2D Mito Plot

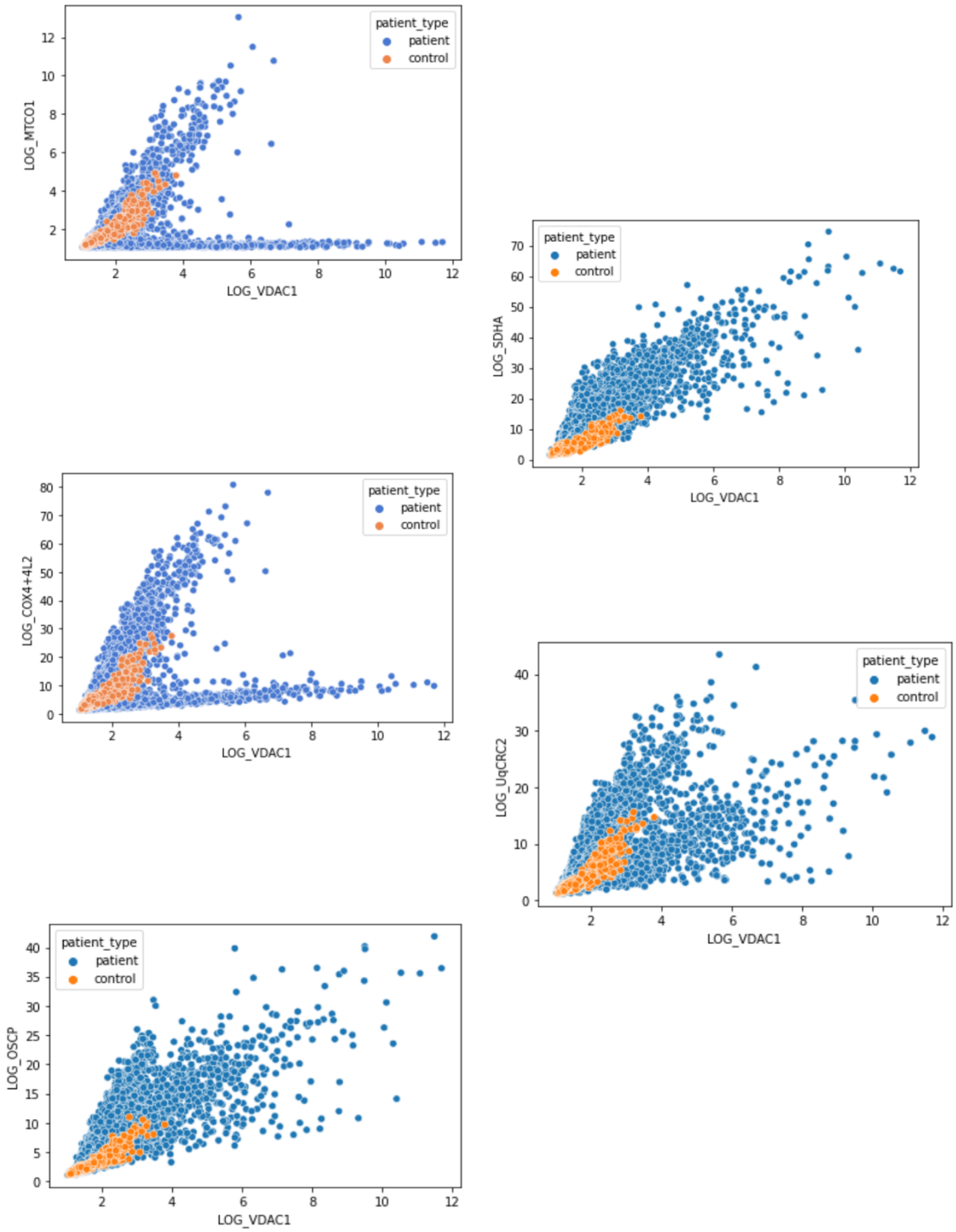
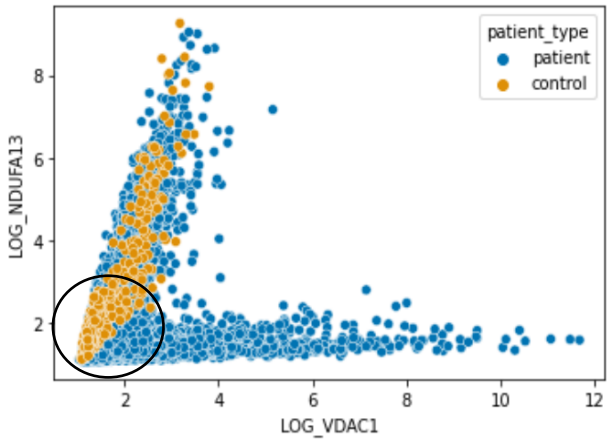


Figure 1. Plots of 5 different proteins against the LOG_VDAC1 protein.

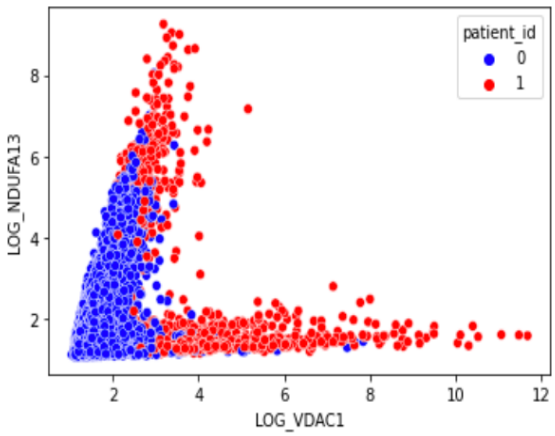
The plots of 5 different proteins against a sixth protein (LOG_VDAC1), all found in the subjects' fibres. The plots identify that the control fibres seem to be in the closer to the bottom left corner, highlighting that all these proteins seem to be in smaller amounts in healthy subjects in comparqison to those with RC deficient fibres.

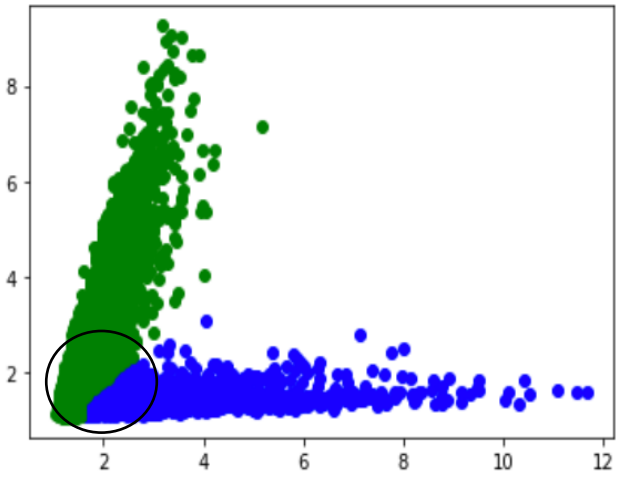
NDUFA13

Raw Data



K-Means





GMM

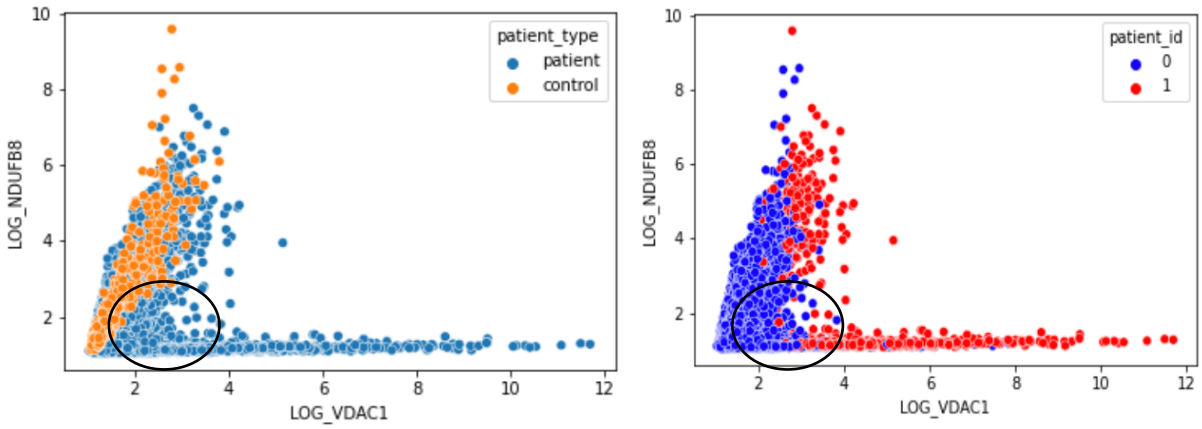
Figure 2. Plots of the amounts of NDUFA13 vs the amounts of LOG_VDAC1 proteins in the raw data, K-Means Model and Gaussian Mixture Model (GMM).

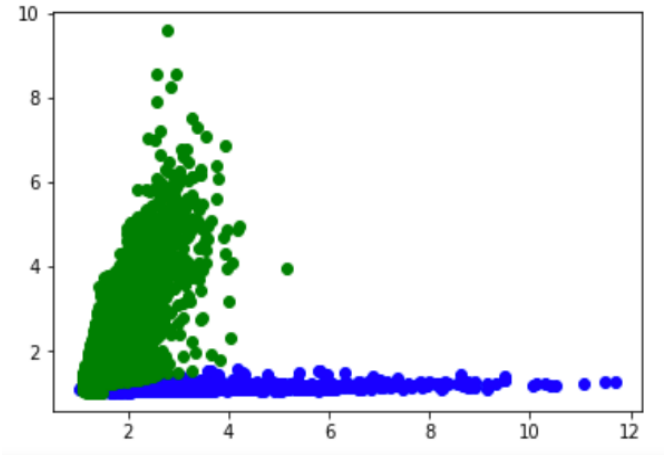
All 8 proteins were used in both the K-Means and GMM algorithms to produce their respective plots shown above. The results of the K-Means algorithm are almost like the raw data and both plots appear to have a very similar structure. However, in comparison to the GMM, there is a clearer separation of the two clusters: green identifying the control, and blue identifying the patients. Additionally, the point on the GMM (circled) at which the clusters divide appears to be at the same point circled in the raw data.

NDUFB8

Raw data

K-means





GMM

Figure 3. Plots of the amounts of NDUFB8 vs the amounts of LOG_VDAC1 proteins in the raw data, K-Means Model and Gaussian Mixture Model (GMM).

All 8 proteins were used in both the K-Means and GMM algorithms to produce the plots above. The K-Means produces a plot like that of the raw data, however, there appears to be more control fibres in the 2nd fork of the K-Means than that of the raw data (circled). The GMM cluster seems to only produce one cluster instead of 2.

K-Mean results

Individual	Proportion of RC deficient Fibres (%)	Disease Type	Proportion of RC deficient Fibres (%)
C01	0%	Control	5.1%
C02	0%		
C03	22.1%		
P01	79.5%	CI	83.7%
P02	89.7%		
P03	0%	Deletion	6.5%
P04	16.6%		
P05	0.6%	MT-TL1	13.3%
P06	33.4%		
P07	23.0%		
P08	15.4%	MT-TG	15.4%
P09	5.7%	MT-TE	5.7%
P10	62.6%	MT-TW	62.6%

Table 1. Proportion of Reactive Chain (RC) Deficient Fibres in each patient and in each disease from the K-Means Result.

The algorithm predicted the control to all have mainly healthy fibres, whereas, for patients there is a mixture of healthy and disease fibres. This led us to investigate whether the type of disease had an impact on the number of the fibres.

The CI mutation disease presents with a much larger proportion of RC deficient fibres in comparison to the other fibres. The control and MT-TE disease were both predicted to have a rather similar proportion RC deficient fibres. MT-TW almost has equal amounts of RC deficient and healthy fibres.

GMM Table of Results part 1: LOG_VDAC1 & LOG_NDUFA13

Individual	Proportion of RC deficient Fibres (%)	Disease Type	Proportion of RC deficient Fibres (%)
C01	0%	Control	0%
C02	0%		
C03	0%		
P01	95.5%	CI	96.7%
P02	98.3%		
P03	0.4%	Deletion	1.6%
P04	3.4%		
P05	18.5%	MT-TL1	21.0%
P06	9.9 %		
P07	39.2%		
P08	79.8%	MT-TG	79.8%
P09	37.2%	MT-TE	37.2%
P10	82.0%	MT-TW	82.0%

Table 1. Proportion of Reactive Chain (RC) Deficient Fibres in each patient and in each disease from the GMM Result.

It was observed that the CI disease type had almost 100% of fibres with RC deficiency. MT-TG and MT-TW both appear to have been predicted to have very similar proportions of RC deficient fibres. Finally, the deletion mutation was predicted to have very few RC deficient fibres based on the patterns identified by the GMM algorithm.

GMM Table of Results part 2: LOG_VDAC1 & LOG_NDUFB8

Individual	Proportion of RC deficient Fibres (%)	Disease Type	Proportion of RC deficient Fibres (%)
C01	0%	Control	0%
C02	0%		
C03	0%		
P01	98.8%	CI	98.9%
P02	99.1%		
P03	2.2%	Deletion	2.1%
P04	1.9%		
P05	20.0%	MT-TL1	25.3%
P06	9.0%		
P07	55.8%		
P08	79.6%	MT-TG	79.6%
P09	35.6%	MT-TE	35.6%
P10	79.8%	MT-TW	79.8%

Table 1. Proportion of Reactive Chain (RC) Deficient Fibres in each patient and in each disease from the GMM Result.

It was predicted for the CI disease type that almost 100% of the fibres had RC deficiency. MT-TG and MT-TW both appear to have been predicted to have an almost identical proportion of RC deficient fibres, both of which are high in percentage. Finally, the deletion mutation was predicted to have very few RC deficient fibres, with its proportion being close to that of the control fibres. The MT-TL1 and MT-TE mutations both had less than half of the fibres predicted to be RC deficient.