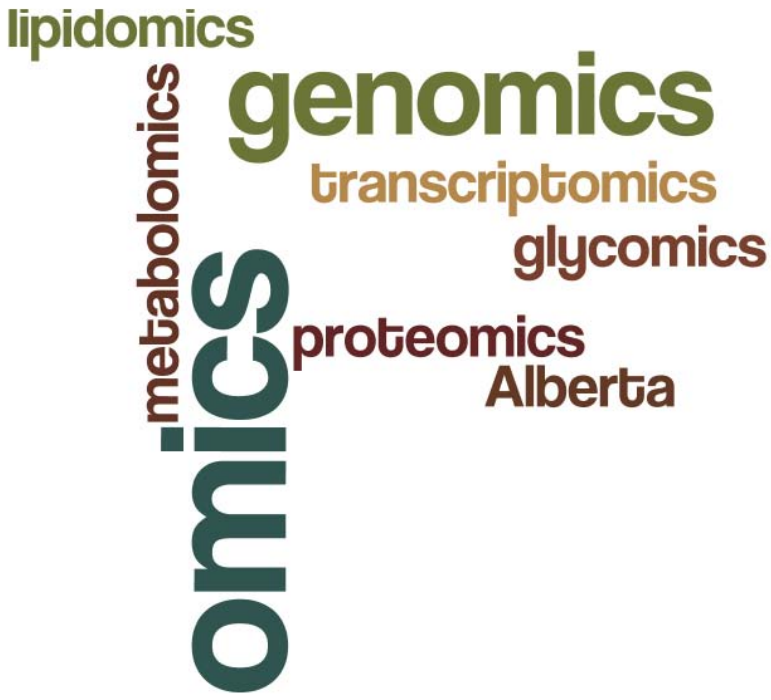
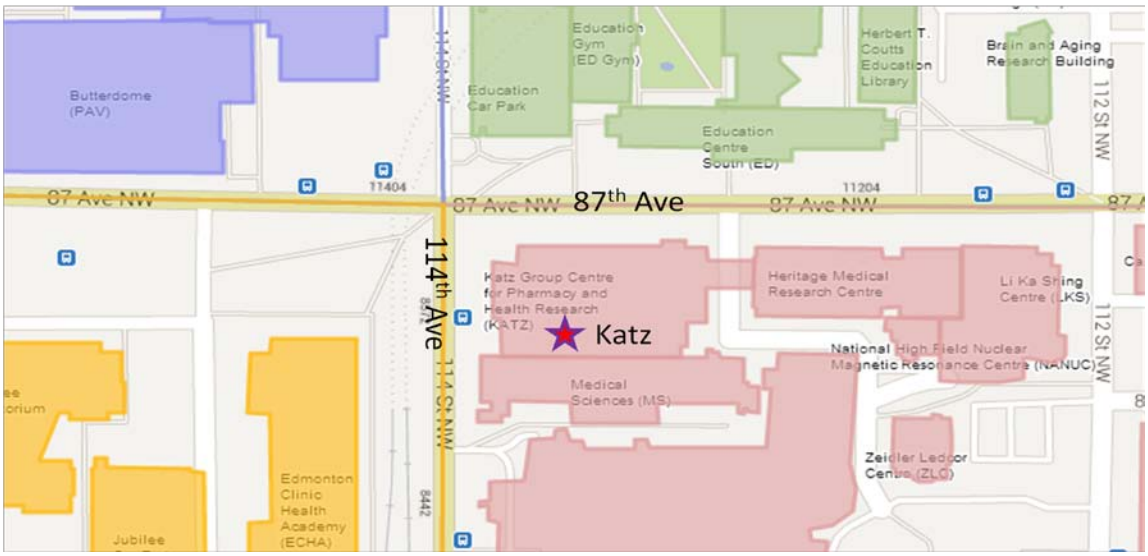


University of Alberta "OMICs" Conference

June 17, 2013

Allard Family Lecture Theatre – Katz
& Katz Foyer

Allard Family Lecture Theatre – Katz
& Katz Foyer



Sponsored By:

ETP Symposium Inc.
Bringing Researchers Together

University of Alberta

Faculty of Science

Faculty of Medicine & Dentistry

Student Awards Provided By:

BRUKER

Thermo SCIENTIFIC

IDT INTEGRATED DNA TECHNOLOGIES

Truin Science

lipidomics genomics transcriptomics glycomics metabolomics proteomics omics Alberta

Alberta OMICs Program

8:50 - 9:00 Introductions

Session 1 Genomics/Transcriptomics

Chair: Joel Weiner

9:00 - 9:30 Gane Wong (Abstract #1)
1KP: An International Consortium Sequencing the Transcriptomes of 1000 Phylogenetically Diverse Plant Species from Angiosperms to Algae.

9:30 - 10:00 Damaraju Sambasivarao (Abstract #2)
Genomic biomarkers for breast cancer predisposition and recurrence: are we there yet?

10:00 - 10:15 Marc St George (Abstract #3)
β Tubulin Isotype and miRNA Expression Profiles in Model Cell Lines Exhibiting Resistance to Paclitaxel

10:15 – 10:45 Coffee Break 1

Session 2 Proteomics

Chair: Liang Li

10:45 - 11:15 Richard Fahlman (Abstract #4)
Comparative Proteomics: Current capabilities and comparative analysis of tumor response to chemotherapy.

11:15 – 11:30 David Stuart (Abstract #5)
Mining the Secreted Proteome of Scheffersomyces stipitis, a xylose fermenting yeast.

11:30 - 12:00 Raymond Lai (Abstract #6)
Applications of Proteomics in Cancer Research

12:00 - 12:15 David Kramer (Abstract #7)
A Feeding-Dependent Analysis of the Hepatic Lipid Droplet Proteome in Mice.

12:15 - 2:15 Lunch & Poster Viewing & Judging

Session 3 Metabolomics

Chair: Richard Fahlman

2:15 - 2:45 Liang Li (Abstract #8)
High-performance Isotope Labeling Liquid Chromatography Mass Spectrometry for Metabolomics Applications.

2:45 - 3:15 David Wishart (Abstract #9)
Metabolomics & Medicine

3:15 - 3:30 David Broadhurst (Abstract #10)
Better Data through Design: the expanding role of the biostatistician in the post-genomic era clinical research.

3:30 - 3:45 John Paul Glaves (Abstract #11)
Pre-symptomatic prion disease metabolite profiles: insights from NMR spectroscopy of mouse urine.

3:45 - 4:15 Coffee Break 2

Session 4 Emerging Omics (glycomics, lipidomics etc)

Chair: Joel Weiner

4:15 - 4:45 Christine Szymanski (Abstract #12)
Advances and challenges in bacterial glycomics

4:45 - 5:15 Jonathan Curtis: (Abstract #13)
The analysis of lipids by LC/MS and related techniques: the Lipid Chemistry Group approach to lipidomics.

5:15 - 5:30 Jessica Sacher (Abstract #14)
Use of genomic expression libraries to screen the bacteriophage proteome for specific carbohydrate-binding molecules

5:30 – Announcement of Presentation Awards

Abstract #1

1KP: An International Consortium Sequencing the Transcriptomes of 1000 Phylogenetically Diverse Plant Species from Angiosperms to Algae.

Gane Ka-Shu Wong[1,2]

[1]Department of Biological Sciences and Department of Medicine, University of Alberta.
[2]BGI-Shenzhen, Beishan Industrial Zone, Yantian District, Shenzhen, 518083, China.

There are approximately eight million species known to science, not counting microbial species. Angiosperms (flowering plants) alone number a quarter million. We do not even know how many more species remain to be discovered. In the field of large-scale DNA sequencing, we have until recently been focusing on a few dozen species of importance to agriculture or medicine. What was left untouched was the majority of the diversity of our planet. 1KP is an international consortium acquiring large-scale gene sequence data for the plant kingdom, incorporating at some phylogenetic or taxonomic level all known species from angiosperms to algae. The project is organized into ~14 subprojects. Some were designed to answer fundamental questions in plant evolutionary (*e.g.*, the role of polyploidy in angiosperm diversification, Darwin's abominable mystery). Others address more practical issues for agriculture (*e.g.*, making crops such as rice more efficient at photosynthesis) or medicine (*e.g.*, creating molecular tools to probe mammalian brains). The latter studies demonstrate in pragmatic terms the value of "completeness" across evolutionary diversity, as opposed to completeness within species, which has long been the defining characteristic of genomics.

Abstract #2

"Genomic biomarkers for breast cancer predisposition and recurrence: are we there yet?"

Sambasivarao Damaraju, PhD

Dept. of Lab Med and Pathology, University of Alberta

Breast cancer is a multifactorial disease, and is influenced by genetic, environmental and lifestyle factors. Breast cancer as a disease is very heterogeneous at both histological and molecular levels. Several decades of efforts at identifying genetic predisposition in affected families (20% of all cases diagnosed) have led to the discoveries of BReast CAncer genes (BRCA1 and 2). BRCA gene mutations are said to be highly penetrant, and follows Mendelian pattern of inheritance. Further efforts to uncover additional gene mutations to explain the genetic risk proved futile prompting search for common variants that confer a low but finite risk, that cumulatively may explain the genetic basis for breast cancer susceptibility; these markers are potentially scattered all over the genome (Polygenic mode of inheritance).

Prognostic markers to guide therapies are traditionally tumor based markers (such as expression status of estrogen, progesterone or human epidermal growth factor receptors) and have served well despite their limitations. There is need for markers with higher specificities to guide treatment decisions; to administer aggressive therapies to those at risk and sparing treatment related toxicities to those who are not predisposed to disease recurrence.

Recently, we have used whole genome association studies (GWAS) scanning 1.8 million polymorphic markers over the entire human genome to uncover germline DNA (constitutive DNA isolated from blood) markers for breast cancer predisposition and prognosis. This area of science is new and emerging and thus far a large focus of the international community was on familial disease. Interest in my lab is on identifying the markers for sporadic breast cancer (80 % of all cases diagnosed) risk in populations, an under-represented area of research and markers to identify risk for disease recurrence and to improve outcomes. Breast cancer related mortality is due to metastatic spread often due to recurrence of the disease (shows poor response to further treatments). I will discuss some recent developments in this field and describe identification of novel markers.

Abstract #3

β Tubulin Isotype and miRNA Expression Profiles in Model Cell Lines Exhibiting Resistance to Paclitaxel

M St.George[1], C Cass[2], O Kovalchuk[3], J Tuszynski[2] and S Damaraju[1]

Dept. of [1]Laboratory Medicine and Pathology, [2]Oncology University of Alberta, [3]University of Lethbridge

Breast cancer is a leading cause of cancer related deaths among women in Canada. Taxane drugs used to treat breast cancer mediate cytotoxic effects by inducing microtubule polymerization which impedes cell division, promote mitotic arrest and leads to apoptosis. Development of resistance to taxane drugs is a mounting clinical problem and mechanisms of differential expression of β-tubulin and their contribution to resistance has been investigated. Our aim was to profile global differential miRNAs and correlate with β tubulin expression and drug resistance in model cell lines. To characterize these mechanisms of drug resistance we sought to develop a model panel of paclitaxel resistant cell lines. A stepwise increase of drug was administered to SKBR-3 cells beginning 1000 fold less than IC_{50}. Drug concentration was increased 3 fold every two passages until a maximum tolerable dose was reached. Western blots were performed to assess β tubulin isotype protein expression. Next Generation Sequencing (NGS) on Illumina platform was carried out to profile small RNAome including miRNAs. miR-200c expression pattern was found to be reversely correlated to β III tubulin expression; loss of expression in resistant cells was observed compared to the sensitive cells. Thirty six additional miRNAs showed significant up or down regulation and correlated with paclitaxel resistance. Model SKBR-3 cell lines are a good starting point to understand β–III tubulin expression, with regulation by miR-200c being confirmed as one of several possible mechanisms of drug resistance. Differential expression of miRNAs in our model system and their potential mRNA target elucidations render insight into the diversity of mechanisms employed by cells to become resistant to chemotherapeutic drugs.

Abstract #4

Comparative Proteomics: Current capabilities and comparative analysis of tumor response to chemotherapy.

Richard P. Fahlman

Department of Biochemistry, University of Alberta

High resolution mass spectrometry has begun to redefine proteomics. The combination of high resolution and high sensitivity is enabling proteomic analysis of relatively small samples at levels not possible a few years ago. Some of the applications offered in the Institute for Biomolecular Design mass spectrometry facility will be discussed.

Using a label-free method, the data from our work on a mouse model system for investigating tumor response to chemotherapy treatment will be presented. In these investigations we have characterized the proteomes of mouse EL4 lymphoma derived tumors with and without chemotherapy treatment. Data reveals numerous changes to the tumor proteome within 48 hours of chemotherapy treatment. Examination of the data reveals a number of observable cellular responses which will be discussed.

Abstract #5

The mining the secreted proteome of *Scheffersomyces stipitis*, a xylose fermenting yeast.

David Stuart, and Thomas Duggan
Department of Biochemistry University of Alberta Edmonton Alberta Canada

A limited number of microbial species display the ability to degrade and digest cellulose and hemicelluose. These carbohydrate polymers make up the most abundant forms of biomass on the planet and substantial effort has been applied to develop methods for conversion of this biomass for the synthesis of biofuels and biochemicals. The major constraint to utilization of cellulose and hemicellulose is cleavage of the β-1,4 glycosidic bonds that link individual molecules of glucose or xylose together. The yeast *Scheffersomyces stipitis* can degrade hemicellulose to xylose and can ferment the resulting monosaccharide to ethanol. Thus, *S. stipitis* can produce fuel ethanol from wheat straw or sawdust. Analysis of the extracellular proteome of *S. stipitis* revealed a set of secreted proteins that based upon sequence were predicted to be involved in hydrolysis of carbohydrate polymers. A predicted β -1,4 endoxylanase was highly abundant in hemicellulose grown cultures whereas a β -1,4 endoglucosidase was more abundant in glucose grown cultures. Interestingly the xylanase was also highly expressed at both the protein and mRNA levels in xylose grown cultures. Comparison of the mature protein sequence with the genomic DNA sequence revealed a pre-sequence that directed secretion from this yeast. Recombinant forms of the mature protein displayed 300% greater catalytic activity than the full-length protein. When the full-length xylanase was expressed in industrial strains of budding yeast the enzyme was secreted and active. Over expression of the xylanase in *S. stipitis* increased the cells growth rate and production of ethanol when grown with hemicellulose as a carbon source. Thus, through proteomic analysis we have demonstrated the expression and secretion of hemicellulose hydrolyzing enzymes by *S. stipitis*, discovered the peptide sequence required for protein secretion by this organism, and shown that altering the expression or activity of the major secreted xylanase can improve the fermentation of hemicellulose to fuel ethanol.

Applications of Proteomics in Cancer Research

Raymond Lai, MD, PhD, FRCP and Fang Wu, PhD
Department of Oncology & Department of Laboratory Medicine & Pathology, University of Alberta

Recent advances in the field of proteomics have provided a powerful tool for cancer researchers to tackle challenges encountered in their studies. In this presentation, we will provide a few examples of how our research teams addressed key questions by using various mass spectrometric-based proteomic techniques. In collaboration with Dr. Liang Li and various members of his mass spectrometry laboratory, we have great successes in the following studies: 1) identification of the binding partners of NPM-ALK, which is a highly oncogenic, fusion gene protein characteristically found in a type of T-cell lymphoma; 2) characterization of the changes in the phosphoproteome induced by NPM-ALK, thereby revealing novel pathways that mediate the oncogenic effects of NPM-ALK; 3) identification of biomarkers implicated in the relapse of diffuse large B-cell lymphomas, by using proteome-wide quantification and identification of differentially expressed proteins between the initially diagnostic and relapsed tumors. In the first study, the use of the hexahistidine-biotin (HB)-tag tandem affinity purification and LC-MS/MS will be highlighted. In the second study, the strategy of combining sequential affinity purification and LC-MS/MS will be highlighted. In the third study, our strategies of quantitative mass spectrometry will be illustrated. In our experience, the establishment of a strong 'omics' infrastructure on campus has proven to be instrumental in facilitating our research.

Abstract #7

A Feeding-Dependent Analysis of the Hepatic Lipid Droplet Proteome in Mice.

David Kramer[1], Ariel Quiroga[2], Richard Lehner[2], and Richard Fahlman[1]

[1]Department of Biochemistry, University of Alberta
[2]Department of Pediatrics, University of Alberta

During fasting, hepatocytes accumulate large amounts of triacylglycerols (TG) and cholesteryl esters (CE) in cellular organelles called lipid droplets (LD). While this is thought to occur as a means of energy storage, it can also lead to a physiological fatty liver condition. Normally hepatic fat accumulation is transient; after re-feeding, the amount of lipid retention in hepatocytes rapidly decreases to a minimum, presumably due to hormonal signaling pathways activated upon feeding and the presence of nutrients in circulation. However, this is not always the case; this rapid decrease in LDs following re-feeding is not observed in individuals afflicted with non-alcoholic fatty-liver disease (NAFLD), causing chronic lipid retention in the liver which can ultimately lead to cirrhosis and liver failure. As the molecular mechanisms regulating the amount of hepatic lipid content observed in fasted/fed conditions is poorly understood, the treatment of patients with NAFLD can be challenging. In light of this, investigations on the enzymes involved in hepatic lipid metabolism are of interest, as the identification of these enzymes (and lack thereof) will provide insight to the regulation of hepatic LD content during fasted and fed states. Triacylglycerol hydrolase (TGH) and esterase-x (ESX) are two lipid hydrolases expressed in the endoplasmic reticulum, and have recently been demonstrated to play key roles in the maturation of LDs within hepatocytes. Here we present, for the first time, a proteomic analysis of the protein composition within LDs purified from WT, TGH -/-, and ESX-/- mouse livers in both fasted and re-fed conditions, illuminating the feeding-dependent recruitment of proteins to hepatic LDs.

Abstract #8

High-performance Isotope Labeling Liquid Chromatography Mass Spectrometry for Metabolomics Applications.

Liang Li

Department of Chemistry, University of Alberta

Metabolome profiling is a key step in applying metabolomics for disease biomarker discovery and systems biology. Due to great diversity of chemophysical properties of metabolites, quantitative analysis of all the metabolites present in a metabolomic sample can be very challenging. Liquid chromatography mass spectrometry (LC-MS) has become a powerful tool for metabolome analysis. It provides high sensitivity for metabolite detection and high specificity for compound identification. Combining with isotope labeling, LC-MS can also perform metabolite quantification with high precision and accuracy. However, at present, the metabolome coverage by LC-MS is still limited. In this presentation, our recent research efforts on developing robust and efficient separation methods aimed at improving the metabolome coverage of LC-MS will be described. Various metabolite extraction and separation methods are being investigated to examine their analytical performances for extracting and separating metabolites in a complicated metabolome sample, such as biofluids and cell extracts. Data processing methods are also being developed for integrating the mass spectrometric results with the metabolome information. These technical developments as well as some applications of the developed methods will be discussed.

Abstract #9

Metabolomics and Medicine

David S. Wishart, Depts. of Computing Science, Biological Sciences, Laboratory Medicine & Pathology, University of Alberta, Edmonton, AB, Canada

Metabolomics is a newly emerging field of "omic" science that is primarily concerned with characterizing small molecule metabolites in cells, organs and organisms. The complete collection of metabolites in a given cell or organism is called the metabolome. Metabolomics is often regarded as the "country cousin" of the better-known omics fields such as genomics and proteomics. Although most of us wouldn't know it, the city of Edmonton is actually home to most of the world's metabolomic companies. Likewise, the University of Alberta houses most of Canada's metabolomics researchers, maintains most of Canada's metabolomic equipment and is the primary host for most of the world's main metabolomic data resources. In this presentation I will briefly describe how Edmonton (and the U of A) got to be in this position. I will then describe how the discoveries, the companies and the unique ideas in metabolomics that have developed here could have a long-term impact on how we practice medicine – both in Alberta and around the world. In particular I will provide a number of examples of how metabolomics is being used to help understand, diagnose, prognose and predict a wide variety of diseases. I will also describe some exciting efforts to make metabolomics far more accessible – both to physicians and the public – through the marriage of nanotechnology and molecular biology. My overall goal is to show that metabolomics has "arrived" and that the U of A is helping to move metabolomics from the lab and into the clinic -- and eventually into your home.

Abstract #10

Better Data through Design: the expanding role of the biostatistician in the post-genomic era clinical research.

David I Broadhurst

Department of Medicine, University of Alberta, Canada

Many clinical metabolomics, and other high-content or high-throughput, 'omic experiments are set up such that the primary aim is the discovery of biomarkers that can discriminate, with a certain level of certainty, between nominally matched 'case' and 'control' samples. However, it is unfortunately very easy to find markers that are apparently persuasive but that are in fact entirely spurious. The main types of danger, which are not entirely independent of each other, include: bias in patient selection; inconsistent biobanking; poor choice in clinical endpoint; inadequate sample size; inappropriate choice of machine learning methods; inadequate model validation; poor experiment design; minimal and ineffective lab based quality assurance protocols. Many studies fail to take these issues into account, and thereby fail to discover anything of true significance, or more seriously report spurious findings which prove impossible to validate. Here I summarise these problems, and provide pointers to assist in the improved design and evaluation of 'omic experiments, thereby allowing robust scientific conclusions to be drawn from the available data.

Abstract #11

Pre-symptomatic prion disease metabolite profiles: insights from NMR spectroscopy of mouse urine.

Glaves, J.P.[1], Daude, N.[2], Wishart, D.[3], Westaway, D.[1,2], and Sykes, B.D.[1]

1- Department of Biochemistry, University of Alberta
2- Centre for Prions and Protein Folding Diseases and Department of Medicine (Neurology), University of Alberta
3- Department of Computing Science, Department of Biological Sciences and Department of Laboratory Medicine and Pathology, University of Alberta

Human prion diseases are diagnosed following changes in behavior and/or difficulty in co-ordination. These symptoms rapidly progress to severe dementia, the inability to speak, and inevitably death. Post-mortem histological diagnosis is based on neuropathological features in the brain. Despite the incubation period of prions prior to symptom onset, the practical difficulties of early diagnosis are quite numerous, including the invasive nature of obtaining brain tissue and the low levels of infectious prions in accessible body fluids.

We hypothesized pre-symptomatic prion disease processes would be reflected as metabolic changes in host biofluids. As a starting point, we investigated the mouse model of prion disease to reduce variability caused by genetic factors and diet. Urine was studied as it is easily accessed in a longitudinal manner during the course of disease. Both intra-cerebral and oral routes of infection were studied with 10 animals per infected and mock-infected group. Urine samples were collected at 20-day intervals from 30 days post-infection (d.p.i.) to termination. Small molecule metabolites from urine samples of individual mice were analyzed by NMR spectroscopy and multivariate statistical methods.

Classifications of infected and mock-infected samples were 97% (oral) and 100% (intra-cerebral) accurate using model-building and test (model-independent) samples. Samples as early as 70 d.p.i. (intra-cerebral) and 90 d.p.i. (oral) were properly classified, suggesting urinary profiles of mice are perturbed prior to observed behavioural changes (115 and 196 d.p.i. for intra-cerebral and oral, respectively). Using a semi-targeted approach, we quantified metabolites that discriminate between infected and mock-infected mice. The resulting metabolite profiles point to a potentially widespread perturbation of metabolic pathways during the early stages of prion disease. This work highlights the potential for early prion disease diagnostics and metabolic pathways for investigating therapeutic intervention.

Abstract #12

Advances and challenges in bacterial glycomics

Christine M. Szymanski

Alberta Glycomics Centre and Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada, T6G 2E9

e-mail: cszymans@ualberta.ca

It is now well recognized that bacteria are capable of both N-linked and O-linked protein glycosylation. Glycosyltransferases add nucleotide-activated carbohydrates sequentially to serine and threonine residues of a single protein in the classical bacterial O-linked glycosylation pathways. However, in some O-linked pathways, bacteria assemble oligosaccharides first onto membrane-bound lipids (undecaprenylphosphate, UndP), then flip these lipid-linked oligosaccharides across a membrane and use an oligosaccharyltransferase to attach these sugars to one or more proteins. As in their eukaryotic counterparts, bacteria also have classical N-linked glycosylation pathways where oligosaccharides are assembled onto a lipid (UndP), flipped across a membrane and added to asparagine in the extended sequon (D/E)-X-N-X-S/T (where X cannot be proline). Recently, an alternate N-glycosylation pathway has been described in several bacterial species where glycosyltransferases in the cytoplasm sequentially add nucleotide-activated sugars to the N-X-S/T sequon of a single protein. Thus, this domain of life exploits all four possible combinations for modifying proteins through N- and O-linkages.

Campylobacter jejuni is the first bacterium demonstrated to possess a classical *N*-linked protein glycosylation (*pgl*) pathway involved in adding heptasaccharides to asparagine-containing motifs of >60 proteins[1], and also releasing the same glycan into the periplasm as free oligosaccharides[2]. Comparative genomics of 29 sequenced *Campylobacter* taxa revealed *pgl* gene clusters are conserved in this genus[3]. Structural, phylogenetic and immunological studies showed that the *N*-glycosylation systems can be divided into two major groups. Group I includes all thermotolerant taxa and produce *C. jejuni*-like glycans. The non-thermotolerant campylobacters comprise group II and produce an unexpected diversity of *N*-glycan structures varying in length and composition. The terminal non-reducing sugar residues of these *N*-glycans are the determinants for adaptive and innate immune recognition. Since, both group I and II oligosaccharides are immunogenic and cell surface exposed, these structures are also attractive targets for glycoconjugate vaccines and diagnostics.

[1] Harald Nothaft and Christine M. Szymanski. *Nat Rev Microbiol* **2010**, 8, 765-778.

[2] Harald Nothaft, David J. McNally, Jianjun Li and Christine M. Szymanski. *PNAS* **2009**, 106, 15019-15024.

[3] Harald Nothaft, Nichollas E. Scott, Evgeny Vinogradov, Xin Liu, Rui Hu, Bernadette Beadle, Christopher Fodor, William G. Miller, Jianjun Li, Stuart J. Cordwell and Christine M. Szymanski. *Mol Cell Proteomics* **2012**, 11, 1203-1219.

Abstract #13

The analysis of lipids by LC/MS and related techniques: the Lipid Chemistry Group approach to lipidomics.

Jonathan M. Curtis and the Lipid Chemistry Group (LCG)

Dept. of Agricultural, Food and Nutritional Science,

University of Alberta

Lipidomics means different things to different people. Typical approaches are strongly influenced by extraction methods and instrumental response factors. This presentation will focus on more targeted methods for the quantitative and structural analysis of complex lipids. Examples will include a global LC-MS/MS method to quantify all choline containing and related compounds. Analysis of complex isomeric lipid structures using a novel on-line ozonolysis method will be explained and a reinvestigation of the traditional methylation procuedure for GC analyses will be described.

Abstract #14

Use of genomic expression libraries to screen the bacteriophage proteome for specific carbohydrate-binding molecules

Jessica Sacher, David Simpson, Denis Arutyunov, M. Afzal Javed and Christine M. Szymanski.

Alberta Glycomics Centre and Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

The growing recognition that glycans play central roles in many cellular processes has resulted in a surge of research in glycomics. Glycans, or sugars, are used by cells to decorate proteins and lipids in order to modify their functions. This process serves to substantially increase the diversity of gene products expressed by cells, and has important implications for the study of the molecular interactions involved in human health and disease. In glycomics, precise and reliable methods of detecting and discriminating between different glycans are required. Currently, carbohydrate-binding proteins known as lectins are employed, however problems with broad specificity as well as low diversity of binding targets highlight the limits of lectin-based strategies. There exists a particular need for specific carbohydrate binding proteins when examining bacteria, where surface-displayed glycans show immense diversity. In order to fill this gap, we are investigating the feasibility of exploiting bacteriophages, which are viruses that specifically target bacteria, as sources of carbohydrate-binding proteins. Phages target their bacterial hosts through receptor binding proteins (RBPs), which are usually carbohydrate-specific. RBPs are highly stable, retaining their structure and function when recombinantly expressed, and display receptor binding affinities and specificities comparable to those of antibodies. However, in spite of the rapid progress in DNA sequencing technologies, phage genome sequencing remains a challenge due to the existence of extensive modification systems, and RBPs are difficult to identify through homology searches due to their wide ranges of host receptors. As an alternative, we are developing a high-throughput whole-phage genomic expression library approach to screen for putative RBPs in any phage. Our preliminary results indicate that it is possible to identify recombinantly expressed RBPs with glycan-binding abilities, including *Salmonella enterica* serovar Typhimurium lipopolysaccharides and *Campylobacter jejuni* flagellar sugars. We are now optimizing a method for whole-phage genome expression and RBP detection. Carbohydrate-binding proteins would be useful for any glycomics study where an antibody is not available and could also be exploited for diagnostic applications of bacterial pathogens.

Poster Abstracts

Abstract #15

LC-MS ion trap workflows for glycan and glycopeptide analysis using CID and ETD fragmentation

Marx, K; Kiehne, A; Meyer, M

Bruker Daltonik GmbH, Fahrenheitstrasse 4, 28359 Bremen, Germany, Kristina.marx@bdal.de

Glycosylation of proteins is the most complex form of post-translational modifications and found in over 50% of all eukaryotic proteins. Glycoproteins have diverse functions and play a crucial role in many biological processes e.g. such as enzymatic or hormone activity. The glycan moieties are either directly involved in biological processes or at least contribute by their effect on the physicochemical properties of a given glycosylated protein. Changes in glycosylation patterns can influence glycoprotein properties and are associated with a number of human diseases. Main fields for the analysis of glycosylation patterns are therefore monitoring of therapeutic glycoproteins or identification of biomarkers e.g. for cancer.

For mass spectrometric analysis of therapeutic glycoproteins, glycans are generally cleaved, e.g. *N*-glycans with PNGase F, and analysed separately yielding the overall glycan pattern. Biomarker analysis on the other hand focuses on the analysis of proteolytic glycopeptides. This is necessary to distinguish the glycosylation at different glycosylation sites, since changes might not occur global and may influence the glycoprotein properties in different ways.
The amaZon speed ETD instrument provides the desired workflows to cover both glycan and glycopeptide analysis on one platform.

The analysis of released glycans in positive mode by CID (generation of Y- and B-ions) provides identification of their composition and the corresponding sequence of the constituent monosaccharide residues. Further structural characterization can be performed with MS^n and/or fragmentation in negative ionization mode. Especially diagnostic and cross-ring fragment ions, which are generated in negative mode allow the interpretation of e.g. the position of fucoses [1] or provide the respectice linkage information [2].

The identification of glycan compositions and the unambiguous assignment of the corresponding glycosylation sites can be carried out by analyzing proteolytic glycopeptides. Here, the combination of CID and ETD fragmentation represents a powerful tool for characterization of glycan structures on the one hand (CID) and the glycopeptide sequence on the other hand (ETD).
A general overview over the different analysis workflows performed on the amaZon speed ETD instrument is demonstrated on different samples.

[1] Harvey, D. J.; *J Am Soc Mass Spectrom* **2005**, *16,* 622–659.
[2] Tang, H., Mechref, Y. and Novotny, M. V.; *Bioinformatics* **2005***; 21(Suppl 1),* i431–i439.

Abstract #16

Isoniazid reinforces host defense system: a novel complementary mode of action and The elucidation of isoniazid-induced agranulocytosis mechanism in HL-60 cells

Saifur R. Khan[1], Argishti Baghdasarian[1], Richard P. Fahlman[2], and Arno G. Siraki[1]

[1] Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, Edmonton, Canada.
[2] Department of Biochemistry, Faculty of Medicine & Dentistry, University of Alberta, Edmonton, Canada

Despite isoniazid being the most successful drug against tuberculosis, the complete mode of action and mechanism of isoniazid-induced agranulocytosis (IIA) are yet to be comprehended. We employed SILAC with LC-MS/MS to investigate the differential expression of HL-60 cell's proteome upon a relatively higher but non-toxic dose regimen of isoniazid. In this study, we found 49 significantly expressed proteins of which 15 were up-regulated and 34 were down-regulated. Our results revealed a complementary mode of action of isoniazid which is to fortify the host defensive mechanisms through boosting the immune system and enhancing the longevity of HL-60 cells. Furthermore, we also found the relation of IIA and oxidative stress. However, IIA in TB patients may not be related with its pharmacological dose rather than other factors or its idiosyncratic nature.

Abstract #17

^{1}H-NMR based metabolomic profiling of Experimental Autoimmune Encephalomyelitis (EAE): a novel approach to understanding multiple sclerosis.

S.N. Reinke, D.I. Broadhurst, C. Power

Department of Medicine, University of Alberta, Edmonton AB

Multiple Sclerosis (MS) is a debilitating neurodegenerative disease affecting as many as one in five hundred Canadians. Metabolomics is the systematic study of low molecular weight (bio)chemicals characterizing the convergence of gene expression and environmental stimuli. Experimental Autoimmune Encephalomyelitis (EAE) initiates a T-cell mediated immune response to myelin, thereby mimicking MS; EAE is the most widely used animal model of MS. HYPOTHESIS: A metabolomic time-course study of central nervous system (CNS) tissue from EAE mice will reveal pathogenic insights into immune system mediated demyelination and identify novel therapeutic targets. METHODS: C57BL/6J mice were induced with EAE, using Complete Freund's Adjuvant (CFA), pertussis toxin, and Myelin Oligodendrocyte Glycoprotein $(MOG)_{35-55}$. Animals were euthanized at 4 time points during disease: 2 days prior to clinical onset, day of onset, post-onset/pre-peak disease, and peak disease. Disease severity was defined by weight and degree of paralysis. Two sets of control mice were also observed at all time-points: (1) healthy controls receiving no induction and (2) CFA controls receiving CFA and pertussis toxin. ^{1}H-NMR metabolomic analyses were performed on spinal cords - the primarily affected CNS tissue. RESULTS: Univariate data analyses identified several metabolites as being significantly altered in EAE animals; these included energy metabolites (pyruvate, lactate, succinate, fumarate), lipid metabolites (acetylcarnitine, carnitine, choline, phosphocholine), amino acids, and biogenic amines (GABA, N-acetylasparate, myo-inositol, taurine). Changes in energy metabolism were corroborated by molecular analyses. Multivariate data analyses revealed that the spinal cord metabolomes of EAE-induced animals were similar to those of the control groups prior to disease onset. After disease onset, EAE-induced animals had a divergent temporal metabolite trajectory when compared to the two controls groups, which followed similar trajectories. CONCLUSIONS: These data reveal that the immune-mediated demyelination process is associated with energy metabolism changes, identifying these central pathways as potential therapeutic targets.

Abstract #18

SERS detection of free thyroxine biometabolite in plasma

Shereen A. Elbayomy and Mark T. McDermott

Department of Chemistry and National Institute for Nanotechnology, University of Alberta. Edmonton, Alberta, Canada.

The detection and quantification of small molecule metabolites is being targeted as a promising diagnostic in disease assessment. Surface-enhanced Raman scattering (SERS) spectroscopy and Raman labeled gold nanoparticles can be used for the determination and quantification of small biometabolites simultaneously on the same immunoassay patterned chip that are hard to identify by conventional immunological methods. We report a new designed indirect competitive SERS immunoassay developed for the analysis of the thyroid hormone thyroxine (T_4) using stable water soluble prepared T_4-conjugated nitrobenzene labeled gold nanorods. SERS signals from nitrobenzene Raman label have been collected after competing of T_4-conjugated nitrobenzene modified gold nanorods and free T_4 standard solutions prepared in plasma for T_4-monoclonal antibodies which immobilized on alkane thiol modified gold chips. Free T_4 in the range from 5.79-105.55 pM has been detected by SERS proposed method of good correlation results with commercial ELISA kit. This proof-of-concept experiment shows that groups of clinically related small molecules biomarkers can be simultaneously analyzed for single sample by microfluidic patterned SERS platform.

Abstract #19

GENETICALLY ENCODED FRAGMENT-BASED DESIGN FROM LOW AFFINITY CARBOHYDRATE FRAGMENT

Simon Ng, Mohammad R. Jafari, Wadim Matochko, Ratmir Derda*

Alberta Glycomics Centre, Department of Chemistry, University of Alberta, Edmonton, AB T6G 2G2, Canada

Phage display is a powerful technology that enables the discovery of peptide ligands for many targets. Chemical modification of phage libraries has allowed the identification of ligands with properties not encountered in natural polypeptides. Recently, we have selectively introduced a mannose moiety to the N-terminus of a heptapeptide phage library through a two-step chemical reaction. Up to 10^8 unique glycopeptides are generated within 1.5 hours. In this present work, we applied the synthetic phage libraries for their selection against a model protein target (Concanavalin A). We have improved the selection by incorporating several controls (negative target and irrelevant library) during the selection process. Most importantly, our criteria for identifying the true "hits" were based on a data-driven approach. With the wide coverage of library DNA that we can sequence (>million at a time), we could easily identify the target-unrelated glycopeptides from the control selection and subsequently rule them out. Through this strategy, we were able to identify glycopeptides that inhibit or bind to Concanavalin A with IC_{50} or K_d at least 10 times better than methyl α-D-mannopyrannoside. The sequence of the selected peptides suggests that Concanavalin A, other than its principal mannose-binding site, holds a secondary binding site which is hydrophobic in nature.

Abstract #20

Novel Spectroscopic Probes for Detecting UV-Induced DNA damage

A. F. El-Yazbi and G. R. Loppnow

Department of Chemistry, University of Alberta, Edmonton, AB T6G2G2

The goal of this work is to design probes that are superior to conventional molecular beacons (MBs) in detecting UV-induced DNA damage. The first approach was to design MBs with modified DNA backbones. So we tried the locked nucleic acid (LNA) MBs and the chimeric RNA-DNA MB (chMB). Our results show that chMBs are more sensitive and selective for DNA damage than LNA MBs that have comparable selectivity to conventional MBs. However, these probes all show signals that are inversely proportional to the amount of damage. Therefore, we designed probes that give signals directly proportional to damage. Probes with 2-aminopurine (2AP) as a fluorescent base show no fluorescence for undamaged DNA and fluorescence for damaged DNA. 2AP probes offer high sensitivity and selectivity comparable to MBs, but are expensive, especially with an increasing number of 2APs in the probe. Thus, a hypochromism probe was designed. The hypochromic effect arises from the formation of a double-stranded target-hairpin hybrid. With accumulated UV exposure, the target-hairpin hybrid concentration decreases and the absorbance increases. This probe is more selective and is more than ten times cheaper than MBs but is less sensitive. The need for a sensitive, selective and inexpensive probe was the motivation to design a Tb^{3+}/hairpin probe. Single-stranded DNA greatly enhances the Tb^{3+} emission, but duplex DNA does not. The Tb^{3+}/hairpin probe proves to be the most sensitive and selective probe for the quantification of DNA damage of all probes presented here.

Abstract #21

Environmental Multilocus Sequence Typing of Natural *Vibrio cholerae* Populations

Tania Nasreen, Paul C. Kirchberger and Yan Boucher

Department of Biological Sciences, University of Alberta, Edmonton, Canada

Abstract: *Vibrio cholerae*, the causative agent of the severe life-threatening diarrheal disease "cholera", naturally inhabits aquatic environments. It is a waterborne pathogen, which infects human hosts ingesting contaminated water. Despite intensive research in characterizing *V. cholerae*'s physiology and mode of infection in the laboratory and in animal model systems, we still have a very limited understanding of the ecological niche it occupies in its coastal aquatic environmental reservoir. This is even more limited for *Vibrio metecus*, the newly discovered close relative of *V. cholerae*. The knowledge obtained from this study will help us to determine the environmental factors that affect the prevalence of these two species. Identifying and understanding *V. cholerae'*s niche in the natural environment and its interaction with other species is critical in managing and predicting outbreaks of this lethal disease

Abstract #22

Quantitative Proteomic Analysis (SILAC) Reveals Aminoglutethimide Induced the Expression of Apoptotic & Oxidative Stress Proteins: A Potential Insight into Drug-Induced Agranulocytosis in HL-60

Argishti Baghdasarian[1], Saifur R. Khan[1], Richard Fahlman[2], Karim Michail[1], and Arno G. Siraki[1]*

[1]Faculty of Pharmacy & Pharmaceutical Sciences, University of Alberta, Edmonton, AB, Canada

[2]Faculty of Medicine & Dentistry, University of Alberta, Edmonton, AB, Canada

Aminoglutethimide (AG), a drug used for the treatment of breast and ovarian cancers is known for its various toxicities such as agranulocytosis. Here, we use quantitative proteomic analysis to gain insight into the differential expression of proteome in Human Leukemia 60 (HL-60) treated with AG. We identified 43 proteins significantly changed upon AG treatment among which 18 (42%) and 25 (58%) of the 43 proteins were up and down-regulated respectively. The quantitative proteomics data showed that among the proteins up-regulated were histones, heat shock proteins, and oxidative stress proteins. Among the down-regulated, we found proteins involved in replication machinery, transcription, pre-mRNA processing proteins, cytoskeleton proteins, nuclear envelope proteins, and anti-apoptotic proteins. These results suggest that AG induced agranulocytosis might be mediated through down-regulation of anti-apoptotic proteins as well as protein synthesis, DNA replication, and structural proteins such as cytoskeleton and nuclear envelope.

*corresponding author: e-mail: Siraki@ualberta.ca; address: Faculty of Pharmacy and Pharmaceutical Sciences, Katz Group Centre for Pharmacy and Health Research, Room 2-020H, University of Alberta, Edmonton, AB T6G 2E1. Tel: +1 780 248 1591; Fax: +1 780 492-3475

Abstract #23

Investigating the Protein Stability Profiling in the Eubacterial N-end Rule Pathway

Angela W.S. Fung[1], Devin K. Ulvestad[1], Aaron Madryga[1], Jack Moore[2], Richard P. Fahlman[1,3]

[1]Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7
[2]Institute for Biomolecular Design, University of Alberta, Edmonton, Alberta, Canada T6G 2H7
[3]Department of Oncology, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

The N-end rule pathway is a regulated proteolytic pathway where the N-terminal amino acid identity determines the *in vivo* stability of a protein. Variations of the N-end rule pathway are present in prokaryotes, fungi, plants and mammals. In eukaryotes these enzymes play critical roles in cardiac development, apoptosis, and G-protein signalling, meanwhile the biological role in prokaryotes remains elusive. The biological role of the eubacterial N-end rule has remained elusive, partly due to the lack of characterization of its *in vivo* substrates and the incomplete understanding of the initiation of the pathway. If the pathway is found to be involved in infectious processes it could be a novel antibacterial target.

Here we aim to identify potential N-end rule substrates by comparing proteomic changes between *E. coli* wild type and N-end rule deletion mutants (lacking specific components of the pathway) under different growth phases. We have utilized both metabolic labelling with stable isotopes and Gel-LC-MS/MS, where cultures are grown in defined minimal media with the substitution of heavy (^{13}C) or light labelled lysines. Raw mass spectrometry data are analyzed with SEQUEST for both protein identification and quantification. Proteins identified as more abundant or stabilized in N-end rule deletion mutants compared to wild type are potential N-end rule substrates. Here we will present our current data to date. The identification of N-end rule dependent proteomic changes may provide novel insights into the biological role of the bacterial N-end rule pathway.

Abstract #24

Proteomic Analysis of Chemotherapy Effectiveness on EL4 Lymphoma tumors in wild-type mice

Leslie D. Shewchuk, David Kramer and Dr. Richard P. Fahlman.

Department of Biochemistry, University of Alberta

Chemotherapy is the use of chemical agents to combat the growth of a tumor in a patient. Unfortunately chemotherapeutic agents are not specific to cancer cells but instead they typically toxic to rapidly growing cells. This lack of specificity is what is responsible for causing many of the side-effects associated with chemotherapy, which include decreased blood cell count, hair loss, nausea and many other side effects. Due to these harmful side-effects it would be beneficial to patients to have an early detection of a treatments efficacy in order to determine whether the full course of treatment should be continued or not. Currently the only clinical evaluation of chemotherapy effectiveness is after the multi-week treatment has been completed.

My research is part of an investigation to identify potential early markers of tumor death upon successful chemotherapy treatment. My specific project was to compare the proteomes of EL 4 Lymphoma derived tumors from mice that were treated with a chemotherapeutic cocktail for 48 hours or untreated controls.

Preliminary analysis reveals an enrichment of proteins involved in proteins synthesis in the untreated control while a number of apoptosis related and DNA damage repair proteins are observed in the treated tumor sample.

Abstract #25

Regulation of PKD2L1 channel function: potential roles of protein kinase C (PKC)

Shaimaa Hussein and Xing-Zhen Chen

Membrane Protein Disease Research Group, Department of Physiology, Faculty of Medicine and Dentistry, University of Alberta, Canada T6G 2H7

Polycystic kidney disease 2 like 1 (PKD2L1), also called TRPP3 and polycystin-L, is a member of the transient receptor potential (TRP) superfamily. PKD2L1 is a Ca-regulated cation channel, permeable to small positively charged ions and is inhibited by amiloride and large monovalent cations. PKD2L1 channel also exhibits a unique response to acid stimuli by an evoked current after the removal of the acid, known as the off response. PKD2L1 is largely involved in sensing extracellular pH, a critical determinant for many fundamental physiological processes. Physical and functional interaction of PKD2L1 has been reported with the receptor for activated protein kinase C (RACK1), a scaffolding/anchoring protein implicated by various cellular processes. This highlighted the possible involvement of PKC in channel function. Exploring the roles of protein kinase C (PKC) in channel function is an important step to unveil the mechanism underlying PKD2L1 pH sensing. We hypothesize that the role of PKC could be direct through the phosphorylation of PKD2L1 itself and/or through acting on a PKC downstream signaling molecule. Experiments were performed utilizing *Xenopus* oocyte expression system, two microelectrode voltage clamp, Western blotting and site directed mutagenesis techniques. Our preliminary results showed that activation of PKC reduces PKD2L1 channel function. This will constitute an essential step towards elucidating the molecular mechanisms underlying acid and calcium sensing of PKD2L1. Supported by CIHR, NSERC and Alberta Innovates Technology Futures.

Abstract #26

An Analysis of Predicted and Identified Targets Using Label-Free Proteomic Analysis of miR-23~24 Gene Regulation

Julian C. A. Sacher,[1] Braden T. Milan,[1] Steven G. Chaulk,[1] Richard P. Fahlman[1,2]

[1]Department of Biochemistry, [2]Department of Oncology, University of Alberta, Edmonton, 114 Street and 87 Avenue, Alberta, Canada T6G 2H7

MicroRNAs (miRNAs) are short endogenous RNAs that repress gene expression by targeting mRNAs. Approximately one half of human miRNAs are organized into clusters and are expressed as a polycistronic primary miRNA (pri-miRNA). miR-23a~24-2 is a miRNA cluster that encodes 3 mature miRNAs and that has been shown to be involved in gene regulation in both development and disease. The coordinated expression of miRNAs within a cluster suggests that their biological functions are intimately intertwined. We are using label-free proteomic analysis to investigate gene regulation by the miR-23~24 cluster. miRNA target prediction webservers are frequently used tools to predict and understand miRNA based gene regulation. TargetScan, the most widely used miRNA target prediction algorithm, predicts ~2500 targets total for the 3 miRNAs of the miR-23a~24 cluster. Our comparison of proteomic data to TargetScan predictions reveals a poor correlation between predicted and experimental data.

Abstract #27

IDENTIFYING NOVEL BIOTINYLATED PROTEINS IN ESCHERICHIA COLI K-12

Michal E. Gozdzik, Angela W.S. Fung, and Richard P. Fahlman

Biotin is an important coenzyme involved in CO_2 transfer as well as assisting in fatty acid and amino acid synthesis, gluconeogenesis, and catabolism. Due to its high binding affinity to avidin proteins ($K_d = 10^{-15}$M), biotin has been used to facilitate purification and enrichment from complex protein mixtures. Presently there has only been a single biotinylated protein identified in *Escherichia coli*, Biotin Carboxyl Carrier Protein (BCCP). Biotin is attached to BCCP via biotin protein ligase (BirA). However, if there are other biotinylated proteins this could prove problematic in applications involving biotin tags. Initially the *E. coli* cells were lysed and examined by Western Blot with an IRDye 680®-Streptavidin. The western blot showed what appeared to be other biotinylated proteins in addition to the known protein BCCP. The purpose of our project is to identify these additional biotinylated proteins. To determine whether these proteins were biotinylated, several BirA mutants were used in the investigations. Biotinylated proteins were isolated using three different types of agarose beads: Captavidin, Streptavidin, and Neutravidin. *E. coli* lysate was incubated with the beads, the beads were washed and then biotinylated proteins eluted. The eluted proteins were examined by SDS-PAGE and Western Blot analysis. Between the three different types of beads, Streptavidin was found to be the most efficient and provided the strongest signal. The analysis of the concentrated lysates revealed the expected BCCP band as well as several other higher molecular weight biotinylated proteins. Optimization was achieved by changing the buffers and volumes**.** The results from the streptavidin pulldown were submitted to the in-house orbitrap mass spectrometer to identify the additional biotinylated proteins. Analysis revealed several potentially novel biotinylated proteins such as 50S ribosomal binding protein L14, GTP-binding protein TypA/BipA, Acriflavine resistance protein A, Cytochrome d ubiquinol oxidase subunit 1, Ribosome-binding factor A and others. Currently we are verifying whether or not these additional proteins are in fact biotinylated.

Abstract #28

DISCOVERY OF CELL-BINDING PEPTIDES USING PH.D.-7, DEEP-SEQUENCING AND EMULSION AMPLIFICATION

Wadim L. Matochko,[1] Simon Ng,[1] Mohammad R. Jafari,[1] Sindy K.Y. Tang,[2] Ratmir Derda[1*].

[1] *Department of Chemistry and Alberta Glycomics Centre, University of Alberta, Edmonton, Alberta, Canada, T6G 2G2.*
[2] *Department of Mechanical Engineering, Stanford University, Stanford, CA 94305, USA*

Our group is interested in the discovery of all ligands to cell surface receptors. In order to accelerate the discovery process we use phage display to select for all binding ligands. Standard selection involves multiple rounds of panning and amplification in a bulk solution to converge on strong binding sequences. However, phage libraries contain clones with high growth rates, which outcompete slow growing clones and leads to collapse in diversity of the phage library. We employed amplification in monodisperse droplets as described in previous reports to preserve library diversity[1,2]. We used next-generation Ion Torrent sequencing to analyze and identify all ligands for cell surface receptors of MDA-MB-231 (breast cancer) cells. We characterized a new population of cell-binding peptide sequences that could not have been identified using the standard bulk amplification method. These sequences have come from part of the phage library that constitutes slow growing phage. Incorporating emulsion amplification will allow for the selection of all binding sequences that originate from phage display libraries.

References.

(1) Derda, R.; Tang, S. K. Y.; Whitesides, G. M. *Angew. Chem. Int. Ed.* **2010**, *49*, 5301.
(2) Matochko, W.L.; Ng, S.; Jafari, M.R.; Romaniuk, J.; Tang, S.K.Y.; Derda, R. *Methods* **2012,** *58*, 18.

Abstract #29

Proteomic Analysis of Gene Regulation by the miR-23~24 miRNA Cluster

Steven G. Chaulk[1], Braden Millan[1], Richard P. Fahlman[1,2]

[1]Department of Biochemistry, [2]Department of Oncology, University of Alberta

Edmonton, Alberta, Canada, T6G 2H7

Small RNA species have been found to regulate gene expression in an increasing number of biological processes. One class of small RNAs, miRNAs (microRNAs), are initially expressed in much larger primary miRNAs (pri-miRNA) that contain one or several miRNA sequences that are located within predicted ~70 nucleotide stem-loop (hairpin) structure. The first processing step in miRNA biogenesis relies on the RNase III enzyme Drosha to release the ~70 nucleotide miRNA containing hairpins from the pri-miRNA. These precursor miRNAs (pre-miRNA) are then exported out of the nucleus to the cytoplasm by Exportin 5/Ran GTPase and are processed by another RNase III, Dicer, into mature ~22 nucleotide miRNAs. The miRNAs are subsequently incorporated into RISC (RNA-induced silencing complex) which directs RNAi mediated gene regulation by targeting a complementary mRNA. Approximately 50% of human miRNAs are located within polycistronic clusters, where two or more miRNAs encoded on a single pri-miRNA transcript are co-expressed. Proteomic investigations of miRNA mediated gene regulation have largely considered only one miRNA of a cluster, ignoring the co-expression of other miRNAs from the cluster. The miR-23~24 miRNA cluster encodes 3 miRNAs that in addition to being involved in B-cell and myeloid cell development, have also been shown to have oncogenic or tumor suppressive activities in various cancers. We have done a label-free quantitative proteomic investigation of gene regulation by the intact miR-23~24 miRNA cluster. Expression of this miRNA cluster in HEK293T cells results in dozens of proteins having increased or decreased expression. Pathway analysis of the targeted genes suggests that the miRNAs from miR-23~24 function cooperatively to regulate their targets.

Abstract #30

Applied Genomics at AITF

Susan Koziel, Jim Davies, Karen Budwill, John Vidmar and Jian Zhang

Alberta Innovates Technology Futures is involved in four areas of Genomics and Metagenomics works that contribute to the application of applied genomics.

1. Bacterial Metagenomics of methane and hydrogen producing unconventional gas wells,
2. Bacterial Genomics of Listeria,
3. Environmental Genomics,
4. Genomics of Hemp species and Flax.

Abstract #31

A computational approach to identify PAX3 target genes in the neural crest

Kirby A. Ziegler[1] and D. Alan Underhill,[1,2] Departments of [1]Oncology and [2]Medical Genetics, Faculty of Medicine & Dentistry, University of Alberta, Edmonton AB, Canada

The transcription factor PAX3 plays critical roles in the neural crest (NC) cell lineage where its dysfunction is associated with birth defects and cancer. This lineage exists transiently during embryonic development and gives rise to most of the skeletal structures of the face, the majority of the peripheral nervous system, melanocytes of the epidermis, cochlea and retinal pigment epithelium, and portions of the cardiac outflow tract. The significant pathogeneses caused by PAX3 loss or gain-of-function reflect its ability to make executive decisions on cell division, differentiation and fate by regulating discrete gene expression programs. Nevertheless, we have only a rudimentary knowledge of the target gene networks that underlie normal and pathogenic PAX3 activity in the NC lineage. To this end, we have used a bioinformatics approach to identify candidate PAX3 target genes on a genome-wide scale using enhancer annotation data. PAX3-specific position weight matrices (PWMs) were used to query chromatin immunoprecipitation-nextgen sequencing (ChIP-seq) datasets that demarcate active and poised enhancers in the NC. The sequences recovered were annotated for all genes within 100kb and then filtered for conservation in placental mammals using PhastCons probabilities. In addition, associated genes were assessed for relative expression in NC compared to neuroectoderm and embryonic stem cells using RPKM (reads per kilobase per million mapped reads) data. This identified transcripts that were specifically up or down-regulated in the NC, which is another important criteria for their regulation by PAX3. Collectively, these analyses provided a prioritized list of potential PAX3 target genes, including the transcription factors TFAP2A and TFAP2B, which are important modulators of the NC lineage. Further validation of these targets will give key insight into how PAX3 governs cell fate decisions in the NC and how these are altered in disease.

Abstract #32

Investigating the Biological Function of ClpP in *E. coli* via Stable Isotope Metabolic Labeling and Mass Spectrometry

Devin K. Ulvestad, Angela W.S. Fung, and Richard P. Fahlman

There are two aims to this project: 1.To determine the rate of metabolic labeling in *Escherichia coli*. 2.To develop a quantitative metabolic labeling method for determining the specific substrates degraded by the ClpP protease in *E. coli*. In this experiment, the metabolic labels are 'heavy' [$^{13}C_6$] lysine and 'light' [$^{12}C_6$] lysine. To investigate the rate of metabolic label incorporation, *E. coli* cells are grown overnight in a metabolically labeled media and then transferred to a differentially labeled media. The rate at which the newly introduced metabolic labels incorporate into the proteome is monitored with Liquid Chromatography in combination with Tandem Mass Spectrometry (LC-MS/MS). It is found that the majority of proteins reach a level of adequate label incorporation after ~180 minutes. A select few proteins did not show this trend and are required to go through all growth phases to reach adequate label incorporation. Differentially labeled Δ*clpP* and wild-type K-12 *E. coli* proteins are identified and quantified by LC-MS/MS to identify proteins that increase in abundance in the Δ*clpP* knockout strain. Proteins that increased in abundance are identified and provide potential substrates for the ClpP protease.

Abstract #33

Neuroendocrine mechanisms and potential post-transcriptional level of regulation as contributing factors for cachexia in mouse model system

A Narasimhan[1], D Marks[3], T Braun[3], V Baracos [2] and S Damaraju[1]

Depts. of [1]Laboratory medicine and pathology, [2]Oncology University of Alberta, [3]Oregon Health &Science University

Cancer cachexia is a multifactorial syndrome characterized by involuntary weight loss, skeletal muscle depletion and negative protein balance. Action of inflammatory cytokines on hypothalamus-pituitary-adrenal axis (HPA) contributes to skeletal muscle atrophy. Insights in to alternative splice variants (ASV), their ratios relative to full length transcripts as well as mRNA stabilities, are some of the post-transriptional levels of gene regulation potentially contributing to the observed phenotypic variations. Up to 70% of the mouse genes were shown to be alternatively spliced, and higher proportion has been implicated in humans but their relevance to health and disease states is not well understood. We hypothesized that post-transcriptional regulation contributes to skeletal muscle atrophy when HPA axis is activated by inflammatory cytokines. Objectives are: i) to identify ASVs in rodents treated with Interleukin-1β; and (ii) to identify if the ASVs harbor AU rich elements (or AREs in 3' un-translated regions), vis-à-vis, further regulation of genes through mRNA stability Methods- Treatment group was given intracerebroventricular injection of IL-β1 and saline was given for the control group. Total RNA was isolated from Gastrocnemius muscle and the expression profiles were captured using Affymetrix mouse exon expression array. Alternative spliced variants were identified using GeneSpring Gx 12.5. AREs were identified using ARE database 3.0. Pathways of the spliced variants were identified using Ingenuity Pathway Analysis (IPA). Results and conclusions- We identified 40 differentially spliced variants showing > 1.5 fold change (treatment vs. the control group, statistical significance at $p<0.05$ and a False Discovery Rate at <0.05). We observed that exon skipping (62.5 %) was the dominant mechanism for ASVs. Further, 60% of the ASVs also showed AREs, implying that the stability of mRNAs, presumably from the early response genes may contribute to gene regulation and phenotypic variance. As expected, we identified inflammatory and glucocorticoid signaling pathways as highly significant ($p<0.05$). We will use RT-PCR to validate the findings. Similar detailed molecular and mechanistic understanding of human muscle catabolism may help develop targeted interventions.

Abstract #34

Profiling of small RNAome by next generation sequencing: Identification of potential prognostic factors to guide treatments in breast cancer

P Krishnan[1], A Narasimhan[1], J Mackey[2], S Ghosh[2], O Kovalchuk[3] and S Damaraju[1]

Depts of [1]Laboratory Medicine and Pathology, and [2]Oncology,University of Alberta; [3]Biological Sciences, University of Lethbridge

Background: Breast cancer is the second leading cause of cancer related death among women. Despite several treatment interventions, up to 30% of breast cancer patients show disease recurrence. Molecular biomarkers to stratify breast cancer patients into treatment responders vs. non-responders if available would help strategize treatments and potentially improve outcomes. We hypothesize that small RNAome dysregulation contributes to inter-individual variation in response to therapy. Objective of the work is to profile small RNAs from treatment naïve breast tumors and to identify candidate small RNAs as potential prognostic markers. Biospecimens and methods: Patients treated with adjuvant taxotere, adriamycin and cyclophosphamide were considered for this study. Formalin fixed, paraffin embedded specimens were used for total RNA isolation from 11 normal (reduction mammoplasty) breast tissues and 57 breast tumors. Next generation sequencing (NGS) was carried out using Illumina GA IIx platform to profile small RNAome. GeneSpring GX software from Agilent Technologies was used to identify differentially expressed ($p<0.05$, FDR cut off 0.05, fold change ≥ 2) small RNAs. Univariate analyses for recurrence-free survival (RFS) or Overall survival (OS) were carried out using Cox's proportional hazards model. Results and conclusions: We identified 86 miRNAs, 82 tRNAs, 150 snoRNAs, 15 snRNAs and 5 tRNA pseudogenes to be differentially expressed. Of these, 17 miRNAs, 15 snoRNAs, 10 tRNAs and 1 snRNA also showed significant association with both Overall survival (OS) and recurrence free survival (RFS) and a subset may be novel and have not been previously linked to outcomes. Further validation using RT-PCR is underway.

Abstract #35

IDENTIFICATION OF LIGHT-RESPONSIVE LIGANDS THROUGH SCREENING OF A LIGHT-RESPONSIVE PHAGE DISPLAY LIBRARY

Reza Jafari[1], Lu Deng[1], Simon Ng[1], Wadim L. Matochko[1], Anthony Zeberoff[2], Anastasia Elias[2], Pavel Kitov[1], John S. Klassen[1], Ratmir Derda[1]

[1]Department of Chemistry and Alberta Glycomics Center, University of Alberta, Edmonton, Alberta, T6G 2G2, Canada
[2]Department of Chemical and Material Engineering, University of Alberta, Edmonton, Alberta, T6G 2G2, Canada

Light responsive ligands are useful tools to study cellular mechanisms because their function can be turned ON and OFF with a high temporal and spatial resolution. Light responsive ligands can be designed by attaching a small photo-switchable core molecule (e.g. an azobenzene structure) to a specific ligand of a receptor. The structural change in the photo-switch that occurs in response to irradiation with light (cis/trans isomerization in the case of azobenzene) can perturb ligand's structure and change its binding properties. The conventional route to design such ligands, however, is challenging, because it needs information about the structure of the ligand, or receptor, or both. Another approach for identification of light-responsive ligands is screening of a light-responsive library against a specific receptor to yield light-responsive binders. The advantage of this combinatorial approach is the potential to screen the library against any target, independent of the target's structure.

A phage display library consists of a library of peptides or proteins that are displayed on the coat protein of the bacteriophage M13. Screening of such library against protein targets has found vast applications in identification of binders for various types of receptors. We generated of a light-responsive phage display library by assembling a photo-switchable light-responsive core to the phage library coat protein. Using this strategy, we generated a library with the complexity of $>2\times10^8$ light-responsive macrocycles. We developed a method for characterization of chemicals reaction on phage particles which can quantify the yield of the reaction and can be used to optimize the chemical modification on phage. We showed that the light responsive library can be screened against a protein target (i.e., streptavidin) to yield photo-switchable ligands that can change their binding affinity by exposing to 370 nm light. Since the screening resulted in several ligands with different binding properties, a heterodimer of two ligands can potentially show improved binding affinity. The method can be used to identify light-responsive ligands, against any target.

Abstract #36

One-Stop Shop for Glycomics Research and Development

Elizabeth Nanak, Alberta Glycomics Centre, Department of Chemistry, University of Alberta, Edmonton

The Alberta Glycomics Centre is an Alberta Innovates funded research and commercialization centre focused on the development of carbohydrate-based solutions with medical applications. Building on the expertise and the capacity developed in the last ten years, the Centre now offers core services for glycomics research and development, including carbohydrate synthesis, glycan structure-function analysis, glycoengineering and glycan screening. The Centre provides services and works with clients and partners to interpret data from even the most complex glycoconjugates.

Abstract #37

Genomic and Proteomic Studies on the Solvent Tolerant *Staphylococcus warneri*

Victor Cheng[1], Glen Zhang[1], Feifei Fu[2], Steve Oyedotun[1], Doug Ridgway[1], Mike Ellison[1], Liang Li[2] and Joel H. Weiner[1]

Departments of Biochemistry[1] and Chemistry[2], University of Alberta, Edmonton, Alberta, Canada

Staphylococcus warneri is a Gram-positive bacterium commonly found in human skin flora. This bacterium is tolerant to a variety of organic solvents including alkanes, short-chain alcohols, and cyclic aromatic compounds. The genome of a laboratory *S. warneri* isolate, strain SG1, was sequenced to explore its mechanism of solvent tolerance and its potential as a chassis for biofuel production. A two-dimensional liquid chromatography mass spectrometry (2D-LC-MS) shotgun approach, in combination with quantitative 2-MEGA (dimethylation after guanidination) isotopic labeling, was carried out to compare the proteomes of SG1 grown under butanol-free and butanol-challenged conditions. In total, 1585 unique proteins (representing 65% of the predicted open reading frames) were identified, covering all major metabolic pathways. Of the 967 quantifiable proteins by 2-MEGA labeling, 260 were differentially expressed by at least 1.5-fold. These proteins are involved in energy metabolism, oxidative stress response, lipid and cell envelope biogenesis, or have chaperone functions.

Abstract #38

Understanding Quantitation Variability in Metabolomics 1D-^{1}H-NMR Spectra: Synthetic Mixtures, Urine, and Insights on Pulse Sequences and Robotic Sampling

Stanislav Sokolenko[1], **Ryan M^{c}Kay**[2], Eric JM Blondeel[1], Michael J Lewis[3], David Chang[3], Ben George[1], and Marc G Aucoin[1]

[1] Waterloo Institute for Nanotechnology, Department of Chemical Engineering, University of Waterloo, Waterloo, Ontario, N2L3G1, [2] **Presenter**: Department of Chemistry, University of Alberta, Edmonton, Alberta, T6G 2G2, [3] Chenomx Inc., Edmonton, Alberta,

The growing use of 'targeted profiling' approaches for the deconvolution of 1D-^{1}H-NMR spectra by comparing to an isolated compound library has created the need for an in-depth characterization of quantitation variability. We have explored the underlying source of variability in ^{1}H-NMR spectroscopy (*e.g.* tube/insertion, spectra acquisition, and profiling) as well as a number of other factors that have not been dealt with in the past, such as temporal consistency of repeated NMR scans, human consistency in repeated profiles, and human vs. machine sampling. We have examined the effect of different NMR pulse sequences and the differences between acquired spectra and the peak reference library. Our investigation into variability sources revealed that apart from profiling, sample insertion and/or shimming can play a significant role final quantification, a finding that is equally applicable to all methods of quantification. Both sources of error were also found to have temporal relationships, with bias identified as a function of both scan and profiling order.

Abstract #39

Dependence of quantitative NMR signal variation due to common spectrometer parameters changes and different probes.

Paige Lacy[1]*, **Ryan M[c]Kay**[2]*, Michael Finkel[3], Alla Karnovsky[4], Scott Woehler[5], David Chang[6], and Kathleen A. Stringer[3,4]

[1]Pulmonary Research Group, Department of Medicine and [2]**Presenter** Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada; [3]Department of Clinical Sciences, College of Pharmacy, [4]Department of Computational Medicine and Bioinformatics, School of Medicine, [5]Department of Medicinal Chemistry and the Biochemical Nuclear Magnetic Resonance Core, College of Pharmacy, University of Michigan, Ann Arbor, MI, USA; and [6]Chenomx, Edmonton, Alberta, Canada. *These two authors contributed equally to the work.

To our knowledge, multicentre validation of 1D-^{1}H-NMR metabolomics data, whether purified or derived from complex biological samples such as human urine, has not been reported. To test the reproducibility of quantified metabolite data using multiple sites via a targeted profiling approach, technical replicates of urine samples from 19 healthy subjects were analyzed on 600 MHz and 500 MHz magnets at the University of Alberta and University of Michigan, respectively. Subsequent analysis of metabolites using standard statistical techniques not only revealed that quantitative data can be achieved from human urine samples measured at multiple sites, but also demonstrated that spectrometer parameters can have dramatic and widespread perturbation of results. In particular, modification of saturation power settings resulted in a hitherto unrecognized scale of solvent suppression, even with purified metabolites prepared in D_2O. We present here a validation of NMR analysis at multiple sites, supporting the use of NMR for metabolomics on complex human biological samples. We also report the range and magnitude that common NMR parameters involved in solvent suppression can have on quantified results, and make recommendations so that future studies may avoid these relatively easily made errors in data collection and analysis.

Abstract #40

Strategies in the Detection of Metabolites using Surface Plasmon Resonance

Lars Laurentius and Mark T. McDermott

Department of Chemistry and National Institute for Nanotechnology, University of Alberta, Edmonton, Alberta, Canada.

The detection and quantitation of metabolites in bodily fluids are of importance in the screening for diseases and toxicity. Due to the small size and sheer number of metabolites, detection strategies typically are time-consuming and involve costly instrumentation such as mass spectrometry and nuclear magnetic resonance spectroscopy. The goal is to utilize other detection methods that allow for cost effective and rapid detection.

We are developing a surface-based bioassay that allows for the simultaneous detection of multiple metabolites. A combination of different capture agents such as antibodies, aptamers and periplasmic binding proteins will facilitate the needed specificity to screen for a large number of metabolites. The detection is based on surface plasmon resonance imaging and in order to compensate for the low molecular weight of metabolites a competitive assay is employed. Metabolite modified gold nanoparticles are used to compete for binding to the capture agents. The strategy, design and preliminary results will be discussed.

Abstract #41

The Evolution of the Indole-3-acetic Acid Pathway in Algal and Bacterial Genomes

Harjot Atwal, Leen Labeeuw, Rebecca Case

Department of Biological Sciences. University of Alberta

Auxins are a class of plant hormones, responsible for growth and development, such as cell division and enlargement, as well as tissue differentiation. One of the most abundant and important members of this family is Indole- 3-acetic acid (IAA). This compound was originally thought to exist only in plants, however the production pathways have since been characterized in bacteria, such as marine bacteria of the roseobacter clade. Interestingly the IAA pathway may have analogues in algae. The objectives of this project were threefold: to investigate the presence of the IAA production pathways in the roseobacter clade; to investigate the presence of the IAA production pathways in algae; and to search for evidence of lateral gene transfer between algae and roseobacters. These objectives were investigated using a bioinformatics approach. The pathways were found to be widely spread in the roseobacter clade, and were also found to be present in algae. Evidence was found indicating both lateral gene transfer between algae and roseobacters and that roseobacters were responsible for the transfer to algae.

Abstract #42

DRUG DISCOVERY FOR INFLAMMATORY BOWEL DISEASE

Yahya Fiteih, Marilyn Gordon, Le Luong, and Shairaz Baksh

Department of Pediatrics, Faculty of Medicine and Dentistry
University of Alberta

Inflammatory bowel diseases (IBD) (Ulcerative colitis and Crohn's Disease) are chronic intestinal diseases characterized by inflammation of the gastrointestinal area resulting in abdominal pain, chronic diarrhea and weight loss. The Ras association domain family protein 1A (RASSF1A or 1A) is a tumour suppressor protein that regulates TLR signaling by restricting NF-κB activity, protecting against intestinal inflammation and stimulating repair following inflammation-induced damage. *Rassf1a*$^{-/-}$ mice have < 20% survival, elevated NF-κB activity and exacerbated colitis following inflammation-induced injury in a dextran sulphate sodium (DSS) model of colitis. 1A can form complexes with Toll receptor (TLR) components (such as MyD88 and TRAF6) to restrict the activation of NF-κB to suggest a novel role of RASSF1A in protection against DSS-induced inflammation by restricting NF-κB pathways.

1A is a component of Rassf/Salvador/Hippo pathway that can modulate cell death and proliferation. In addition to modulating NF-κB activity, we have uncovered a novel tyrosine phosphorylation (pY) of Yes-associated protein (pY-YAP) in colonic sections from DSS-treated *Rassf1a*$^{-/-}$ mice resulting in increased apoptosis and autophagy. Phosphorylation of YAP proceeds independent of MST1/2 and possibly LATS. Inhibition of the tyrosine phosphorylation of YAP with gleevec (a protein tyrosine kinase [PTK] inhibitor) or inhibition of the autophagic response with 3-methyladenine (3-MA) resulted in > 70% survival of DSS-treated *Rassf1a*$^{-/-}$ mice.

We will present data to propose the use of PTK inhibitors and autophagy inhibitors to promote the efficient recovery in animals with a compromised inflammatory response pathway, such as in the *Rassf1a*$^{-/-}$ mice. We speculate that both gleevec and 3-MA may be novel therapeutic approaches to enhance recovery from inflammation-induced injury and be useful in treating IBD patients.

FUNDING: WCHRI, CFI, Stollery Children's Hospital Foundation/Hair Massacure Fund.

Abstract #43

Characterizing Human Diseases and Human Biofluids using Multi-Platform Metabolomics Techniques

Constance Sobsey[1], Rupasri Mandal[1], Ram Krishnamurthy[1], Souhaila Bouatra[1], Igor Sinelnikov[1], An Chi Guo[1], Philip Liu[1], Kruti K. Chaudhary[1], Edison Dong[1], Farid Aziat[1], Faizath S. Yallou[1], Tom Blydt-Hansen[2], Ray Bahado-Singh[3] and David S. Wishart[1]

[1]Departments of Biological and Computing Sciences, University of Alberta, Edmonton, AB Canada T6G 2E8
[2]Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, MB Canada R3E 3P4
[3]Department of Obstetrics and Gynecology, Wayne State University, Detroit, MI, USA 48201

TMIC (The Metabolomics Innovation Centre) is Canada's national metabolomics platform. It was officially launched in the summer of 2011. As a national platform, TMIC has several mandates including the provision of low-cost metabolomics services to academic and industrial researchers, the maintenance of freely available metabolomics databases and web servers (HMDB, DrugBank, T3DB, MetaboAnalyst) and the development of improved or more comprehensive metabolomics assays. TMIC specializes in performing quantitative metabolomics assays on human, animal, plant and microbial samples using a wide range of technologies including NMR, GC-MS, LC-MS/MS, LC-FT/MS, HPLC-UV/FD, ICP-MS and HPLC-ELSD-FAMES-MS. In order to keep pace with the rapid technology developments in metabolomics, TMIC is constantly working towards developing, acquiring, testing and implementing new metabolomic technologies.

TMIC is also involved in the systematic characterization of human biolfuids. This effort, which is an extension of the Human Metabolome Project, is aimed at providing reference values for all identifiable or quantifiable metabolites in blood, urine, cerebrospinal fluid and saliva. Using multi-platform metabolomic techniques along with extensive literature surveys, TMIC reported the measurement of 3687 metabolites in the human serum metabolome. We recently finished updating the human cerebrospinal fluid metabolome and were able to identify and quantify a total of 419 metabolites. In an ongoing study for human urine metabolome, more than 813 metabolites have been identified, quantified and catalogued. These efforts represent the most comprehensive metabolomic characterization of any human biofluid achieved to date. The biofluid data are now publicly available through the TMIC website. Using the same multi-platform metabolomic methods used in these biofluid studies, TMIC has also undertaken several disease biomarker studies including a study aimed at developing non-invasive monitoring techniques to identify patients at risk for organ rejection and another aimed at first trimester prediction of early- and late-onset preeclampsia (Figures 1 and 2). These applications will be presented in detail. Here we intend to show the potential that multi-platform, quantitative metabolomics holds in contributing to our understanding of health, disease development and human physiology.

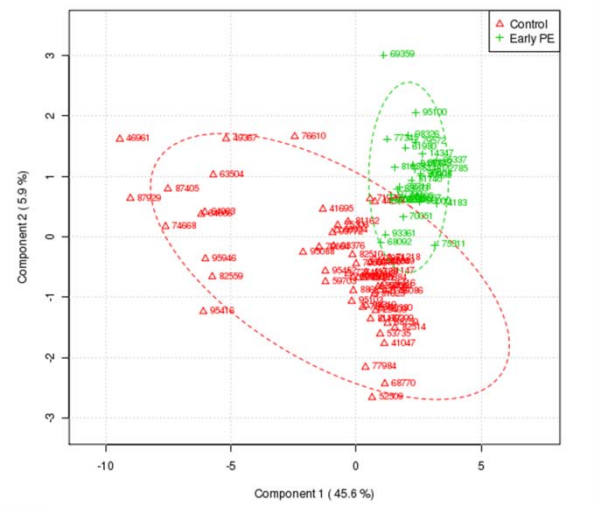


Figure 1. PLS-DA plot showing the separation between the cases of early preeclampsia (in green) an d controls (in

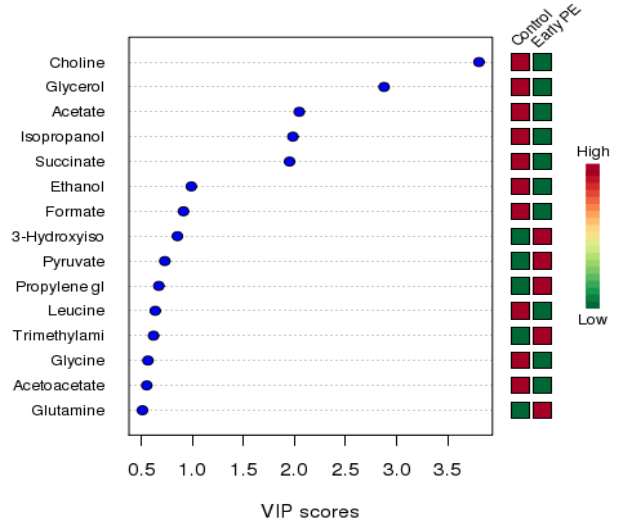


Figure 2. Variable Importance in Projection (VIP) plot listing the most important discriminating metabolites in descending order of importance.

Abstract #44

Isotope Labeling Liquid Chromatography-Mass Spectrometry in Metabolite Biomarker Discovery for Discriminating Normal Aging and Mild Cognitive Impairment Disease

Ngoc Tran Tran (ttt1@ualberta.ca) [1], Roger A. Dixon[2] and Liang Li[1]

Departments of [1]Chemistry and [2]Psychology, University of Alberta, Edmonton, Alberta, Canada

Introduction

Mild Cognitive Impairment (MCI) is defined as a clinical state characterized by significant cognitive impairment in the absence of dementia. MCI is known as the transition between normal aging and the prodromal phase of Alzheimer's disease (AD). Early diagnosis of MCI and AD can assist in management and treatment of the diseases. There is a pressing need for an improved, rapid and sensitive screening method for diagnosing MCI patients. In this work, we apply the strategy of differential $^{12}C/^{13}C$-dansylation labeling of human salivary samples, in conjunction with Liquid Chromatography-Mass Spectrometry (LC-MS), for comprehensive and quantitative profiling of the salivary metabolome with an objective to discover metabolite biomarkers. Herein we present a preliminary work for discovering the discriminant metabolites for MCI using a set of 82 clinical saliva samples from MCI and AD patients and healthy aging volunteers.

Methods

Each saliva sample was labeled with ^{12}C-dansyl chloride and the pooled saliva from all 82 samples was labeled with ^{13}C-dansyl chloride. A ^{12}C-dansyl labeled pooled saliva mixture at various dilutions was utilized to generate a calibration curve for total metabolite concentration determination through a fast gradient UPLC-UV run. The ^{12}C- and ^{13}C-labeled saliva samples were amount-normalized prior to mixing to ensure a 1:1 concentration ratio. They were subsequently analyzed using LC-MS. The MS data was extracted by an in-house peak picking program to generate the peak pairs detected. Statistical analysis was performed on the obtained data set to compare the two classes of samples and to generate the potential MCI metabolite biomarkers.

Preliminary data

The new protocol took into account of different solvents used for protein precipitation as well as omitting the precipitation step due to non-existent protein amounts present in saliva samples collected. The LC gradient and sample injection amount were also optimized for better ion pair detection. Five μL of saliva sample was used as the starting material in each labeling reaction, and about 2006 putative metabolites on average have been found, which demonstrates a significant improvement compared to past publication. Various statistical models (PCA, OPLS-DA, etc) have been applied to the LC-MS data to draw a comparison among different groups. Using OPLS-DA, separations among the healthy control (normal aging) group, MCI diseased group and AD diseased group were observed, thus leading to the discovery of discriminant metabolites. We will continue analyses of the present and larger data sets in order to discover, verify, and validate key metabolite biomarkers for discriminating normal aging, MCI, and AD groups.

Novel aspect

An improved isotope labelling LC-FTICR-MS method is developed for metabolite biomarker discovery using human salivary samples.

Abstract #45

Abstract #45

An improved isotopic labeling protocol for LC-MS metabolomic profiling of carboxylic acids in biofuids and cell extracts

Jun Peng and Liang Li

Departments of Chemistry, University of Alberta,

Introduction

Our lab has developed a high performance isotopic labeling reagent, ^{12}C- and ^{13}C-coded *p*-dimethylaminophenacyl (DmPA) bromide, that can enhance the detection and quantification of carboxylic acids by LC-MS significantly. The current acid labeling protocol has some limitations for metabolomic profiling; for example, some high abundant amine compounds in the sample could also be labeled which may suppress the signals of some low abundant acids. The presence of water in the sample could affect the labeling efficiency. Here we report an improved protocol for this labeling reaction by applying liquid-liquid extraction to separate the high abundant amine compounds and remove the water content in the sample.

Method

Urine samples can be used directly for processing and labeling. Serum, saliva, and cell extracts need to be pretreated by protein precipitation, then concentrated and re-dissolved into aqueous solution. A aliquot of 90 uL urine sample, or pretreated serum, saliva or cell extracts, was mixed with 10 uL 6M HCl,10 uL saturated NaCl, and then extracted using 300 uL ethyl acetate. The organic phase was transferred into another reaction vial. The sample was adjusted pH to 8 and was dried down by using SpeedVac. An Agilent HPLC coupled with a Bruker Apex-Qe 9.4-T FT-ICR-MS was used to collect the metabolomic profiling data. XCMS and our in-house written software were used to pick the $^{12}C/^{13}C$ ion pairs.

Preliminary Data

We first demonstrated the performance of the new protocol by using two standard mixture of ethanolamine and hippuric acid. The new protocol by liquid-liquid extraction effectively removed the peak of ethanolamine at an accurate mass of 384.2233 in the mass spectrum, and the hippuric acid peak at an accurate mass of 341.1148 was increased by at least 2-fold. We also tested other standards individually, such as asparagine, lactic acid and malic acid. The data showed that asparagine signals could be largely removed, and lactic and malic acid signals could be improved significantly. The LC-FTICR-MS results showed that the new extraction protocol could detect 1424 ion pairs in a urine sample; in contrast, the existing protocol could detect 776 ion pairs under the same instrument condition and data processing method. By comparing the new protocol to the old protocol, it was shown that amine metabolites, such as ethanolamine, ethylamine and glycine, in the urine were mostly removed, while at the same time for most acids, such as citic acid, hippuric acid and lactic acid, peak intensity was significantly increased. In addition, the peak from hydrolyzed acid labeling reagent was much reduced. The UPLC-UV results also confirmed that the peak intensities of the standards and the urine sample were much improved using the new protocol. We also applied the new protocol to the human serum sample and preliminary results indicated that the new protocol could detect 443 ion pairs with high intensity, compared to 392 ion pairs detected by the old protocol; optimization of acid extraction from the serum samples is still underway. In the near future, this new protocol will be applied for the metabolomic profiling of carboxylic acids in the saliva and cell extracts, and the results will be presented. We envisage that this improved acid labeling protocol for carboxylic acids, in combination with isotopic dansylation labeling for amine metabolites, will greatly enhance the metabolome coverage; up to several thousands of metabolites may be profiled using a single 25-min LC-MS run.

Novel aspect

An improved isotopic labeling protocol for metabolomic profiling of carboxylic acids in biofluids and cell extracts was developed.

Abstract #46

Comprehensive Phosphorylation Site Analysis of α- S2 Casein Using Microwave-Assisted Acid Hydrolysis and Phosphopeptide Enrichment.

Zhendong Li, Nan Wang and Liang Li
Department of Chemistry, University of Alberta

Introduction

Traditional bottom-up phosphorylation site analysis of a protein uses trypsin and other enzymes to produce short peptides for tandem MS analysis. Due to the specificity of enzymes, only a small number of peptides are usually generated which may not cover the entire sequence of a protein. In addition, sometimes multiple phosphorylation sites are present in a single peptide making the exact localization of these sites very difficult. Microwave-assisted acid hydrolysis (MAAH) can generate a large number of non-specific peptides from a protein. Combined with phosphopeptide enrichment, this technique can be used to generate many phosphopeptides with sequence overlaps which greatly increases the chance of isolating individual phosphorylation sites for a detailed mapping of phosphorylation sites. In this work, we demonstrate the use of a combination of bottom-up techniques for mapping all the phosphorylation sites of α-S2 Casein.

Methods

α-Casein sample with a low percentage of α-S2 Casein was digested with MAAH and the phosphopeptides were then enriched by Fe-based immobilized metal affinity chromatography (IMAC) and TiO_2 beads. The phosphopeptides were analyzed using a quadrupole time-of-flight (QTOF), and an Orbitrap analyzer. To enhance the detection of predicted phosphopeptides, inclusion lists were used. Further enhancement was done by enriching α-S2 Casein with weak anion exchange followed by the same analysis as the non-enriched sample. To analyze one specific site that was missing, trypsin digestion was used; the tryptic peptides were analyzed by a MALDI TOF/TOF.

Preliminary data

Without IMAC or TiO2 enrichment only 108 peptides belonging to α-S2 Casein was found, and only 5 out of the 13 known phosphorylation sites were covered. Following separate enrichments with IMAC and TiO2 the number of phosphopeptides greatly increased and 10 out of the 13 known phosphorylation sites were found. Due to the low abundance of α-S2 Casein in the α-Casein sample, not all peptides containing isolated phosphorylation sites were detected. The use of an inclusion list containing calculated masses of low abundance phosphopeptides proved ineffective.

Weak anion exchange was then performed at the protein level, and α-S2 Casein was greatly enriched. After MAAH of the enriched α-S2 Casein, followed by phosphopeptide enrichments and QTOF and Orbitrap analysis, 800 α-S2 Casein peptides were found compared to 160 α-S1 Casein. One previously unreported site was found in the middle of the protein.

Redundant phosphopeptides, showing multiple phosphorylation states, were found for all but three phosphopeptides. These three phosphopeptides were missing from all analyses. The redundant phosphopeptides show a complicated phosphorylation pattern, with most sites being variably modified.

In order to find the remaining phosphopeptides trypsin digestion on the α-S2 Casein enriched sample was performed, and analyzed on the MALDI-TOF/TOF. Tryptic peptides containing the three missing phosphorylation sites were found with multiple phosphorylation states.

Finally, a detailed phosphorylation map was compiled with data from all the complementary techniques employed in this study, covering all the known sites and one unknown site.

Novel aspect

A detailed analysis of phosphorylation on α-S2 Casein with MAAH, phosphopeptide enrichment and MS analysis with QTOF, Orbitrap, and MALDI-TOF/TOF.

Abstract #47

IsoMS: A High-Throughput Data Analysis Software for Extracting Quantitative Information from Data Generated by Differential Isotope Labeling LC-MS

Ruokun Zhou, Chiao-Li Tseng, Liang Li
Department of Chemistry, University of Alberta

Introduction

Differential isotope labeling LC-MS has been used to quantify a certain class of metabolites such as all amine-containing metabolites or amine-containing-metabolome in biological samples. After mixing two comparative samples separately labeled by light or heavy isotopic tags, the peak intensity ratio of the isotope labeled analyte pair can provide relatively or absolutely quantitative information on the analyte. Although high quality quantitative data can be obtained by this method, data processing for handling a large number of LC-MS runs from a large scale metabolome profiling work with many comparative samples can be very challenging. We have developed a data processing tool (IsoMS) to peak the ion pairs, group the isotopic peak pairs, as well as filter the noises. IsoMS can rapidly process hundreds of LC- MS datasets with better than 95% accuracy in ion pair picking.

Method

Chemicals: Amine standards and reagents were purchased from Sigma-Aldrich (Oakville, ON). The stock solutions of the standards (10 mM each) were prepared in water: ACN (50:50 (*v/v*)) and stored at -20 °C.

Human Urine samples**:** The second/third morning urines before drinking coffee and the fourth/fifth morning urines after drinking coffee were collected on a daily basis from one healthy volunteer for 5 days. The samples were stored at -80 °C.

Dansylation: The amines standards and urine samples were derivatized by the method described by Guo and Li (Anal. Chem. 2009, 81, 3919)with some modification.

Instruments: Agilent Zorbax Eclipse Plus C18 (2.1 mm × 100 mm, 1.8 μm particle, New Castle, DE), Agilent 1100 series (Palo Alto, CA), Bruker 9.4T Apex-Qe FTICR (Billerica, MA), Agilent ESI-TOF and Waters QTOF MS.

Preliminary Data

IsoMS was developed to accurately pick up the ion pairs detected in differential isotope labeling LC-MS, calculate the peak ratios of individual metabolites, and convert the quantitative data into a form readable by various biostatistics tools. Only CSV format can be recognized by IsoMS and thus a VB script has been developed to convert the Bruker FTICR-MS data before processing. For LC-TOF and LC-QTOF data, ProteoWizard was used to transform the data formats into the one recognized by IsoMS. After the conversion, there are two filters prior to peak picking and pairing for the FTICR-MS data. The first filtration is used to remove defective spectra, and the second one is used to remove homogeneous peaks. IsoMS takes advantage of unique accurate mass differences and isotopic patterns which are determined by the labeling reagent used to search peak pairs formed between the light and heavy isotopic tag labeled metabolites in each spectrum. To evaluate the possibility of the mismatched pairs or false positive picking, each pair is marked as level one to level three based on its isotopic pattern. For instance, if IsoMS can find the first nature isotopic peak for each peak of a pair, the pair is marked as level one.

In LC-MS, metabolites may not be detected only as the protonated ions. Adducts, dimers and in-source fragments may be formed as well. In IsoMS, several filters are built in to eliminate the effect of multiple forms of the same metabolite on the dataset,. All the ion pairs derived from the same metabolite in continuous mass spectra found over a period of retention time are grouped into one pair, and the m/z and retention time of the most abundant light-chain labeled peak is chosen to represent the pair.

To test the accuracy and utility of IsoMS, five amine standards labeled by $^{12}C_2$-/$^{13}C_2$-DnsCl were detected by LC-FTICR-MS, and the data files were analyzed by IsoMS. In all samples, IsoMS correctly detected the isotopic peak pairs and filtered out the interference peaks. The performance of the IsoMS program was then examined using real world samples including biofluids, plant and cell extracts labeled with dansylation chemistry and it was found that the rate of false positive ion-pair pick was less than 5%. The software was then used to process the datasets of the labeled human urine samples collected before and after drinking coffee. In each sample, IsoMS could detect more than 600 ion pairs with high confidence. The pair list along with their relative ratios was used for statistical analysis from which the urine samples collected before and after drinking coffee were clearly differentiated. The effect of food and drinks on urine metabolome and how the food and drink metabolomes present in biofluids influence the discovery of disease biomarkers are currently under investigation.

Novel Aspect

A software, IsoMS, has been developed to rapidly and accurately process the data generated by differential isotope labeling LC-MS.

Abstract #47

Metabolomic analysis of a Staphylococcus variant SG1 cultured in the absence and presence of butanol

Feifei Fu[1]; Yiman Wu[1]; Victor Cheng[2]; Joel Weiner[2]; Liang Li[1]
[1]Deptartment of Chemistry, University of Alberta
[2] Department of Biochemistry, University of Alberta

Introduction

Microorganisms that can thrive in the presence of toxic organic compounds are of great interest in biofuel production and bioremediation. Our study focuses on the Gram-positive bacterium *Staphylococcus warneri* (strain SG1) which can survive up to 2.5% butanol. Our previous analyses on the genome and the proteome of SG1 offered insight into the butanol tolerance mechanism of this bacterium on the protein level. In order to obtain information about metabolic changes upon butanol challenge, we applied isotope-coded dansylation chemistry for profiling amine- and phenol-containing metabolites and isotope-coded p-dimethylaminophenacyl (DmPA) bromide chemistry for profiling carboxylic acid-containing metabolites, in combination with LC-MS, for metabolome analysis.

Method

SG1 cells were grown, in triplicate, in the absence and presence of 1.5% butanol and harvested at stationary phase. Whole cell lysates were acetone precipitated to remove proteins and the resultant supernatants were concentrated. Metabolite samples from $BtOH^-$ and $BtOH^+$-grown cells were combined to form the internal standard. All samples were then labeled with ^{13}C-dansyl chloride, and labeled metabolites were quantified by LC-UV using step gradient. ^{12}C-labeled individual samples were then mixed with ^{13}C-labeled standard in equal amounts and loaded on a liquid chromatography tandem fourier-transform ion-cyclotron resonance mass spectrometer (Bruker 9.4 Tesla Apex-Qe FT-ICR-MS) for analysis.

Preliminary Data

Biotriplicates of SG1 grown in the absence and presence of butanol were analyzed in technical duplicate, thus yielding a total of 12 data files from LC-FTICR-MS analysis. XCMS was used to pick up ^{12}C-/^{13}C-ion pairs of the same metabolites in the mixed sample. After data alignment of the 12 processed files, a total of 1848 peak pairs were extracted. Statistical analysis was done by online software Metaboanalyst 2.0 (http://www.metaboanalyst.ca). Principal component analysis shows a clear discrimination between $BtOH^-$ and $BtOH^+$ samples. We were able to extract 88 peak pairs from volcano plot at p-value < 0.01 and fold change >1.5 or <0.67 were considered significantly changed and their accurate masses were searched against HMDB for identification.

We were able to match approximately half of the most discriminant peak pairs in the $BtOH^-$ and $BtOH^+$ samples to HMDB compounds. Metabolites such as ornithine, cadarverine, spermine, methionine, methionine sulfoxide were unambiguously identified by matching both the accurate mass and retention time to the authentic standard under the same experimental conditions. Butanol is known to have a chaotropic effect on cell membranes and to cause oxidative stress to the cell. The up-regulation of cadarvarine and spermine were in accordance as they were reported to be antioxidants. In addition, methionine sulfoxide was reported to be a biomarker of oxidative stress. The up-regulation of the enzyme peptide-methionine (R)-S-oxide reductase that catalyzes the formation of methionine sulfoxide from methionine, as well as the increased ratio of methionine sulfoxide/methionine under the butanol-challenge condition, all indicate an oxidative stress to the cells in the presence of butanol.

Future work will focus on the comparative analysis of the carboxylic acid-containing metabolites using p-dimethylaminophenacyl (DmPA) bromide chemistry to study specific metabolic pathways which showed altered protein expression levels upon butanol challenge , such as the TCA cycle and glucuronate interconversion pathways.

Novel aspect

Applying quantitative metabolomics to better understand butanol tolerance mechanism of the Gram-positive bacterium *Staphylococcus warneri*.

Abstract #48

Optimization and applications of in-gel microwave-assisted acid hydrolysis (MAAH) of proteins for mapping protein sequences

Difei Sun, Nan Wang, Liang Li*
Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

Introduction

Protein sequence mapping is widely used for unambiguous identification of proteins, study of post translational modifications and discovery of amino acid substitutions from point mutations in the genome.

In-solution microwave-assisted acid hydrolysis (MAAH) is an effective method to generate high protein sequence coverage. However, it requires a relatively pure protein sample which can be difficult to prepare because of limited resolution in solution-based separation methods. In contrast, SDS-PAGE is commonly used for protein separation because of its high resolution.

In this work, we report a method based on the use of in-gel MAAH of proteins to produce peptides for LC-ESI MS/MS analysis. We demonstrate that this method can be used for full sequence analysis of various types of proteins separated by gel electrophoresis.

Method

For method development, 40 µg of standard protein (BSA) was loaded to each lane of 12% polyacrylamide gel. After separation and fixation, the gel was stained with coomassie blue and then destained with water to get clear gel bands. The BSA band was excised into pieces and ready for in-gel MAAH.

The gel pieces were dehydrated and then dried down. Twenty-five percent TFA was added and the protein in gel was acid-hydrolyzed by using microwave irradiation. After in-gel MAAH, peptides generated were extracted by using the solution which contained 0.1% TFA, 85% acetonitrile and 15% water. The reduction and alkylation of peptides were performed by using DTT and IAA. After desalting, the purified peptide mixture was analyzed by LC-ESI MS/MS.

Preliminary data

By using a standard protein, BSA, as the model, the in-gel microwave-assisted acid hydrolysis method was developed and optimized. First, by using 40 µg BSA, the irradiation time was varied from 6 min, 8 min to 10 min to compare the unique peptide number and protein sequence coverage. The results showed that, by using optimized 10 min irradiation time, the number of identified unique peptides and sequence coverage were 1107 ± 42 (n=3) and 99% ± 1% (n=3), respectively.

Another set of experiments were performed to examine the effect of protein loading amount on the detectability of the peptides generated from in-gel MAAH. In this case, the standard protein BSA was used as the model with a sample loading of 4 µg, 2 µg, 1 µg and 0.5 µg. These results show that the number of peptides identified decreased significantly as the protein loading amount changed from 4 µg to 0.5 µg and the protein sequence coverage obtained was 99% ± 1% (n=3), 94%±3% (n=3), 76%±5% (n=3) and 32%±2% (n=3) for 4 µg, 2 µg, 1 µg and 0.5 µg of protein loading, respectively. For 4 µg protein loading to the gel, 689 ± 54 (n=3) unique peptides were identified. Thus, the in-gel MAAH LC-MS/MS method can generate near 100% sequence coverage using a small amount of proteins.

We have successfully applied this method for sequencing proteins in various research areas. One example was the use of this method to identify a novel protein form that was formed from truncation of a mature form during the signal transduction related to heart disease. Another example was for characterizing human plasma proteins where nineteen high abundant proteins including IgG could be identified with high sequence coverage.

Novel aspect

The full sequence of a gel-separated protein can be mapped by using in-gel MAAH and LC-ESI MS/MS.

Abstract #49

Effective Extraction Method and Stable-Isotope Dansylation Labeling Combined with RPLC-FTMS for the Analysis of Arabidopsis Thaliana Metabolome

Chiao-Li Tseng; Michael Deyholos; Liang Li
Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

Introduction

Plant metabolomics has become an important area of research for understanding plant biology. Arabidopsis thaliana has a small size of genome and rapid life cycle and thus is a widely used model system for understanding the genetic, cellular, and molecular biology of flowering plants. The objective of this work is to develop a LC-MS method for studying the metabolome of arabidopsis thaliana. In our group, ^{12}C-/^{13}C-dansylation reaction has been successfully developed as a stable isotope derivatization method for analyzing amine- and phenol-containing metabolites. In the present work, we have developed a robust metabolite extraction method from arabidopsis thaliana and applied the dansylation labeling method to improve the detectability of the arabidopsis thaliana metabolome by LC-MS.

Method

Arabidopsis thaliana was extracted by the microwave-assisted process. Briefly, 40 mg of sample was mixed with 80% organic solvent. The sample was subjected to microwave irradiation at 120 watt for 45 min. After centrifugation, the supernatant was evaporated to less than 0.1 ml. The final volume was adjusted to 0.5 ml by adding the deionized water.

The extracts were mixed with the Na_2CO_3/$NaHCO_3$ buffer and labeled by ^{12}C- or ^{13}C-dansyl chloride at 60℃ for 60 min. The labeled mixture was separated by the Agilent 1100 HPLC with the Zorbax XDB-C18 column and analyzed by the Bruker 9.4 T Apex-Qe FTICR mass spectrometer. The data were processed and analyzed to determine the putative metabolites and their relative abundances by our in-house developed software written with R language.

Preliminary data

Extracting metabolites is an important step for plant metabolome analysis. The conventional solvent extraction with vortex was too time-consuming and may not be very efficient. We focused on the development and application of the microwave-assisted extraction method. The optimal results were obtained with a simple, 45 minute microwave-assisted extraction of 40 mg arabidopsis thaliana in 1 mL of 80% of methanol.

A reproducible and high-quality separation method has been developed for analyzing the isotope-tagged metabolites from arabidopsis thaliana. After labeling by the stable isotope dansylation reagents, the very polar metabolites changed their chromatographic retention behavior. As a result, the reversed-phase chromatographic separation can be easily done in 31 min and the sample throughput can be improved due to the use of one mode of separation, i.e., RPLC, instead of multiple modes of separation to handle metabolites of different ionic strength and polarity.

In general, over 400 ion pairs or putative metabolites were detected from a arabidopsis thaliana extract. We are currently in the process of applying this method to generate the quantitative information on the metabolome profiles of arabidopsis thaliana. Statistical analysis, such as unsupervised principal component analysis (PCA) or supervised partial least square (PLS), will be applied to profile the metabolome among different parts of arabidopsis thaliana or among arabidopsis thaliana grown or harvested under different conditions.

Novel aspect

A stable isotope labeling (SIL) LC-MS method is developed and applied to profile amine- and phenol-containing metabolomes of arabidopsis thaliana.

Abstract #50

Off-line high-pH and low-pH 2D-LC separation combined with isotope labeling MS for comprehensive metabolome profiling of saliva, serum and urine samples

Tao Huan, Wei Han, Ruokun Zhou and Liang Li

Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

Introduction

Metabolomics focuses on the study of small molecules present in a biological system and is a crucial tool in systems biology. Over the last ten years, metabolomics has lived up to its potential and shown capability in many biological applications. Currently there are more than 8000 metabolites found in urine, blood, saliva, and other human biofluids. Studying the metabolomes as a whole to reveal the metabolic relations and differences among these biofluids should expand our knowledge of systems biology. In this work, high-pH and low-pH two-dimensional reversed phase liquid chromatography (RPLC) mass spectrometry (MS) combined with isotope labeling of metabolites has been developed and applied for an integrated analysis of urine, serum and saliva metabolomes.

Methods

All saliva, serum and urine samples were collected from the same volunteer on the same day, and were labeled by $^{12}C_2$-dansyl (Dns) chloride or $^{13}C_2$-Dns chloride. First-dimension separation was performed on these labeled samples using a high-pH RPLC. Fractions were collected every minute for 24 minutes. These 24 fractions were dried down and then reconstituted before loading on a second dimension for separation and then detected by Fourier Transform Ion Cyclotron Resonance (FTICR) MS. Definitive identifications of metabolites were performed against a dansylation compound library consisting of 188 metabolites while putative identifications were carried out by matching against the Human Metabolome Database (HMDB). Data processing and interpretation were performed to get a comprehensive understanding of these three biofluids.

Preliminary data

The limited resolving power of one-dimensional LC has prompted the development of a two-dimensional technique for the purpose of better separation of complex biosamples before detection and analysis by mass spectrometry. In our work, a high-pH RPLC and low-pH RPLC 2D strategy is applied to separate various metabolites in the complex biofluid system. The orthogonality of the 2D-LC separation approach was confirmed by testing the dansyl labeled amino acids standard solution containing 18 amino acids under both high- and low-pH separation conditions.

The optimized 2D-LC FTICR-MS system was applied to the urine samples as a pilot study. The preliminary results indicate that more than 12000 ion pairs or putative metabolites could be detected, which is about 6-fold more ion pairs than that obtained from 1D-RPLC FTICR-MS. While most of the ion pairs detected in the 1D experiment can be consistently detected in the 2D experiment, the 2D experiment provides a great opportunity for us to detect more metabolites that are of low concentration or low ionization efficiency. In addition, a dansylation compound library consisting of 188 authentic standards of amine- or phenol- containing metabolites were used for the definitive identification. Results show that there were 170 endogenous metabolites definitively identified in the urine samples using retention time and accurate mass as matching criteria. Another 2199 metabolites were putatively identified by searching against the HMDB by their accurate mass. These preliminary results indicate the possibility of using our approach to do comprehensive study of saliva, serum and urine metabolomes.

All the serum and saliva samples have been processed and will be analyzed. Data from the three different biofluids will be used to compare their metabolic profile differences. Statistics analysis of these comparisons will be presented and biological significances of the differences and similarities found will be discussed.

Novel aspect

Isotope labeling 2D-LC MS is used to profile the metabolic differences in different types of human biofluids of an individual.

Abstract #51

Development of Isotope Labeling LC-MS for Metabolic Profiling of Bacterial Cells and Its Application for Bacterial Differentiation

Yiman Wu; Liang Li
Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

Introduction

Cellular metabolomics plays an increasingly important role in systems biology, such as interrogation of the enzymatic pathways and network within a cell. However, quantitative and comprehensive profiling of cellular metabolites is currently a challenging task. Both the sample preparation method and the analytical platform have to be carefully evaluated in order to achieve high-throughput, sensitive and robust analysis.

In this work, we report a differential isotopic labeling LC-MS method for the metabolic profiling of bacterial cells which enables relative quantification of each individual metabolite. The sample preparation methods were optimized for detection of the labeled metabolites. Application of this method has been demonstrated on the differentiation of three bacterial species in cultured media and spiked human urine samples.

Method

For method optimization, *E. coli* cells were grown in nutrient broth at 37°C for ~24 h. All cultures were harvested at OD600 of 1.5. For cell differentiation, cells were grown at 37°C on nutrient agar plates for ~24 h. The cell pellets were washed with 0.9% NaCl and extracted with 50% MeOH followed by 10 min ultra-sonication.

The extracted metabolites were mixed with Na_2CO_3/$NaHCO_3$ buffer and labeled with ^{12}C-dansyl chloride for individual samples and ^{13}C-dansyl chloride for pooled sample. The ^{12}C- and ^{13}C-labeled samples were combined and analyzed by LC-MS on an Agilent 1100 HPLC system and a Bruker 9.4 T FTICR mass spectrometer. The resulting MS data were processed using our in-house peak pair picking software.

Preliminary Data

The sample preparation methods were first evaluated. We used 0.9% NaCl to wash the cells and compared the first and second wash solutions. We saw a significant decrease in signal intensity in the second wash, suggesting that the washing step is effective to remove the extracellular metabolites from the cells while not causing observable cell lysis.

We then compared the performance of three extraction solvents and four disruption methods according to three criteria: the number of peak pairs detected, the relative intensity of each peak pair and the reproducibility of each extraction method. We found that the three solvents studied (1:1 MeOH/H2O, 1:1 ACN/H_2O and 2:2:1 MeOH/ACN/H_2O) gave a similar number of peak pairs and gave similar number and intensity reproducibility, but 1:1 MeOH/H2O gave higher amounts of metabolites and was therefore chosen as the extraction solvent for all the subsequent experiments. Similarly, we determined the optimal disruption method to be ultra-sonication.

By using this extraction method coupled with the isotope labeling strategy, we were able to detect over a thousand peak pairs or putative metabolites from bacterial cells in one experiment and near 3000 putative metabolites can be found in one bacterium from combined results of multiple analyses.

Application of this method was then demonstrated on the differentiation of three model organisms, *Escherichia coli*, *Bacillus subtilis and Bacillus megaterium*, as well as identification of bacteria cells spiked in urine samples. The score plot obtained by principal component analysis clearly showed that the three different bacteria can be well separated, and top discriminating metabolites were identified. *We have also successfully identified the bacteria strain in the spiked urine samples by comparing with the standard bacterial cultures, illustrating the prospect of our method to practical applications of identifying unknown bacterial species isolated from real samples.*

Novel Aspect

We report a robust and sensitive method for profiling bacterial metabolites, with the prospect to clinical applications.

Abstract #51

Using transcriptomics to investigate temperature adaptation in *Kosmotoga olearia*, a bacterium with a 60°C growth range

Stephen Pollo[1], Julia Foght[1] and Camilla Nesbø[1,2]

1 University of Alberta Edmonton, Alberta Canada
2 University of Oslo, Norway

Microbes living at high temperatures (thermophiles) have a multitude of traits for survival and maintaining function of their biological molecules that are not found in organisms growing at moderate temperatures (mesophiles). At the cellular level, thermophiles have mechanisms for maintaining their membranes and nucleic acids, as well as many other organelles and structures. At the protein level, each protein remains stable and retains activity at temperatures that denature their mesophilic homologs. These fundamental differences between thermophiles and mesophiles presumably present a barrier to transitioning between the two lifestyles, yet, according to the thermophilic ancestor of life hypothesis, this has happened independently many times. The bacterial order Thermotogales represents a deep branching lineage comprising hyperthermophiles (growing up to 90°C), moderately thermophilic lineages (50-70°C) and mesophiles (<45°C), and thus presents an excellent opportunity for studying temperature adaptation. Among these bacteria, *Kosmotoga olearia* grows over an extraordinarily broad temperature range of 20–80°C. To investigate how *K. olearia* can tolerate such a large temperature range, we are examining transcription of its global cellular responses to temperature by generating RNA-seq data from cultures grown at temperatures spanning its growth range. Briefly, total RNA isolated from each growth condition is depleted of rRNA and processed for next-generation sequencing on the IonTorrent PGM system. The resulting sequences are then trimmed and filtered for quality before they are aligned to the *K. olearia* genome. Once aligned, normalized counts of all transcripts detected are used to calculate expression levels of every gene. These data will be used to generate a model to understand how *K. olearia* can grow over such a huge range of temperatures by determining what changes to cell structure, gene expression and regulation are occurring.

Abstract #52

Phylogeography of Thermotoga isolates; Extensiv gene flow across geographical and ecological borders.

Camilla L. Nesbø[1], Kristen Swithers[2], Thomas Haverkamp[1], Olga Zhaxybayeva[3]

[1]Univ. of Oslo and Univ. of Alberta, Edmonton, AB, Canada; [2]Smith College, Northampton, MA, USA; [3]Dartmouth College, Hanover, NH, USA

The *Thermotoga* are anaerobic hyperthermophilic bacteria commonly detected in hot ecosystems around the world. Earlier analysis of fosmid clones revealed frequent recombination between two *Thermotoga maritima* isolates from geothermally heated sea floors at the Azores (strain RQ2) and Italy (MSB8) and two Japanese isolates from an oil well, *T. petrophila* RKU1 and *T. naphthophila* RKU10. Not only were these bacteria isolated from opposite sides of the globe, they also occupy different types of hot biotopes and, by most definitions, represent different "ecotypes" and different "species". Seven additional isolates were collected from oil wells in the North Sea and from the geothermally heated sea floors near the Kuril Islands, North of Japan. In this study we investigate the evolutionary patterns recorded in the genomes of these 11 *Thermotoga* isolates to determine how geographic barriers and ecological differences affect their evolutionary histories. Genomes from the same geographical location are almost identical (>99% nucleotide identity), with the exception of *Thermotoga* sp. Mc24 from the Kuril Islands, which, based on SNP data, show higher divergence than any of the other isolates. Phylogenetic analyses of typical marker genes resulted in highly incongruent trees. Hence, quartet decomposition (QD) analysis was used to assess phylogenetic histories of all gene families present in at least four of seven representative Thermotoga genomes. This analysis showed that there is no unique bifurcating phylogenetic history relating the seven strains that is supported by a majority of the gene families, suggesting a history of numerous recombination events among the groups represented by these strains. More gene families supported clustering of the strains based on a shared ecology (oil reservoir v. geothermal beach) than based on geography (Atlantic v. Pacific). High level of recombination was also confirmed by recombination detection programs and more recombination events are suggested to have occurred between geographically close isolates. Taken together, this suggests that an underlying phylogenetic signal with clustering according to ecology, which is obscured by gene flow and homologous recombination between neighbors.

Abstract #53

Evaluation of Different Dansylation Reaction Conditions for Isotope Labeling of Metabolites in Metabolome Profiling of Biological Samples

Jared Curle and Liang Li

University of Alberta - Department of Chemistry, Edmonton , CANADA

NOVEL ASPECT: Investigation of the effects of reaction time, temperature, and heating method on dansylation labeling of metabolites for comprehensive metabolome profiling.

INTRODUCTION:

Metabolomics is a field that endeavors to characterize small molecules and their interactions in biological systems for understanding and clinical application. Liquid chromatography mass spectrometry (LC-MS) is one of the main tools used for metabolome profiling. Many metabolites exist in very low concentrations and are obtained with limited amounts. To detect these compounds we have developed a number of isotope labeling methods such as danyslation with which these compounds can be separated, sensitively detected, identified, and quantified. The number of metabolites detected by MS will depend on the number that are labeled. The goal of this work is to explore the effect of changing reaction conditions, such as temperature and time, on results of metabolic profiling experiments of biological samples.

METHODS:

For amine/phenol labelling, amino acid standards or biofluids were mixed with acetonitrile and Na_2CO_3/$NaHCO_3$ (pH 9.4) and reacted with ^{12}C- or ^{13}C-dansyl chloride (DnsCl). The reaction mixture was then heated with a conventional oven or microwave oven or left at room temperature. All reactions methods were carried out at varying temperatures and times. Microwave-assisted heating (MAH) was applied using a microwave reactor (CEM, Matthews, NC). NaOH was added and the mixture was incubated for 10 minutes at the same temperature as the reaction, after which 50:50 Acetonitrile/H2O with formic acid was added. Samples were analyzed by LC-FTICR-MS (Bruker, Billerica, MA) and the number of peak pairs discovered and Human Metabolome Database (HMDB) metabolite hits were used to assess reaction conditions.

ABSTRACT:

In order to analyze the effects of different temperatures and time for labeling, human urine samples were labeled and analyzed using LC-FTICR-MS. Because samples are split and then labeled using both heavy and light DnsCl tags (^{13}C and ^{12}C) we are able to detect those metabolites by specific mass differences detected using MS; these are known as peak pairs. By averaging the number of peak pairs found in replicate samples labeled under differing conditions we are able to determine differences in the conditions used for metabolite labeling. Beyond simple peak number comparison, sample sets can be compared by aligning data and using retention time and exact mass. This method should expose any differences in the type of putative metabolites labeled under the differing conditions.

Currently the labeling protocol calls for a 60 minute reaction at a temperature of 60 ºC. In lower temperature reactions (40ºC) many more peak pairs could be found. It should also be noted that there seemed to be less variability in the replicate samples labeled using microwave as the heating method. The higher reaction temperature of 80 ºC showed a further decrease from that of the 60 ºC. We hypothesize that an increase in temperature may decrease the labeling efficiency or that it causes overall sample degradation and therefore reduced amounts of intact metabolites. Samples with shorter reaction times have a larger variability in peak pair number. This could be due to the labeling reaction being incomplete. It was found that reactions of 60 minutes produced small relative standard deviation in replicate samples.

Further work will focus on the statistical analysis of the data already obtained in order to elucidate further differences in reaction methods for metabolite labeling with an objective of determining the optimal conditions for comprehensive and quantitative labeling of metabolites using dansylation chemistry.

Abstract #54

The Evolution of the Indole-3-acetic Acid Pathway in Algal and Bacterial Genomes

Authors: Harjot Atwal, Leen Labeeuw, Rebecca Case

University of Alberta

Auxins are a class of plant hormones, responsible for growth and development, such as cell division and enlargement, as well as tissue differentiation. One of the most abundant and important members of this family is Indole- 3-acetic acid (IAA). This compound was originally thought to exist only in plants, however the production pathways have since been characterized in bacteria, such as marine bacteria of the roseobacter clade. Interestingly the IAA pathway may have analogues in algae. The objectives of this project were threefold: to investigate the presence of the IAA production pathways in the roseobacter clade; to investigate the presence of the IAA production pathways in algae; and to search for evidence of lateral gene transfer between algae and roseobacters. These objectives were investigated using a bioinformatics approach. The pathways were found to be widely spread in the roseobacter clade, and were also found to be present in algae. Evidence was found indicating both lateral gene transfer between algae and roseobacters and that roseobacters were responsible for the transfer to algae.