^{1}H NMR based Metabolomics of CSF and blood serum: a metabolic profile for a transgenic rat model of Huntington disease

Kim A Verwaest[1§], Trung N Vu [2], Kris Laukens[2], Laura Clemens[3], Huu P Nguyen[3], Bjorn Van Gasse[4], José C. Martins[4], Annemie Van Der Linden[4], Roger Dommisse[1]

[1]Department of Chemistry, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium

[2]Department of Mathematics & Computer Science, University of Antwerp, Middelheimlaan 1, 2020 Antwerp, Belgium

[3]Department of Medical Genetics, University of Tuebingen, Calwerstrasse 7, 72076 Tuebingen, Germany

[4]Department of Organic Chemistry, University of Ghent, Krijgslaan 281 S4, 9000 Gent, Belgium

[5]Department of Biomedical Science, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium

[§]Corresponding author: telephone: 003232653267

Email addresses:

KAV: kim.verwaest@ua.ac.be

TNV: TrungNghia.Vu@ua.ac.be

KL: kris.laukens@ua.ac.be

LC: Laura.Clemens@med.uni-tuebingen.de

HPN: hoa.nguyen@med.uni-tuebingen.de

BVG: Bjorn.vangasse@UGent.be

JCM: jose.martins@UGent.be

AV: Annemie.VanDerLinden@ua.ac.be

RD: roger.dommisse@ua.ac.be

Abbreviations: ^{1}H NMR, proton nuclear magnetic resonance; 2CV, double cross validation; AUROC, area under the receiver operating characteristic; CPMG, Carr-Purcell-Meiboom-Gill; CSF, cerebrospinal fluid; HD, Huntington disease; LOOCV, Leave-one-out cross validation; MS, mass spectrometry; NAA, N-acetylaspartate; PCA, Principal component analysis; PLS-DA, Partial-least squares discriminant analysis; SVM, support vector machine; WT, wild-type.

Abstract

Huntington Disease (HD) is a hereditary brain disease. Although the causative gene has been found, the exact mechanisms of the pathogenesis are still unknown. Recent investigations point to a metabolic and energetic dysfunction in HD neurons.

In this study, proton nuclear magnetic resonance spectroscopy (^{1}H NMR) was used to analyse serum and cerebrospinal fluid (CSF) taken from presymptomatic HD transgenic rats and their wild-type littermates. Both univariate and multivariate analyses were used to compare these metabolic profiles. N-acetylaspartate (NAA), an indicator of neuronal function, was found to be significantly decreased in the serum of HD rats compared to wild-type littermates. In addition, levels of glucose, lactate, glutamine and succinic acid are significantly increased in the serum of HD rats. Moreover, the increased concentration of glucose and lactate can also be found in the CSF of transgenic rats. There is a 1:1 stoichiometry coupling glucose utilization and glutamate cycling. The observed increase in the glutamine concentration, which indicates a shutdown in the neuronal-glial glutamate-glutamine cycling, results therefore also in an increased glucose concentration. The elevated succinic acid concentration might be due to an inhibition of succinate dehydrogenase, an enzyme linked to the mitochondrial respiratory chain and TCA cycle. Moreover, reduced levels of NAA may reflect an impairment of mitochondrial energy production. In addition, the observed difference in lactate supports a deficiency of oxidative energy metabolism in rats transgenic for HD as well.

The observed metabolic alterations seem to be more profound in serum in comparison to CSF in presymptomatic rats. All findings suggest that even in presymptomatic rats, a defect in the energy metabolism is already apparent. Moreover, these results support the hypothesis of a mitochondrial energy dysfunction in HD.

Keywords: Huntington disease, CSF, serum, metabolomics, transgenic rats, ^{1}H NMR spectroscopy.

1. Introduction

Huntington Disease (HD) is an autosomal dominant neurodegenerative disorder, characterized by progressive motor, cognitive and psychiatric dysfunctions. Although the disease can occur at any age, the median age of onset is 40 years. Most patients die one to two decades after disease onset.

HD is caused by a mutation in the gene coding for the protein Huntingtin. The exact function of this protein is still unknown. Normal individuals have a CAG-trinucleotide repeat length of less than 37 [1]. People whose CAG repeat length exceeds this number will develop HD. The length of the CAG expansion is proportional to the severity of the disease and inversely proportional to the age of onset [2]. At present, the exact mechanism of the HD pathology is still unknown. However, recent evidence points to metabolic and energetic dysfunction in HD neurons [2-4]. Better knowledge about this mechanism is expected to lead to more appropriate analyses, which in turn should result in a better understanding of the disease. Moreover, this will help to direct the quest for improved treatments in the proper direction.

According to Nicholson et al [5], metabolomics is defined as ‘the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modifications’. Since 1999, applications of this technique have emerged in other fields (e.g. environmental sciences)[6]. Metabolites in biofluids are in dynamic equilibrium with those in cells and tissues. A healthy individual attempts to retain the concentration of metabolites in cells and tissues constant by homeostasis. Abnormal cellular processes due to sickness, toxins, etc.

result in altered biofluid compositions. Various analytical techniques are available to detect low molecular weight metabolites. The ones most commonly applied in metabolomic studies are nuclear magnetic resonance spectroscopy (NMR) and mass spectrometry (MS). The latter is frequently used in combination with a chromatographic separation technique (liquid chromatography, gas chromatography, etc.) [7]. Although NMR is less sensitive than MS, this technique offers some important advantages. It is a high-throughput analysis that does not require any preceding chromatographic separation or purification procedure. Additionally, sample preparation is very straightforward. Furthermore, it is a non-destructive method capable of detecting all non-exchangeable protons, provided they are present above a certain threshold concentration (µM concentrations). Moreover, the measurements are very reproducible and inexpensive on a per sample basis [8-10]. All these advantages have contributed to the application of NMR spectroscopy in clinical diagnoses of diseases and in the follow-up of subjects participating in medical treatment studies [11].

The NMR spectrum of a biofluid is generally extremely complex in nature, since each molecule present in the biofluid gives rise to one or several peaks in most cases. Moreover, many peaks show some fine structure and splittings which causes a prominent overlap of signals originating from different molecules [12]. Visual inspection of the spectra will therefore only reveal a small percentage of the available information [13].

Spectral interpretation can be simplified by applying automatic data reduction methods. Subsequently, these reduced datasets can be analyzed with univariate [2,14-22] or multivariate [13,23-27] statistical methods.

In this study we applied ^{1}H NMR spectroscopy in order to disentangle the metabolic profile of serum and cerebrospinal fluid (CSF) from a rat model for HD. Both univariate and multivariate statistical analyses were applied in a search for pathology specific differences. It is anticipated that these investigations lead to a gain in knowledge about the pathogenesis of HD and make it possible to identify some potential biomarkers.

2. Materials and methods

2.1 Animals

Transgenic HD rats carrying a truncated huntingtin cDNA fragment with 51 CAG repeats under the control of the native rat huntingtin promoter, and their wild-type littermates were used [28]. The expressed gene product was 75 kDa, corresponding to 22% of the full-length huntingtin (cDNA position 324-2321, amino acid position 1-709/825, corresponding to exon 1-16), which are under the control of 886 bp of the rat huntingtin promoter (position 900 to 15). Genotyping was carried out as previously described [28]. The number of CAG-repeats was analyzed in a subset of transgenic rats using these primers: cggctgaggcagcagcggctgt (forward) and ccttcgagtccctcaagtccttc (reverse). The reverse primer is labelled at the 5'end with the fluorescent dye Cy5. The PCR amplicon length was then analyzed on the Beckman coulter sequencer (CEQ 8000 Cycle Sequencer, Krefeld, Germany). In more than 15 generations only a small variation of ±1-2 CAGs were observed in transgenic rats of HD. Also for this study, in a subset of animals, the CAG-length was checked and it was found to be 51±2 CAGs. After genotyping, rats were housed in gender- and genotyped matched groups of two, according to FELASA recommendations. All rats were kept under a 12:12 hours light-dark cycle with lights on at 06.00 a.m. and food

(Altromin lab chow pallets, Altromin standard diet: 1320; Lage, Germany) and tap water available ad libitum. All research and animal care procedures had been approved by the district government of Tuebingen, Germany, and followed principles described in the European Community's Council Directive of 24th November, 1986 (86/609/EEC).

For this study, we used serum and CSF from female, transgenic HD rats (n = 10 for serum and n = 8 for CSF) at an age of two months and compared these with age-matched wild-type (WT) littermates (n = 12 for serum and respectively 9 for CSF). We are aware of the fact that the number of samples used in this study is rather small. However, this was determined by keeping in mind the well-being of animals and in order to reduce costs. Moreover, in the literature a lot of studies where even less samples per group are investigated, can be found [15,27,29-32].

2.2 Sample collection and preparation

2.2.1 CSF samples

Rats were sacrificed by CO_2 inhalation and immediately exsanguinated. Subsequently, each animal was placed prone onto the stereotaxic instrument. A sagittal incision of the skin was made inferior to the occiput. The subcutaneous tissue and neck muscles through the midline were bluntly separated. Next, the rat was laid down in a way that the body made a 135° angle with the fixed head. In this angle, the dura mater of cisterna magna was exposed sufficiently. The dura was then penetrated with an insulin syringe and CSF was drawn. All samples were transferred into polypropylene tubes, immediately snap frozen in liquid nitrogen and stored at -80 °C until shipped for NMR analysis. Upon arrival, CSF samples were immediately stored at -80 °C until measurement.

All CSF samples (40 µL) were simultaneously lyophilized. In cases where less then 40 µL of CSF was available, the total available volume was lyophilized. The freeze-dried samples were then stored at -20 °C in sealed vials until analysis. In order to keep the time spent at room temperature to a minimum, samples were randomly divided in five batches. Samples gathered in one batch were prepared together. Prior to NMR analysis, each sample was reconstituted in 8 µL sodium phosphate buffer (50 mM dissolved in D_2O, pH = 7.05) containing 0.05 mM sodium azide in order to prevent bacterial contamination. Finally, this volume was transferred to a 1 mm NMR tube.

2.2.2 Serum samples

Samples were taken as done previously [33]. In short: rats were sacrificed with CO_2 and blood was drawn directly from the heart. All rats were tested for diabetics immediately after sample collection. The blood was incubated for one hour at 37 °C, followed by centrifugation for 30 minutes at 3000 x g. The serum was collected and immediately stored at -80 °C until shipped for NMR analysis. Upon arrival, serum samples were immediately stored at -80 °C until measurement.

In order to keep the time spent at room temperature to a minimum, samples were randomly divided in three groups. Samples assembled in one group were prepared together. Just before their acquisition, samples were thawed to room temperature. An aliquot of serum (250 µL) from each sample was mixed with 300 µL of an aqueous saline solution. In this way, each sample contains 0.1 M NaCl and 10 % D_2O. The addition of an internal standard like DSS is prohibited by its interaction with proteins present in the sample. Subsequently, the total volume was transferred to a 5 mm NMR tube.

2.3 ^{1}H NMR spectroscopy

Prepared samples were analyzed on a Bruker Avance II-700 spectrometer, operating at a proton frequency of 700.13 MHz. The spectrometer is equipped with a BACS-60 automatic sample changer and a 5 mm inverse TXI-Z probe when measuring serum samples. Since the brain volume of rats is rather small, the amount of CSF that can be collected at once from one animal is too small to perform a classical NMR analysis (> 100 µL is necessary, prior to dilution to the final 550 µL measuring volume, in order to get a good S/N). Combining CSF taken from different animals within a population is not advisable, since this will remove the possibility to detect individual differences. Recently, Bach and co-workers [10,34] demonstrated that the use of a 1 mm probe setup provides an alternative to address the measurement of such volume restricted samples. Therefore, CSF samples were measured with a 1 mm inverse TXI-Z probe, requiring only 8 µL of sample volume. The samples were held at room temperature while on the sample changer and at 303 K during acquisition. Carr-Purcell-Meiboom-Gill (CPMG) experiments were executed in order to get 1D ^{1}H NMR spectra of the samples. The individual CMPG spin echo (in total 1 ms) was repeated 20 times, resulting in a total spin-spin relaxation delay of 20 ms. Suppression of the water signal was performed using 2.22 s presaturation during the relaxation delay. For each serum/CSF sample, 64 respectively 1024 scans, 32 K data points each, were accumulated. Each scan had an acquisition time of 0.78 s. To keep variation due to sample handling to a minimum, tuning and matching as well as shimming were performed automatically for each sample. Digital filtering was performed using the ‘baseopt’ option available in Topspin 2.0 (Bruker Biospin Corporation) which results in flat baselines. Prior to Fourier transformation, the Free Induction Decays (FIDs) were zero-filled to 64 K and an exponential window function with a line broadening factor of 0.3 Hz was applied. The acquired NMR spectra were consecutively

processed for phase and baseline correction. In the absence of an external reference (vide supra 2.2.2), all serum spectra were referenced to the methyl doublet of lactate at 1.33 ppm. After initial referencing of the CSF spectra using the anomeric glucose ^{1}H signal at 5.225 ppm, the spectra were still shifted one against the other. They were aligned using an in house developed spectral alignment algorithm (manuscript in preparation).

2.4 Data reduction of the NMR spectra

The ^{1}H NMR spectra were automatically reduced into consecutive integrated spectral regions (buckets) of an equal width (0.05 ppm) applying R (version 2.7.2, http://www.r-project.org). In order to eliminate spectral variations due to different quality of water suppression from one spectrum to another, the region containing the water resonance (4.5-5.0 ppm) was not included in the analysis. In order to account for differences in concentration between samples, the subdivided spectra were scaled to their total intensity.

2.4.1 Multivariate analysis

Prior to multivariate analysis, the data were scaled to unit variance. Besides unit variance scaling, also pareto scaling and no scaling were tested on the dataset. However, these did not generate significant differences. Since all the variables possess an equal weight, irrespective of their absolute magnitude in the original data when unit variance scaling is applied, this method was preferred.

Principal component analysis (PCA) is used to project the data in two or three dimension in order to be able to discern outliers. This method transforms the original variables in new, uncorrelated variables or principal components (PC) in such a way that the first PC represents most of the variation present in the original dataset. Higher PC will subsequently contain less amount of the variance [5,7,26].

Partial least-squares discriminant analysis (PLS-DA) is a multivariate classification method based on PLS, a regression extension of PCA. This technique uses a priori knowledge about the data to maximize the separation between samples belonging to different classes [7,35]. Here, it concerns information about the pathological condition (healthy or HD) of each sample. The obtained scores plots are used to visualize the separation between the samples based on their class membership. The corresponding loadings plots contain information about the variables responsible for the observed separation. Both PCA and PLS calculations were performed with R (R version 2.7.2, http://www.r-project.org, packages *FactoMineR* version 1.14, *caret* version 4.76). A double cross-validation (2CV) strategy and permutation test [36] were applied to estimate the quality of the developed PLS-DA model. The data points were randomly divided into 5 fractions, of which one was used for testing, whereas the others were used for training and validation. Leave-one-out cross validation (LOOCV) [36] is used for validation in order to select the optimal model. LOOCV uses all but one data point as the training set. The left-out data point is used for validation. The data were randomly ordered 100 times. The average of area under the ROC curve (AUROC) was used as the criterium. In order to select the optimal number of component for PLS-DA, different numbers were tested (from 1 to the number of samples). Furthermore, the number of misclassifications and the Q^2 value were collected. A distribution was generated for the H0 hypothesis that no difference exists between the two classes. In the permutation test, the class labels of control and disease are permutated by randomly assigning them to different individuals. With 'wrong' class labels, a classification model is again calculated. The data were permutated 1000 times. For each permutated set, the cross validation was repeated 20 times and the average parameter values were computed.

2.4.2 Univariate analysis

The bucket table created for the multivariate analysis was exported to Excel (Microsoft Office 2003) in order to carry out a Student's t-test. P levels smaller than 0.05 were considered to be significant. An advanced classifier, Support Vector Machine (SVM) [37], was applied to classify the samples based on the significantly differential buckets. Each sample contains one out of two possible labels, and is represented by a feature vector based on the values of the significant buckets. The data samples were randomly divided into 5 fractions, of which one was used for testing, whereas the others were used for training and validation. The data were randomly ordered 100 times. The average of area under the ROC curve (AUROC) was used as the criterium. To estimate the quality of the developed SVM model, the double cross-validation (2CV) strategy and permutation test described in [36] used for PLSDA were applied again. In all experiments, the SVM function in the package e1071 (version 1.5-22) in R (R version 2.7.2, http://www.r-project.org) was used with the default setting parameters, e.g. linear kernel and cost equals 1.

3. Results

3.1 Visual inspection of NMR spectra

There is a lot of information comprised in ^{1}H NMR spectra of serum and CSF samples. Visual inspection of the spectra can already reveal some interesting features in some cases. A representative ^{1}H Carr-Purcell-Meiboom-Gill (CPMG) serum spectrum of both a wild type and a transgenic HD rat is represented in additional file A (see additional file A). The projection of all available serum spectra onto each other makes a clear differentiation between wild-type (WT) and HD rats possible. Peaks representing glucose and succinic acid seem to be elevated in spectra from HD rats

(figure 1). This indicates that both glucose and succinic acid are increased in concentration in the serum of transgenic animals. Visual examination of an overlay of the CSF spectra does not indicate an immediate difference between WT and HD rats.

3.2 PCA to detect outliers

Principal component analysis (PCA) was applied on the data in order to uncover inherent similarities and differences potentially present within the spectral profiles. The scores plot of the first two principal components gives a representative overview of the data (figure 2).

3.2.1 Serum samples

The first principal component (PC1) accounts for 76.42 % of the total variation present in the dataset. The second principal component (PC2) accounts for an additional 10.02 % of the total variation. Data points lying outside the 95 % confidence ellipse (Hotelling T^2) can be considered as outliers, i.e. observations that have an aberrant behaviour. As indicated in figure 2, two data point are laying outside the Hotelling T^2 plot. Reconsidering these spectra reveals that the water suppression technique, which has been used while acquiring the spectra, has caused a shift of the baseline in comparison with the other spectra. This slightly different prospect is revealed by this multivariate statistical tool. However, these two spectra are not discarded for further analysis, as their differences are not from a biological nature.

3.2.2 CSF samples

The scores plot of the first two principal components accounts for 45.57 % of the total variation present in the CSF dataset. Possible outliers present in the dataset were uncovered by spotting the data points lying outside the 95% confidence ellipse of this scores plot. Consecutive PCA indicate three data points (one WT and two HD samples) with an aberrant behaviour (see additional file B, additional file C and

additional file D). Those spectra correspond with low available volume samples (< 40 µL lyophilized). For this reason, these spectra were excluded from the dataset before further analysis.

3.3 Discrimination between normal and pathological samples using PLS-DA

A partial least squares - discriminant analysis (PLS-DA) model was developed in order to discriminate samples according to their class membership, i.e. healthy or HD. The results of this analysis are depicted in figure 3. The scores plot shows a reasonable separation between both sample groups. The corresponding loadings plot reveals the metabolites that are responsible for the observed separation. Annotation of the loadings was confirmed by a comparison with reference spectra of an in house database recorded at 700 MHz. A double cross-validation (2CV) strategy and permutation test were applied to estimate the quality of the developed PLS-DA model, as described in [36]. This yielded an area under the ROC curve (AUROC) of 84.90 % and 72.73 % for serum and CSF respectively. For all assessment parameters (Q^2 value and the number of misclassifications), there is a clear distinction between the permutation distribution and the original classification. This shows that the specific classification is significant (see Additional File E1 for serum and E2 for CSF).

3.3.1 Serum samples

The loadings coefficients responsible for the separation of the serum samples according to their pathological condition correspond to the metabolites glucose (δ3.28, δ3.38, δ3.43, δ3.48, δ3.53, δ3.73, δ3.78, δ3.83, δ3.88, δ3.93 and δ5.23), lactate (δ1.33 and δ4.13), lipids (δ0.83, δ0.88, δ1.23, δ1.28, δ1.58, δ2.03, δ2.23, δ2.73, δ2.78, δ5.28 and δ5.33), succinic acid (δ2.43), glutamine (δ2.13 and δ2.43) and NAA (δ2.03). An overview of these results is given in table 1.

3.3.2 CSF samples

The most relevant loadings in the separation of CSF samples according to their pathological condition are situated around δ3.38, δ3.43, δ3.48, δ3.73, δ3.78, δ3.83 and δ3.88 (glucose) and δ1.33 (lactate).

3.4 Discrimination between normal and pathological samples using a univariate analysis

A Student's t-test is used to uncover the metabolites that show a statistically significant concentration difference between the two observed populations. Thus, these metabolites can be used to differentiate HD from healthy samples. The bucket table, which was created for the multivariate analysis, was used as input variable for the univariate analysis. For each bucket, the corresponding p-value was calculated. A p-value smaller than 0.05 was considered as being significant.

3.4.1 Serum samples

The bucket comprising both the singlet peak of the methylene groups of succinic acid and the multiplet peak of CH_2CO of glutamine was found to be significantly different between WT and TG rats. In order to assign this statistically significant difference unambiguously to the correct metabolite, both signals were integrated referenced to the methylene singlet of creatine at 3.93 ppm. A visual representation of these results is displayed in figure 4. These boxplots reveal that both metabolites contribute to the differentiation of the samples according to their genotype. Both glutamine ($p < 0.05$) and succinic acid ($p < 0.005$) were found to be significantly increased in HD rats compared to WT littermates. Glucose was found to be elevated in HD rats as well ($p < 0.05$). Furthermore, NAA was decreased in HD rats ($p < 0.05$). Moreover, the bucket comprising the area between 4.30 and 4.25 ppm is found to be significantly decreased

in HD rats. At present, we have not been able to identify the component(s) present in this bucket. These results are illustrated in figure 4. The thirteen most significant buckets for whom the p-value is smaller than 0.05, able in discriminating between healthy and HD samples, were used to build an SVM classifier. An area under the ROC curve (AUROC) of 71.44 % could be obtained. For all the assessment parameters, a clear distinction between the permutation distribution and the original classification is obtained. This demonstrates that the specific classification is significant (see Additional file E3).

3.4.2 CSF samples

In CSF, a student's t-test could not reveal any statistically significant metabolite concentration difference between WT and HD rats.

4. Discussion

In this study we have shown that rats transgenic for the HD mutation can be discriminated from healthy, wild-type littermates even prior to the onset of motor deficits [38]. A comparable study was already executed by Tsang and co-workers on a different animal model [27]. However, there are some inconsistencies between the figures and text described in their paper. The metabolites they indicated in the loadingsplot of the serum samples correspond to those mentioned for urine samples in the text and vice versa.

We used three different methods for the analysis. Visual inspection of the serum spectra already suggests that glucose as well as succinic acid levels are elevated in HD rats. These results are statistically confirmed by means of a univariate analysis. Moreover, the Student's t-test discloses that glutamine and NAA are able to discriminate between HD and WT rats in serum samples. PLS, a supervised multivariate statistical technique, was performed in order to reveal the discriminative

power of different metabolites between the studied groups. For the serum samples, the loadings which are able to separate healthy from pathological samples correspond to the metabolites glucose, glutamine, lipids, lactate, succinic acid and NAA. Regarding CSF spectra, visual inspection and univariate analysis are not able to discriminate WT from HD rats. PLS, on the other hand, distinguishes both glucose and lactate as discriminating metabolites.

As well in CSF as in serum samples, glucose is indicated as a biomarker to detect Huntington Disease presymptomatically. Although the literature [2,4,39] exposes some evidence of diabetes in subjects with HD, plasma glucose measurements in the same transgenic rat model at different ages does not show elevated glucose levels [28]. Thus, the observed difference in glucose must have been caused by a perturbation in another metabolic pathway. Some evidence of a disturbed glucose metabolism in HD can be found in the literature [4,40,41]. A possible explanation for the observed increased concentrations of glutamine and glucose is also documented: according to Sibson et al. [32], there is a 1:1 stoichiometry between the consumption of glucose and the formation of glutamate. The observed increase in glutamine can reflect a decreased glutaminase activity. Glutaminase is a mitochondrial enzyme found in neurons that takes care of the conversion of glutamine in glutamate. A perturbation of the glutamate-glutamine cycling may indicate an impairment of energy metabolism and mitochondrial respiration [27]. Thus, a decrease in glutaminase activity precludes the conversion of glutamine into glutamate. Owing to this, the neurons will have a lack of glutamate for neurotransmission. Finally, this will result in an elevated concentration of glutamine and glucose (due to decreased glucose utilization) as observed in our experiments [4]. This theory can be affirmed by previous studies which also point to a disruption of energy metabolism in HD [42,43].

In addition, the observed difference in lactate supports the theory of a deficiency of oxidative energy metabolism in rats transgenic for HD. The conversion of glucose to pyruvate during glycolysis involves the concomitant reduction of 2 moles of NAD^+ to NADH. In cells undergoing an aerobic respiration, pyruvate is subsequently oxidized to acetyl-Coenzyme A (Acetyl-CoA), which in turn, enters the citric acid cycle. In this way, the NADH produced during glycolysis is reoxidized through the mitochondrial electron transport chain. However, a perturbation in the Krebs cycle or mitochondrial electron transport chain might prevent pyruvate to enter the oxidative energy metabolism. Since NADH needs to be reoxidized to NAD^+ to maintain a steady state condition, an alternative pathway to transfer electrons is accomplished, i.e. the reduction of pyruvate to lactate by the catalyzed reaction of lactate dehydrogenase. As a result, an increased concentration of lactate, the end product of anaerobic glycolysis, indicates a defect in energy metabolism as well [44].

An elevated concentration of succinic acid is probably due to an inhibition of succinate dehydrogenase. This enzyme has two important functions in the mitochondria. First, it supports the oxidation of succinic acid to fumarate in the tricarboxylic acid (TCA) cycle. Second, it is linked to the respiratory chain, i.e. complex II of the mitochondrial electron transport chain [45,46]. A decrease in mitochondrial complex II activity has already been observed in HD [3,47,48]. Our finding of an elevated succinic acid concentration is therefore in line with these previous results.

A decreased NAA concentration is known to be a marker of neuronal dysfunction [27,49]. Reduced NAA levels have been well documented for several neurological diseases, e.g. Alzheimer disease [50], Huntington Disease [4,27,51] and Parkinson Disease [17,50]. Moreover, in HD it clearly correlates with the duration of symptoms

and the CAG repeat length [16,52]. The enzyme responsible for the synthesis of NAA, L-*N*-acetylaspartyl transferase, is exclusively found in mitochondria [50]. Since inhibitors of the mitochondrial respiratory chain seem to decrease the concentration of NAA [4], reduced NAA levels may reflect an impaired mitochondrial energy production [27].

Obviously, a change in the concentration of each of these metabolites can be linked to a variety of conditions. However, the changes observed for all metabolites combined could allow for a more narrow correlation, i.e. with HD.

5. Conclusions

In conclusion, this study shows that metabolic alterations already emerge in the serum and CSF of these rats before overt symptoms of HD are manifested. In an early stage of the disease, metabolic alterations seem to be more pronounced in serum compared to CSF. All metabolites found being able at discriminating WT from HD rats point to a disruption of the mitochondrial respiratory chain and energy metabolism. These findings suggest that compounds which act on mitochondria might have a positive effect on HD pathology. Moreover, this study shows that ^{1}H NMR spectroscopy is an appropriate tool to investigate the metabolic disturbance occurring in early stages of HD. A more profound knowledge of the disease pathology can be obtained by performing a metabolomic analysis of samples originating from more mature rats. Such an analysis could give more insight into the development of this disease, which in turn may lead to improved therapeutic interventions.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

KAV thanks the University of Antwerp (UA) and the FWO-Vlaanderen for funding. AVDL, JCM and RD thank the FWO-Vlaanderen for research funding (G.0064.07). The 700 MHz NMR spectrometer (Ghent, Belgium) of the interuniversity NMR facility was financed by Ghent University, the Free University of Brussels (VUB) and the UA via the 'Zware Apparatuur' FFEU Incentive of the Flemish Government. KL and TNV are supported by an SBO grant (IWT-600450) and a BOF interdisciplinary grant of the University of Antwerp. LC and HN are supported by the European Community 'RATstream STREP project'.

References

[1] The Huntington's Disease Collaborative Research Group, A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's Disease chromosomes, Cell 72 (1993) pp. 971-983.

[2] O.A. Andreassen, A. Dedeoglu, R.J. Ferrante, B.G. Jenkins, K.L. Ferrante, M. Thomas, A. Friedlich, S.E. Browne, G. Schilling, D.R. Borchelt, S.M. Hersch, C.A. Ross, M.F. Beal, Creatine increases survival and delays motor symptoms in a transgenic animal model of Huntington's disease, Neurobiology of Disease 8 (2001) pp. 479-491.

[3] S.J. Tabrizi, M.W.J. Cleeter, J. Xuereb, J.-W. Taanman, J.M. Cooper, A.H.V. Schapira, Biochemical Abnormalities and excitotoxicity in Huntington's Disease Brain, Annals of Neurology 45 (1999) pp. 25-32.

[4] B.G. Jenkins, P. Klivenyi, E. Kustermann, O.A. Andreassen, R.J. Ferrante, B.R. Rosen, M.F. Beal, Nonlinear decrease over time in N-acetyl aspartate levels in the absence of neuronal loss and increases in glutamine and glucose in transgenic Huntington's disease mice, Journal of Neurochemistry 74 (2000) pp. 2108-2119.

[5] J.K. Nicholson, J.C. Lindon, E. Holmes, 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data, Xenobiotica 29 (1999) pp. 1181-1189.

[6] J.G. Bundy, M.P. Davey, M.R. Viant, Environmental metabolomics: a critical review and future perspectives, Metabolomics 5 (2009) pp. 3-21.

[7] J.C. Lindon, J.K. Nicholson, E. Holmes, The Handbook of Metabonomics and Metabolomics, Elsevier, Amsterdam, 2007.

[8] M.R. Viant, B.G. Lyeth, M.G. Miller, R.F. Berman, An NMR metabolomic investigation of early metabolic disturbances following traumatic brain injury in a mammalian model, NMR in Biomedicine 18 (2005) pp. 507-516.

[9] O.A.C. Petroff, R.K. Yu, T. Ogino, High-resolution proton magnetic resonance analysis of human cerebrospinal fluid, Journal of Neurochemistry 47 (1986) pp. 1270-1276.

[10] P. Khandelwal, C.E. Beyer, Q. Lin, L.E. Schechter, A.C. Bach, Studying rat brain neurochemistry using nanoprobe NMR spectroscopy: a metabonomics approach, Analytical Chemistry 76 (2004) pp. 4123-4127.

[11] J.C. Lindon, J.K. Nicholson, J.R. Everett, NMR spectroscopy of biofluids, Annual Reports on NMR Spectroscopy 38 (1999) pp. 1-88.

[12] A.D. Maher, D. Crockford, H. Toft, D. Malmodin, J.H. Faber, M.I. McCarthy, A. Barrett, M. Allen, M. Walker, E. Holmes, J.C. Lindon, J.K. Nicholson, Optimization of human plasma ^{1}H NMR spectroscopic data processing for high-throughput metabolic phenotyping studies and detection of insulin resistance related to type 2 diabetes, Analytical Chemistry 80 (2008) pp. 7354-7362.

[13] J.C. Lindon, E. Holmes, J.K. Nicholson, Pattern recognition methods and applications in biomedical magnetic resonance, Progress in Nuclear Magnetic Resonance Spectroscopy 39 (2001) pp. 1-40.

[14] R. Hewer, J. Vorster, F.E. Steffens, D. Meyer, Applying biofluid ^{1}H NMR-based metabonomic techniques to distinguish between HIV-1 positive/AIDS patients on antiretroviral treatment and HIV-1 negative individuals, Journal of Pharmaceutical and Biomedical Analysis 41 (2006) pp. 1442-1446.

[15] M.A. Constantinou, A. Tsantili-Kakoulidou, I. Andreadou, E.K. Iliodromitis, D.T. Kremastinos, E. Mikros, Application of NMR-based metabonomics in the investigation of myocardial ischemia-reperfusion, ischemic preconditioning and antioxidant intervention in rabbits, European Journal of Pharmaceutical Sciences 30 (2007) pp. 303-314.

[16] B.G. Jenkins, H.D. Rosas, Y.C.I. Chen, T. Makabe, R. Myers, M. MacDonald, B.R. Rosen, M.F. Beal, W.J. Koroshetz, ^{1}H NMR

spectroscopy studies of Huntington's disease - Correlations with CAG repeat numbers, Neurology 50 (1998) pp. 1357-1365.

[17] H.M. Baik, B.Y. Choe, H.K. Lee, T.S. Suh, B.C. Son, J.M. Lee, Metabolic alterations in Parkinson's disease after thalamotomy, as revealed by ^{1}H MR spectroscopy, Korean Journal of Radiology 3 (2002) pp. 180-188.

[18] A.R. Tate, S.J.P. Damment, J.C. Lindon, Investigation of the metabolite variation in control rat urine using ^{1}H NMR spectroscopy, Analytical Biochemistry 291 (2001) pp. 17-26.

[19] A. Bender, D.P. Auer, T. Merl, R. Reilmann, P. Saemann, A. Yassouridis, J. Bender, A. Weindl, M. Dose, T. Gasser, T. Klopstock, Creatine supplementation lowers brain glutamate levels in Huntington's disease, Journal of Neurology 252 (2005) pp. 36-41.

[20] S. Kasparová, Z. Sumbalová, P. Bystrický, J. Kucharská, T. Liptaj, V. Mlynárik, A. Gvozdjáková, Effect of coenzyme Q10 and vitamin E on brain energy metabolism in the animal model of Huntington's disease, Neurochemistry International 48 (2006) pp. 93-99.

[21] P. Verbessem, J. Lemiere, B.O. Eijnde, S. Swinnen, L. Vanhees, M. Van Leemputte, P. Hespel, R. Dom, Creatine supplementation in Huntington's disease - A placebo-controlled pilot trial, Neurology 61 (2003) pp. 925-930.

[22] I. Tkác, C.D. Keene, J. Pfeuffer, W.C. Low, R. Gruetter, Metabolic changes in quinolinic acid-lesioned rat striatum detected non-invasively by in vivo ^{1}H NMR spectroscopy, Journal of Neuroscience Research 66 (2001) pp. 891-898.

[23] E. Holmes and H. Antti, Chemometric contributions to the evolution of metabonomics: mathematical solutions to characterising and interpreting complex biological NMR spectra, The Analyst 127 (2002) pp. 1549-1557.

[24] T.M. Alam and M.K. Alam, Chemometric analysis of NMR spectroscopy data: a review, Annual Reports on NMR Spectroscopy 54 (2005) pp. 41-80.

[25] E.M. Lenz, J. Bright, I.D. Wilson, A. Hughes, J. Morrisson, H. Lindberg, A. Lockton, Metabonomics, dietary influences and cultural differences: a ^{1}H NMR-based study of urine samples obtained from healthy British and Swedish subjects, Journal of Pharmaceutical & Biomedical Analysis 36 (2004) pp. 841-849.

[26] W. El-deredy, Pattern recognition approaches in biomedical and clinical magnetic resonance spectroscopy: a review, NMR in Biomedicine 10 (1997) pp. 99-124.

[27] T.M. Tsang, B. Woodman, G.A. Mcloughlin, J.L. Griffin, S.J. Tabrizi, G.P. Bates, E. Holmes, Metabolic characterization of the R6/2 transgenic mouse model of Huntington's disease by high-resolution MAS ^{1}H NMR spectroscopy, Journal of Proteome Research 5 (2006) pp. 483-492.

[28] S. von Hörsten, I. Schmitt, H.P. Nguyen, C. Holzmann, T. Schmidt, T. Walther, M. Bader, R. Pabst, P. Kobbe, J. Krotova, D. Stiller, A. Kask, A. Vaarmann, S. Rathke-Hartlieb, J.B. Schulz, U. Grasshoff, I. Bauer, A.M.M. Vieira-Saecker, M. Paul, L. Jones, K.S. Lindenberg, B. Landwehrmeyer, A. Bauer, X.J. Li, O. Riess, Transgenic rat model of Huntington's disease, Human Molecular Genetics 12 (2003) pp. 617-624.

[29] I.F. Duarte, B.J. Goodfellow, A. Barros, J.G. Jones, C. Barosa, L. Diogo, P. Garcia, A.M. Gil, Metabolic characterisation of plasma in juveniles with glycogen storage disease type 1a (GSD1a) by high-resolution ^{1}H NMR spectroscopy, NMR in Biomedicine 20 (2007) pp. 401-412.

[30] E. Brouillet, P. Hantraye, R.J. Ferrante, R. Dolan, A. Leroy-Willig, N.W. Kowall, M.F. Beal, Chronic mitochondrial energy impairment produces selective striatal degeneration and abnormal choreiform movements in primates, Proceedings of the National Academy of Sciences of the United States of America 92 (1995) pp. 7105-7109.

[31] I. Tkác, J.M. Dubinsky, C.D. Keene, R. Gruetter, W.C. Low, Neurochemical changes in Huntington R6/2 mouse striatum detected by *in vivo* ^{1}H NMR spectroscopy, Journal of Neurochemistry 100 (2009) pp. 1397-1406.

[32] N.R. Sibson, A. Dhankhar, G.F. Mason, D.L. Rothman, K.L. Behar, R.G. Shulman, Stoichiometric coupling of brain glucose metabolism and glutamatergic neuronal activity, Proceedings of the National Academy of Sciences of the United States of America 95 (1998) pp. 316-321.

[33] P. Conforti, C. Ramos, B.L. Apostol, D.A. Simmons, H.P. Nguyen, O. Riess, L.M. Thompson, C. Zuccato, E. Cattaneo, Blood level of brain-derived neurotrophic factor mRNA is progressively reduced in rodent models of Huntington's disease: restoration by the neuroprotective compound CEP-1347, Molecular and Cellular Neuroscience 39 (2008) pp. 1-7.

[34] P. Khandelwal, C.E. Beyer, Q. Lin, P. McGonigle, L.E. Schechter, A.C. Bach, Nanoprobe NMR spectroscopy and in vivo microdialysis: new analytical methods to study brain neurochemistry, Journal of Neuroscience Methods 133 (2004) pp. 181-189.

[35] L. Eriksson, E. Johansson, H. Antti, E. Holmes, Multi-and Megavariate Data Analysis: Finding and Using Regularities in Metabonomics

Data, in: D.G.Robertson and J.Lindon (Eds.), Metabolomics in toxicity assessment, Taylor & Francis Group, Boca Raton, 2005, pp. 263-335.

[36] J.A. Westerhuis, H.C.J. Hoefsloot, S. Smit, D.J. Vis, A.K. Smilde, E.J.J. van Velzen, J.P.M. van Duijnhoven, F.A. van Dorsten, Assessment of PLSDA cross validation, Metabolomics 4 (2008) pp. 81-89.

[37] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-Based Learning Metods, Cambridge University Press, Cambridge, 2000.

[38] H.P. Nguyen, P. Kobbe, H. Rahne, T. Wörpel, B. Jäger, M. Stephan, R. Pabst, C. Holzmann, O. Riess, H. Korr, O. Kántor, E. Petrasch-Parwez, R. Wetzel, A. Osmand, S. von Hörsten, Behavioral abnormalities precede neuropathological markers in rats transgenic for Huntington's disease, Human Molecular Genetics 15 (2006) pp. 3177-3194.

[39] M.S. Hurlbert, W. Zhou, C. Wasmeier, F.G. Kaddis, J.C. Hutton, C.R. Freed, Mice transgenic for an expanded CAG repeat in the Huntington's Disease gene develop diabetes, Diabetes 48 (1999) pp. 649-651.

[40] S.E. Browne and M.F. Beal, The Energetics of Huntington's Disease, Neurochemical Research 29 (2004) pp. 531-546.

[41] A. Ciarmiello, M. Cannella, S. Lastoria, M. Simonelli, L. Frati, D.C. Rubinsztein, F. Squitieri, Brain white-matter volume loss and glucose hypometabolism precede the clinical symptoms of Huntington's Disease, Journal of Nuclear Medicine 47 (2006) pp. 215-222.

[42] B.G. Jenkins, W.J. Koroshetz, M.F. Beal, B.R. Rosen, Evidence for impairment of energy metabolism in vivo in Huntington's disease using localized ^{1}H NMR spectroscopy, Neurology 43 (1993) pp. 2689-2695.

[43] M. Gu, M.T. Gash, V.M. Mann, F. Javoy-Agid, J.M. Cooper, A.H.V. Schapira, Mitochondrial defect in Huntington's disease caudate nucleus, Annals of Neurology 39 (1996) pp. 385-389.

[44] C.K. Mathews, K.E. van Holde, K.G. Ahern, Carbohydrate Metabolism I: Anaerobic Processes in Generating Metabolic Energy, Biochemistry, Addison Wesley Longman, San Francisco, 2000, pp. 446-482.

[45] B.R. Underwood, D. Broadhurst, W.B. Dunn, D.I. Ellis, A.W. Michell, C. Vacher, D.E. Mosedale, D.B. Kell, R.A. Barker, D.J. Grainger, D.C. Rubinsztein, Huntington disease patients and transgenic

mice have similar pro-catabolic serum metabolite profiles, Brain 129 (2006) pp. 877-886.

[46] C. Dautry, F. Condé, E. Brouillet, V. Mittoux, M.F. Beal, G. Bloch, P. Hantraye, Serial 1H-NMR spectroscopy study of metabolic impairment in primates chronically treated with succinate dehydrogenase inhibitor 3-nitropropionic acid, Neurobiology of Disease 6 (2006) pp. 259-268.

[47] S.E. Browne, A.C. Bowling, U. MacGarvey, M.J. Baik, S.C. Berger, M.M.K. Muqit, E.D. Bird, M.F. Beal, Oxidative damage and metabolic dysfunction in Huntington's Disease: selective vulnerability of the basal ganglia, Annals of Neurology 41 (1997) pp. 646-653.

[48] M. Gu, M.T. Gash, V.M. Mann, F. Javoy-Agid, J.M. Cooper, A.H.V. Schapira, Mitochondrial defect in Huntington's Disease caudate nucleus, Annals of Neurology 39 (1996) pp. 385-389.

[49] C. Demougeot, C. Marie, M. Giroud, A. Beley, N-Acetylaspartate: a literature review of animal research on brain ischaemia, Journal of Neurochemistry 90 (2004) pp. 776-783.

[50] JK. Choi, A. Dedeoglu, B.G. Jenkins, Application of MRS to mouse models of neurodegenerative illness, NMR in Biomedicine 20 (2007) pp. 216-237.

[51] A. van Dellen, J. Welch, R.M. Dixon, P. Cordery, D. York, P. Styles, C. Blakemore, A.J. Hannan, N-Acetylaspartate and DARPP-32 levels decrease in the corpus striatum of Huntington's disease mice, NeuroReport 11 (2000) pp. 3751-3757.

[52] J.C.H. van Oostrom, P.E. Sijens, R.A.C. Roos, K.L. Leenders, ^{1}H magnetic resonance spectroscopy in preclinical Huntington disease, Brain Research 1168 (2007) pp. 67-71.

Figures

Figure 1 - Visual comparison of serum spectra

An overlay of all acquired ^{1}H NMR 700 MHz serum spectra, **a** some representative peaks of glucose **b** signal of succinic acid. Spectra of wild-type rats are coloured black; those of HD rats blue

Figure 2 - PCA to detect outliers

PCA scores plot of the first and second principal component mapping the serum spectra of twelve wild-type and ten HD rats (females, 2 months old). The Hotelling T^2 ellipse is used to elucidate outliers. Colour coding: black = wild type rats, blue = HD.

Figure 3 - PLS, clustering of samples according to their pathological condition

a PLS scores plot of ^{1}H NMR serum spectra of twelve wild-type (black squares) and ten HD (blue triangles) rats (females, 2 months old). **b** The corresponding loadings plot represents the variables attributable to the observed class separation. **c** PLS scores plot of ^{1}H NMR CSF spectra of nine wild-type and eight HD rats (females, two months old). **d** The loading coefficient plot corresponding to the scores plot (c).

Figure 4 - Visualisation of the results obtained by univariate analysis

A boxplot is used to visualize the data of the student's t-test. The boxplot is based on the five-number-summary, i.e. the median (black line), 25 % quartile (lower box line), 75 % quartile (upper box line), minimum (lower whisker) and maximum (upper whisker). This means that 50 % of the data points are situated in the box. The range between the 25 % quartile (Q25) and 75 % quartile (Q75) is called the interquartile range (IQR). Subsequently, the whiskers are calculated by the formula f_{lower} = Q25 – 1.5*IQR for the lower whisker and respectively f_{upper} = Q75 + 1.5*IQR for the upper whisker. **a** glucose, **b** glutamine, **c** NAA, **d** the unknown component in the bucket comprising the area between 4.30 and 4.25 ppm and **e** succinic acid. The boxplot of WT samples is coloured grey whereas the boxplot of TG samples has a blue colour.

Tables

Table 1 - Overview of the metabolites responsible for the separation of serum samples according to their pathological condition.

Additional Files

Additional file A - Typical serum spectrum
additional file A.pdf

A representative ^{1}H NMR 700 MHz CPMG serum spectrum of a wild-type (black) and HD (blue) rat.

Additional file B – First PCA analysis of the CSF samples
additional file B.pdf

PCA scores plot of the first and second principal component mapping the CSF spectra of wild-type and HD rats. The Hotelling T^2 ellipse is used to elucidate outliers. Colour coding: black = wild type rats, blue = HD.

Additional file C – Second PCA analysis of the CSF samples
additional file C.pdf

A second PCA was performed after that the outlier visual in additional file B was removed from the dataset. The PCA scores plot of the first and second principal component mapping the CSF spectra of wild-type and HD rats resulting from this analysis is displayed. The Hotelling T^2 ellipse is used to elucidate outliers. Colour coding: black = wild type rats, blue = HD.

Additional file D – Third PCA analysis of the CSF samples
additional file D.pdf

A third PCA was performed after that the outliers visual in additional file B and additional file C were removed from the dataset. The PCA scores plot of the first and second principal component mapping the CSF spectra of wild-type and HD rats

resulting from this analysis is displayed. The Hotelling T^2 ellipse is used to elucidate outliers. Colour coding: black = wild type rats, blue = HD.

Additional file E – Assessment of the classification results based on cross model validation and permutation tests.

additional file E.pdf

The parameters used to quantify the classifications are the number of misclassifications, Q2 and the area under the ROC curve. The three cases mentioned in the paper are in Additional file E1 (PLS-DA prediction results of serum spectra), Additional file E2 (PLS-DA prediction results of CSF spectra) and Additional file E3 (SVM prediction results of serum spectra). The solid circles are the actual values from the original classifications.

Table 1

	metabolites
higher concentration in TG animals	glucose lactate succinic acid glutamine
lower concentration in TG animals	NAA lipids

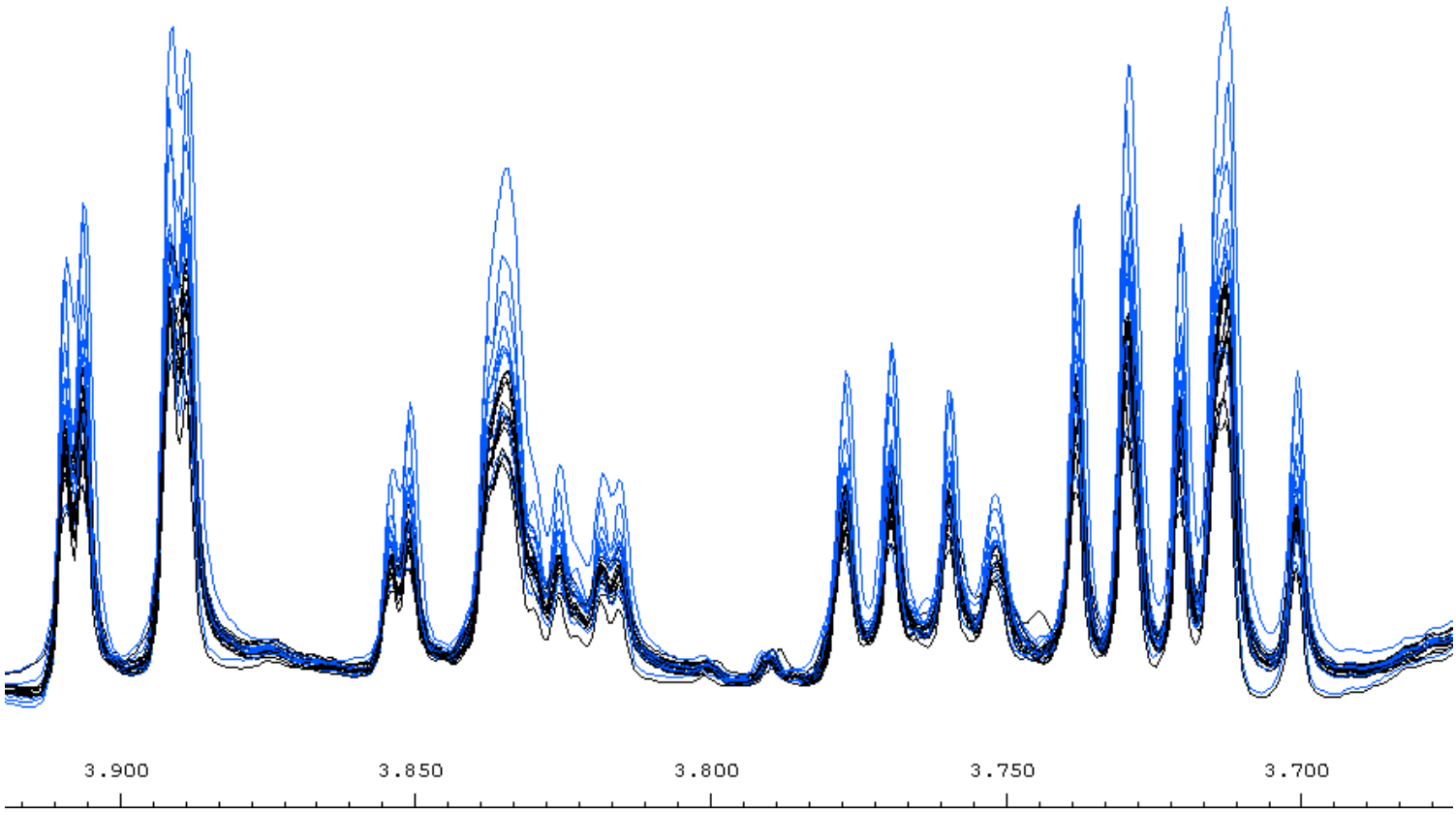


Figure 1a

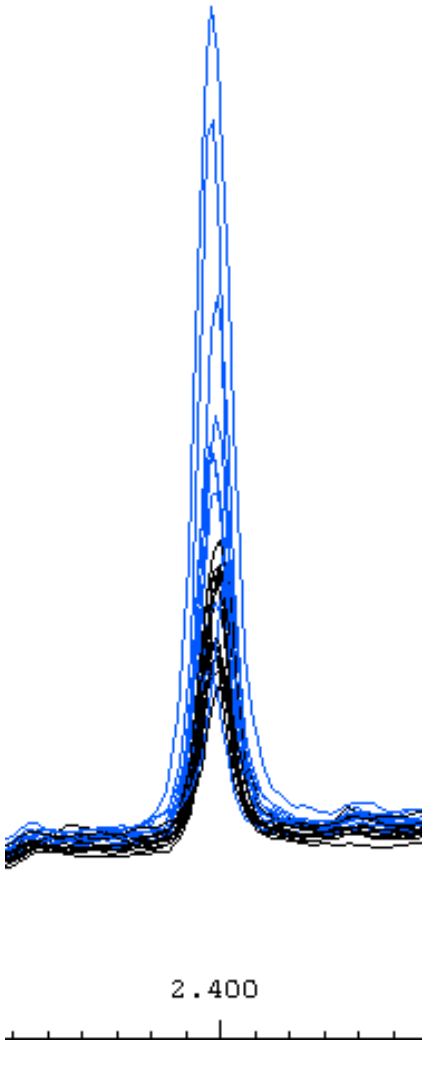
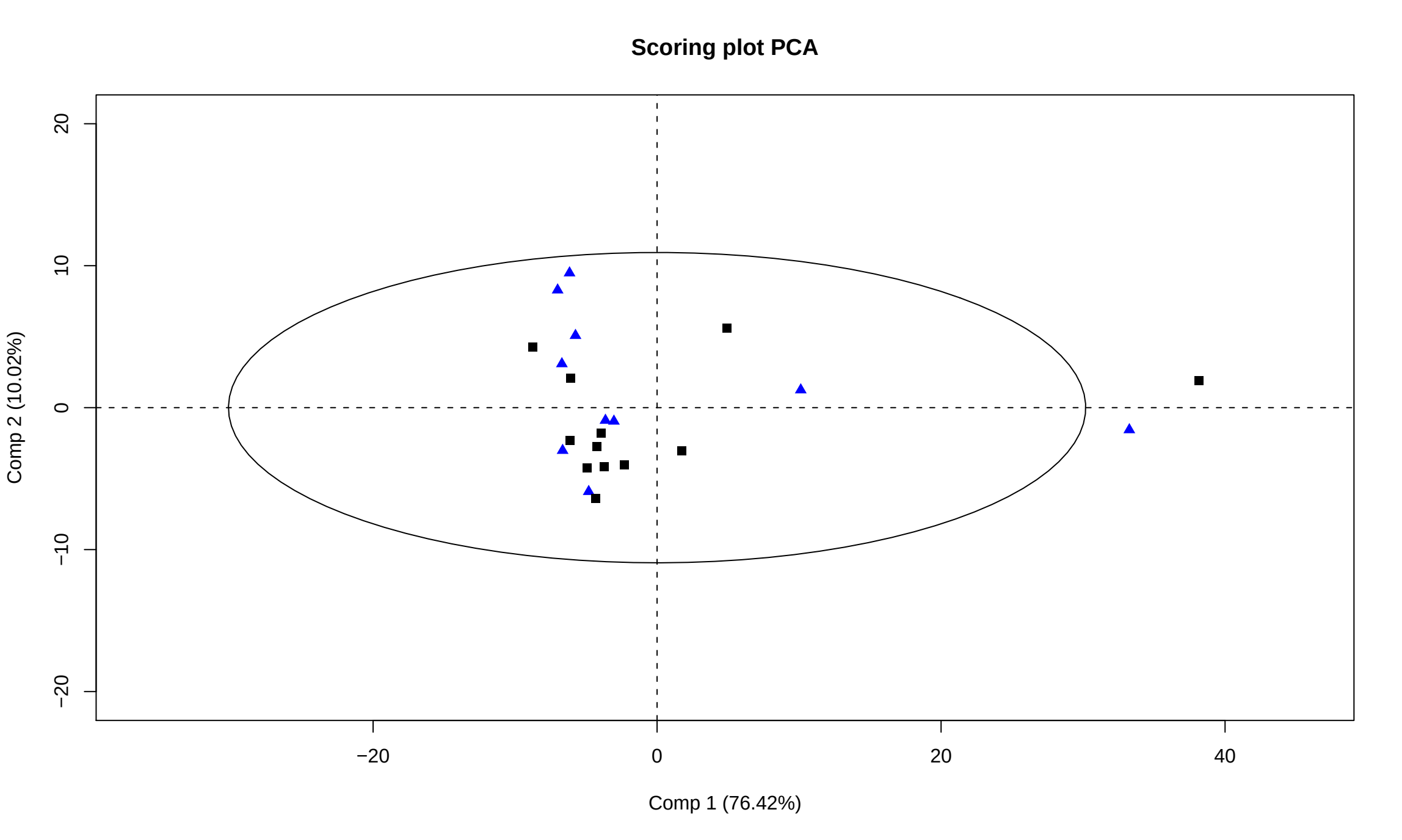


Figure 1b


Scoring plot PCA
Comp 2 (10.02%)
Comp 1 (76.42%)
20
10
0
−10
−20
−20
0
20
40

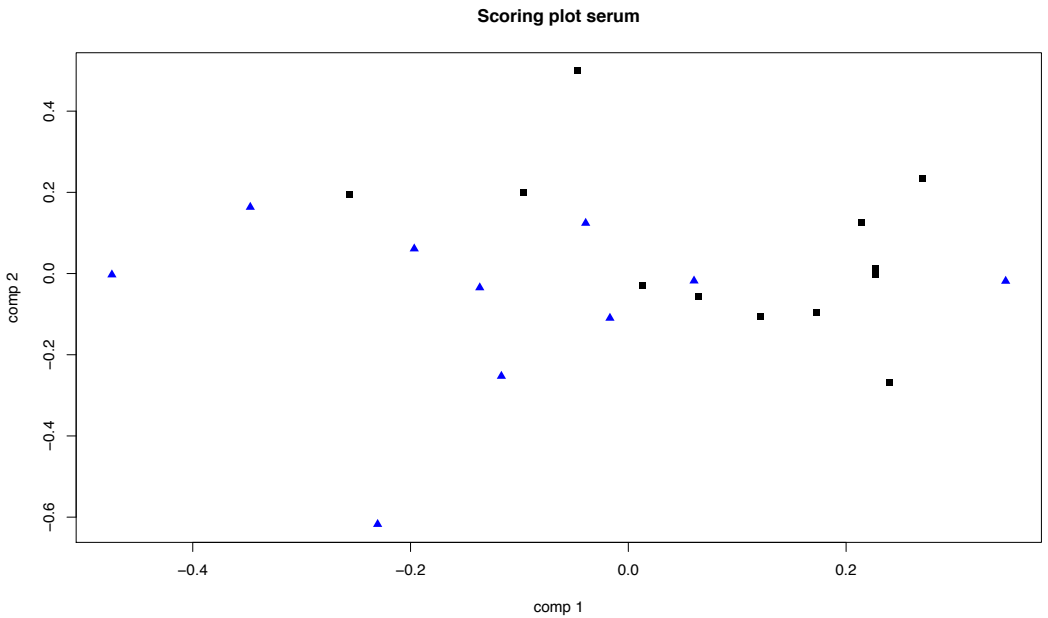


Figure 3a

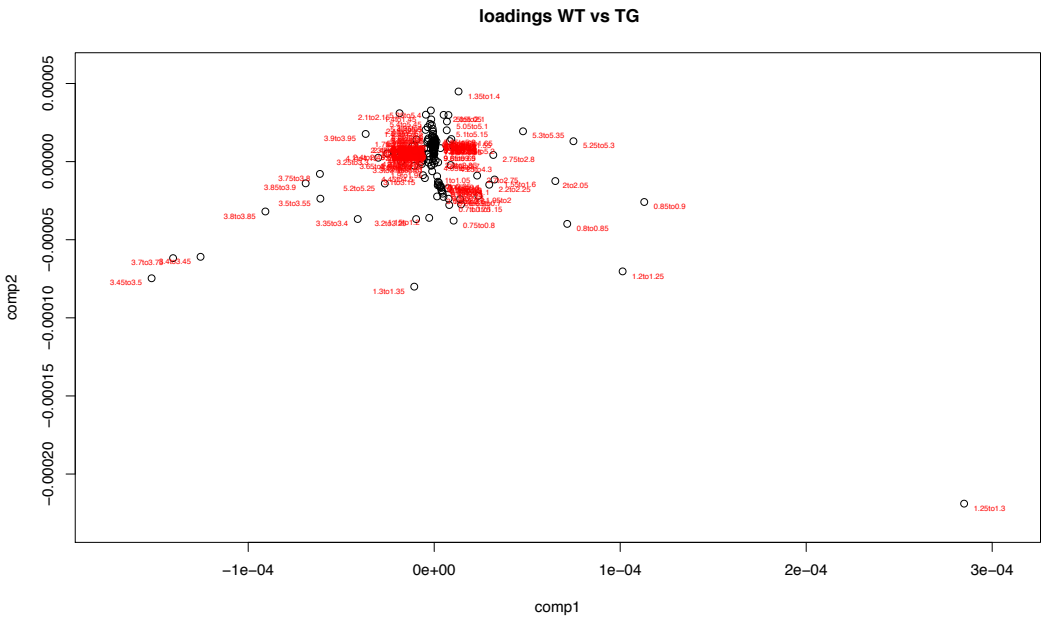


Figure 3b

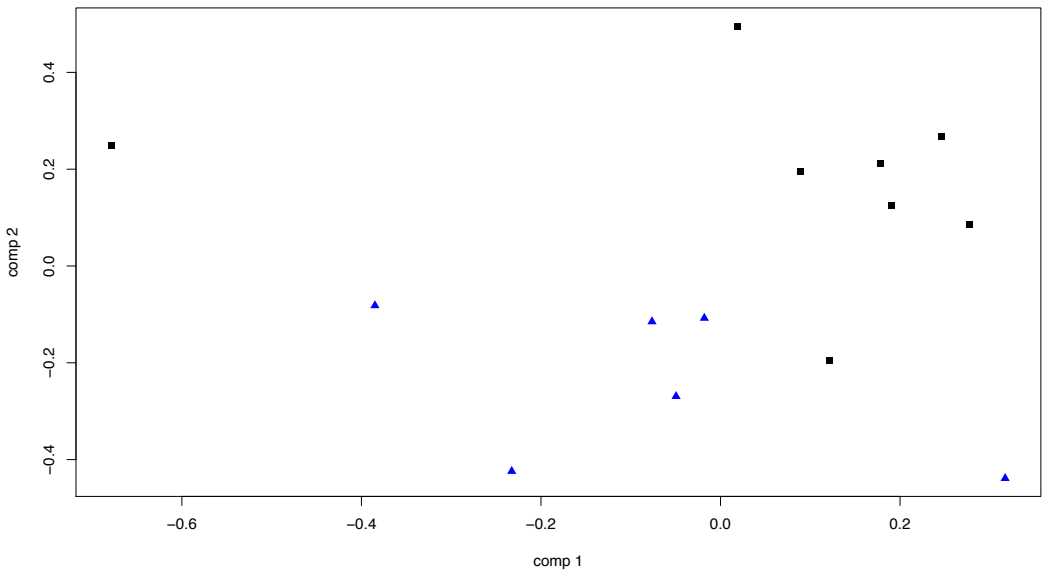


Figure 3c

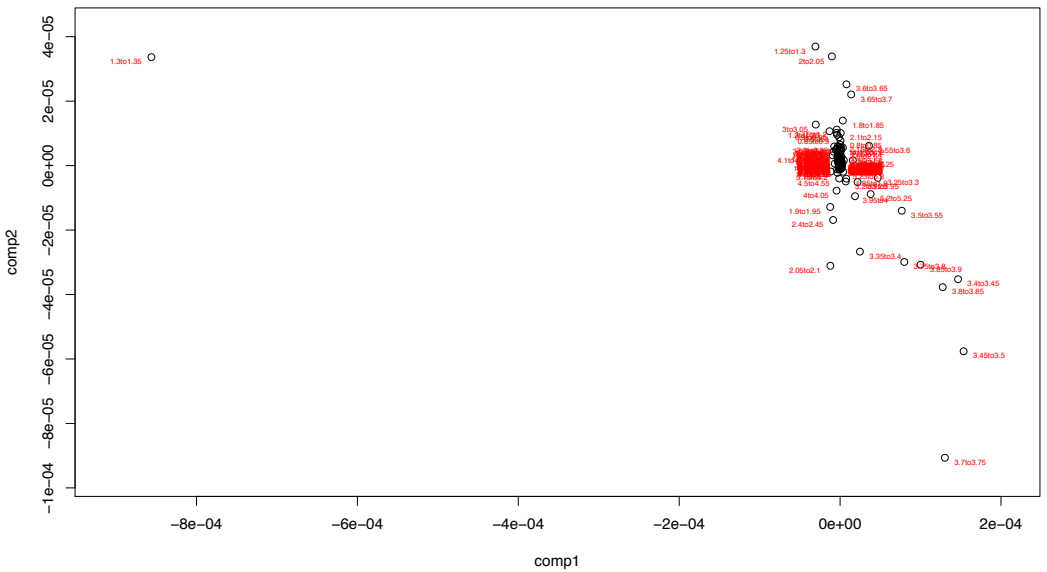


Figure 3d

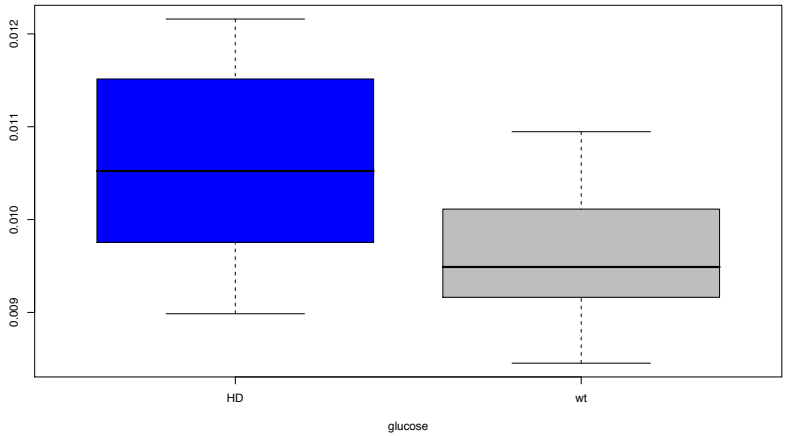


Figure 4a

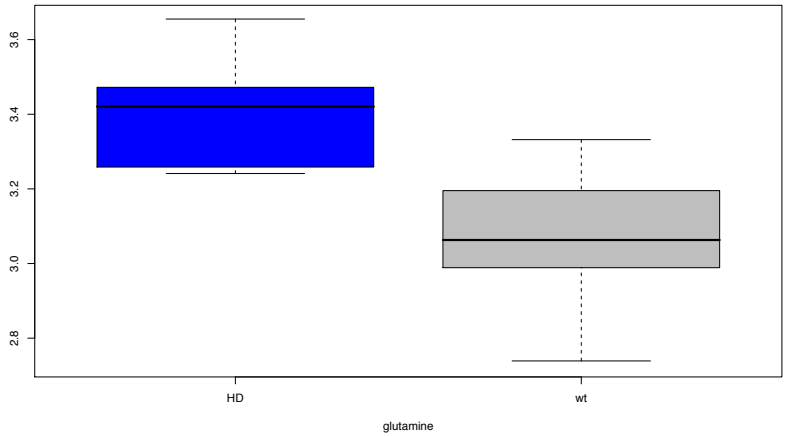


Figure 4b

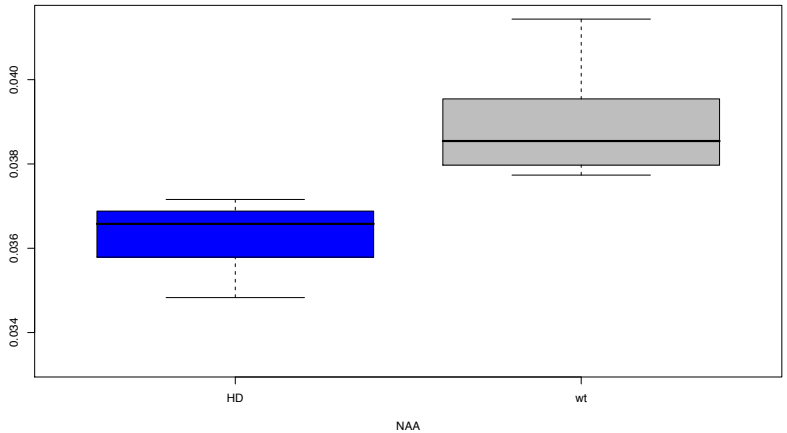


Figure 4c

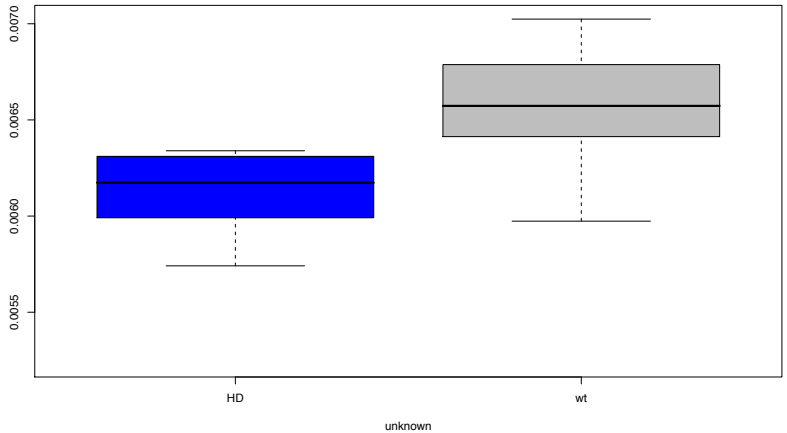


Figure 4d

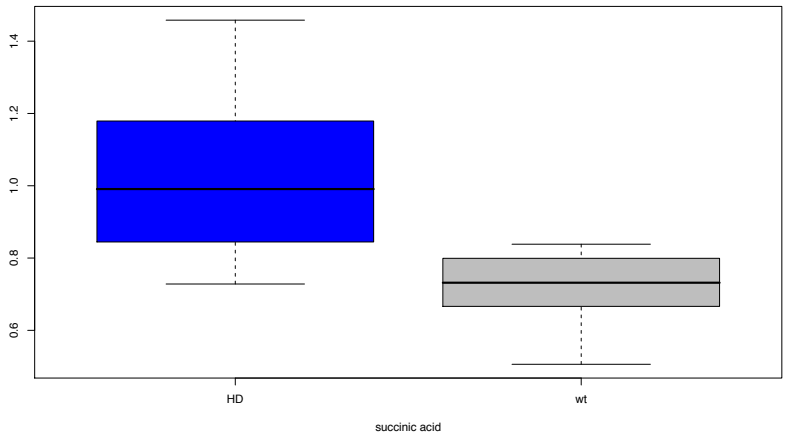
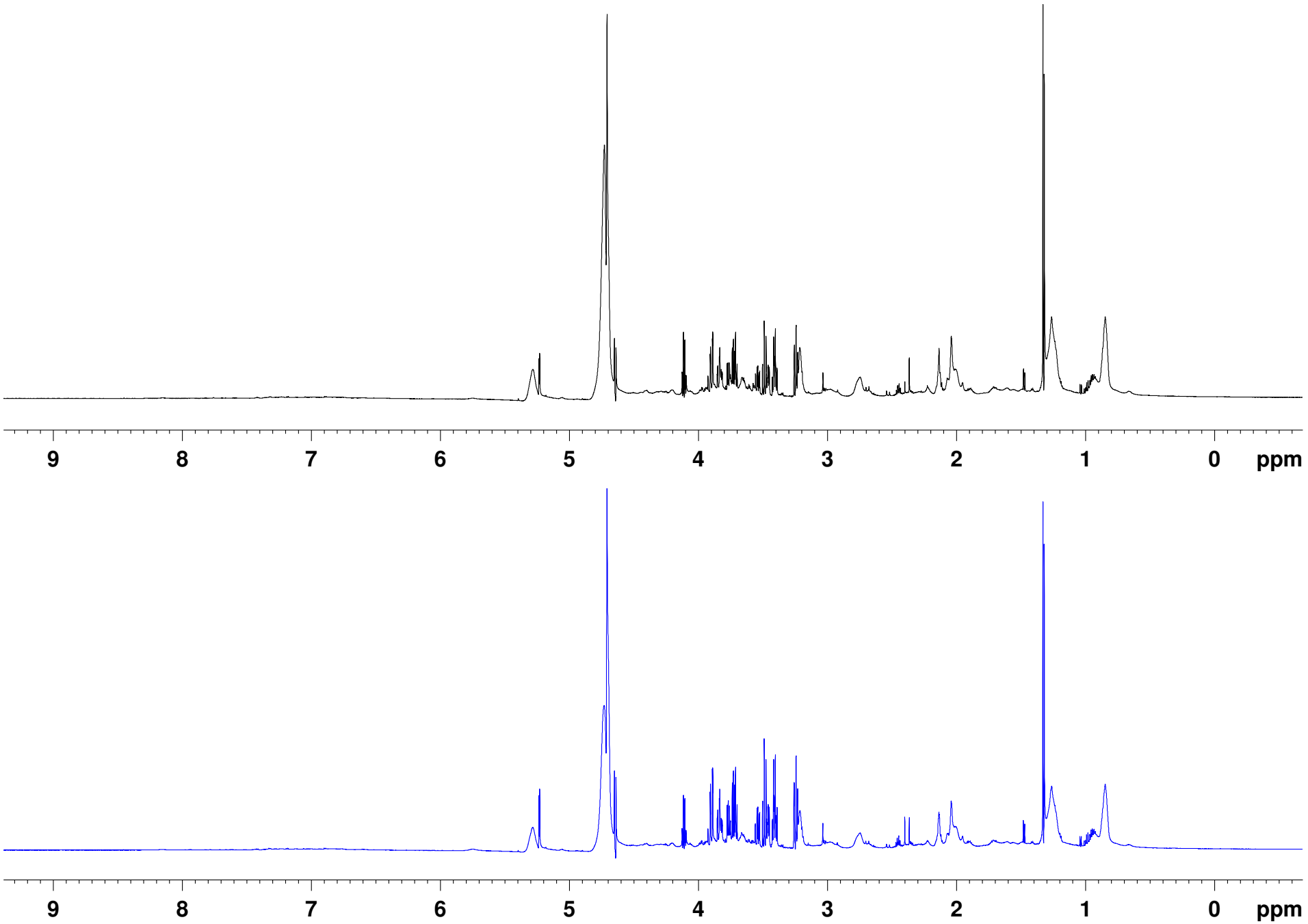
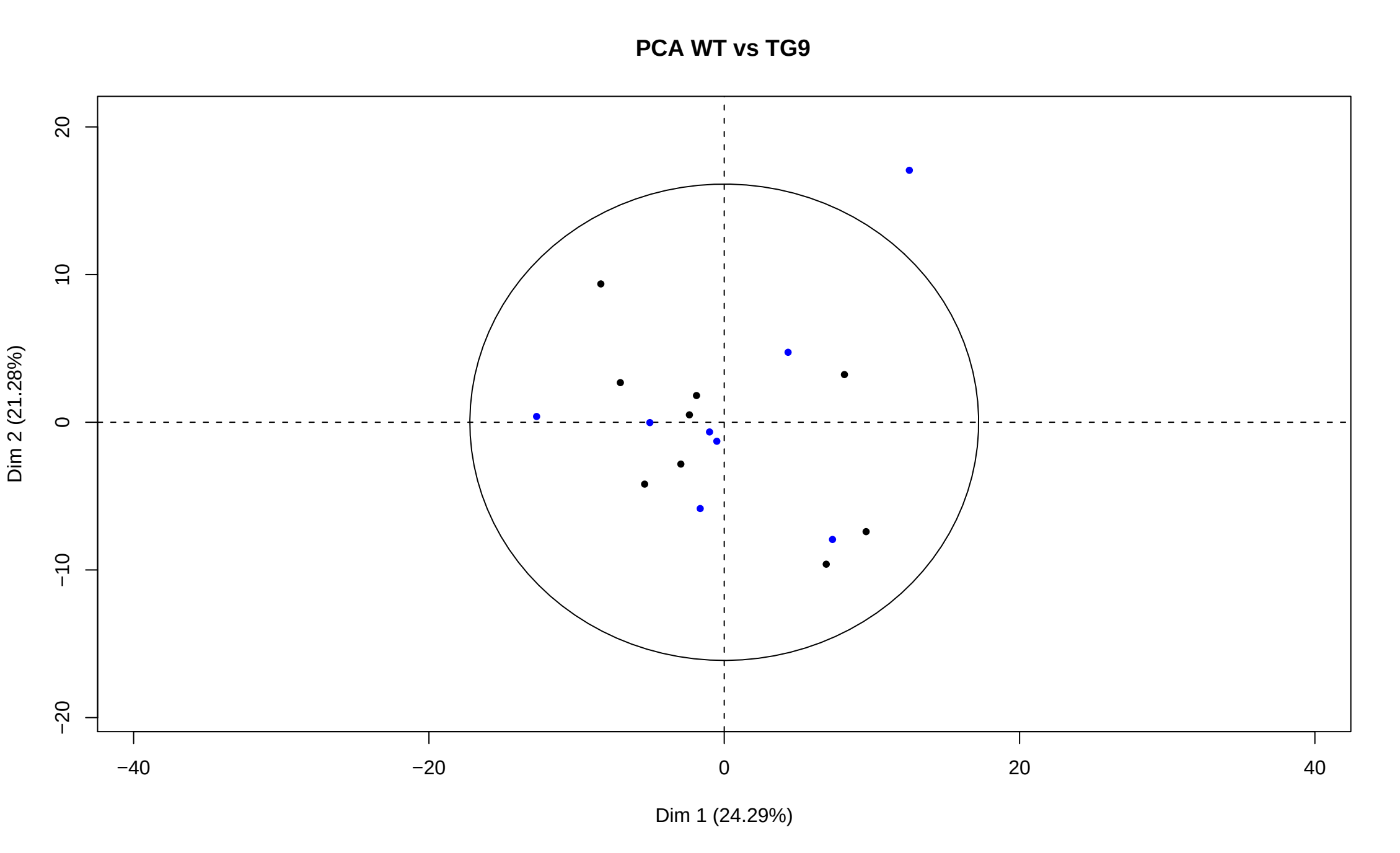
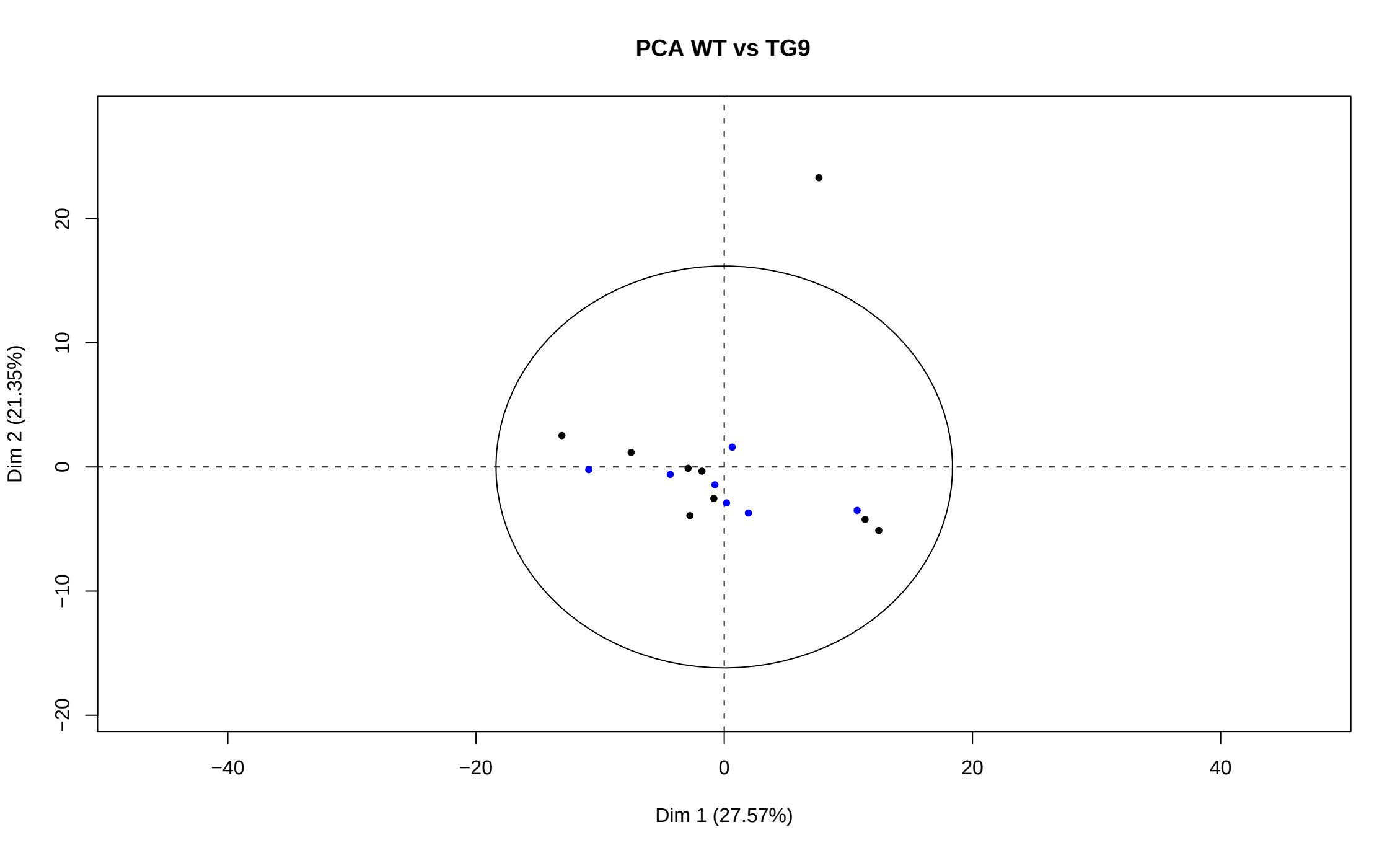


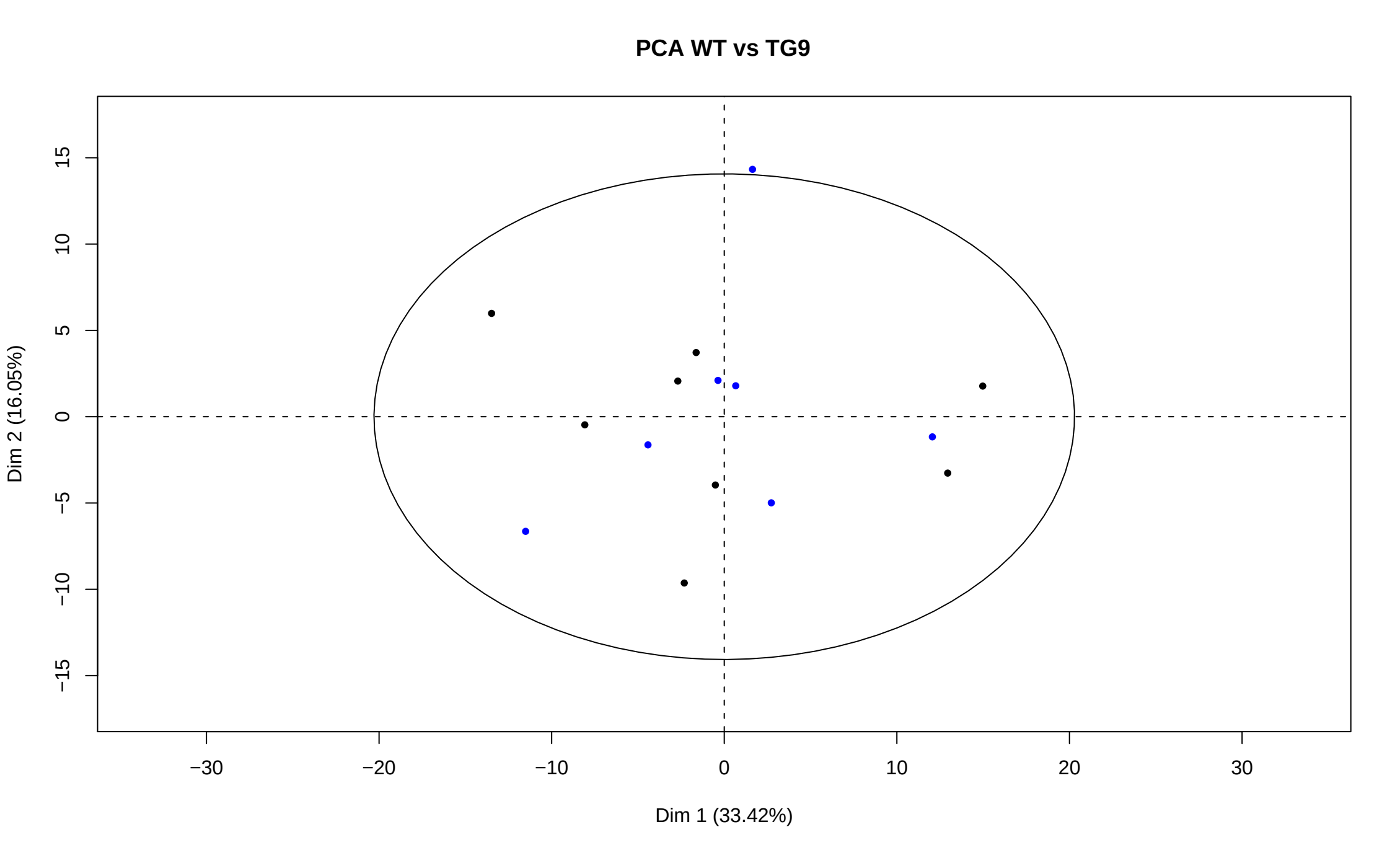
Figure 4e

	metabolites
higher concentration in TG animals	glucose lactate succinic acid glutamine
lower concentration in TG animals	NAA lipids


9
8
7
6
5
4
3
2
1
0
ppm
9
8
7
6
5
4
3
2
1
0
ppm


PCA WT vs TG9
Dim 2 (21.28%)
Dim 1 (24.29%)
20
10
0
−10
−20
−40
−20
0
20
40


PCA WT vs TG9
Dim 2 (21.35%)
20
10
0
−10
−20
−40
−20
0
20
40
Dim 1 (27.57%)


PCA WT vs TG9
Dim 2 (16.05%)
15
10
5
0
−5
−10
−15
−30
−20
−10
0
10
20
30
Dim 1 (33.42%)

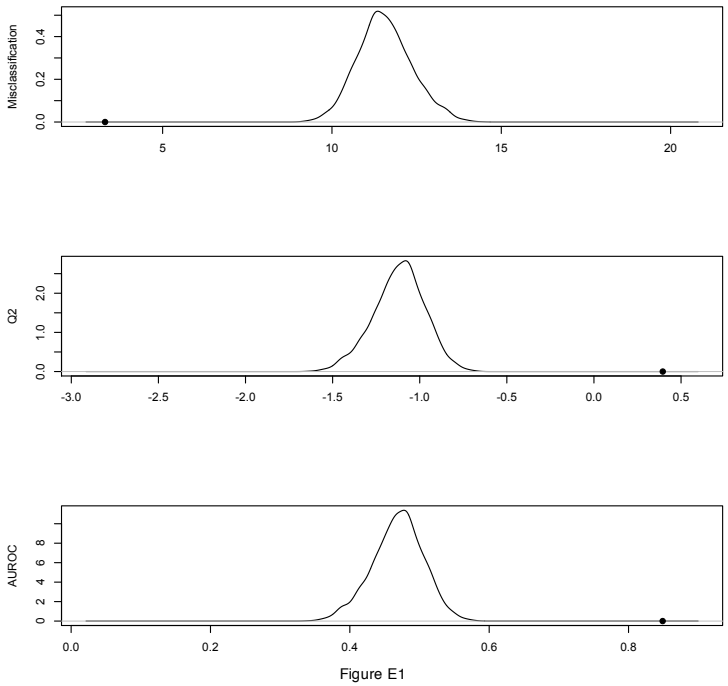


Figure E1

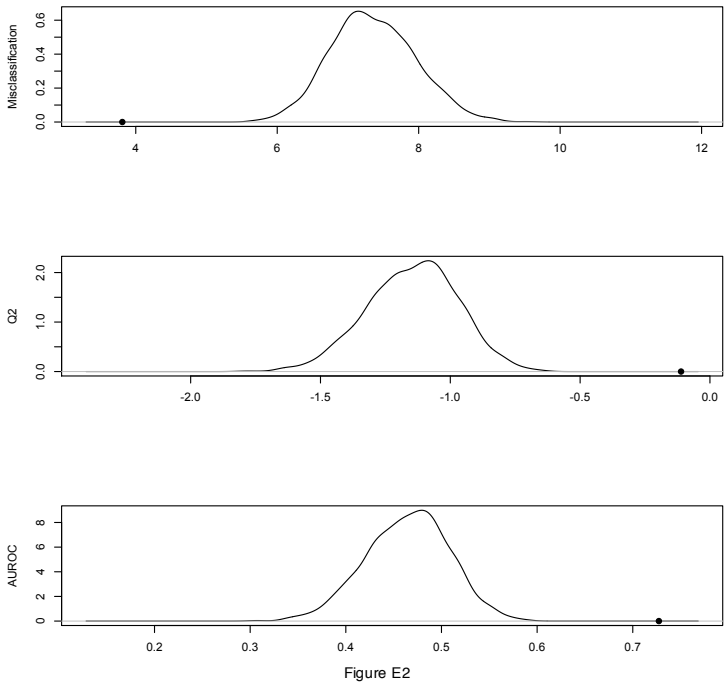


Figure E2

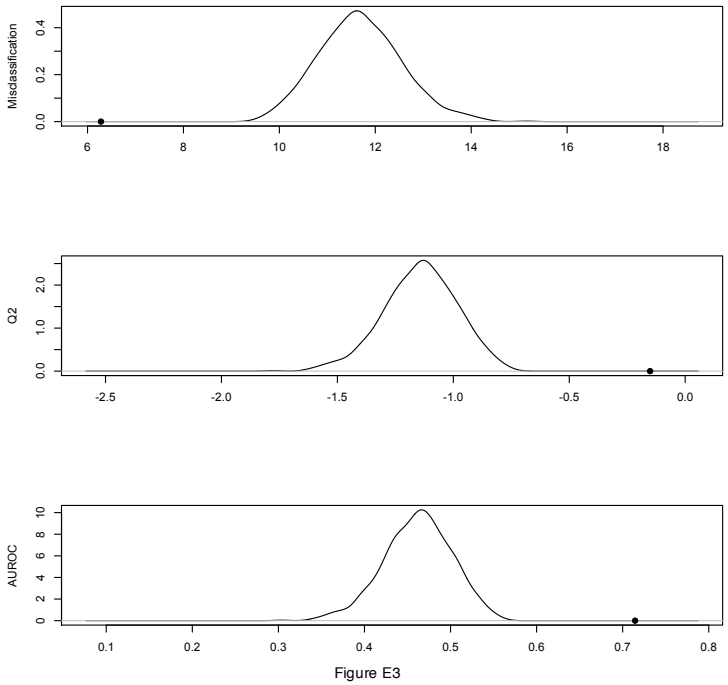


Figure E3