Laminitis in horses; NMR-based metabonomics

Jacqueline M. Dragon[1], Pedro Fernandez, Jr.[2], David Garcia[3], Daniel Kim[1], Kenith Conover[1], Pablo A. Hoijemberg[4], Michelle Coleman[5], Sarah Ralston[6], István Pelczer[1]*

[1] *Department of Chemistry, Frick Chemistry Laboratory, Princeton University, Princeton, NJ 08544, USA (*ipelczer@princeton.edu)*

[2] *Harold Washington College, City Colleges of Chicago/NSF REU Molecular Biophysics Program 2018 at Department of Chemistry, Princeton University*

[3] *Department of Environmental Sciences, Rutgers University/ NSF REU Molecular Biophysics Program 2018 at Department of Chemistry, Princeton University*

[4] *NMR Group, Centro de Invest. en Bionanociencias, CIBION-CONICET, Polo Científico Tecnológico, 1425 Ciudad Autónoma de Buenos Aires, Argentina*

[5] *College of Veterinary Medicine & Biomedical Sciences, Texas A&M University (CVM-TAMU)*

[6] *Rutgers University, Emerita*

Introduction

Laminitis is a degenerative inflammatory condition of the lamina in the hoofs of horses. It is expected to be connected to metabolic disorders. We are conducting an NMR-based metabonomics investigation to identify, and characterize the metabolic reflection of this condition, its onset, and possibility for treatment.

CVM-TAMU biobank serum samples from laminitic versus non-laminitic horses were run using simple water suppressed and also relaxation filtered experiments, and T2 measurement of the bulk water by TD-NMR. All data have been analyzed by multivariate statistics.

We have been facing a multitude of diversity issues regardless of the careful selection of samples matching age, gender, etc., as closely as possible. Yet issues remained, such as (i) geographical dispersion, (ii) variations of diet and housing, (iii) sample management, collection and storage times, (iv) extreme sample content, (v) different information from the experiments. This level of diversity can be largely avoided in laboratory arrangements, but is quite common in case of real life situations, especially for human biological samples. For meaningful multivariate statistical models and a valid metabolic interpretation, as demonstrated, judicious exclusions of selected outliers and careful inspection of individual spectra were conducted. Outliers were carefully weeded out to find reasonable consistency and acceptable statistical models. Back calculated univariate OPLS-DA coefficient plots provide information about metabolites and other components, which contribute significantly to distinction of the healthy and diseased cohorts.

The three experiments obviously present different information; the water suppressed spectrum is dominated by the large molecules (lipids, lipoproteins), the relaxation filtered version highlights the small molecule metabolites. The TD-NMR spectra tell about the bulk water's behavior mostly as the function of the protein/lipoprotein/lipid composition, and were processed using inverse Laplace Transform (iLT).

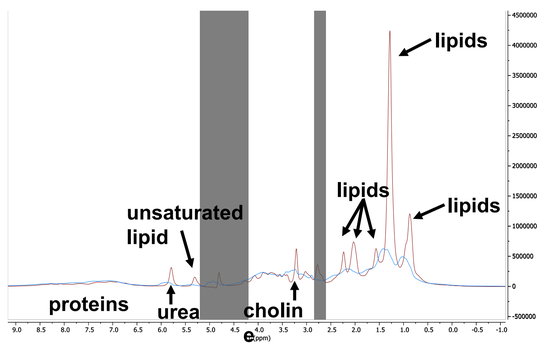


Figure 1: High-lipid outlier (red) overlaid on sum of remaining full spectra (blue). Spectra with 20Hz Gaussian apodization. Broad peaks are tentatively identified.

Experimental

Sample Preparation, data acquisition

All serum samples were stored at -80ºC, then gradually thawed to ambient temperature (22ºC) for recording the NMR spectra on a 500 MHz Bruker Avance-III instrument at 295.2K, equipped with a TCI cryoprobe. Approximately 300 uL of the sample was transferred into a 4 mm OD tube plugged with a teflon plug. No buffer was added. This tube was inserted into a 5 mm OD container tube with ca. 35 uL D_2O, which lined up between the walls to provide the lock signal.

The same 4 mm OD tube samples were inserted into the Bruker mq10 Minispec TD-NMR to measure the composite T2 decay.

Data processing

NMR spectral data processing and visualization was conducted using MNova, up to v.14.0 (MestreLab Research S.L., Santiago de Compostela, Spain). TD-NMR data were processed using iLT in MATLAB.

Multivariate statistical analysis was done in SIMCA (v.14, Umetrics, Umea, Sweden). Back calculated OPLS-DA coefficient plots have been generated by an in-house written script (PAH).

Samples had been collected in the CVM-TAMU Laminitis Project from geographically diverse regions, by different people, between 2011 to 2015. Some samples were obvious outliers just by looking (extreme lipid content, residual hemoglobin). An extreme example with tremendous lipid content is shown in Figure 1, where blinded regions (residual water and a contaminant resonance) are also shown. Blood (serum) samples are often very homogeneous and need no particular normalization. In our case it was necessary as it is illustrated in Figure 2.

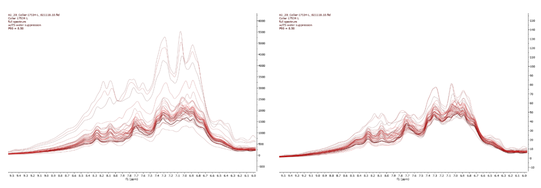
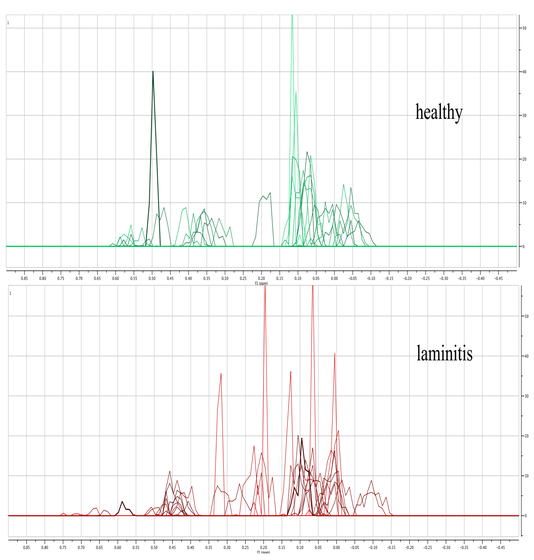


Figure 2: Non-normalized (*left*) and normalized spectra (*right*) showing protein peaks between 6.0-9.5 ppm, excluding the high-lipid outlier.



TD-NMR

T2 measurement of the serum samples (10 MHz) – only the water peak response is shown. This data is still under evaluation. The most unusual observable feature is that there are two major water peaks (T2 relaxation times) for most samples, some with additional fine structure. The healthy cohort (*top*) has more or less two distinct clusters, while the laminitis spectra (*bottom*) are much more diverse without clear clustering. The lipid peaks often have fine structure as well (not shown). All this is under further investigation.

SIMCA modelling full and relaxation filtered spectra

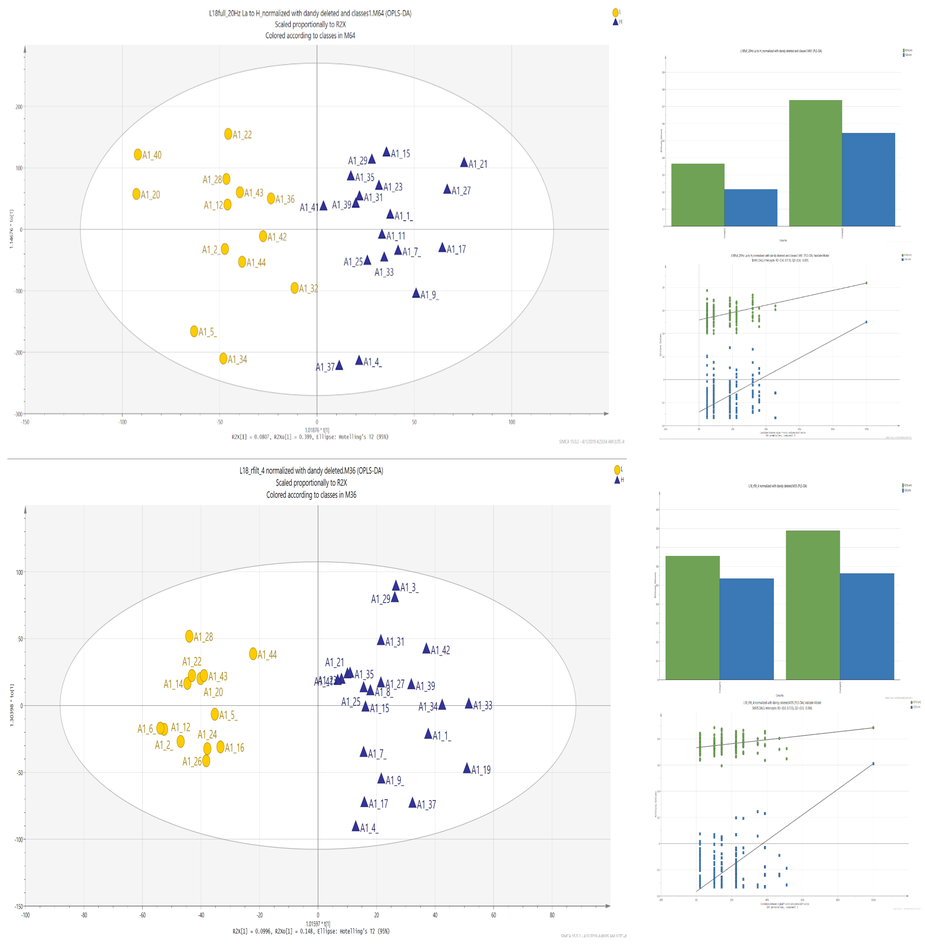


Figure 3: Outliers of the ellipse and central points were excluded, and central points were reclassified to optimize PLS-DA Q2(cum) values. A.) This OPLS-DA full spectrum model had the highest PLS-DA value seen, 0.545 for two components, and excluded ten outliers and three central points, and reclassified one point. B.) This OPLS-DA relaxation filtered spectrum model had the highest PLS-DA value seen, 0.562 for two components, and excluded eight outliers and one central points, and reclassified two points.

Back-calculated OPLS-DA coefficient plots

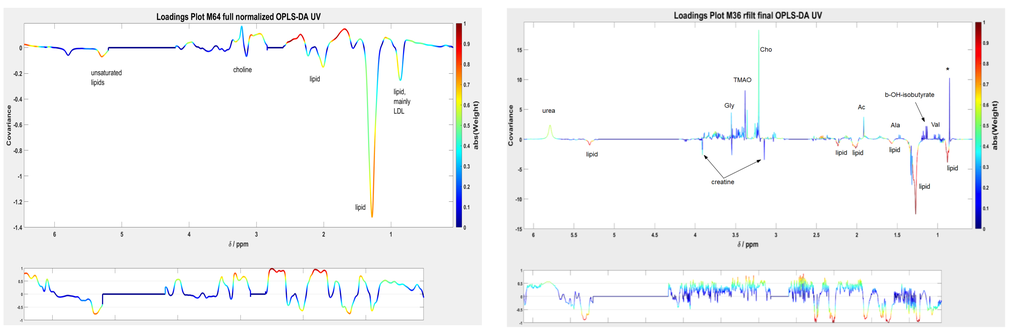


Figure 5: Coefficient plots of the normalized full spectra model with tentative assignments (protein peaks 6-8ppm not shown), and the normalized relaxation filtered spectra model, both excluding the high-lipid outlier.

Metadata & Clustering

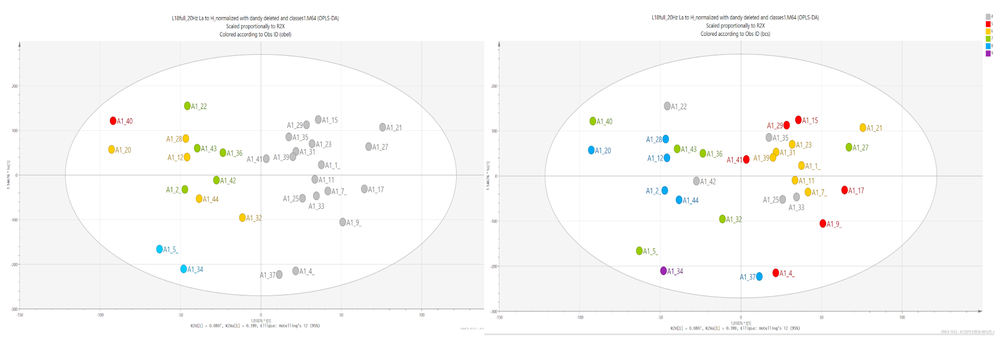


Figure 4: Metadata was used to explore clustering on the full spectra model. A.) The Obel pain scale, which is used for classifying laminitis pain, is shown. Possible clustering of level 4. B.) Body condition score of 1-9 on the Henneke scale scores the horse's condition. Clustering of 6 and 8 seems possible.

Conclusions

- Techniques to manage sample diversity include visual exclusion, and SIMCA PCA-X, PLS-DA, and OPLS-DA modelling where exclude outliers of the tolerance ellipse, exclude central points, and reclassify central points to maximize PLS-DA Q2(cum). Metadata can be used to probe clustering. Successfully reclassified samples may identify misdiagnosed individuals.

- Our SIMCA models and back-calculated coefficient plots can differentiate the significance of different isoforms of a molecule at greater resolution than is possible with only spectra. Back-calculated coefficient plots can identify the relative contribution of peaks to class determination (laminitic vs. healthy). Our results suggest lipid peaks, choline, and unsaturated lipids are the most significant.

- T2 relaxation analysis using TD-NMR is an interesting alternative approach to classify sample composition and disease conditions.

Future directions include: STOCSY (statistical total correlation spectroscopy) to determine correlations between peaks and connections in the metabolic pathway network; continued analysis with the second group of 30 samples separately and in comparison with the 44 samples.

Acknowledgements

J.D. gratefully acknowledges the Office of Undergraduate Research and the Hewlett and Godfrey Funds, the Global Health and Health Policy program in the Center for Health and Wellbeing, and the Princeton Chemistry Department and NMR Facility for their funding and support.

Bruker-Biospin (Billerica, MA) kindly provided the Minispec mq10 tabletop TD-NMR machine.

Frank Bosco (New Era Enterprises, Inc., Vineland, NJ) is our constant supplier of glassware and willing collaborator for devising new tube constructs (such as the 5/4 mm OD combination).

Umetrics (Umea, Sweden) generously offered 20 SIMCA licenses for use of our students.