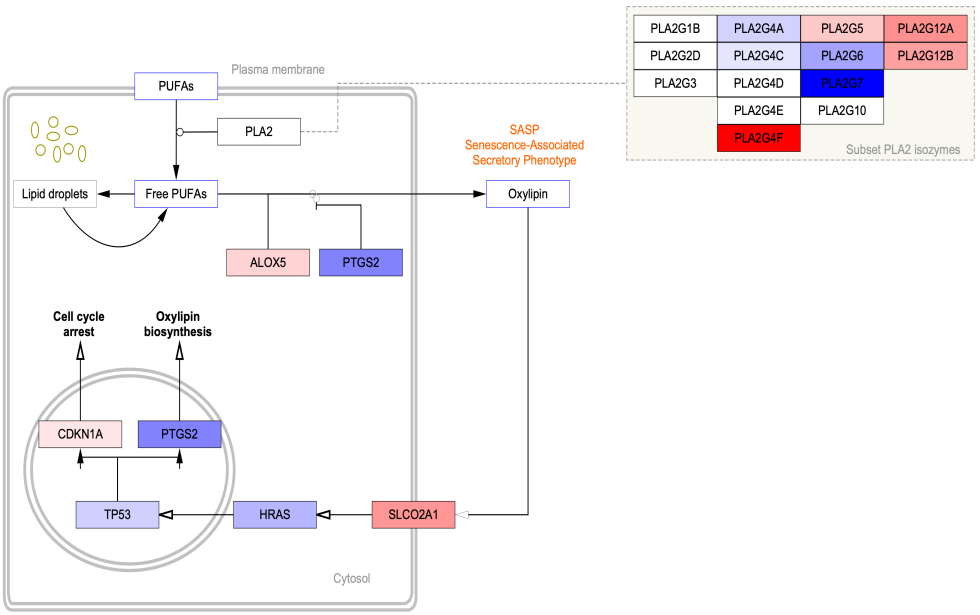


Fig. D.9. Lipid metabolism in senescent cells imported from WikiPathways (id=WP5149) with log-fold changes between DCM and NF mapped to the nodes, blue corresponding to up regulated and red to down regulated

Appendix E. FKBP5 in the farsenoid X receptor pathway

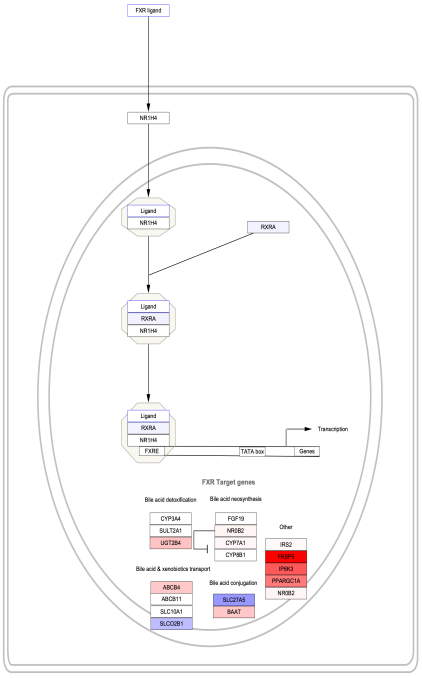


Fig. E.10. Farsenoid X receptor pathway imported from WikiPathways (id=WP2879) with log-fold changes between DCM and NF mapped to the nodes, blue corresponding to up regulated and red to down regulated