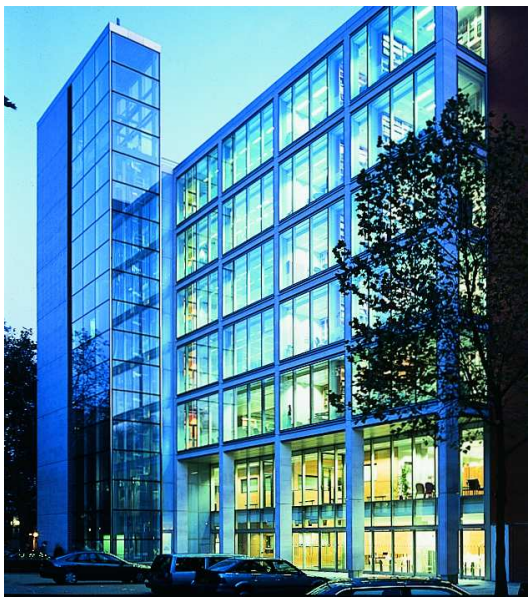
Imperial College London

Metabolic Profiling Forum

Welcome to Metabomeeting 3

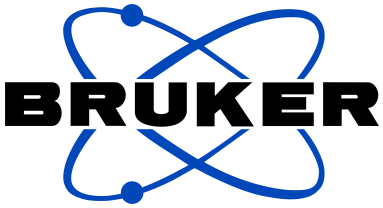
A two-day meeting on metabolic profiling, and a UK/European forum for practitioners from academia, government and industry interested in the techniques and applications of metabolic profiling (metabolomics/metabonomics). It is intended to be a friendly and informal opportunity to present and discuss the latest developments in the field, and follows on from the two highly successful meetings held at the University of Cambridge in July 2005 and January 2006.



Organizing committee:

Dr Tim Ebbels, Imperial College London
Dr Hector Keun, Imperial College London
Dr Jake Bundy, Imperial College London
Dr Andrew Nicholls (GSK & the MPF)

We are very grateful to the following organizations for their sponsorship of Metabomeeting 3:



Practical details

Location.

The conference will be located in the Sir Alexander Fleming Building. All talks will take place in Lecture Theatre 1, entrance on the mezzanine level up the stairs from the main foyer.

Posters.

Posters will be located in the rooms opposite the entrance to the lecture theatre, on the mezzanine level. They are arranged in alphabetical order of the presenting author. If you are putting up a poster, please look for your name on the poster board. Posters may be put up at any time during the day. The main poster session will be from 6 to 7.30 on Monday, but they may be viewed during coffee and lunch breaks as well, so put up your poster early to give as many people as possible a chance to view it.

Speakers.

If you have emailed us your presentation beforehand, it will be ready for you at the appropriate time. If not, please make sure you put it onto the laptop before your session, thank you.

Sponsors' displays.

The tables with the sponsors' displays are located in the mian foyer. We urge you to take the time to talk to the sponsors, without whose generous support this meeting would have been impossible.

Toilets.

These are located on the ground floor, on the left hand side if entering from the front.

Lunch and refreshments.

These will be served in the main foyer, ground floor.

Conference dinner.

This will be held at the Rembrandt Hotel, 5 minutes walk from the conference venue.

> Exit the Sir Alexander Fleming Building, turn right, and walk to the end of Imperial College Road (100m, past the traffic barriers). Turn right down Exhibition Road, and walk to the first set of traffic lights. Turn left down the Cromwell Road; the Rembrandt Hotel is 200m along this road, on the right hand side.

Web access.

If you have a laptop with a wireless card, you can access the network in the foyer of the Sir Alexander Fleming Building, using the username **conf-10220**. Full instructions are given in Appendix 1 at the end of this programme.

Pharmaceutical interest group meeting.

On Tuesday 19th, 10-11 am, there will be a meeting in room 164 (MDL1 seminar room) of the Sir Alexander Fleming Building open to representatives from the pharmaceutical industry to discuss a response to the US FDA call for guidance on establishing a Grand Round on Metabolomics. The meeting will be chaired by Dr John Haselden (Chairman, The Metabolic Profiling Forum).

Programme

All lectures will be held in Lecture Theatre 1 of the Sir Alexander Fleming Building, South Kensington campus.

Monday December 18

10.00	REGISTRATION AND COFFEE
11.00–11.10	Welcome and Introduction Dr. Tim Ebbels
Session 1	Systems Biology and Microbial (Chair: Dr. Tim Ebbels)
11.10-12.00	Top-down systems biology – from personalized healthcare to molecular epidemiology **Professor Jeremy Nicholson, Imperial College London**
12.00-12.30	What should be our expectations for cellular metabolite concentration changes? **Professor David Fell, Oxford Brookes University**
12.30-13.00	Metabolomics investigations of abiotic perturbations to microbial systems **Professor Roy Goodacre, University of Manchester**
13.00-13.30	What does the TB bacillus eat? Systems approaches to microbial metabolism **Professor Johnjoe McFadden, University of Surrey**
13.30-14.30	LUNCH
Session 2	Plant and Environmental (Chair: Dr. Jake Bundy)
14.30-15.00	Metabolomic detection of novel signalling compounds- tackling the problem of low abundance unknowns **Dr Colin Turnbull, Imperial College London**
15.00-15.30	Metabolomics of aquatic wildlife: Trials, tribulations and the need for optimised data processing **Dr Mark Viant, University of Birmingham**
15.30-16.00	Tracking the plant metabolome by NMR **Professor Mike Beale, Rothamsted Research**
16:00-16.30	TEA
Session 3	Clinical and Human Health I (Chair: Dr. Andrew Nicholls)
16.30-17.00	Metabolomics by NMR in Cancer Research **Professor John Griffiths, Cancer Research UK, Cambridge**
17.00-17.30	Metabolomics and diabetes: reaching the PPARts other `omes cannot reach **Dr Julian Griffin, University of Cambridge**
17.30-18.00	LC/MS Profiling of Lipids and Small Molecules Provides an Insight into Type II Diabetes **Dr Robert Plumb, Waters Corporation**
18.00-19.30	Poster Session and Drinks Reception
20.00-late	Conference Dinner (Rembrandt Hotel, South Kensington)

Tuesday December 19

08.30	**COFFEE**
Session 1	**Data Analysis and Metabolomics Standards (Chair: Prof. John Lindon)**
09.00-09.50	**Metabolomics Standardisation Initiative (MSI) – updates from committees** Various speakers
09.50-10.10	**Piecewise multivariate modelling of short time-series in Metabonomics** Mattias Rantalainen, Imperial College London
10.10-10.30	**DOC and DOPY: two dwarfs set to become giants of the NMR world** Leon Smith, Imperial College London
10.30-11.00	**Phenotyping using metabolite fingerprint data: model robustness, interpretability and variable significance** Professor John Draper, University of Wales, Aberystwyth
11.00-11.30	**COFFEE**
Session 2	**Analytical (Chair: Prof. Elaine Holmes)**
11.30-12.00	**Metabonomics and LC-MS: mass spectrometry not mass speculation** Professor Ian Wilson, AstraZeneca
12.00-12.30	**Targeted metabolomics – why do we need it?** Dr William Griffiths, School of Pharmacy
12.30-13.00	**An optimised strategy to increase the dynamic range of detected metabolites when using direct infusion FT-ICR-MS based metabolomics** Andrew Southam, University of Birmingham
13.00-13.20	**Metabolome profiling in steady-state *Escherichia coli* culture in response to different nutrient limitations** Dr Catherine Winder, University of Manchester
13.20-14.20	**LUNCH**
Session 3	**Clinical and Human Health II (Chair: Dr. Hector Keun)**
14.20-14.50	**Clinical Metabolomics: a measurement of ease** Dr Andrew Nicholls, GlaxoSmithKline
14.50-15.10	**Metabonomics and integrative genomics of insulin resistance in rodents** Dr Marc Dumas, Imperial College London
15.10-15.30	**Metabonomic Profiling of Gastro-oesophageal Cancer** Mr Danny Yakoub, Imperial College London
15.30-16.00	**Application of ^{1}H NMR metabolomics biopattern: Inflammatory Arthritis** Dr Aalim Weljie, Chenomx Inc, Canada
16.00-16.30	**Optimising metabolomics analyses via evolutionary computing: why and how** Professor Douglas Kell, University of Manchester
16.30-16.40	**Closing Remarks**
16.40	**CLOSE OF CONFERENCE AND TEA**

Speaker abstracts

Session 1. Systems biology and microbial

Top-down systems biology – from personalized healthcare to molecular epidemiology?

Jeremy Nicholson (Imperial College London)

What should be our expectations for cellular metabolite concentration changes?

David Fell (Oxford Brookes University)

An implicit assumption in metabolomics is that measuring changes in metabolite concentrations in different physiological states, or for different genotypes, will yield information about the causes of the differences. In order to argue from the changes back to the causes, however, it is useful to understand the types of metabolic effect produced by simple, known alterations in metabolism. In the case of changes in the activity of a single enzyme, Metabolic Control Analysis offers a theoretical framework for understanding why, usefully for metabolomics, changes in metabolite concentrations are often much larger than changes in metabolic rate. This will be illustrated with experimental results from enzyme over-expression. However, experimental data will be presented that shows that the same effect is not always seen when flux changes are induced physiologically. Here there are cases where the flux changes are relatively much larger than the changes in metabolite concentration. Theoretical mechanisms that could cause this will be considered and illustrated with a specific example.

Metabolomics investigations of abiotic perturbations to microbial systems

Felicity Currie, Emma S. Wharfe, Soyab Patel, Warwick Dunn, David Broadhurst, Roger M. Jarvis, Catherine L. Winder and Royston Goodacre (University of Manchester)

During the last twenty years the rate at which new antimicrobial agents are produced has decreased dramatically, with concomitant increase in the number of pathogens that are becoming multidrug resistant. Together these have created a patient healthcare risk and this is of great concern. A crucial aspect for the discovery of new antibiotics is the development of new techniques that allow rapid and accurate quantitative characterisation of the mode-of-action of the pharmacophore.
This increase in drug resistant microbes also reminds us that microbes are metabolically versatile organisms and respond rapidly to changes in their natural environment. The prediction and understanding of how microbes react to abiotic perturbations would be beneficial as this could allow the use of sentinel microorganisms allowing the assessment of environmental pollution, with (for example) active pharmaceutical ingredients.
This presentation will give an overview of how metabolic fingerprinting and profiling, via vibrational spectroscopy and GC-MS, can be used to assess the mode-of-action of antibiotics, how microbes adapt to API exposure, and we shall also introduce spatial metabolic fingerprinting as a means of generating detailed chemical maps from algae.

What does the TB bacillus eat? Systems approaches to microbial metabolism

Johnjoe McFadden (University of Surrey)

Metabolic networks are complex systems that function as integrated units rather than isolated parts so are best studied using a Systems Biology approach. A genome-scale metabolic model of *Mycobacterium tuberculosis* was constructed consisting of 849 unique reactions, 739 metabolites and involving 726 genes. The model includes central biosynthetic pathways and pathways for most of the mycobacterial-specific lipids and carbohydrates such as mycolic acid, arabinogalactan and phenolic glycolipid. The BCG vaccine strain of the TB bacillus was grown in continuous culture and steady state growth parameters were measured and used to calibrate the model. Substrate consumption rates were shown to correspond closely to experimentally-determined values. Predictions of gene essentiality were compared with global mutagenesis data for in vitro-grown *M. tuberculosis* and generated a prediction accuracy of 78%. Known drug targets were predicted to be essential by the model. The model successfully simulated many of the growth properties of M. tuberculosis, such as its ability to grow in microaerophilic

but not anaerobic conditions. The model was also used to interrogate global mutagenesis data for *M. tuberculosis* grown in a mouse model of infection to attempt to deconstruct the metabolic state of in vivo M. tuberculosis. The analysis indicated that virulence factors, such as phthiocerol dimycocerosates (DIM), were essential for mouse infection whereas other virulence factors, such as phenolic glycolipid, PGL, were not required. Our aim is to generate a realistic model that simulates the growth of *M. tuberculosis* inside the host and which can be used to study phenomena, such as persistence, and identify novel drug targets.

Session 2. Plant and environmental.

Towards a comprehensive low abundance metabolome

Colin Turnbull (Imperial College London)

We have attempted to develop mass spectrometry methods to enable detection of low abundance unknown metabolites in complex mixtures. Our motivation is combined genetic, molecular and physiological data that predict existence of a novel plant hormone which regulates shoot branching. The compound is probably a carotenoid derivative but in planta biochemical evidence is lacking.

Most current metabolomic methods take either 'unbiased' approaches that in fact are selective for the most abundant signals (eg NMR), or targeted approaches to profile groups of related known metabolites at high sensitivity. Using an Applied Biosystems QTrap hybrid mass spectrometer, we have instead devised a means to generate MS-MS signatures for essentially every compound in a sample regardless of abundance, down to the instrument detection limit. Capillary LC enables higher sensitivity and lower sample consumption, coupled to an electrospray (TurboIon) source and the MS operating in ion trap (enhanced product ion, EPI) mode. Central features include script routines within the Sciex Analyst software that force collection of MS-MS data for every parent mass within a list, and enable multidimensional data export. Given the constraints of peak width (~0.3 min) and instrument scan rate (4000 amu/s), a minimum cycle time of 2 s was set which allowed 12-15 parents to be scanned within a single run. By re-injecting each sample 15 times, the whole desired mass range (parents 150-508) could be covered, yielding a total of 180,000 EPI MS-MS scans per sample. Comparison of wild-type samples with mutant ones lacking the putative signal was sufficient to generate datasets for candidate discovery. Using d-labelled internal standards and signals for known endogenous plant hormones, the detection limit was estimated at 10 pg, equivalent to low ng/g levels typical of hormones in tissue. Further enrichment of putative signal compounds could be achieved prior to analysis by sampling cryodissected vascular tissue known to express the relevant carotenoid cleavage genes at high levels, and by simple fractionation according to polarity.

Datasets were processed via custom Matlab scripts to handle file sorting, chromatogram alignment, noise filtering, subdivision of chromatograms into overlapping time windows, and ratiometric comparison of MS-MS intensity values within each time window for each sample. Example outputs will be presented, and options for further method refinement will be discussed.

Metabolomics of aquatic wildlife: Trials, tribulations and the need for optimised data processing

Mark Viant (University of Birmingham)

The important role that metabolomics can play in areas such as drug discovery and chemical risk assessment has been well demonstrated. Such studies typically involve genetically similar animals that are housed in pristine conditions with a highly controlled environment and diet. Clinical metabolomics is considerably more challenging from the perspective of metabolic variability, but at least human subjects can not only give a medical history but can control their diet prior to sampling. The application of metabolomics to free living wildlife has no such advantages, and the metabolomes of these organisms can be affected by a plethora of environmental factors that are typically not controllable by the researcher. Furthermore, even the most fundamental of genotypic and phenotypic (e.g., sex, age) traits of wildlife studied using metabolomics may not be known. This presentation will highlight the challenges of working with specimens obtained from the environment, specifically the marine mussel and the 3-spined stickleback. For the invertebrate species, NMR based metabolomic studies that have led to optimised methods for field sampling will be discussed, as will the importance of knowing key genotypic and phenotypic traits in order to interpret the metabolic data. In addition, the vital contribution made by the generalised logarithmic transformation in processing the mussel data will be described. Similarly, for the fish species, which were collected from the field and then exposed to toxicants in the laboratory, the challenges of dealing with considerable metabolic variability will be discussed.

Tracking the plant metabolome by NMR

Jane L. Ward, John Baker, and Michael H. Beale (National Centre for Plant and Microbial Metabolomics, Rothamsted Research)

We will describe how the [^{1}H]-NMR-PCA screening methodology, developed in the clinical arena, has been adapted and extended for application to plant metabolomics.
The model plant *Arabidopsis thaliana* is small and has a rapid growth cycle and the availability of collections of ecotypes, knockout mutants, and transgenic lines and a fully sequenced genome makes this an ideal system for large-scale metabolomic analysis both in the context of systems biology and in gene function analysis. The ability to produce large numbers of genetically identical subjects in controlled environments leads to a distinct advantage in experimental statistical design for plant metabolomic experiments. Experimental protocols for the uniform growth and harvesting of Arabidopsis have been extensively researched and developed into a high-throughput screening operation on a 600MHz NMR instrument. Results from the analysis of mutants as well as basic studies on the variation in metabolites during development and diurnal variation in metabolites over light-dark growth cycles will be drawn on to describe how [^{1}H]-NMR-PCA provides a comprehensive fingerprint that reports rapidly on the status of the more abundant polar metabolites in the plant. The resultant fingerprint datasets can be used to classify plant lines by PCA, HCA and PLS-DA cluster analysis. Further pattern matching of PCA loadings and/or contribution plots from single-gene knockouts is showing a potential as a rapid tool for mutant screening and classification by biosynthetic pathway.

Session 3. Clinical and human health I.

Metabolomic studies on Akt and the PI-3-kinase pathway

JR Griffiths, Y-L Chung, A Schulze, SL Leevers, M Wu, T Porstmann (CR UK Cambridge Research Institute)

Akt (Protein Kinase B), a proto-oncogene implicated in several human cancers, is a downstream effector of PI-3-kinase that upregulates cell proliferation and growth, and is also involved in mediating the effects of insulin on glucose uptake, lipogenesis and synthesis of fatty acids, cholesterol etc, all via activation of mTOR. We studied (i) metabolic changes associated with Akt activation and the role of mTOR in mediating them; (ii) whether these changes could be related to Akt's role in the insulin pathway. Akt activation in human epithelial cells was controlled by tamoxifen administration to a constitutively activated Akt gene fused to the oestrogen receptor. A dual-phase extraction procedure was used to extract both lipids and water-soluble metabolites, and metabolomic studies were then performed by 1H NMR at 600MHz. Akt activation significantly increased lipid metabolites (saturated and unsaturated fatty acids, phosphoethanolamine, acetate, phosphocholine, glycerophosphocholine and phosphoglycerates), amino acids (leucine, isoleucine, valine, alanine, glutamate, glycine and tyrosine), products of glycolysis (lactate and alanine) and in [ADP+ATP]. The cells showed significantly higher uptake of leucine, isoleucine, valine, succinate, glucose and tyrosine from the culture medium, and also higher production of lactate, alanine and acetate. Inhibition of mTOR by rapamycin suppressed all these effects. These results are consistent with the putative roles of Akt and mTOR in increasing glucose transport, glycolysis and synthesis of sterols, fatty acids and proteins, all of which would tend to upregulate cell growth. In a second study on drosophila cells, insulin treatment caused very similar metabolomic effects to Akt, consistent with the role of the Akt/PI-3-K pathway in mediating the effects of insulin.

Metabolomics and diabetes: reaching the PPARts other 'omes cannot reach

Julian Griffin (University of Cambridge)

In many respects NMR based metabolomics is an ideal tool for functional genomics by combining high throughput metabolite profiling with computer assisted pattern recognition approaches for a rapid phenotyping tool [1, 2]. This versatile tool is set to increase in popularity as functional genomic approaches produce more genetically modified models for phenotyping. However, one central issue is that the sensitivity of the approach limits the detection of metabolites within tissue extracts and intact tissues to a few high concentration metabolites. Many of these metabolites are found at central hubs of metabolism [2], and a perturbation at one point in the network can be transferred to other pathways through these highly connected hubs, making them poor biomarkers for specific diseases. To expand the coverage of the metabolic pathways which are perturbed in metabolic disorders such as type II diabetes and obesity, we have examined the cross correlation of data sets from NMR metabolomics with similar data obtained using mass spectrometry, as well as proteomic and transcriptomic data sets. This has greatly aided the interpretation of the NMR derived data by better defining the initial metabolic perturba-

tions which affect the hubs of metabolism. In particular this talk will focus on a recent study to investigate the influence the peroxisome proliferator activated receptors (PPARs) have on mouse metabolism. These receptors operate as nutritional sensors for metabolism, allowing the transcriptional switch between fed and fasted states in the whole organism. Our combined 'omic' approach demonstrates the versatility of NMR based metabolomics as a tool for triage, prior to more detailed studies using mass spectrometry, transcriptomics and proteomics.

1. Griffin JL et al., Metabolic profiles of dystrophin and utrophin expression in mouse models of Duchenne Muscular dystrophy. FEBS Letts. 2002, 530(1-3), 109-16.
2. Griffin JL, Sang E. Can NMR spectroscopy contribute to proteomic analysis? in Metabolic profiling book 2002 ed. Harrigan GG & Goodacre R. "Metabolic Profiling: Its role in Biomarker Discovery and Gene Function Analysis." Kluwer Academic Publishers.

LC-MS profiling of lipids and small molecules provides an insight into Type II Diabetes

Robert Plumb (Waters Corporation)

One of the major scientific challenges of the 21st century will be determining the relationship between the human genome and the risks of developing major diseases such as cancer, diabetes, arthritis etc. Understanding the relationship between the genome, the proteome and the expressed endogenous metabolites (metabolic phenotyping or metabotyping) requires the development of ever more powerful analytical chemistry techniques and data interpretation tools. In the Metabonomics arena, this task is normally carried out by proton NMR, GC/MS or LC/MS, the later with exact mass analysis, and multivariate statistical analysis. Interesting ions in the LC/MS data are highlighted by the multivariate analysis are subsequently reanalyzed by using LC/MS/MS for structural ID. Traditionally, mass spectrometers lacked the MS/MS duty cycle required to collect product ion scan information for thousands of metabolite ions simultaneously. In this paper; however, we will describe how the application of Ultra Performance LC ™ with the simultaneous acquisition of both precursor and product ion MS data has been combined with novel Statistical heterospectroscopy, has been applied to this task of the analysis of lipids and small molecules in the Zucker rat as a model for Type II diabetes. We will show the simplicity of the methodology and the ability to identify biomarkers of disease state progression.

Session 4. Data analysis and standards

Phenotyping using metabolite fingerprint data: model robustness, interpretability and variable significance

John Draper, David Enot, Manfred Beckmann, David Overy (University of Wales, Aberystwyth)

Although metabolome fingerprints are high dimensional and contain significant levels of unwanted variance they provide useful data for 'first pass' phenotyping where high throughput is required. Classification accuracy is often excellent when analysing fingerprint data using a range of supervised multivariate analysis methods but unfortunately subsequent model interpretation can be very difficult. Using previously well characterised genotypes with predictable biochemical differences we describe a strategy to validate the interpretability potential of models generated from mass spectrometry fingerprint data. Linking the vast majority of signals in metabolome fingerprint data to their progenitor metabolites is rarely a trivial task and thus it is important to be able to validate that any highlighted variables in models with apparently good classification accuracy are truly explanatory. We describe a strategy to validate the explanatory potential of Random Forest decision tree models and also explore approaches to develop significance metrics appropriate for different types of experimental situations. One outcome is the definition of a baseline indicative of a significant difference in models involving pairwise comparisons of sample classes. Building on this information we develop a rationale for the detection of models with potentially sufficient explanatory power to guide deeper investigation of any significant metabolic phenotype. Finally, based on these model interpretability measures we suggest a strategy for the future de novo assessment of phenotype class membership in larger multi-class populations.

Piecewise multivariate modelling of short time series in metabonomics

Mattias Rantalainen, Olivier Cloarec, Johan Trygg, Jeremy Nicholson, Elaine Holmes (Imperial College London)

Modelling time series of biological systems is essential for understanding their dynamics and their responses to perturbations. In metabonomic studies the time series are commonly usually fairly short (< 20 time points), the sampling rate is many times restricted due to experimental factors of the biological systems studied. Such data have previously often been

modelled and visualised by Principal Component Analysis (PCA) as time-trajectories. Although this approach provides a general overview of the data, PCA does not explicitly model time related changes, but is rather maximising the modelled variance in the data, which is not necessarily time related.
Here a robust multivariate approach with the objective to explicitly model the time-related variation in the data is described, aiming to increase the time resolution and opportunity to interpret and predict time related events.
Piecewise multivariate models are estimated to describe time related changes between neighbouring time frames over the time series. The set of local models describe the changes over the full time series. Although linear models are used in the local time frames, the framework encompasses the description and prediction of non-linear time related changes over the full time series. Recent development in the area of chemometrics [1] has opened up for estimation of multivariate models with predictive performance identical to Partial Least Squares, but reducing the number of components to interpret to a single component for a rank one problem, which we have here. To decrease the impact on the model of potential outliers in the data, robust estimates of model parameters are employed.
Applying the proposed method for analysis of a short time-series from a NMR metabonomic study of Mercury Chloride toxicity in rat illustrates how time related changes can be successfully modelled and interpreted by the proposed method. In addition we demonstrate how time predictions can be made.
[1] Trygg, J.; Wold, S., Orthogonal projections to latent structures (OPLS). Journal of Chemometrics 2002, 16, (3), 119-128.

DOC and DOPY: two dwarfs set to become giants of the NMR world

Leon M. Smith, Anthony D. Maher, Mattias Rantalainen, Olivier Cloarec, Huiru Tang, Paul Elliott, Jeremiah Stamler, John C. Lindon, Elaine Holmes, Jeremy K. Nicholson (Imperial College London)

The characterization of endogenous and xenobiotic metabolites in complex biological matrices presents a continuing analytical challenge. We show a novel approach Diffusion Ordered Correlation (DOC) analysis to complex mixture analysis which combines Diffusion Ordered (DO) ^{1}H NMR spectroscopy with Statistical Total COrrelation SpectroscopY (STOCSY) and demonstrate the application of these methods to help characterise novel urinary biomarker metabolites and investigate the distribution of lipids in intact human blood plasma. We also show a new type of visualization tool in which the apparent diffusion coefficients from DO spectra are projected onto 1D NMR spectrum (Diffusion-Ordered Projection Spectroscopy, DOPY). Both methods either alone or in combination have the potential for general applications to any complex mixture analysis where the analytes in question have a range of diffusion coefficients.

Session 5. Analytical methods

Metabonomics and LC-MS: mass spectrometry not mass speculation

Eleni Gika, Georgios Theodoridis, Ian Wilson (AstraZeneca)

LC-MS is a powerful emerging technique for producing global metabolite profiles in metabonomics and metabolomic studies on sample types including biofluids and tissue extracts, Here the role of HPLC and UPLC-MS for the study of metabonomics in animals and man will be discussed and compared, when appropriate, with other analytical techniques such as NMR spectroscopy and GC-MS. Examples of the use of LC-MS- based strategies for the detection of biomarkers of toxicity or disease in rodents will be described. The advantages of using approaches based upon HPLC-MS include the widespread availability of the instrumentation, the utility of MS for biomarker identification and the potential for high throughput analysis. However, there are also pitfalls awaiting the unwary and metabonomics using LC-MS needs to be performed with care if valid biomarker data are to be generated. This requires rigorous control of the assay from sample collection, through sample preparation and sample analysis. Emerging experience of method validation for LC-MS-based metabonomics assays will be shared, discussed and illustrated. The likely direction of future developments in LC-MS, and how this may affect the use of the technique for metabonomics/metabolomic studies will be considered.

Targeted metabolomics: why do we need it?

William Griffiths, Yuqin Wang, Kersti Karu (School of Pharmacy)

While the goal of the "global" metabolomic approach is to observe as many metabolites as possible in a unbiased fashion, and to identify which of these metabolites change in abundance between sample groups, the idea of "targeted" metabolom-

ics is focused on a single group of metabolites. Our interest is in cholesterol metabolites in brain, which function as transport forms of cholesterol and as ligands for nuclear receptors involved in lipid homeostasis. Using a "global" metabolomic approach such metabolites are invisible, for the simple reasons that they are present at levels about 1000 - 100 000 lower than background cholesterol, and that they ionise very poorly by ESI or MALDI and decompose in GC-MS. Thus to observe these important metabolites we must employ a targeted approach. In this presentation we will discuss our approach to the identification of novel metabolites in brain.

An optimised strategy to increase the dynamic range of detected metabolites when using direct infusion FT-ICR mass spectrometry based metabolomics

Andrew D. Southam, Tristan Payne, Helen J. Cooper and Mark R. Viant (University of Birmingham)

Due to its ultra high mass resolution and mass accuracy, Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) is a particularly powerful tool for metabolomics analysis. In principle, it enables the empirical formulae of low molecular weight metabolites to be unambiguously identified based upon mass alone. Direct infusion (DI) MS analysis is ideal for high throughput metabolomics as it is more rapid than liquid chromatography MS analysis. However when analysing complex mixtures, DI MS typically suffers from ionisation suppression, arising when particular analytes preferentially ionise over less polar metabolites. Decreasing the delivery flow rate using a fully automated chip-based DI nanoelectrospray delivery system (NanoMate, Advion BioSciences) can vastly reduce ionisation suppression compared with conventional electrospray ionisation. Also, an ideal MS detector will have a high dynamic range, enabling the simultaneous measurement of both the low and high concentration metabolites in a complex mixture. Although this can be achieved using FT-ICR MS, the large numbers of ions that must enter the ICR cell experience significant space-charge effects, which degrades mass accuracy to such an extent that identification of unique empirical formulae for the peaks observed is no longer possible. Here we report a strategy for chip-based nanoelectrospray DI FT-ICR MS that effectively increases the overall dynamic range of the mass spectrum, enabling observation of both low and high concentration biological metabolites, whilst maintaining high mass accuracy. The approach is based upon the collection of multiple small m/z segments, each composed of relatively few ions to minimise space-charge effects, that are subsequently stitched together using novel algorithms. This segment-stitching technique can provide elevated metabolite detection over conventional wide scan data acquisition, which when combined with a NanoMate multi-sample delivery system can provide an ideal platform for high-throughput metabolomic analysis of complex biological mixtures.

Metabolome profiling in steady-state Escherichia coli culture in response to different nutrient limitations

Catherine L Winder, W. Dunn, S. Schuler, R. Jarvis, D. Broadhurst, G. Stephens, R. Goodacre (University of Manchester)

Traditional studies investigating the physiology of *Escherichia coli* have tended to be driven by perceived hypothesis and generally focus on specific biological pathways. By contrast, we have adopted a global metabolic profiling approach to investigate the physiological changes that occur when *E. coli* MG1655 is cultivated under different nutrient limitations.
Metabolomics requires the acquisition of reproducible, robust, reliable and biological homogenous datasets. The versatility and reproducibility of continuous culture provides a valuable tool for post-genomic studies since cellular metabolism is held in steady state conditions.
To generate reproducible metabolomics data we have developed standard operating procedures (SOPs) in order to validate the quenching and efficient extraction of metabolites from Gram-negative bacterial cells. The metabolic profiles have been generated using GC-ToF-MS. The results that shall be presented will illustrate the differences in the metabolome profiles observed in three different nutrient limitations of *E. coli*, these will including two carbon-limitations and one nitrogen limitation. We also performed continuous culture under anaerobic conditions. We have used univariate and multivariate statistical analyses to determine the most significant metabolite differences between the growth conditions.

Session 6. Clinical and human health II

Clinical metabolomics: a measure of ease

Andrew Nicholls (GlaxoSmithKline)

Although metabolomics has been used in a number of disciplines, the major area that remains to be proven is in the application to the clinical situation. Notable examples exist of the use of metabolomics for the study of kidney transplantation, pre-eclampsia, cardiovascular and vascular disease. These all represent a powerful application of metabolomics, but from the pharmaceutical viewpoint, although compelling, they provide a frustration due to the extreme (i.e. disease-related) modifications to the metabolome. Indeed, measurement of the 'extreme' has been the basis of much of the work published in the discipline of metabolomics, which in turn has lead to questions about the biological sensitivity of the method in less extreme situations.

From the pharmaceutical perspective, the use of metabolomics falls into broad camps, the application to the discovery of novel mechanisms (and hence potential drug targets), application in Phase I volunteers trials and Phase II to IV patient studies. In the former situation extreme cases are used, but patients with minimal disease development are usually focused on since they offer the best understanding from a treatment development perspective and from a longitudinal viewpoint allow for the potential to offer novel treatment as well as to monitor any amelioration of the mechanistic dysfunction. The latter situation represents the greatest complexity due to the lack of a genuine appreciation of the extent of human variability from the metabolomics viewpoint. Complications include: a definition of the precise composition of the human metabolome, the concentration variations that exist for each metabolite, the localisation and global interplay of numerous metabolites, the stability of the metabolome under various sources of normal physiological variation, the innate variation resulting from genetic differences within the human population, the metabolic consequences of one's lifestyle, the effects of the wider environment, and the effects from parasitic/symbiotic/commensal/nullistic relationships. This presentation aims to exemplify areas that need to be monitored for clinical studies, shows some results from an in-house volunteer study, whilst highlighting the projects already in progress, to assess humanity whilst at ease.

Metabonomics and integrative genomics of insulin resistance in rodents

Marc Dumas, Dominique Gauguier, Jeremy K Nicholson (Imperial College London)

The study of human multifactorial diseases like insulin resistance is a real healthcare challenge for the western and developing world.
Animal models of human disease can be used to to decipher genetic from environmental variation. In that regard, high-throughput "-omics" biotechnologies like genomics, transcriptomics and metabonomics are invaluable tools for investigating insulin resistance-related pathologies (type 2 diabetes, obesity, non-alcoholic fatty liver disease).
In this lecture, I will show how metabolic profiling can be used as a structuring tool in mathematical models to integrate genotyping and gene expression profiling data.
I will then show that integration of metabonomics with other "omics" extends the spectrum of rodent genetics to the study of transgenomic interactions between the symbiotic gut flora and the mammalian host.

Metabonomic profiling of gastro-oesophageal cancer

Danny Yakoub (Imperial College London)

Gastro-oesophageal cancer is a highly lethal disease with a 5-year mortality >80%. The shortcomings in current diagnosis and staging tools result in unreliable patient selection for multimodality treatment and uncertainty in the assessment of their outcome. Furthermore, much attention has recently been devoted to predicting progression from intestinal metaplasia (Barrett's disease) to dysplasia and low-grade to high-grade dysplasia/ adenocarcinoma with relatively disappointing results in terms of identifying useful markers for daily practice. Recent work in magnetic resonance spectroscopy techniques and multivariate data analysis has shown that the pathology of biological samples can be characterised by metabolic profile data. Altered metabolism in tumours is already exploited diagnostically by techniques such as PET and profiling metabolism ('metabonomics') is means by which new metabolite biomarkers can be identified and validated. 600 MHz ^{1}H NMR spectroscopy was used to analyse biofluid samples & intact tissue from patients with and without gastro-oesophageal cancers. Based on their metabolic profiles, NMR spectroscopy of tissue samples correctly classify them according to their tissue of origin,

presence of malignancy, pathological types and chemotherapy status. NMR spectroscopy of serum samples could differentiate between normal and malignant samples and between pre and post chemotherapy samples. The presence of malignancy affected predominantly intermediates of carbohydrate metabolism, while samples were differentiated according to chemotherapy status based on changes in lipid signals. The data presented in this report demonstrates the potential of NMR-based metabonomics in the diagnosis and management of gastro-oesophageal cancers.

Application of a ^{1}H NMR metabolomics biopattern: inflammatory arthritis

Aalim M. Weljie, Reza Dowlatabadi, B. Joan Miller, Hans J. Vogel, Frank R. Jirik (University of Calgary / Chenomx Inc.)

Rheumatoid arthritis is a debilitating, systemic inflammatory joint disease impacting 1-2% of the population, accompanied by alterations in specific metabolites. As an initial approach to investigating this possibility in a well-defined system we selected a murine model of inflammatory arthritis, the K/BxN mouse. Sera from arthritic populations of K/BxN mice with a genetic predisposition to arthritis (N=15), as well as healthy parent strain population (N=19) were ultra-filtered and subsequently analyzed using 1H NMR spectroscopy. A "Targeted Profiling" approach [1] was used to identify and quantify 59 metabolites, and track 29 unknown resonances for >95% spectral coverage. Subsequent multivariate analysis was performed to build a highly predictive model based on cross-validation A highly significant subset of 18 spectral features (15 known compounds and 3 unknown resonances) was identified (p=0.00075 using MANOVA analysis), which we term a "biopattern" or "bioprofile". Specific markers were identified relating to nucleotide synthesis, fatty acid metabolism, protein synthesis, inflammation, and methylation and were not influenced by gender. Interestingly, the most significant biomarker, uracil (p=2.4 x 10-5), was identified in spite of overlap with the intense urea peak, hampering identification using traditional approaches. The results attest not only to the complexity of systemic inflammatory responses, but also the power of the experimental approach in being able to reveal such a wide variety of markers. The implication of monitoring a spectrum of metabolic events simultaneously is discussed with respect to potential for new treatments and individualized medicine in light of other biomarkers revealed from ongoing studies of human rheumatologic disease. (Supported by Genome Alberta/Canada; and the Arthritis Society of Canada)

Optimising metabolomics analyses via evolutionary computing: why and how?

Douglas Kell, Steve O'Hagan, Warwick B. Dunn, David Broadhurst, Marie Brown & Joshua D. Knowles (University of Manchester)

The optimisation of scientific instrumentation involves many settings or parameters, and makes this a combinatorial optimisation problem of high dimensionality. Heuristic methods [1] such as those based on genetic search are appropriate for tackling such problems [2], while closed loop methods [3] make these still-large search spaces tractable. Because there are several things to optimize (such as number of peaks, run time and signal:noise), multi-onbjective methods are appropriate [4; 5]. We have combined these ideas and developed closed-loop optimisation methods for metabolomic instrumentation [6; 7]. Such strategies have served to increase tenfold the number of metabolite peaks we can detect, leading (when coupled to appropriate numerical techniques [8] to considerable benefits for biomarker detection and other metabolomic purposes [9; 10].

[1] Corne, D., Dorigo, M. & Glover, F. (1999). New ideas in optimization. McGraw Hill, London.
[2] Vaidyanathan, S., Broadhurst, D. I., Kell, D. B. & Goodacre, R. (2003). Explanatory optimisation of protein mass spectrometry via genetic search. *Anal. Chem.* 75, 6679-6686.
[3] King, R. D., Whelan, K. E., Jones, F. M., Reiser, P. G. K., Bryant, C. H., Muggleton, S. H., Kell, D. B. & Oliver, S. G. (2004). Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* 427, 247-252.
[4] Corne, D., Jerram, N. R., Knowles, J. & Oates, M. (2001). PESA-II: Region-based selection in evolutionary multiobjective optimization. In *GECCO 2001: Proceedings of the Genetic and Evolutionary Computation Conference* (ed. L. Spector, E. D. Goodman, A. Wu, W. B. Langdon, H.-M. Voigt, M. Gen, S. Sen, M. Dorigo, S. Pezeshk, M. H. Garzon and E. Burke), PP. 283-290. Morgan Kaufmann, San Francisco, CA.
[5] Knowles, J. (2006). ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Trans Evol Comput* 10, 50-66.
[6] O'Hagan, S., Dunn, W. B., Brown, M., Knowles, J. D. & Kell, D. B. (2005). Closed-loop, multiobjective optimisation of analytical instrumentation: gas-chromatography-time-of-flight mass spectrometry of the metabolomes of human serum and of yeast fermentations. *Anal Chem* 77, 290-303.
[7] O'Hagan, S., Dunn, W. B., Broadhurst, D., Williams, R., Ashworth, J. A., Cameron, M., Knowles, J. & Kell, D. B. (2006). Closed-loop, multi-objective optimisation of two-dimensional gas chromatography (GCxGC-tof-MS) for serum metabolomics. *Anal Chem*, in the press.
[8] Broadhurst, D. & Kell, D. B. (2006). Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics*, in press; DOI: 10.1007/s11306-006-0037-z.
[9] Kenny, L. C., Dunn, W. B., Ellis, D. I., Myers, J., Baker, P. N., The GOPEC Consortium & Kell, D. B. (2005). Novel biomarkers for pre-eclampsia detected using metabolomics and machine learning. *Metabolomics* 1, 227-234 - online DOI: 10.1007/s11306-005-0003-1.
[10] Kell, D. B. (2006). Systems biology, metabolic modelling and metabolomics in drug discovery and development. *Drug Disc Today* 11, 1085-1092.

Poster abstracts

A FUNCTIONAL GENOMIC APPROACH TO DEFINE THE INFLUENCE OF PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS ON THE MOUSE

Helen J. Atherton, Melanie K. Gulston, Kieran Clarke, Nigel J. Bailey and Julian L. Griffin (University Of Cambridge)

The metabolic syndrome describes the coexistence of a number of cardiovascular risk factors including obesity, hyperglycaemia, type II diabetes mellitus, dislipidaemia and hypertension. The peroxisome proliferator activated receptors (PPARs) are a group of ligand activated transcription factors important in lipid metabolism and energy homeostasis. Three isoforms have been identified (PPAR-α, PPAR-δ and PPAR-γ), where PPAR-α agonists (e.g. fibrates) lower levels of circulating triglycerides, increase HDL cholesterol and reduce blood pressure, and PPAR-γ agonists (e.g. thiozolidinediones) ameliorate insulin resistance and induce adipocyte differentiation.
In this study the role of PPAR-α in the heart, has been investigated in the PPAR-α null mouse using a combined metabolomic and transcriptomic approach. Metabolism was monitored by a combination of NMR, GC-MS and LC-MS analyses, in conjunction with multivariate statistics. At a range of time points between 1 and 13 months there was a profound decrease in glucose and a number of amino acids including glutamine and alanine, and an increase in lactate demonstrating that a failure to express PPAR-α results in perturbations in glycolysis, gluconeogenesis and the citric acid cycle. As PPAR-α is known to control the expression of a number of enzymes in the β-oxidation pathway, the observed changes may be a consequence of perturbed fatty acid metabolism. To establish such a link, the metabolic changes were cross-correlated with transcriptional changes as measured by RT-PCR. These experiments showed loss of PPAR-α function caused a compensatory increase in PPAR-γ which in turn increased the expression of a number of genes associated with fatty acid metabolism including uncoupling protein 2 (Ucp2), which increases fatty acid utilisation, and glycerol-3-phosphate dehydrogenase 1 (Gpd1) which is required for lipogenesis and adipogenesis. In an additional study, high resolution- magic angle spinning ^{1}H-NMR spectroscopy was used to simultaneously analyse and quantify concentrations of both small molecule metabolites and lipids in intact tissue samples. The results show cardiac samples from PPAR-α null mice contain increased saturated fatty acids, and decreased unsaturated fatty acids relative to control samples, which further suggest that fatty acid metabolism is severely perturbed in the absence of PPAR-α. In conclusion, loss of PPAR-α function leads to perturbations in fatty acid metabolism, as well as glycolysis, gluconeogenesis and the citric acid cycle. The loss of PPAR-α function is in part compensated for by an increase in PPAR-γ and other enzymes that increase β-oxidation of fatty acids in cardiac tissue.

THE COMPARISON OF LIPID EXTRACTION METHODS FOR THE 'LIPIDOMIC' ANALYSIS OF BLOOD PLASMA AND LIVER TISSUE

Ali Awwad, Florian Wulfert, Dave Barrett, Clare A. Daykin (University of Nottingham)

In recent years, the concepts of metabolomics and metabonomics have demonstrated enormous potential in furthering the understanding of disease processes, toxicological processes and the phenotypic outcome of gene expression. However, although metabolomics is and should remain an integrated approach by itself, its complexity requires analogous integrated approaches focused on its components such as lipidomics. Lipidomics is an emerging technology and can be defined as 'systems-level analysis of lipids and factors that interact with lipids'.
Whilst lipids can be studied in whole blood plasma and liver using various NMR spectroscopy-based spectral editing methods without sample extraction, if these methods are used in isolation, valuable information may be lost. Hence, efficient and reproducible extraction of lipids from blood plasma and liver is critical. In the literature, many different protocols for lipid extraction are reported, however; there is a severe lack of data comparing their efficiency and reproducibility. This study has therefore investigated the variations among different methods by extracting lipids from plasma and liver tissues. The results presented here show that these lipid extraction methods vary considerably in the recovery of lipids and whilst the chloroform/methanol based Bligh and Dyer extraction method appears to be the most reproducible method, it requires large amounts of sample and is known to underestimate lipid concentrations in samples where lipids constitute > 2% of the sample. On the other hand, the Folch method allows the detection of more lipids and at higher concentrations than other methods whilst retaining the advantage of relatively good reproducibility.

^{1}H NMR METABOLIC PROFILING FOR BLOOD PLASMA BIOMARKERS OF CERVICAL CANCER

Alexandra Backshall, Hector Keun (Imperial College London)

The aim of this study was to determine if NMR-detectable metabolites could be used as a marker of disease status in patients suffering with either cervical intraepithelial neoplasia (CIN) or cervical cancer. ^{1}H NMR spectroscopy of blood plasma from 76 women (27 with CIN, 34 with cancer, 15 relapse) was analysed to identify metabolites that show a difference between the disease groups and the disease stages as defined by the CIN and the International federation of gynaecologists and obstetricians (FIGO) clinical staging systems. Ketone body concentrations are shown to be increased in patients with malignant disease whereas choline, and cholesterol dominated lipoprotein/lipid signals show depletion. This indicates that there is an NMR-detectable relationship between the metabolites that are indicative of nutritional status and the stage of cervical disease. Cancer cachexia-anorexia syndrome is a candidate for the changes in metabolism that have been observed. ^{1}H NMR detection of these metabolic altera-

tions has the potential to aid cervical cancer disease staging and management of treatment.

METABOLOMIC INVESTIGATION OF THE BURKHOLDERIA CEPACIA COMPLEX

Volker Behrends, Huw Williams, Jake Bundy (Imperial College London)

The bacteria of the *Burkholderia cepacia* complex (Bcc) are a group of at least nine closely-related species. Bcc members are both commercially interesting because of their metabolic capacity and of clinical concern as some strains are among the most prominent causes of death for cystic fibrosis (CF) patients. The pathogenicity of Bcc bacteria is complex and multi-factorial, and is strain- rather than species dependent. The mechanisms of pathogenesis are not fully understood.
To investigate metabolic differences of a set of Bcc strains, the strains were grown under different conditions and subjected to exo- and endometabolome analysis using NMR spectroscopy. Exometabolome analysis revealed a strict carbon source dependent channelling of metabolism on minimal media whereas species-independent metabolic production and consumption patterns were observed when the cells were grown on rich medium. Endometabolome analysis was performed with strains of two Bcc species (*B. cepacia* and *B. cenocepacia*) grown in liquid cultures to either mid-log or stationary phase.
Our initial results have shown that there are clear and reproducible metabolic differences between individual isolates, but it is not yet clear what the causes of these differences are. The metabolomic data do not simply cluster the samples into their different species. In our future work we will aim to investigate what physiological factors are causing the observed differences – in particular, whether metabolomics can tell us something about pathogenic potential.

FIXED-WIDTH BINNING, VARIABLE-WIDTH BINNING OR NO BINNING: A STUDY OF DIFFERENT BINNING METHODS IN NMR-BASED METABOLOMICS ANALYSIS

Yann Bidault, Chen Peng, Gregory Banik, Scott Ramos, Brian Rohrback, Ian (Bio-Rad Laboratories, Inc.)

In multivariate analysis of NMR-based metabolomics data, small variations in the resonance position of the individual peaks caused by experimental and instrument induced variations can adversely impact the PCA results. Techniques to address this NMR peak misalignment issue commonly include binning or bucketing and employing more advanced algorithms that aim at shifting the individual peaks to reach a better alignment across the spectra.
Fixed-width binning (usually at 0.04 ppm) is used to alleviate the impact of the peak misalignment by averaging the data points that fall inside the bin width. However, since this drastically reduces the data resolution, it becomes more difficult to interpret the PCA results.
We have recently developed the following informatics tools to tackle the global and local peak misalignment problems:

- Global spectra alignment for multivariate analysis using full-resolution without binning.
- Fixed width binning, with graphical interface for the researcher to visualize and manually adjust the bins on top of the original spectra.
- Technology for variable-width binning by which the width of a bin is automatically adjusted in order to place bin boundaries at the local minima in an overlap density consensus spectrum.
- Automatic Filtering of NMR Spectra (AFNS), a novel method that uses a rolling binning algorithm, multiple bin widths, and ANOVA-based filtering as a means of identifying significant features in complex spectra.

In this study, the pros and cons of each of these methods are compared and discussed using two spectral datasets, one with and one without local chemical shift variations.

BIAS, OVERFITTING, SAMPLE SIZE, AND OTHER DANGERS IN THE DESIGN & ANALYSIS OF METABOLOMICS EXPERIMENTS

David Broadhurst, Douglas Kell (University of Manchester)

Many metabolomic experiments are designed as binary discrimination studies such that their primary aim is the discovery of biomarker metabolites that can classify, with a certain level of certainty, between 'case' and 'control' conditions. For example, in human metabolomics this discrimination could be between individuals suffering from (or likely to get) a particular disease. Experience has shown that, even with a nominally matched control group, it is very easy to find persuasive but entirely spurious markers in this way. There are several contributing factors to this problem including: inadequate sample size, in which case it is always easy to find random multivariate correlations when the number of variables greatly exceeds the number of samples; ignorance of Type I statistical errors, i.e. high false discovery rate due to massively parallel significance tests on single peaks, or combinations thereof; hidden bias, particularly prevalent in human metabolomics, where confounding factors, including sampling, transport, storage and analysis are distributed differentially between 'case' and 'control' populations; overfitting, where inadequate model validation is implemented, typically by failing to use independent/blind 'test' samples which are held back from model optimization and simply used to test the robustness of prediction in the final phase of the study; myopic approach to model selection, in which the wrong statistical test, or multivariate model is chosen for a given study, either due to lack of theoretical understanding, or familiarity with 'favourite' algorithms.
We draw attention to these problems by reviewing key standard statistical literature, together some recent papers addressing these problems. Few published papers in the metabolomics literature withstand this scrutiny.

THE INTROGRESSION OF NOVEL BIOCHEMICAL TRAITS INTO TOMATO.

Robert Coe (Sheffield University)

Wild species are a good source of genetic variation by with which to improve the agronomic value of cultivates crops, but traditional breeding techniques are time consuming and often unsuccessful. Populations of introgression lines carrying wild species alleles afford an opportunity to identify traits associated with the introgressed regions, and facilitate characterisation of the biochemistry and genetics underlying these phenotypes.

Metabolite profiling provides a snapshot of the metabolome in samples which differ in a known factor such as genetic background. Differences between the metabolite profiles can identify those metabolites and hence metabolic pathways affected by the introgression of defined genes or genomic regions and allow genetic maps for metabolic alterations to be established. Whilst also providing a much more sensitive analysis of the introgression lines.

We are using the genetically well characterized family of introgression lines of *L. esculentum* and *L. pennellii* to identify metabolic traits associated with the introgressions. We are also using a second family (*L. esculentum* and *L. hirsutumi*) to test whether the phenotype observed was dependant on the introgressed DNA or the silencing of *L. esculentum* genes.

A major problem with this type of analysis is the identification and statistical validification of metabolic differences, especially in very large multivariate data sets. Therefore the data generated is being analysed by a range of chemometric approaches ranging from PCA to supervised machine learning algorithms.

A METABOLOMICS APPROACH TO ASSESS *IN VIVO* HEPATIC INDUCTION IN THE RAT

Vincent Croixmarie, E Werner, T Umbdenstock, F Dartiguelongue, C Boursier-Neyret, B Walther (Technologie Servier)

Metabolomic approaches provide an essential tool to optimize early metabolism and safety studies. Our prime objective in this work is to identify indirect markers of induction enabling to classify compounds in regards the mechanism of induction but also to identify possible cross-talks between endogenous pathways triggered by the administration of xenobiotics.

Well-known xenobiotic inducers such as Phenobarbital (PB), isoniazid (INZ), 3-methylcholanthrene (3MC), β-naphtoflavone (BNF) and pregnenolone 16 α-carbonitrile (PCN) were selected for their induction capabilities towards CYP families. Treatment is realised for 5 days and each group, including vehicle, is composed of 20 rats.

The collected bio fluids (serum and urine) have been analyzed using Time of Flight Mass Spectroscopy (ToF-MS) on samples separated by Ultra-Performance Liquid Chromatography (UPLC).

The raw data were treated with Markerlynx for peak extractions and the multivariate analysis have been carried out with SIMCA-P software using PCA for classification purpose and O-PLS or PLS-DA for group separation and identification of specific biomarkers (discriminant and common) linked with each mechanism of induction. Special routines developed for data reduction (exclusion of metabolites and adducts) will also be discussed.

DISCRIMINATION OF PHARMACEUTICAL EXPOSURE MEASURED BY FT-IR AND GC-MS OF PSEUDOMONAS PUTIDA KT2440 (UWC1)

Felicity Currie, Warwick B. Dunn, David Broadhurst, Royston Goodacre (University of Manchester)

Human pharmaceuticals have been readily detected in the UK in waste water treatment plants (WWTPs), rivers and estuaries at the ng L-1 or low μg L-1 level. Although levels are not currently thought to be high enough to cause immediate harm to aquatic life, it is widely acknowledged that there is insufficient information available to determine whether exposure to low levels of these substances over long periods of time is having an impact on the environment.

In this study we used metabolite fingerprinting and profiling approaches of FT-IR and GC-MS coupled with univariate and multivariate analysis techniques to discriminate between cultures of *P. putida* exposed to six pharmaceuticals. While little effect was observed for exposure to four of the drugs, metabolic effects were observed for cells exposed to Propranolol and Ibuprofen. Univariate and multivariate analyses were also used to identify endogenous metabolites contributing to discrimination between cells exposed to Propranolol and Ibuprofen and the other four drugs.

NOVEL APPROACHES FOR HANDLING AND VISUALISATION OF NMR METABOLOMIC DATA.

James Donarski, A J Charlton, J Godward, R Stones, J Wilson, R Davis, S Poulding (Central Science Laboratory)

NMR spectroscopy is a powerful technique that is increasingly being used for metabolomic studies. Large libraries of 1D ^{1}H spectra are often produced which are then examined using uni- and multivariate statistical analysis tools to decipher categorical variation between sample sets.

To aid NMR metabolomic investigation we have developed a suite of methods and software to aid analysis. These include:

- Adaptive binning routines utilising the undecimated wavelet transform for data reduction, peak picking and alignment, baseline correction and noise reduction.
- Two stage genetic programming algorithms(1) enhanced for spectral feature selection.
- Metabolab, a software application for statistical analysis (univariate and multivariate), and data visualisation.

- A database application for compound identification.
- An algorithm to reduce T1 noise in 2D spectra to aid automated spectral assignment.

In this poster we will present these methods and their utilisation in the detection of biomarkers.

CLOSED-LOOP, MULTIOBJECTIVE OPTIMISATION OF TWO-DIMENSIONAL GAS CHROMATOGRAPHY-MASS SPECTROMETRY (GCXGC-TOF-MS) FOR SERUM METABOLOMICS

Warwick B. Dunn, S. O' Hagan, J. Knowles, D. Broadhurst, R. Williams, J. Ashworth, D.B. Kell (University of Manchester)

In metabolomic studies we aim to detect and quantify all metabolites within a biological system to obtain the maximum information achievable. Analytical systems providing high chromatographic resolving powers are the most appropriate for studying the 100-1000s of metabolites in any one sample. The optimization of such analytical methods and instrumentation is essential to ensure these large metabolite numbers can be detected routinely. The presentation will describe the extension of the closed-loop (iterative, automated) optimisation system that we had previously developed for 1-dimensional GC-tof-MS (O'Hagan et al, Anal. Chem 77, 290-303 (2005)) to comprehensive two-dimensional (GCxGC) chromatography-mass spectrometry. The heuristic approach used was a multi-objective version of the Efficient Global Optimisation algorithm. In just 300 automated runs we improved the number of metabolites observable relative to those in 1D GC by some 3-fold. The optimised conditions allowed for the detection of over 4000 raw peaks, of which some 1800 were considered to be real metabolite peaks and not impurities or peaks with a signal:noise of less than 5. Data were examined statistically to understand the basis of these improvements.

THE USE OF DIRECTED GRAPHS IN THE REPRESENTATION OF METABOLIC NETWORKS

John M. Easton, Mark R. Viant, Andrew C. Peet, Theodoros N. Arvanitis (University of Birmingham)

Metabolic networks have become a commonly used analysis technique in bioinformatics, offering a complementary approach to the traditional biological method for studying metabolism (the metabolic pathway). While the techniques used for the construction and analysis of metabolic networks have been inspired and informed by well-established tools and metrics from fields such as physics and communications engineering (where networks have been studied for many years), the important question of how best to construct them from field-specific experimental data and represent the associated biological knowledge remains largely unanswered. Correlation-based techniques offer a quick, easy and largely unbiased way of discovering the network topology but fail to provide the directional information required for anything other than a cursory analysis of the system. Methods based on the layering of metabolite data onto the pre-existing metabolic pathway diagrams found in repositories, such as KEGG, can force a modular structure onto the resultant network, a process which may mask other important topological features hidden within the data. In this work, we present a method of building directed metabolic networks from experimental datasets which allows a priori information on known metabolic reactions to be included, while at the same time displaying the dynamic architectural discovery that can be seen in correlation-based networks. The approach uses an enhanced version of the reaction component of the KEGG LIGAND database, which contains not only definitions for each reaction but also pathway specific directional information which would allow the networks produced to be tailored to specific areas of metabolism should such information be available. The resulting networks are directed graphs, derived from the association between array probes and enzymes in the KEGG database. The directionality of the network is determined by the flow of catalytic reactions within the associated pathways. We demonstrate this approach on existing data produced from a study of circulating tumour cells (CTCs) by Smirnov et al. (2005), which consists of microarray data for three different types of cancers: breast, colorectal and prostate. We show the novelty of this method in the sense that the derived directed metabolic networks are produced without the bias introduced by mapping metabolite data onto the pre-existing topology of a pathway, while demonstrating its usefulness by supporting related biochemical hypotheses.

FAST 2D NMR METHODS IN METABOLOMICS

C. Ludwig, D. A. Emwas, S. Tiziani, D. Ward, C. Bunce, P. Johnson, M. Wakelam, M. R. Viant, U. Günther (CR UK Institute for Cancer Studies, University of Birmingham)

Sensitivity and overlap of signals in one-dimensional spectra represent the largest limitations for NMR based metabolomics. We have recently explored different two-dimensional NMR methods to reduce signal overlap in NMR spectra. J-resolved NMR spectra have been used to deconvolute spectra by eliminating scalar couplings in the direct dimension. Using new methods to analyse the spectra the result of principle component analysis (PCA) can be significantly improved. Here J-resolved spectra were used to obtain a signature for rare head and neck cancers and to analyse cancer cell lines treated with different protein inhibitors.

Two-dimensional Hadamard-encoded TOCSY spectra represent another method to deconvolute chemical shift degeneracies in biological samples at a very modest increase of recording time. By encoding selected frequencies in the spectrum spin systems representing individual metabolites are obtained as subspectra with virtually no overlap between signals. By applying this approach to serum samples from colon and cancer patients we were able to differentiate between samples from patients with different forms of colon cancer and healthy control samples. For a

subset of samples it was also possible to differentiate between different stages of colon cancer. This result is of particular importance as there is currently no good diagnostic method for colon cancer.

DIRECT SAMPLING OF ORGANISMS FROM THE FIELD IS A VIABLE APPROACH FOR METABOLOMICS-BASED ENVIRONMENTAL MONITORING

Adam Hines (University of Birmingham)

NMR-based metabolomics has been used extensively in mammalian toxicology studies but rarely in ecotoxicology. Critical questions must be addressed to evaluate its potential for studying free-living wildlife, e.g. can metabolomics identify stress-induced phenotypes outside of a controlled laboratory in animals whose normal phenotype is largely uncharacterized? And to what extent is knowledge of organism genotype (e.g. species) and phenotype (e.g. sex, age) required to be able to adequately interpret metabolomics data? To address these questions we have characterized the natural variability of the metabolome in the most widely studied marine bivalve, the mussel, and determined if this variability masks the metabolic response to a hypoxic stress. Specifically, we have compared the NMR metabolic fingerprints of four groups of mussels from Port Quin, UK – those collected directly from the field with and without a 2-hr hypoxia stress, and those collected and then stabilized in a marine laboratory for 2 days, also with and without a 2-hr hypoxic stress immediately prior to dissection. Although we hypothesized that laboratory stabilization would reduce overall metabolic variability, the variability was similar between the four groups and the effects of hypoxia were significantly more discernible in mussels directly sampled from the field. NMR metabolomics data was recorded from two different tissue types, adductor muscle and mantle. The overwhelming metabolic variability in mantle tissue was found to be due to sex based differences and a number of sex dependent metabolic changes were recorded. The results from an additional study investigating the metabolic profiles of mussels from two different sites in the UK over a 12 month period will also be discussed. Overall our recommendations include that direct sampling from the field is the preferred method for metabolomics as it allows stress-induced phenotypic changes to be considerably more readily detected. Furthermore, we recommend that fundamental genotypic and phenotypic traits of the study organism must be known for meaningful interpretation of the metabolic measurements.

METABOLITE REGULATION OF GENE EXPRESSION DURING THE HETEROTROPHIC TO AUTOTROPHIC TRANSITION IN DEVELOPING SEEDLING OF ARABIDOPSIS

Elizabeth Allen, Annick Moing, Timothy M. Ebbels, Mickael Marcourt, A. Deri Tomos, Dominique Rolin & Mark A. Hooks (University of Wales Bangor)

Seed germination and development represents a unique stage of plant development where metabolic programs that are geared initially towards catabolism of stored carbon reserves (heterotrophy) undergo a transition to permit fixation of inorganic carbon into organic compounds by photosynthesis (autotrophy). This transition involves a transcriptional reprogramming to dismantle catabolic and produce the photosynthetic machinery. Because this transcriptional programming revolves around metabolic changes, a key question is the role of metabolic signals in regulating this process. Expanding on our work on acetate regulation of gene expression during seedling development, we are taking a holistic approach to identify other potential signalling metabolites. Using 1H-NMR, we have profiled a group of metabolites in samples from imbibed seeds to seedlings 8 days after imbibition. We have also profiled expressed genes at corresponding time points using microarrays. We have calculated correlations among differentially expressed genes and metabolite levels in order to determine potential metabolic signals for future study. Interestingly, Spring Embedding models of metabolite and gene expression networks have shown that metabolite profiles from day 2 samples group with profiles from day 0 (imbibed seeds) and day 1, but that gene expression profiles from day 2 samples group with those from day 3 to day 8. This suggests that metabolic programming of embryos within seeds establish a pattern that may precede reprogramming of gene expression.

PYCHEM – A MULTIVARIATE ANALYSIS PACKAGE FOR PYTHON

Roger M. Jarvis, David Broadhurst, Helen Johnson, Noel O'Boyle and Royston Goodacre (University of Manchester)

We have implemented a multivariate statistical analysis toolbox, with an optional standalone graphical user interface (GUI), using the Python scripting language. This is a free and open source project that addresses the need for a multivariate analysis toolbox in Python. In contrast to tools like MATLAB, PyChem[1] is easily accessible and free, allows for rapid extension using a range of Python modules, and is part of the growing amount of complementary and interoperable scientific software in Python based upon SciPy. One of the main aims of PyChem is to continually evolve a user-friendly platform that has applicability across a wide range of analytical and post-genomic disciplines.

The software includes a range of tools for data reduction, multivariate discriminant analysis and calibration, as well as unique implementations of genetic algorithms for spectral feature selection. To demonstrate the utility of this package in metabolomic and post-genomic studies, we present the results of analysis in PyChem for a range of spectroscopic and transcriptomic data.

For more information and downloads please go to http://pychem.sf.net/, http://sourceforge.net/projects/pychem/ or http://biospec.net/

[1] Jarvis, R.M.; Broadhurst, D.; Johnson, H.E.; O'Boyle, N. & Goodacre, R. (2006) PyChem - a multivariate analysis package for Python. Bioinformatics 22(20):2565-2566.

COMBINING METABOLOMIC AND TRANSCRIPTOMIC TOOLS TO IMPROVE CELLULAR PROCESSES IN INDUSTRIAL MAMMALIAN CELL CULTURE

Soo Hean Gary Khoo, Mohamed Al-Rubeai (University of Birmingham)

As the biopharmaceutical industry increasingly uses of mammalian cells as a source of therapeutic proteins, several metabolic engineering strategies have been devised to increase the overall productivity of industrially important cell lines. These include the increase of specific cell productivity or the delay of cell death in bioreactors. To understand the effects of cellular processes that affect specific productivity, we have used cDNA microarrays and NMR metabolic profiling techniques to study the effect of increased productivity during proliferation arrest. Our results show the consistency in the gene expression and the resulting metabolic network shifts that result with this phenomenon.

STRATEGIES FOR METABOLOMICS STUDIES INVOLVING HUMAN DIETARY INTERVENTIONS: THE USE OF GENETIC ALGORITHMS TO SELECT DISCRIMINATING VARIABLES

Gwen Le Gall, L.J. Harvey, J.R. Dainty, E.K. Kemsley, I.J. Colquhoun (Institute of Food Research)

Our aim was to assess whether an increase in dietary copper intake could be detected by NMR urinalysis and, if so, to identify possible novel biomarkers of copper status. Six healthy male volunteers (age range 34-57) were recruited from the Norwich region and given a daily copper supplement of 6mg/d for 6 weeks. They provided 24h urinary samples for eight consecutive days in two separate periods prior to the supplement intervention and in one period immediately after. The study was performed with no other dietary control; subjects consumed their normal diet over the entire study period but filled in food diaries to provide a snapshot of their eating habits.
600 MHz 1H NMR spectra of the urines were acquired and multivariate methods were used to discriminate between samples and to identify the signals responsible. Identification of markers via loadings was problematic and a different approach was taken. A Genetic Algorithm-Linear Discriminant Analysis (GA-LDA) method was found effective for selection of the most discriminatory signals.
A total discrimination between all six individuals over the combined collection periods was achieved using very few spectral features. Signals could be selected for discrimination of pre- and post-intervention samples for individuals, but not when all volunteers were considered together. We are cautious in attributing changes to Cu intervention as (a) different metabolites were identified for different individuals; (b) the same metabolite was selected for two individuals but with opposing differences between pre- and post-intervention periods. More careful selection of subjects or stricter dietary control may be needed to discern systematic changes in metabolism following micronutrient interventions.

^{1}H MAS NMR-BASED METABONOMIC STUDY OF SCHISTOSOMA MANSONI INFECTION IN THE MOUSE

Jia Li, Yulan Wang, Jürg Utzinger, Jeremy K. Nicholson, Jacques Chollet, Jasmina Saric and Elaine Holmes

Schistosomiasis is a chronic and debilitating parasitic disease caused by the digenetic trematode flatworms (adult blood flukes) belonging to the genus of *Schistosoma*. Here we report the metabonomic fingerprint of a *Schistosoma mansoni* infection in the mouse, using ^{1}H Magic Angle Spinning (MAS) NMR spectroscopy and multivariate data analysis techniques. 30 female mice were involved into the study. 20 of them were infected with 80 *S. mansoni* cercariae and the other 10 mice as an uninfected control group (CV). At 54 days post-infection, 10 mice from the infected group (TP) were administrated with the praziquantel drug and the other 10 infected mice (TV) and CV group were treated with the drug vehicle only (7% Tween 80 and 3% ethanol). After 74 days of the infection, all mice were sacrificed and intestinal tissues were removed. NMR metabolic profiles of ileum were obtained from a 600 MHz magic-angle-spinning NMR spectrometer. The data were analyzed by principal component analysis (PCA), orthogonal signal corrected-projection to latent structure-discriminant analysis (O-PLS-DA). The results from O-PLS-DA revealed a significant separation between the infected group and uninfected control group with an elevation of leucine, isoleucine, alanine, lysine, glutamine, creatine, GPC, tyrosine, cytosine and phenylalanine in infected mice with the treatment of drug vehicle compared with the uninfected control group. The following work will be focus on investigating jejunum and colon and combine all metabolic changes to identify biomarkers for the diagnosis of schisotomiasis in human.

INTEGRATING NMR AND MS DATASETS USING HIERARCHICAL MODELING

Jack Newton, Mark Earll, Erik Johansson, John Shockcor (Chenomx Inc.)

NMR and MS techniques offer complimentary views of metabolomic data. Here we present the results of analyzing pre- and post-dose serum samples obtained from 16 rats. NMR and MS data were acquired in parallel. We analyze the NMR data using the targeted profiling technique to obtain metabolite identification and quantification. We validate the targeted profiling results against the MS dataset, and present a comparison of metabolite coverage with the MS dataset. We also demonstrate how to integrate both sources of information in a statistically coherent manner by using hierarchical PCA/PLS-DA models. By combining both approaches, we leverage the complimentary information contained in the NMR and MS data.

COMBINING TARGETED AND GLOBAL 1D NMR SPECTRAL PROFILING TECHNIQUES

Jack Newton, David Chang, Aalim Weljie (Chenomx Inc.)

Targeted profiling is a technique for directly recovering quantitative compound information from complex mixtures acquired via 1D NMR experiments. Targeted profiling uses a predefined library of individual compound signatures that are scaled quantitatively and summed together using a linear combination model to create a "simulated" NMR spectrum. In an ideal scenario, this "simulated" spectrum would precisely match the experimental spectrum. However, in all but the most contrived examples, there will be unknown compounds present in the mixture that are not captured using the targeted profiling methodology. Global, compound-agnostic techniques, such as spectral binning or full spectrum analysis, are commonly used to model NMR spectra when the compounds present in a mixture are not known a priori. However, these global profiling techniques can be confounded by a variety of factors, such as the "usual suspect" metabolites. Here we present a method for fully integrating targeted and global profiling techniques. By combining both approaches, we gain the benefits of targeted profiling for known compounds, and retain the extensive spectral coverage of global profiling techniques.

VARIANCE STABILISING TRANSFORMATIONS FOR NMR METABOLOMICS DATA

Helen Parsons, Mark Viant (University of Birmingham)

Classifying the fingerprint of an NMR spectrum is a crucial step in many metabolomics experiments. Since many classification techniques such as principal component analysis (PCA) depend upon variance discrepancies, it is important to first maximise any contribution from wanted class variance between biological samples and minimise any contribution from unwanted technical variance arising from the preparation of the samples and measurement of the NMR metabolic fingerprints. The generalised logarithm (glog) transform was developed to stabilise the variance between technical replicates in a two component error model [1] and has also been applied to NMR spectra previously [2]. To increase the effectiveness of the transform on NMR spectra, the glog was extended to include a baseline offset term. This decreases the unwanted noise contribution on the transformed spectra.

Here we have applied the extended glog transform to technical replicates of NMR spectra of tissue extracts from marine mussels, to determine the optimised transformation parameters used in the glog transform. Next we applied the optimised transformation to a data set comprised of two classes of NMR spectra from stressed and unstressed mussels. Following transformation, the results show significantly better separation of the classes on a PCA scores plot than can be achieved with both untransformed data and also data transformed using Pareto scaling, a widely used method in NMR metabolomics [3]. In conclusion, we have demonstrated the value of the extended glog transformation to stabilise the technical variance in an NMR metabolomics dataset and have achieved significantly improved classification of NMR fingerprints from stressed and unstressed animals.

[1] D.M. Rocke and S. Lorenzato, "A two-component model of measurement error in analytical chemistry," Technometrics, vol. 37, no. 2, p176-184, 1995.
[2] P.V. Purohit, D.M. Rocke, M.R. Viant and D.L. Woodruff, "Discrimination Models Using Variance-Stabilizing Transformation of Metabolomic NMR Data," Omics, vol. 8, no. 2, p118-130, 2004.
[3] H.C. Keun, T.M.D. Ebbels, H. Antti, M.E. Bollard, O. Beckonert, E. Holmes, J.C. Lindon and J.K. Nicholson, "Improved analysis of multivariate data by variable stability scaling: application to NMR-based metabolic profiling" Analytica Chimica Acta, vol. 490, no. 1, p265–276, 2003.

MICROPOROUS AND MESOPOROUS MATERIALS FOR LASER DESORPTION IONISATION MASS SPECTROMETRY FOR METABOLOMICS AND PROTEOMICS

Ketan Patel, Mike Anderson, Roy Goodacre (University of Manchester)

Work undertaken in 1999 by Siuzdak and colleagues provided evidence that small molecules could be analysed by laser desorption ionisation (LDI)-MS. They showed that porous silicon (Si) had optical properties which enabled the absorption of UV radiation at 337 nm (typically applied in MALDI-MS experiments), which lead to the desorption and ionisation of small peptides from Si wafers electrochemically etched with HF. A process termed desorption ionisation on Silicon (DIOS).

The current aims of this project will be complementary to the above, but also contain novel development of microporous and mesoporous materials for LDI-MS. An important project objective will be the synthesis and subsequent characterisation of a wide range of nanoporous materials, specifically variation on the mesoporous range of nanoporous materials, i.e. pore sizes greater than 50.00 nm. The aim is to fully demonstrate the credentials of the mesoporous materials for utilisation as an inorganic matrix platform, thus replacing the presently used organic matrix accompanied with an extra need for a proton source (H+). Subsequently, making the inorganic mesoporous matrix approach suitable for high-throughput, accurate and instrument sensitive analysis of both metabolomics and proteomics datasets. The characterisation techniques to be employed in the analysis of the nanoporous materials, are techniques such as powder X-ray diffraction (XDR), N2 adsorption/desorption hysteresis loops analysis, UV-Visible diffuse reflectance, UV resonance Raman (UVRR), Fourier transform infrared (FT-IR), transmission electron microscopy (TEM) analysis and finally silicon graphics simulation work (Cerius 2 Silicon) will be utilised to analysis the 3D construction of the mesoporous materials. Subsequently, MALDI-ToF-MS studies will be conducted on the synthesised mesoporous materials to test their sensitivity as viable matrix enhancing platforms.

IMAGE-BASED 'OMICS – UNRAVELING SPATIAL CHANGES IN COMPLEX BIOLOGICAL SYSTEMS

Soyab Patel, Felicity Currie, Nalin Thakker, Phil Sloan, Alan Barnes, Roy Goodacre (University of Manchester)

We have been investigating metabolic fingerprinting via Fourier transform infrared (FT-IR) spectroscopy and spatial proteomics via imaging mass spectrometry for the generation of information rich chemical maps. This poster will give details of our progress in these two areas to date.
Microalgae are the major primary producers in most aquatic ecosystems, accounting for approximately 50% of total planetary primary productivity and could be good indicators of insult by active pharmaceutical ingredients (APIs) released into the environment. In this study Micrasterias hardyi cells were exposed to combinations of propanolol (a beta-adrenergic receptor agonist drug), Mefanamic (non steroidal anti inflammatory drug), Metoprolol (analgesic). FT-IR results showed that by adding these drugs had a dramatic effect on the rate of protein synthesis, which leads to the reduction of total protein and lipid content of the cell, particularly with propanolol.
MALDI (matrix-assisted laser desorption/ionization) imaging mass spectrometry (IMS) is a recently developed technique that generates molecular profiles and two-dimensional ion density maps of peptide and protein signals directly from the surface of thin tissue sections. This allows specific information to be obtained on the relative abundance and spatial distribution of proteins. The objective of our current study was to evaluate printing of MALDI matrix with a novel matrix delivery device, Chemical Inkjet Printer (CHIP). Biostatistical analyses of the resulting protein profiles reveals a spectral feature at 4500 and 8360 Da that strongly correlates with the squamous cell carcinoma (SCC) region of the tongue.

AN INTEGRATED CALIBRATION AND STITCHING ALGORITHM FOR OPTIMISING MASS ACCURACY AND DYNAMIC RANGE IN FT-ICR MASS SPECTROMETRY BASED METABOLOMICS

Tristan G. Payne, Andrew D. Southam, Mark R. Viant, Helen J. Cooper, Theodoros N. Arvanitis (University of Birmingham)

High resolution mass spectrometry is an increasingly favoured method of identifying metabolites in complex biological samples, particularly with ultra-high resolution afforded by the recent developments in Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry. However, it has become apparent that the software tools for analysing the resulting data are not able to maximise the returns from the advances in instrumentation. This work presents a project aimed at optimising the processing of mass spectra and maximising the benefits of high resolution, high accuracy mass spectrometry, traits that are critical for the unambiguous identification of metabolites. Recent work in our laboratory has identified that the acquisition of multiple single ion monitoring (SIM) scans provides better mass accuracy than a single wide scan. Described here is a novel tool that enables multiple SIM spectra (that are overlapping) to be simultaneously calibrated and stitched to form a single wide range spectrum. Consideration is given to the optimal selection of raw transient data, subsequent pre-processing and transformation into mass spectra, through to alignment, calibration and stitching in both m/z and abundance dimensions. The features and benefits of this new algorithm will be presented, and its robustness will be demonstrated using data acquired on an FT-ICR mass spectrometer.

PRINCIPAL COMPONENT ANALYSIS (PCA) OF SIX HERBAL EXTRACTS OF THE SAME PLANT SPECIES USING DATA GENERATED BY A COMBINATION OF ANALYTICAL TECHNIQUES

P.J.Russell, E. van Velzen, P.J. Hylands and D.J.Barlow (Unilever Safety & Environmental Assurance Centre and Kings College London)

Principal Component Analysis (PCA) has been investigated for use in determining similarities and differences in six herbal extracts of the same plant species. Analytical fingerprints were obtained using ^{1}H nuclear magnetic resonance (NMR) spectroscopy, fourier transform infrared spectroscopy (FT-IR), gas chromatography with mass spectrometric detection (GC-MS) and liquid chromatography with mass spectrometric detection (LC-MS). Results from each of the analytical techniques were subjected to multivariate analysis as individual and combined datasets.
PCA was shown to be suitable for discriminating between two different extraction techniques. Using PCA it was also possible to identify analytical techniques and regions of chromatograms and spectra responsible for creating variation from very large combined data sets. This could have a useful application when identifying components in complex mixtures on which to set product specifications.

A METABONOMIC ASSAY ON FASCIOLA HEPATICA DAMAGED RAT LIVER, INVESTIGATED BY MAGIC ANGLE SPINNING

Jasmina Saric, Mansour Taghavi Azar Sharabiani, Jia Li, Yulan Wang, Elaine Holmes, Jürg Utzinger

Fasciolosis, which is caused mainly by *F. hepatica*, affects an estimated 2.4 million humans. 180 million people live at risk of an infection with a liver fluke and the geographical extension of *F. hepatica* is worldwide. The main hosts of the worm parasite are usually vertebrates, especially mammals, and Fasciolosis represents a serious health problem to humans and domestic animals.
Once the ingested metacercariae reached the liver, the immature flukes move trough the liver tissue for 5-6 weeks, eat and grow. The migration of the larval stages causes massive mechanical damages to the liver and lead to extensive fibrosis and haemorrhage. In this study, the focus is on the mechanical damages in the liver tissue, caused by the juvenile worms. We wanted to complete the pathogenic overall picture on a molecular level with high resolution MAS ^{1}H-NMR spectroscopy in combination with multivariate statistical analysis methods, and relate the metabolic pattern from the MAS acquired spectra to the mechanical damages.
Partial least squares as well as the principal component analysis demonstrated a clear separation in the metabonomic hyperspace, between liver tissue of controls and infected animals. The metabolic fingerprint of the rat-liver after damage by juvenile *F. hepatica* worms resembles the metabolic state of a liver cirrhosis and is characterized

by an increase of the branched chain amino acids (isoleucine, leucine and valine), phenylalanine, tyrosine, lysine, cholin and glutamate, and a decrease in glucose and glycogen.

INVESTIGATION OF INTERCELLULAR COMMUNICATION IN CULTURES OF SACCHAROMYCES CEREVISIAE BY MALDI-QTOF MASS SPECTROMETRY.

Jan Schripsema, Martijn Pinkse, Arthur Kroon and Peter Verhaert (Delft University of Technology)

In many multicellular organisms, short polypeptides have been identified as signaling molecules in intercellular communication. In recent years it has become evident that also unicellular organisms employ peptides as extracellular messengers. Examples, however, remain scarce, with quorum sensing peptides in bacteria and peptide pheromones in yeast being among the best studied cases.
Recent advances in the fields of mass spectrometry (MS) / peptidomics now allow for a straightforward analysis of secretory peptides directly from biological material. At the Department of Biotechnology of Delft University of Technology a big interest exists in industrially important unicellular eukaryotes, such as *Saccharomyces cerevisiae* and *Penicillium chrysogenum*. Our research, therefore, is focused on the MS analysis of (derivatives of) biological peptides and other secreted small biomolecules of these microorganisms.
To evaluate our microbial peptidomics approach we selected the well described peptide mating factor signaling pathway in haploid yeast cells as a model system. In *Saccharomyces cerevisiae*, mating involves the secretion of and response to peptide pheromones in the extracellular matrix. The pheromone peptide secreted by alpha-type cells, the alpha-factor, is a 13-amino-acid peptide; that of the a-type cells, the a-factor is a posttranslationally modified dodecapeptide. Using a very simple and fast extraction procedure we were able to detect both mating factors individually from minimal amounts of culture medium in one single MALDI MS experiment. These yeast peptidome MALDI MS profiles indicate that various peptide species, derived from the mating factor, but also from other yeast proteins, show up in various (relative) quantities. Their appearances seem to depend on the physiological condition of the yeast culture, such as its growth phase. In order to inventorize these subtle differences in the yeast secretion peptidome.

NMR-BASED ANALYSIS OF RESPONSES TO MODEL TOXINS IN A SENTINEL ORGANISM, THE EARTHWORM LUMBRICUS RUBELLUS.

Jasmin Sidhu, David Spurgeon, Claus Svendsen, Peter Kille, Jake Bundy (Imperial College London)

The environmental implications of soil contamination are a significant issue as perturbations in soil health can have negative effects on soil biota with subsequent effects on human health through direct exposure or bioaccumulation. Earthworms are susceptible to such soil contamination and are known to accumulate metals and pesticides. Therefore, earthworms can be utilised as pollution indicators. As such the metabolic effects of three specific soil contaminants in the earthworm *Lumbricus rubellus* were examined through metabolomic analysis. These contaminants were atrazine, fluoranthene and cadmium, representing a pesticide, a typical organic pollutant, and a toxic heavy metal. Metabolic changes in earthworm extracts were identified through ^{1}H NMR spectroscopy and chemometric analysis. Dose related metabolic changes were examined with a view to identify potential biomarkers of soil toxicity.

NMR MANAGER – METABOLOMICS SOFTWARE FOR INTERPRETING NMR SPECTRA

Robert Stones, Adrian Charlton, James Donarski (Central Science Laboratory)

^{1}H NMR (Proton Nuclear Magnetic Resonance) spectroscopy is one of a number of complimentary techniques used to detect presence of metabolites within complex matrices. Regions where chemical groups are located within NMR spectra can be used to identify metabolites.
NMR Manager is a metabolomics analysis software application for storing, retrieving and viewing NMR spectra and associated peak information. The software allows comparison and interpretation of 1D NMR to identify metabolites in complex mixtures, using peak detection/assignment algorithms and statistical data analysis methods.The software includes a database, which is designed to hold parameter and peak information, and NMR ^{1}H spectra. All associated files generated from an NMR spectrometer can be imported. Spectra can be viewed within a graphical viewer, along with peak information and acquisition data.
NMR Manager has a toolkit, which has the following features:

- Interactively and automated setting of reference peak positions
- Perform visual comparisons of spectra
- Built-in functions for spectral alignments - database entries versus experimental NMR
- Search engine allowing retrieval of database NMR via inputting peak ranges
- Automated peak detection/assignment algorithms
- Statistical data analysis tools
- 3D graphical view for comparing peak area clusters in different populations of experimental NMR

NMR Manager is designed to aid comparison and interpretation of complex NMR spectra. Development of a knowledge base of NMR data, analysis and data mining tools, will reduce the bottleneck in data interpretation in metabolomics.

METABOLIC PROFILING OF ANDROGEN INSENSITIVITY IN PROSTATE CANCER CELL LINES

Orla Teahan, Simon Gamble, Charlotte Bevan, Jonathan Waxmann and Hector Keun (Imperial College London)

Despite the initial success of androgen ablation therapy, all prostate tumours eventually recur. Currently there is no way of predicting which prostate tumours are likely to progress to a hormone-resistant stage, or when a hormone-responsive tumour will develop resistance.
Tumour cells share a common metabolic phenotype characterised by increased lactic acid production, reduced aerobic metabolism and increased glucose uptake and glycolysis (Warburg effect). It has been proposed that this altered glucose metabolism emerges as a survival mechanism in response to the intermittent hypoxic episodes that result from rapid cell proliferation in pre-malignant lesions. We are using the powerful techniques of metabonomics, primarily 1H Nuclear Magnetic Resonance (NMR) spectroscopy and subsequent chemometric analysis, to obtain spectra of metabolites and other small molecules in medium and cell extracts from various androgen-sensitive and insensitive prostate cancer cell lines.
Our aim is to explore the impact of altered metabolism on tumour evolution and also to investigate disease mechanisms leading to androgen-independent prostate cancer by examining differences in metabolic profiles from androgen-sensitive and insensitive cell lines. Our results suggest that as prostate cancer cells progress to the more advanced malignant phenotype of androgen insensitivity they become adapted to acidosis and more effective at removing lactic acid. The androgen sensitive LNCaP cell line secreted less lactate than other cell lines, and was found to have more intracellular lactate. In contrast, the androgen insensitive LNCaPR subline appeared to develop increased tolerance to lactate production and was more similar metabolically to the other androgen insensitive cell lines than to the LNCaP line. Further investigation will determine whether these changes are causal or consequential of androgen receptor function or mechanism.

^{1}H NMR AND LC-MS BASED METABONOMICS FOR THE DISCOVERY OF URINARY BIOMARKERS OF CARCINOGENESIS IN THE TRAMP MOUSE MODEL OF PROSTATE CANCER

Friederike Teichert, R Verschoyle, P Greaves, P Farmer, W Steward, H Keun, I Wilson, A Gescher (University of Leicester)

The TRAMP (TRansgenic Adenocarcinoma of the Mouse Prostate) model is a well established and widely used animal model for human prostate cancer. The prostate specific expression of the simian virus 40 (SV40) early-region tumour antigens leads to the development of preneoplastic lesions in the prostate which progress to invasive tumors. Human prostate cancer has a long latency period before clinical symptoms occur. Reliable and easy to measure biomarkers for early detection would be of great benefit. Metabolic profiling of urine combines the advantages of an easily accessible biofluid with the comprehensive analysis of the metabolome, rendering it suitable for biomarker discovery. In this study, urine from 8 weeks and 28 weeks old mice (TRAMP and wild type mice) was subjected to ^{1}H NMR and LC-MS (negative mode electrospray ionization) analysis, followed by multivariate statistical analysis. The predominant differences in the urinary metabolic profile were related to age, as disclosed by Principal Component Analysis (PCA). Changes related to disease within one age group were much more subtle. Partial Least Square-Discriminate Analysis (PLS-DA) of the LC-MS data revealed three distinct highly discriminatory ions, which were individually elevated in 8 weeks old TRAMP mice, 28 weeks old TRAMP mice and 8 weeks old wild type mice, respectively, when compared to all other groups. Visual inspection of NMR spectra revealed mostly spectral regions elevated in 28 weeks old TRAMP mice though differences were also seen between 8 weeks old TRAMP and wild type mice.
These results suggest that both metabonomics platforms are suitable for the detection of urinary biomarkers of disease, already in 8 weeks old TRAMP mice, where the pathological changes in the prostate tissue do not exceed minimal focal glandular hyperplasia. As the information provided by both analytical methods is complementary, the chances for revealing disease affected pathways are remarkably increased.

WATER SIGNAL SUPPRESSION IN NMR SPECTRA OF METABOLOMICS SAMPLES

Stefano Tiziani, Christian Ludwig, Ulrich Guenther (University of Birmingham)

The identification of compounds of low abundance can be essential for the analysis of the metabololite composition in biofluids. Whilst the availability of high magnetic field strengths and cryogenically cooled probes increases the signal-to-noise ratio of NMR spectra and therefore allows to resolve resonances in crowded regions of proton spectra containing hundreds of compounds; the suppression of the solvent signal requires increasingly sophisticated pulse sequences. The use of highly sensitive cryogenically cooled probes at high magnetic field strengths introduces the problem of radiation damping caused by an induced radiofrequency field arising from the large signal of the solvent acting as an on-resonance flip back pulse on the solvent. Highly effective water suppression techniques are therefore paramount to circumvent this problem.
In this study we compare several water suppression techniques for their suitability to identify metabolites and to quantify metabolite concentrations in human blood serum at 800 MHz proton frequency using a cryogenically cooled probe. Two different classes of water suppression methods have been examined: (i) methods which saturate the water resonance and (ii) methods which destroy the water resonance using pulsed field gradients. Advantages and possible sources of error are reported for each technique to obtain more accurate and comprehensive metabolomic information from NMR data.
Excitation sculpting has several advantages over presaturation techniques, WET and SWET, and several versions of

WATERGATE. Spectra recorded with excitation sculpting were characterised by minimal phase and baseline distortions, a uniform excitation profile with a narrow suppression of resonances around water and is therefore highly suitable for qualitative and quantitative analysis of metabolites in human blood serum.

APPLICATION OF HIGH-FIELD NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY (NMR) IN A METABOLOMIC STUDY OF AQUATIC ECOTOXICOLOGY AND METABOLISM

Will Tuffnail, Richard Greenwood, Graham Mills (University of Portsmouth)

The metabolites extracted from the foot tissue of the marine mollusc Mytilus edulis were used to investigate the biochemical perturbations caused by exposure of this non-target species to the insecticide Lindane (g-HCH). The metabolomic analysis was carried using 600 MHz ^{1}H NMR spectroscopy and multivariate data analysis. Lindane is a GABA-gated chloride channel antagonist, and mussels were exposed two concentrations (5 ug L^{-1} and 1 mg L^{-1}) in artificial seawater, controls were maintained under identical conditions in artificial seawater containing the carrier solvent acetone (1 ml L^{-1}). The use of principal component analysis provided a clear discrimination between controls and the two treated groups. The major metabolites that contribute to the separation include alanine, glycine, aspartate, formate and homarine.

LONGITUDINAL VARIATIONS IN FUNCTIONAL METABOLIC SIGNATURES OF HUMAN GUT EPITHELIAL BIOPSIES REVEALED BY HIGH-RESOLUTION MAGIC-ANGLE-SPINNING ^{1}H NMR SPECTROSCOPY

Yulan Wang, E. Holmes, J. Lindon, H. Tang, L. Fay, S. Kochhar, J. Nicholson (Imperial College London)

The variation in metabolic profiles of human gut biopsies from antrum, duodenum, jejunum, ileum and transverse colon has been revealed using high-resolution magic-angle-spinning (HRMAS) 1H NMR spectroscopy and pattern recognition. Samples from antrum, duodenum, jejunum, ileum and transverse colon were taken from 8 male and 8 female individual and were found to have highly characteristic metabolic profiles consistent with different biochemical functions at various longitudinal levels of the gut. The mucosa from the antrum contained higher levels of choline, glycogen, phosphorylethanolamine, taurine and acetyl glycoproteins than the mucosa from other regions. The duodenum and jejunum were similar in terms of their biochemical composition, and consisted of higher levels of choline, glutathione, GPC, lipids and glycoproteins. Ileal mucosa showed strong signatures from amino acids but the GPC level was low. The colon sample showed elevated levels of acetate, glutamate, inositols, lactate and reduced levels of creatine, GPC and taurine compared to the small intestine. The metabolic variations observed in the different mucosa of a human GI tract could be associated with specific functions in each part of the human GI tract including glucose homeostasis, osmoregulation, microbiotal activity, free radical protection and digestive processes.

INVESTIGATING THE FATE OF CHIRAL APIS IN AQUATIC ENVIRONMENTS

Emma S. Wharfe, Catherine L. Winder, Roger Jarvis, Royston Goodacre (University of Manchester)

Active pharmaceutical ingredients (APIs) and their metabolites are ubiquitous in the environment and their occurrence in the aquatic environment is of growing concern. Whilst, little is currently known about the effects of APIs in the aquatic environment; it is likely that there will be unexpected harmful effects caused by such compounds on some of the aquatic biota. In addition, there is the potential for mixture effects with multiple APIs, and European guidelines now require testing on new APIs before their release.
Chiral pharmaceuticals are of particular concern as the enantiomers may be metabolized differently, with the potential for the production of harmful compounds. The biodegradation of chiral pharmaceuticals in the aquatic environment has not previously been extensively studied, and certainly a metabolomics approach has not been attempted yet.
We have exposed a variety of aquatic microorganisms to the chiral pharmaceuticals atenolol and propranolol. The effect of these APIs on the metabolic fingerprint has been investigated through the use of FT-IR spectroscopy with appropriate chemometrics. In addition, GC-MS has been employed to investigate the fate of these APIs. This poster will give an update of our work in this area.

FOURIER-TRANSFORM INFRARED SPECTROSCOPY AS A METABOLIC FINGERPRINTING TOOL FOR MONITORING WHOLE CELL BIOTRANSFORMATIONS

Catherine L. Winder, R. Cornmell, S. Schuler, R. Jarvis, G. Stephens & R. Goodacre (University of Manchester)

Redox biotransformations should principally provide many new routes to the manufacture of both chiral and non-chiral speciality chemicals. However, the majority of redox reactions require the use of whole cells as biocatalysts to provide a means of cofactor regeneration.
The problem with using whole cells as biocatalysts is that the product yields are limited by substrate and/or product toxicity. The objectives of this study are to improve the stability of whole cell biocatalysis by monitoring the transcriptome, proteome and metabolome changes in whole cell biotransformation systems using Escherichia coli MG1655 harbouring a plasmid carrying toluene dioxygenase. However, it is not readily possible to perform these 'omic methods across the entire time course of the biotransformation growth, as this is time consuming, labour intensive and expensive.
Therefore we sought to use Fourier- transform infrared (FT-IR) spectroscopy since it is a rapid phenotypic typing technique which may be used to generate a metabolic fin-

gerprint. When coupled to multivariate statistical analysis this provides a powerful combination to differentiate bacteria to sub-species levels, and to monitor any stress they may be undergoing.
We have employed E. coli MG1655 pDTG601A as a biocatalyst to convert toluene to toluene cis-glycol (TCG). The culture was grown under fed-batch conditions using two methods to supply the toluene to the culture (either vapour or syringe feed). Product concentrations were recorded at 83mM (vapour feed) and 22mM (syringe feed). Multiple samples were taken at hourly intervals across the time course of the biotransformations and metabolic fingerprints were generated. Multivariate statistical analysis of the data was performed to determine the most suitable time points to perform the 'omic analyses.

NMR-BASED METABOLOMIC STUDY ON THE TOXICITIES OF ENVIRONMENTAL POLLUTANTS IN THE THREE-SPINED STICKLEBACK

Huifeng Wu, Jonathon Ball, Eduarda Santos, Charles Tyler, Ioanna Katsiadaki, Mark Viant (University of Birmingham)

Adult male three-spined sticklebacks (*Gasterosteus aculeatus*) were treated with various doses of chemicals, including dibenzanthracene (DbA, 10, 100, 1000, 10000 and 50000 ng/L), copper (Cu. 3.2, 10, 32, 64 and 128 μg/L) and ethinylestradiol (EE2, 0.1, 1.0, 10, 32 and 100 ng/L). Fish liver tissues were extracted using an optimised methanol and chloroform solvent system and the polar metabolites were analysed using 1D ^{1}H NMR spectroscopy. Following spectral processing, principal components analysis (PCA) was applied to each of the three sets of spectra. For DbA treated groups, there existed a limited separation between lower (0, 10 and 100 ng/L) and higher (1000, 10000 and 50000 ng/L) dose groups. However, there was no obvious clustering between the different dose groups for both Cu and EE2.
From the PC loadings plots we discovered that one particular metabolite – a derivative of cholic acid – made a major contribution towards the pattern of samples in the PCA scores plots. Unfortunately, because of the complexity of the NMR spectrum of this metabolite, much of the fish NMR spectra below ca. 2 ppm was not analysable. One possible solution that is now being investigated is to subtract the NMR spectrum of the pure metabolite from the spectra of the fish livers. This depends upon the identification of the metabolite (likely taurocholic acid). PCA could then be repeated on the modified datasets, which would then hopefully show improved clustering of the data in terms of the dose groups.

CLOSED LOOP OPTIMIZATION OF UPLC-TOF-MS PERFORMANCE FOR MODERN METABOLOMICS

Eva Zelena, K. Carrol, W. Dunn, S. O'Hagan, J. D. Knowles, D.B. Kell (University of Manchester)

Several different analytical methods are being employed for metabolomics, including mass spectrometry, vibrational spectroscopy and NMR. One relatively new and promising method is Ultra Performance Liquid Chromatography coupled to orthogonal acceleration time of flight mass spectrometry (UPLC-ToF-MS) which offers the potential to detect and identify large number of metabolites in a biological sample. To provide the highest possible number of detected metabolites by this method, the optimization of instrument settings has to be considered very carefully. Even comparatively small changes in conditions could substantially influence the effectiveness of a particular separation and subsequent detection of analytes in complex biological samples. In our analyses of human serum using UPLC-ToF-MS, we are also applying closed loop optimization which proved very useful for GC-ToF-MS performance optimization [1]. This approach utilises an evolutionary algorithm for parameterising optimal instrument settings and requires no human intervention, which leads to substantial improvement in the instrumentation performance. The values of fourteen parameters influencing chromatographic separation and spectrometric detection were selected and the preliminary results of this optimization are discussed in this poster.
1. O'Hagan, S., Dunn, W.B., Brown, M., Knowles, J.D. and Kell, D.B. (2005) Closed-loop, multiobjective optimization of analytical instrumentation: Gas chromatography/time-of-flight mass spectrometry of the metabolomes of human serum and of yeast fermentations. Analytical Chemistry 77, 290-303