ANTIBODY TECHNOLOGY

1 Sensitive and High Throughput ChIP Assays Enable Characterization of Chromatin State

Michael Sturges, Z. Li, K. Taganov, B. Thyagarajan and J. Rosenfeld
Epigenetics Research and Development, Bioscience BU, EMD Millipore, Temecula, CA

Expression of eukaryotic genes during development requires complex spatial-temporal regulation. This complex regulation is often achieved through the coordinated interaction of transcription regulatory elements in the promoters of the target genes. The identification and mapping of regulatory elements in genome scale is crucial to understand how gene expression is regulated. Chromatin immunoprecipitation is a standard method for assessing the occupancy of DNA binding proteins in vivo in their native chromatin context using antibodies. However, standard chromatin immunoprecipitation procedure is time consuming, labor intensive and not suited for analyzing many samples simultaneously.

Recently, we have developed a simple ChIP protocol that requires fewer steps and less hands-on time. This protocol is compatible with both 96-well plate and single tube formats, and enables higher sensitivity and more reliable performance, as compared to conventional approaches.

We have successfully used this protocol to map various clinically relevant chromatin marks and controls across several cell types to quantitatively measure chromatin states. This analysis included a variety of marks corresponding to repressed, poised and active promoters, strong and weak enhancers, putative insulators, transcribed regions, as well as large-scale repressed and inactive domains. This study demonstrates the utility of this approach for the characterization of model cellular systems in perturbation studies with chemical probes.

2 Analysis Of Dialyzed Flask Technology For Antibody Production

Brian Canna1, D Phichith1, C Miller1
1Wheaton, Millville, NJ, 2Rockland Immunochemicals

Efficient and cost effective hybridoma culture is essential to small and large scale monoclonal antibody production for research purposes. This study evaluates multiple aspects of two existing production methods and a new dialyzed cell culture flask method for culturing hybridoma cells. The flask separates the cell cultivation compartment from the cell culture media via a 10kDa cut off limit membrane. This method allows for multiple harvests, longer run times and a super-concentrated supernatant.

To determine the advantages of the dialyzed flask technology, an anti-6X histidine epitope tag secreting hybridoma cell line and an anti-AKT3 isoform secreting hybridoma cell line were selected. The current production methods for the anti-6X histidine epitope tag and anti-AKT3 isoform producing clones include an animal method (ascites) and a proprietary suspension culture method. Several criteria were measured including direct labor cost, reagent cost and yield of antibodies. The purified monoclonal antibodies were analyzed by SDS-PAGE 4-20% under denatured conditions and characterized by Western Blot, direct ELISA titration, and immunohistochemistry to assess performance.

Monoclonal anti-6X histidine epitope tag and anti-AKT3 isoform from the three methods were affinity purified by protein A. SDS-PAGE analysis showed that all antibody samples had a purity higher than 95%. As indicated by the study, production of monoclonal antibodies in ascites often leads to contamination by endogenous host proteins. Western Blot analysis demonstrated no decrease in performance from the antibodies produced in the dialyzed flask. Immunohistochemistry on prostate cancer slides exhibited similar results from all production methods. An economic analysis comparing the three methods was completed. A significant yield increase and a cost savings were realized from the new production method. This increase in performance was due to the reusable nature of the dialyzed flask and its ability to concentrate the antibodies.

3 Creating Custom Reagents for Use in Flow Cytometry

Ryan Duggan, Carol McShan, Mandel Davis
University of Chicago

As the field of flow cytometry grows beyond its basic use by Immunologists, an increase in the demand for custom reagents becomes evident. Companies provide antibodies coupled to multiple fluorochromes however, their focus is on investing in reagents that can turn a profit. Those investigators who may have developed their own novel antibodies, or those who work in a niche scientific environment are forced to make due with whatever is available commercially. In addition these more obscure antibodies may only be available in the most common colors (e.g. FITC, or PE) forcing investigators to look for their common antibodies in obscure colors.

The Cytometry and Antibody Technology (CAT) Facility at the University of Chicago is a recent partnership between the long-established Fitch Monoclonal Antibody Facility (FMF) and Flow Cytometry Facility (FCF) with a mission to provide custom tools for investigators representing various departments. Here we present preliminary data showing the efficiency of a service to provide custom reagents that are pre-titered and tested for quality using the flow and image cytometers available on campus. The reagents start at the hybridoma stage where cells are grown and antibody-containing supernatant is purified and coupled with fluorescent tags. These reagents are then titered and aliquoted for investigator use. Documentation is provided with the reagent so that the end-user can know exactly how to use the reagent. This service is provided at a fraction of the cost of commercial antibodies.

As we become aware of commonly used antibodies we can stockpile those that will move quickly. However, we are not limited to only pursue common reagents. We can easily provide the same services for uncommon or niche applications with equally good quality and efficiency. This service provides tremendous added value to our users leading to an increase in usage of associated services and equipment.
Background and Objective: The problem of identifying SNP-SNP interactions in case-control studies has been studied extensively and a number of new techniques have been developed. Little progress has been made, however in the analysis of SNP-SNP interactions in relation to continuous data.

Methods: We present an extension of the two class multifactor dimensionality reduction (MDR) algorithm that enables detection and characterization of epistatic SNP-SNP interactions in the context of Quantitative trait. The proposed Quantitative MDR (Quant-MDR) method handles continuous data by modifying MDR’s constructive induction algorithm to use T Test.

Results: We then applied Quant-MDR to genetic data from the ongoing prospective Prevention of Renal and Vascular End-Stage Disease (PREVEND) study. We identified that BR2_5BCTATR1AC is the top SNP-SNP interaction that associated with Tissue plasminogen activator (tPA) level for male and ACEID&BRR2EXT1 is the top interaction for female tPA expression.

Discussion and Conclusions: Quant-MDR is capable of detecting interaction models with weak main effects. These epistatic models tend to be dropped by traditional linear regression approaches. With improved efficiency to handle genome wide datasets, Quant-MDR will play an important role in a research strategy that embraces the complexity of the genotype-phenotype mapping relationship.

5 Keeping Track of Interactomes Using the ProHits LIMS

Anne-Claude Gingras1,2, Guomin Liu1, Jianping Zhang1, Hyungwon Choi1, Brian Raught1, Alexey I. Nesvizhskii1, Mike Tyers3
1Samuel Lunenfeld Research Institute, Mt Sinai Hospital, Toronto, ON, Canada, 2Department of Molecular Genetics, University of Toronto, 3Saw Swee Hock School of Public Health, National University of Singapore, Singapore, 4Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada, 5Department of Pathology and Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, 6Institute for Research in Immunology and Cancer, Université de Montréal, Montréal, Québec, Canada

Affinity purification coupled with mass spectrometry (AP-MS) is a robust technique used to identify protein-protein interactions. With recent improvements in sample preparation, and dramatic advances in MS instrumentation speed and sensitivity, this technique is becoming more widely used throughout the scientific community. To meet the needs of research groups both large and small, we have developed software solutions for tracking, scoring and analyzing AP-MS data. Here, we provide details for the installation and utilization of ProHits, a Laboratory Information Management System designed specifically for AP-MS interaction proteomics that we distribute freely to the scientific community at ProHitsMS.com, and which is under continuous development. The complete ProHits solution1 performs scheduled backup of mass spectrometry data and initiates database searches (Mascot, X!Tandem, COMET, SEQUEST and the output from the TransProteomics Pipeline are now supported). It stores search results and enables linking the mass spectrometry data to entries in the relational database module called “Analyst”, which is also available as a stand-alone application (including as an easy-to-install virtual machine implementation3). ProHits Analyst is organized in a hierarchical manner by project, bait, experiment and sample and also serves as an electronic notebook. When a sample is created, mass spectrometry search results can be uploaded. Search results can be explored using a series of viewers, filtered based on mass spectrometry quality, frequency of detection or background lists, viewed in Cytoscape-Web or exported to text or as a PSI XML format for deposition in interaction databases. Importantly, however, search results can be further analyzed using the SAINT statistical tool which is seamlessly integrated within ProHits to derive interaction confidence scores3-4. With the integration with a number of open source tools and public repositories, ProHits facilitates transparent analysis and reporting of AP-MS data.

PMID:20944583
PMID:22948730
PMID:20489023
PMID:21131968
PMID:22948729

6 Isoform-level Analysis of Next Generation Sequence Data Highlights Mechanisms of Ewing Sarcoma Progression

Winston Brasor, Jean-Noel Billaud, Carla Bullitt
Ingenuity Systems

Ewing Sarcoma (ES) is a prototypical translocation sarcoma, which harbors characteristic translocations fusing the 5′ portion of the EWS gene with the 3′ region encoding the DNA binding domain of one of five ETS family genes. ES behaves as an aberrant reprogramming of mesenchymal stem cells and is the second most frequent bone tumor in adolescents and young adults. To understand the molecular mechanisms of tumorigenesis and metastatic processes in ES, gene expression profiling ES patient samples was performed using NGS technology. In this study we present the results of using Ingenuity® iReport™, a software application employing bioinformatics and biological interpretation best practices, to demonstrate the differences between metastatic and primary tumor samples from one representative patient with ES. Analysis in iReport highlights the unique value of RNA-Seq (compared to traditional gene expression microarrays) by providing functional information on transcript isoforms produced in metastatic tumors (relative to primary), and generating testable hypotheses as to the mechanisms of ES progression. We characterized and identified signalling and metabolic pathways (such as Glucose and Fatty Acid Metabolism) and biological processes (such as Cellular Invasion) that may be involved in the progression of ES. Visualizing hubs of molecular interactions highlighted transcription factors such as SREBF1 and PPARGC1A that may contribute to the pathological context observed in this patient. In particular, analysis of differentially expressed isoforms of STAT3 and SREBF1 indicate that these 2 transcription factors induced specific transcriptional programs in the metastatic process in ES. The data presented here is from one patient, therefore similar studies are needed to understand if these insights are supported by data from other Ewing Sarcoma’s patients. Nonetheless the insights gleaned from this analysis demonstrate a greatly improved workflow for bench scientists to proficiently handle RNA-Seq data and generate testable hypotheses of mechanisms of cancer progression.
7 Identifying Mutations in Transcriptionally Active Regions of Genomes Using Next Generation Sequencing

N. Eric Olson, Hugh Arnold, Nathan Lazar, Todd M. Smith
PerkinElmer

The pace at which genome references are being generated for plants and animal species is rapidly increasing with Next Generation Sequencing technologies. While this is a major step forward for researchers studying species that previously did not have sequenced genomes, it is only the beginning of the process toward defining the biology underlying the genome. As long as a reference is available, DNA variants can be readily identified on a genome wide scale, often producing lists of 100s of thousands or even millions of variants. Often those occurring in expressed genes are of the most interest; however, if annotation defining where genes exist within a genome is not available or poorly defined, identifying which mutations might affect protein coding may not be possible. To address this challenge we will describe a method whereby RNA-Seq can be readily used to identify transcriptionally active regions which creates transcript annotation for un-annotated or enhanced annotation for any organism. This annotation can then be used in conjunction with whole genome sequencing to annotate variants as to whether they fall within transcriptionally active regions thus facilitating the identification of mutations in larger repertoire of expressed regions of a genome.

8 Sample Size Determination for Clinical Proteomics Experiments: Diagnostic, Prognostic and Predictive Biomarker Discovery and Validation

David Cairns, Jennifer Barrett, Peter Selby, Rosamonde Banks
Division of Cancer Studies and Pathology, Leeds Institute of Molecular Medicine, University of Leeds, Leeds, United Kingdom

Clinical proteomic biomarker discovery experiments have produced few discoveries which have had an impact beyond their initial publication. In addition, there are far fewer validation experiments reported than discovery experiments. It has been suggested that one reason for this slow progress is a lack of application of standard concepts in experimental design. One such concept is the determination of sample size a priori to ensure that powerful investigations are undertaken, where there is a good chance of a difference being observed if one truly exists.

Proteomic experiments raise challenges in sample size determination due to their megavarriate nature and inherent biological and technical variation. However, it is possible to determine sample size for investigations which seek to identify diagnostic, predictive and prognostic biomarkers. This can be undertaken by combining tools to control the expected false discovery rate, appropriate sample size formulae and simulation. A flexible protocol is described that allows the determination of sample size for a variety of proteomic biomarker discovery and validation experiments. These allow the determination of appropriate sample sizes to conduct suitably powerful experiments. It is possible that the application of such techniques could lead to more robust discoveries being validated and enter clinical practice, after appropriate evaluation.

9 Bioinformatics Core Project Management

Heather Driscoll1, Mahesh Vangala1, James Vincent2
1Norwich University, Northfield, Vermont, 2University of Vermont, Burlington, Vermont

Bioinformatics cores that provide fee for service style support encounter a wide variety of projects. The scope of projects varies greatly among investigators. Because of this variety, it is difficult to develop a set of predefined services that fit all project types. While our own core has developed a baseline set of services, we found in practice these often needed significant modification to meet the goals of particular investigator. To overcome this problem we factored common features of all projects and partitioned them into groups: workflow management, data management, user results, and tracking and reporting. We then implemented best practices for each group using commercial and open source software combined with our own management policies. Finally we linked these areas together to produce an overall integrated project management solution that combines workflow management, data management, user results management and reporting capabilities. This system solves the problem of developing well defined services that are trackable and repeatable while simultaneously enabling flexibility that is easily managed. The result improves the effectiveness and efficiency of the bioinformatics core for scientists working within the core, for investigators receiving core support and for external auditors and evaluators.

10 Exomic Sequencing to Identify Germline Variants in Familial Melanoma

Hira Shabbir, Maria Teresa Landi, MD/PhD

Germlines are the source of DNA in all cells. A mutation at the germline level is the first step to developing cancer, and the vast majority of cancer is genetic. Melanoma, the leading cause of skin cancer death, is known to be highly heritable and rare. Using a family model, high risk variants related to melanoma can be identified. The goal of the study is to integrate information from sequencing, epigenetics, and expression to identify functional and regulatory genes that are associated with melanoma. Families with two or more 1st degree relatives with melanoma were considered at high risk and were investigated in this study. Initially, sequencing data of families with 3 or more relatives with the disease were examined and shared DNA variants were selected for further examination. Genetic databases and annotation tools were used to identify genes based on their known gene function and regulation, pathways, and variant conservation. Gene browsers were also used to identify any histone markers, DNA methylation sites, and other epigenetic indicators. Based on our candidate genes, there is a possibility of genetic heterogeneity, in which multiple genes may be responsible for disease susceptibility. Selected candidate genes will undergo fine mapping to further investigate the region and replication in additional families and population studies of melanoma.

11 What Does Take to Identify the Signal From the Noise in Molecular Profiling of Tumors?

N. Eric Olson, Hugh Arnold, Todd Smith
PerkinElmer

Cancer is a complex, heterogeneous disease that is driven by continually evolving genomic changes. Our current efforts to identify and the cure or demise of patients has utilized snap shots of DNA, RNA, proteins, and/or protein/nucleic acid interactions among numerous assays. For example, sequencing genomes or exomes
distinguishes germline variants from somatic mutations as one step toward identifying nucleotide changes that are truly driving mutations. However, these assays identify very large numbers of variants and substantially reducing the noise requires considering the potential impact of variants (missense, non-sense, synonymous), quality of the call, prevalence of mutations in tumor versus normal cells, and whether a gene carrying a mutation is even expressed. Consequently, molecular profiling of tumors, benefits from data obtained from different types of NDA sequencing-based assays. Using data from paired tumor and normal samples we will show an example workflow that combines exome and transcriptome sequencing to identify putative driver mutations that display high signal for being impactful in cancer.

12 Quantitative Analysis of Shotgun Metagenomic Data with the Real Time Genomics Platform

Brian Hilbush, John G. Cleary, Richard Littin, Len Trigg, Sean Irvine
Real Time Genomics, Inc., San Bruno, CA

Real Time Genomics has developed high performance tools for quantitative analysis of metagenomic experiments on the Illumina, Ion Torrent and Roche 454 platforms. The search algorithms and processing pipeline were created around requirements for analysis of shotgun metagenomics samples from the Human Microbiome Project (HMP). With the configurable metagenomics pipeline, data sets can be analyzed in three separate arms for 1) functional or metabolic profiling using translated nucleotide searches of annotated databases such as KEGG with mapx in conjunction with the gene/metabolic pathway analysis program HUMANN (HMP and Huttenhower lab), 2) taxonomic profiling with searches of reference genome databases with map followed by quantitative analysis of species abundance and sample composition from alignment data (SAM files) using RTG species, and 3) analysis of differences in bacterial community structure with a RTG’s similarity matrix generation tool, used in combination with multivariate analysis (PCA or hierarchical clustering algorithms (SVD, singular value decomposition)). We will present analysis of HMP project data to demonstrate the utility of the pipeline for a series of metagenomic investigations.

13 Search Strategies for Glycopeptide Identification

Marshall Bern1, Christopher Becker1, Wilfred Tang1, Yong J. Kil1, Xiaoke Yin2, Manuel Mayer2, Kay-Hooi Khoo3, Rosa Viner4
1Protein Metrics, *King’s College, U. of London,
2Academia Sinica, Taiwan, 3Thermo Fisher Scientific

Byonic is a new proteomics search engine that can identify peptides carrying N- and O-linked glycans. Byonic offers a number of ways to search for glycopeptides, including preset glycan tables and manually entered glycan masses, and the search strategy affects the quality and quantity of spectrum assignments. Here we show how a progression of searches, from wider to narrower in both proteins and glycans, can improve sensitivity and specificity for glycopeptide identification.

We obtained data from the following samples: Glycophorin-A, PSA, human blood serum enriched for glycoproteins, and secreted proteins from human endothelial cells. All data were acquired on various Thermo Orbitrap instruments and included both HCD and ETD fragmentation. We first searched the data with a full human protein database with contaminants and decoys, and later with smaller databases produced by Byonic’s “focused database” option. We started with Byonic’s preset glycan search, which allows only one glycan per peptide, and then, guided by prior search results, augmented or replaced these tables with user-defined glycan modifications with appropriate limits on each type of modification.

We found that focused protein databases containing 10 – 200 proteins greatly improve the sensitivity of glycopeptide search relative to full-database searches. We found a database of likely glycoproteins, determined by PNG-ase release of N-glycans in O18 water, helpful for identifying glycopeptides carrying single N-linked glycans in the endothelial secretome. Focused glycan lists also improve sensitivity, and make possible still more complex searches. We have identified glycopeptides carrying up to two N-glycans, one N-glycan and one O-glycan, and up to four O-glycans, with only minor ambiguities in modification placement and mass distribution. More complex searches, for example, five or more O-glycans, will require improvements in completeness of fragmentation and computational methods.

14 A Core-tailored Platform for Pinpointing Clinically Relevant Variants in Human Genomes and Exomes

Ingenuity

Genomic studies of human disease and drug response aim to find one or a few causal variants among millions of possible candidates. A streamlined platform for quickly and reliably assessing each variant called in such studies can save months of tedious effort, speeding discoveries for core labs’ clinical and research clients, and freeing core staff to solve other data analysis challenges posed by broader clientele. Made to help cores and their clients quickly interpret human genomes, the Ingenuity® Variant Analysis™ platform (www.ingenuity.com/variants) lets users upload, annotate, and thoroughly compare whole or partial human genomes, to smartly shortlist candidate variants, genes, and pathways that may best explain user-specified phenotypes. Leveraging Ingenuity’s deep, structured knowledge base of published findings and robust predictive insight, Variant Analysis runs sophisticated tests of family- and population-scale genetic association, tailored to the user’s study design and phenotype, via a simple, fluid interface that also lets one easily review, revise, and share findings. To meet the distinctive challenges of clinical genome interpretation, we have comprehensively curated published clinical assessments and population incidence of individual human sequence variants, supplementing gene- and pathway-level functional knowledge. To help researchers identify new causal candidates, we have built and validated Variant Analysis to run gene- and pathway-level burden tests ~100x faster than conventional methods. And, by letting users easily and securely share genome data and findings with collaborators, the platform mediates efficient collaborative discovery among researchers studying similar or (as mutual controls) distinct rare diseases. Combining sophisticated functional analytics; statistically robust genetic analysis at the variant, gene, and pathway levels in an intuitive interface, the Variant Analysis platform is built to meet the varied genome interpretation needs of researchers and clinicians studying single and multiple probands, pedigrees, and case-control cohorts, to understand rare inborn diseases, common complex disease, cancers, and drug response.
Technical challenges facing researchers performing next-generation sequencing (NGS) analysis threaten to slow the pace of discovery and delay clinical applications of genomics data. Particularly for core laboratories, these challenges include: (1) Computation and storage have to scale with the vast amount of data generated. (2) Analysis pipelines are complex to design, set up, and share. (3) Collaboration, reproducibility, and sharing are hampered by privacy concerns and the sheer volume of data involved. Based on hands-on experience from large-scale NGS projects such as the 1000 Genomes Project, Seven Bridges Genomics has developed IGOR, a comprehensive cloud platform for NGS Data analysis that fully addresses these challenges:

• IGOR is a cloud-based platform for researchers and facilities to manage NGS data, design and run complex analysis pipelines, and efficiently collaborate on projects.

• Over a dozen curated and peer-reviewed NGS data analysis pipelines are publicly available for free, including alignment, variant calling, and RNA-Seq. All pipelines are based on open source tools and built to peer-reviewed specifications in close collaboration with researchers at leading institutions such as the Harvard Stem Cell Institute.

• Without any command-line knowledge, NGS pipelines can be built and customized in an intuitive graphical editor choosing from over 50 open source tools.

• When executing pipelines, IGOR automatically takes care of all resource management. Resources are seamlessly and automatically made available from Amazon Web Services and optimized for time and cost.

• Collaboration is facilitated through a project structure that allows researchers working in and across institutions to share files and pipelines. Fine-grained permissions allow detailed access control on a user-by-user basis for each project. Pipelines can be embedded and accessed through web pages akin to YouTube videos.

• Extensive batch processing and parallelization capabilities mean that hundreds of samples can be analyzed in the same amount of time that a single sample can be processed. Using file metadata, batch processing can be automated, e.g., by file, library, sample or lane.

The IGOR platform enables NGS research as a “turnkey” solution: Researchers can set up and run complex pipelines without expertise in command-line utilities or cloud computing. From a lab and facility perspective, the cloud-based architecture also eliminates the need to set up and maintain a large-scale infrastructure, typically resulting in at least 50% cost savings on infrastructure. By facilitating collaboration and easing analysis replications, the IGOR platform frees up the time of core laboratories to emphasize and focus on the research questions that ultimately guide them.
and fatty aldehydes. We have developed uHPLC-MS/MS based methodology capable determining oxygen sensitive metabolites using a powerful antioxidant mixture.

We have assembled a team with expertise in analytical chemistry, radiation biology, metabolomics, nutrition, and bioinformatics to address and develop a systems biology approach toward understanding the relationships between the complex biopterin metabolome and radiation injury to specific organ systems. Tetrahydrobiopterin and dihydrobiopterin were stable in the antioxidant mixture for 8 hours. The lower limit of quantitation for all biopterin-related compounds was 3 ng/ml. Quantitative data was obtained using only 10 mg of solid tissue or 10 ul of plasma.

18 **Translating microRNA Discovery in Biofluids into Robust Biomarkers for Disease Using LNA™-enhanced qPCR**

Brian Glassner, Peter Mouritzen, Thorarinn Blonadal, Ditte Andreassen, Niels Tolstrup, Maria Wrang Teilum Exiqon

microRNAs represent the best described class of small RNAs (21-23nt) and have been shown to function as post-transcriptional regulators of gene expression. The high relative stability of microRNA in common clinical source materials and the ability of microRNA expression profiles to accurately classify discrete tissue types and specific disease states have positioned microRNA quantification as a promising new biomarker for a wide range of diagnostic applications.

We have developed a genome-wide LNA™-based microRNA qPCR platform with unparalleled sensitivity and robustness even in biofluids where microRNA levels are extremely low. Only a single cDNA synthesis reaction is required to conduct full miRNome profiling thereby facilitating high-throughput profiling in important clinical sources without the need for pre-amplification. Thousands of biofluid samples have been profiled including blood derived plasma/serum and urine to accurately determine normal reference ranges for circulating microRNAs. Procedures have been developed to control pre-analytical variables such as hemolysis in serum/plasma samples. In addition, a data QC system has been implemented to secure technical excellence and reveal any unwanted bias in the dataset.

We are currently screening for and validating microRNAs as biomarkers for stage II colorectal cancer (CRC). microRNA profiling has been performed on plasma samples from a clinical trial conducted in 7 different hospitals. We show that hemolysis in this sample set correlates with hospital ID, and with the utilization of specific blood sample collection vials. Using a microRNA-based hemolysis signature, we eliminated hemolyzed samples and demonstrated that this step leads to a major improvement of CRC detection (ROC AUC increase from 0.67 to >0.80). We conclude that pre-analytical variables such as hemolysis can be a source of bias in samples of different origin, and that sample and data QC procedures can overcome this challenge and lead to improved miRNA biomarker performance.

19 **Proteomics-based Biomarker Discovery: The Need for Less Complicated Methodologies**

Lawrence Sandberg, Cliff Herrmann, Mia Perez, Ravi Raghavan

**Depts of Biochemistry & Pathology, Loma Linda University Medical Center, Loma Linda, CA**

**Introduction:** Although proteomic technologies hold much promise, the discovery of meaningful tumor biomarkers using this methodology is challenging. One issue is the size of the databases, and in spite of utilizing costly bioinformatics software, tumor marker detection is time-consuming. While engaged in a mass-spectrometry based project for biomarker discovery in head and neck cancer, we found that combining methodologies gave quicker and more meaningful results.

**Methods:** Archived, formalin-fixed, paraffin-embedded tissues from twelve cases of head and neck squamous cell carcinoma, and a matching set of control (normal) tissues formed the basis for the study. Both commercially available (Expression Pathology [EP]), and in-house reagents were used for protein extractions. The EP digestion protocol with Protease Max released peptides from a one-hour trypsin digestion in both instances. A Thermo Deca LCQ was used for mass spectrometric analysis.

**Results:** Data from twenty-four paired samples by each method was entered using spectral quantitative counts obtained from the Mascot/Scaffold/X-Tandem program. Instead of relying on pathway analysis software alone for biomarker analysis, we opted for a two-column MudPIT format. We first derived P-values (Fisher's Exact Test), and the ‘normal’ and ‘tumor’ columns were sorted by ascending values. This brought the ‘zero’ values to the top of the ‘normal’ column, enabling potential biomarker proteins to be expressed only on the ‘tumor’ side. By further analysis utilizing bioinformatics software (Ingenuity Pathway Analysis – IPA), we could narrow the search down to less than ten relevant candidate proteins that were common to both analyses. Those with high enough P-values, and good interactions via IPA were selected for immunohistochemical validation as tumor biomarkers.

**Conclusions:** Using the simpler MudPIT analytical approach in conjunction with IPA, faster discovery times of biomarker candidates may be possible. Further validation awaits application of these techniques on a larger sample of cases.

**CARBOHYDRATE ANALYSIS**

20 **Evaluation of Two Complementary Methods for Quantitative Profiling of PSA N-Glycans and N-Glycopeptides**

Morten Thaysen-Andersen, Nicolle H. Packer

Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, Australia

The analytical methods for the structural characterization of protein glycosylation in glycomics and glycoproteomics are gradually becoming more suitable for biological applications. Over the past 10 years there has been a concerted effort to increase the sensitivity, speed and automation of the analysis of more complex glycoprotein mixtures at the level of biologically relevant dynamic ranges. However, these promising analytical tools, which predominantly are mass spectrometry based, needs thorough comparison and validation, in particular when quantitation is required. As a part of the ABRF Glycoprotein Research Group (gPRG) Quantitative Glycoprotein Study, we have profiled the single N-glycosylation of two sources of human prostate specific antigen (PSA) using two different, but commonly used, analytical methods allowing us to compare their qualitative and quantitative performance: i) Quantitative site-specific analysis of enriched PSA N-glycopeptides and ii) quantitative global analysis of released and reduced PSA N-glycans. For both approaches porous graphitized carbon LC connected directly with ion trap MS/MS (CID) was used for analyte separation and detection, respectively, but with different MS polarity mode detection. Using label-free relative quantitation, the two analytical methods produced very similar PSA glycoprofiles considering their different nature. Of the 40-50 monosaccharide compositions detected from each PSA variant, mostly minor quantitative differences in the glycoprofiles produced by the two approaches were observed, enhancing the confidence of the analysis. The pros and cons of the two glycoprofiling approaches and the intended near-future improvements will be discussed.
**Glycan Quantitation Strategies Based on High Sensitivity NanoLC/MS**

Mellisa Ly, G. Staples, H. Yin, K. Kileeen  
Agilent Labs, Agilent Technologies, Santa Clara, CA

Monoclonal antibodies (mAb), therapeutic proteins for treatments of cancers and immunological diseases, need detailed structural elucidation and improved analytics because of their biopharmaceutical presence. Inherently heterogeneous proteins, mAbs have various protein modifications, and the most important modification for these large glycoproteins is N-glycosylation on their Fc region. In this work, we compare several quantitative techniques; including microfluidic based on-line deglycosylation and integrated nanoLC/MS analysis of the released glycans, nanoLC/MS analysis of intact glycoprotein and/or glycoprotein fragments, as well as UHPLC-FLD analysis of off-line released derivatized glycans.

We have utilized middle-up and bottom-up strategies for glycan quantitative analysis of humanized IgG1. In addition, we have also performed analysis of intact proteins such as RNase B and prostate specific antigen and compared such results with released glycan strategy. Thorough glycan analysis was first performed using the mAb Glyco Chip to identify released free N-glycans, with structures extracted by molecular formula, and quantified by extracted ion chromatograms. IgG1 was also digested with IdeS to form Fc and Fab fragments. The ~25kDa Fc fragment was analyzed using an HPLC-Chip/TOF MS system. For Fc region analysis, a maximum entropy deconvolution algorithm was used and the peak height of the deconvoluted spectra was used for high sensitivity relative quantitation of various glycoforms. An injection of as little as 5ng (30fmol) IgG1 generally was enough to generate high quality spectra. Complementary positive-ion and negative-ion modes on HPLC-Chip/MS with a make-up flow channel provided accurate identification and quantitation of PSA N-glycans. Two different PSA samples with different glycosylation patterns were compared, with about 30 glycans identified in each sample. HILIC UHPLC was used to separate the various glycans form of IgG isofroms. Glycan quantitation results based on HILIC separation paralleled the relative abundances of both the mAb Glyco Chip results and Fc region analysis.

**CORE ADMINISTRATION**

**Administrative Strategies for Increasing Impact**

William Decaneas  
Beth Israel Deaconess Medical Center

At BIDMC we’ve tried a number of strategies to increase the impact of our Research Cores while simultaneously increasing their efficiency. Increased operating efficiency has allowed for a lower average subsidy per core, which has let us increase our core offerings. I hope to use this poster to outline a few of our successes and one or two of our failures. One area we’ve focused on is supply chain; we’ve made great strides in consolidating and leveraging our purchasing power through bundled service awards and supply purchases. We’ve also increased our marketing efforts through a stronger web presence, improved marketing materials and collaboration with local academic institutes. We’ve improved our finances by tightening and standardizing our marketing materials and collaboration with local academic institutes. And we’ve created a Cores Advisory committee to provide strategic guidance as well as help us choose which capital collection processes. We’ve improved our finances by tightening and standardizing our marketing efforts through a stronger web presence, improved marketing materials and collaboration with local academic institutes.

We’ve improved our finances by tightening and standardizing our marketing efforts through a stronger web presence, improved marketing materials and collaboration with local academic institutes. And we’ve created a Cores Advisory committee to provide strategic guidance as well as help us choose which capital collection processes. And we’ve created a Cores Advisory committee to provide strategic guidance as well as help us choose which capital collection processes. And we’ve created a Cores Advisory committee to provide strategic guidance as well as help us choose which capital collection processes. And we’ve created a Cores Advisory committee to provide strategic guidance as well as help us choose which capital collection processes. And we’ve created a Cores Advisory committee to provide strategic guidance as well as help us choose which capital collection processes.

**CORE FACILITIES**

**Manager of Next Generation Sequencing Orders – MANGO**

Catharine Fournier Aquino, Tanguy Le Carrou, Lennart Opitz, Jelena Kühn Georgijevic, Ralph Schlabbach  
Functional Genomics Center, ETHZ/UZH

The Functional Genomics Center Zurich (FGCZ) is a joint state-of-the art research and training facility of the ETH Zurich and the University of Zurich. With latest technologies and expert support in genomics, transcriptomics, and bioinformatics, the FGCZ carries out research projects and technology development in collaboration with the Zurich Life Science research community. The FGCZ offers services for different applications on the Illumina HiSeq2000, Illumina MiSeq, Ion Torrent, Roche 454 and PACBIO RS. At the FGCZ, we handle hundreds of NGS projects a year. A working tool is necessary to monitor and document these sequencing projects. Because of our specialized need, we conceptualized, developed and implemented the MANGO to help manage, track, monitor and document our various and diverse NGS service orders. The MANGO works in multiple levels, first, it is a web accessible sample tracking system. It can be accessed and sample data can be added in real-time through a computer, an android tablet or an iPad. Secondly, it manages multiplexing of sequencing runs because it can detect sub-optimal index combinations from various popular commercial kits and self made indices. Thirdly, the MANGO creates well-formatted sample sheets for the various sequencers available in the FGCZ. Lastly, it can accept data in .csv format from instruments used for QC during library preparation. The MANGO is a reliable and secure cross-platform manager of our NGS service orders.

**A Case Study – The Unique Operating Model of the Research Core Facility (RCF) at the Keenan Research Centre (KRC) in the Li-Ka Shing Knowledge Institute (LKSKI) at St. Michael’s Hospital, Toronto, Canada**

Stephen Barker  
St. Michael’s Hospital, Toronto, Canada

For the past several years we have planned a new building to house our Research Programme. We used this as an opportunity to create centralized core facilities with a unique operating model. Here we describe the design process, operating model and highlight the correlation between a well designed and professionally run core facility that stresses teaching, collaboration, user productivity and satisfaction.

The RCF is a multi-user facility supporting basic science research in the new LKSKI. The underlying design principle for the LKSKI was to foster a collaborative research environment where there was a convergence between scientific investigation, education and patient care. The RCF does this both through an atmosphere of connectivity, as well as through an emphasis on providing training and project design advice with a focus on teaching people to run their own samples rather than running the samples for them.
The facility brought together new and donated equipment, centralizing management through a professional staff of trained scientists. It includes both common and specialized facilities such as: Biol imaging, Flow cytometry, Molecular Biology, and Histology. By reducing duplication and increasing usage, the facility offers significant savings, creating a critical mass where cutting edge equipment and techniques are available to scientists who couldn’t otherwise afford them. The RCF costs are shared by the users and the Institution. Instead of user fees, which can be hard to quantify and administer, the institution requires that all grants include financial support, equivalent to 3% of the operating budget.

The conclusion, amongst the majority of researchers is that after a year of occupation and operation, there is a definite improvement in productivity and user satisfaction. A recent survey indicates that 85% of users are more productive, with 83% rating the facilities and operations as good or excellent.

### 25 Optimizing Project Administrative Workflow with Formstack, Sharepoint, and Vanderbilt CORES Software

Lisa Wright, Paige Vinson, PhD
Vanderbilt High Throughput Screening Facility

Tracking administrative workflow for Core projects is a difficult task. Cores are increasingly required to provide metrics demonstrating productivity, scope of projects, and success rates, yet scientific staff members do not have sufficient access or bandwidth to produce this type of broad spectrum data easily. In an effort to reduce redundancy, automate recurrent tasks and minimize staff labor, the Vanderbilt High Throughput Screening (HTS) Facility has combined readily available web-based software with institutional CORE software. The HTS Facility is striving toward a goal of having common sets of metrics available, as needed, to communicate the institutional impact of the Core to senior leadership and funding agencies. These administrative workflow improvements also increase effective and efficient communication in daily project administration and minimized required labor from scientific staff.

### 26 Sequencing Initiative at the Norris Cotton Cancer Center

Joanna Kerley-Hamilton, S Shipman, H Trask, C Lytle, W Taylor, J Moore and C Tomlinson
Geisel School of Medicine and Norris Cotton Cancer Center

The Dartmouth Genomics Shared Resource recently purchased the Ion Torrent Personal Genome Machine (PGM) and the Ion Proton with contributions from the Norris Cotton Cancer Center (NCCC), Geisel School of Medicine and the Institute for Quantitative Biomedical Sciences. The transition to Ion Torrent deep sequencing was relatively smooth and the workflows easily established. In collaboration with the NCCC, we are offering NCCC investigators an initiative to encourage deep sequencing and translational research. Investigators can choose one of two cancer panels: the Ion Torrent hotspot cancer panel (50 genes), and a custom-designed cancer gene panel (541 genes). The 541-cancer gene panel includes the desired genes from every NCCC investigator, which covers a broad spectrum of cancers and signaling pathways. The 541-cancer gene panel was designed using the Haloplex system (Agilent, Santa Clara, CA). We have validated extraction of DNA from both formalin-fixed paraffin-embedded (FFPE) and fresh frozen tissues to offer clinicians and researchers options for sample collection. Data are presented from the hotspot cancer gene panel using DNA obtained from FFPE and frozen breast cancer tissues.

### 27 Bringing the Pieces Together – Placing Core Facilities at the Core of Universities and Institutions: Lessons from Mergers, Acquisitions and Consolidations

Claudius Mundoma
Florida State University, Institute of Molecular Biophysics

As organizations expand and grow, the core facilities have become more dispersed disconnected. This is happening at a time when collaborations within the organization is a driver to increased productivity. Stakeholders are looking at the best way to bring the pieces together. It is inevitable that core facilities at universities and research institutes have to be integrated in order to streamline services and facilitate ease of collaboration. The path to integration often goes through consolidation, merging and shedding of redundant services. Managing this process requires a delicate coordination of two critical factors: the human (lab managers) factor and the physical assets factor. Traditionally more emphasis has been placed on reorganizing the physical assets without paying enough attention to the professionals who have been managing the assets for years, if not decades. The presentation focuses on how a systems approach can be used to effect a smooth core facility integration process. Managing the human element requires strengthening existing channels of communication and if necessary, creating new ones throughout the organization to break cultural and structural barriers. Managing the physical assets requires complete assessment and this requires direct input from the administration as well as the facility managers. Organizations can harness the power of IT to create asset visibility. Successfully managing the physical assets and the human assets increases productivity and efficiency within the organization.

### 28 Integration of PacBio RS into Massive Parallel Sequencing and Data Analysis Pipelining at the UC Davis Genome Center

Ryan Kim, Rashbrook, Vanessa, Henriette O’Geen, Oanh Nguyen, Siranoush Ashtari, Xiaohong Fan
UC Davis Genome Center

Whole genome sequencing and genomic biology has been widely adopted in many fields of biology as next-generation sequencing technology (NGS) has rapidly improved quality, read length, and throughput to make whole genome sequencing and association studies possible in a very cost effective manner. Continued improvement and development of sample preparation protocols and data analysis tools have been significant in helping to extend genome sequencing technology to genomes that were previously difficult to sequence. Recent arrival of Pacific Biosciences RS (PacBio) contributed in furthering such opportunity by providing options for single molecule long read sequencing in real time and kinetic analysis (methylation). PacBio has been employed successfully for sequencing low complexity genomic region such as extremely high GC, long repeats, rearrangement, gene fusion, etc.

In this poster we present the optimization of PacBio sample preparation that was fine-tuned to meet unique challenges of sequencing through “difficult-to-sequence” template. We discuss the integration of PacBio into the wet lab equipped with other NGS platforms and data pipelining workflow including cloud computing and robotic sample preparation at the Genome Center.

UC Davis Genome Center currently operates NGS technology platforms including HiSeq, MiSeq, PacBio, and has genotyping capacity using Illumina Infinium and GoldenGate technology. UC Davis Genome Center and Bioinformatics Program provides most up-to-date genome technology and informatics support tailored
High Throughput and High Content Screening Capabilities of the University of Cincinnati Drug Discovery Center

Rakesh Rathore¹, Jason Kirby¹, Hong Tang¹, Fadi J Najm³, Paul J Tesar¹, Ken Greisii, William Seibel¹, Ruben Papoian¹
¹Drug Discovery Center and Department of Cancer and Cell Biology, University of Cincinnati, Cincinnati, OH, ²Department of Cancer and Cell Biology, University of Cincinnati, Cincinnati, OH, ³Department of Genetics and Genome Sciences, Case Western Reserve University, Cleveland, OH

The Drug Discovery Center collaborates with a wide range of academic and industrial research centers to facilitate the identification of active small molecules with high potential for use as biological probes or as starting points for drug discovery programs. The DDC operates state-of-the-art high throughput and high content screening instrumentation and a diverse 350,000 compound library. The center's personnel provide collaborators with advice in assay design, analytical technology selection, and library design via cheminformatics and/or structure-based approaches. Typical programs are exemplified by an HTS program targeting the identification of novel atypical PKC inhibitors for potential use in cancer and an HCS program targeting the identification of stimulators of the differentiation of oligodendrocyte progenitor cells to oligodendrocytes for potential use in multiple sclerosis. Activities of atypical PKCs, PKCβ and PKCζ, were measured using both fluorescent detection of ADP production and MALDI-TOF detection of substrate phosphorylation. 30,000 compounds were screened for their effect on PKC activity using these orthogonal detection methods to identify potential inhibitors. Oligodendrocyte differentiation was measured in a High Content Screen by pairing a selective marker of oligodendrocytes with Alexa Fluor 488 secondary antibody for the detection of mature ODS. DAPI stain was used for nucleus detection. Confocal microscope images were acquired and analyzed using an algorithm for neurite outgrowth adapted for the characterization of oligodendrocyte processes. Three measurements: Mean Maximum Process Length per Cell, Mean Process Signal Intensity per Cell, and Percentage Differenced Cells per Well, were quantified by the image analysis script and afforded a statistically significant separation of promoter controls and inhibitor controls from non-treated (neutral) controls with Mean Maximum Process Length as measurement.

Tissue to Library Preparation on the Apollo 324 System for Clinical Sequencing

Monika Tomczyk, Tom Olenic, Stephanie H.I. Yeung and Shanavaz Nasarabadi
IntegenX Inc., Pleasanton, CA

Over the past decade, next generation sequencing (NGS) technologies have revolutionized genetic and biological research. The sensitivity, speed, throughput and continuing reductions in cost per sample of NGS technologies have stimulated broad interest in clinical applications. Successful adoption of NGS technologies for routine and widespread clinical applications requires consistent and reliable sample processing from raw tissue through nucleic acid extraction, size fractionation, and library preparation.

We have explored automation of library preparation from mammalian brain tissue samples for analysis on the Genome Analyzer IIX (Illumina Inc.). Rabbit brain tissue lysate was prepared by grinding frozen tissue with buffer in a mortar and pestle. The process of DNA isolation, enzymatic fragmentation and library preparation was automated on the Apollo 324 system (IntegenX Inc.). In our streamlined automated workflow, we isolated DNA from the tissue lysate by digestion with Proteinase K (New England BioLabs), followed by DNA purification with paramagnetic beads, enzymatic fragmentation using NEBNext dsDNA Fragmentase® (New England BioLabs) and library preparation with the PrepX ILM Library Preparation Kit reagents (IntegenX Inc.). The library was enriched by amplification in a commercial thermocycler and sequenced on the Genome Analyzer IIX. Our automated method produces consistent libraries from eight tissue samples in three hours. We obtained no GC bias, <5% of adapter dimer sequence, and >80% of confidently mapped reads from libraries prepared from 1 μg of brain tissue sample. The automation of all downstream library processing steps after homogenization will be directly relevant to the development of clinical NGS systems.

Visualizing the Transcriptome: A Comparison of Different RNA Library Preparation Methods

Kanhai Khanna, Xiaoping Su, Louis Ramagli, Hongli Tang, and Erika Thompson
The University of Texas MD Anderson Cancer Center

RNA-seq is a powerful tool used to obtain in-depth information on expression profiling, gene annotation, and transcript discovery. With the growing popularity of RNA sequencing, new library preparation techniques are becoming commercially available. These techniques are improvements on the classic poly-A selection and rRNA reduction methods, and in some cases sensitive enough to analyze the transcriptome of a single cell. However, limited information is available on comparative analysis of these methods and their appropriate application for the transcriptome studies. We utilized Illumina’s HiSeq technology to compare the merits of four commercial sample preparation kits: NuGen’s Ovation RNA-seq system v2, Illumina’s TruSeq RNA Sample Preparation kit v2, Epicentre’s ScriptSeq RNA-seq kit v2 and Clontech’s SMARTer Ultra Low RNA kit. We found that the quality of input RNA was critical for optimum performance of SMARTer Ultra Low RNA kit. Ovation and ScriptSeq kits, on the other hand, worked well with moderate quality input RNA as well. Based on analysis of the sequencing data, 12% of reads from ScriptSeq mapped to the mitochondrial genes as compared to 24% reads from Ovation. The library complexity and percentage of reads aligning to non-exonic region was similar between both kits. However, 28% reads aligned to the coding region for ScriptSeq versus 18% for Ovation. While TruSeq and SMARTer kits are designed for Poly-A containing RNAs only, ScriptSeq and Ovation kits provide more global analysis of the transcriptome. Analyzing the differences between these methods provides a better understanding of their specific advantage over the other. This information is especially useful for Sequencing Core Facilities, to recommend and apply appropriate methods to different transcriptome studies.

Tailored Internal Communication Strategies for the Life Science Shared Resource

Peter Lopez¹, Linda C. Lopez²
¹NYULMC, “CUNY Baruch College

In the life-science laboratory management setting, there is a constant need to address the question, “what can make us better?” Although there are many variables that can be examined, a review of internal-communication practices can identify areas of opportunity. People in a shared-resource management position have a number of levels of communication responsibilities: communication with the organization’s
33 Developing a Culture of Research in Vermont: Training and Research Support for Faculty and Students Through Outreach

Kara Pivarski¹, Heather Driscoll², Tim Hunter²; Janet Murray²
¹Norwich University; ²University of Vermont

The Vermont Genetics Network (VGN) Outreach Core works with undergraduate college faculty throughout the state to implement and integrate cutting edge technology into their curricula and research programs. The opportunities afforded to undergraduates exposed to VGN Outreach activities allow them to gain important skills and encourage them to pursue research careers. The VGN Outreach Core, based at the University of Vermont (UVM) and Norwich University, has utilized the technologies and expertise from the three VGN Core Facilities at UVM – Microarray, Proteomics, and Bioinformatics – to support faculty and student research and to create educational modules.

Through these modules, our team has worked with over 20 faculty and 560 undergraduate students from eight baccalaureate partner institutions (BPIs). Seven of the eight colleges have integrated one or more modules into their curricula and all eight BPIs have reported other changes to curricula that were influenced by VGN Outreach interactions. Results from outreach surveys suggest that our outreach programs influence the educational and career goals of undergraduates. Students report an increased interest in studying science and pursuing a career in science or medicine after participating in VGN educational modules. Further results from the outreach surveys will be discussed.

Additionally, the VGN Outreach Core directly supports faculty and student research. The team works with faculty to enhance their research by bringing research into the classroom, integrating novel experiments into the modules, establishing relationships with the core facilities, offering technical support for project design, and providing extensive bioinformatics support. Core members also work with undergraduate students by serving as technical advisors and/or thesis committee members for independent senior projects and by providing research support. Through all the VGN Programs, a stronger culture of research is being developed at our partner institutes and around the state.

This work was sponsored by Grant Number P20 RR16462, from the IDeA Network of Biomedical Research Excellence (INBRE) Program of the National Institute of General Medical Sciences (NIGMS), a component of the National Institutes of Health (NIH).

34 Direct Sequencing of Linear DNA on PacBio RS Without Library Preparation

Savita Shanker, Xiao H. Zhou, Kevin Holland, David Moraga, Ernesto C. Almira, William G. Farmerie
Interdisciplinary Center for Biotechnology Research, University of Florida

All nextGen sequencing technologies require library construction as a first step where platform-specific adapters are ligated to target DNA fragments. Pacific Biosciences RS technology requires 5-10µg of DNA for large insert (~10kb) library construction. However, under circumstances such as when performing long-range PCR on DNA regions containing repetitive sequences or other difficult to amplify targets, only limited product is available. Cloning, which poses sequence stability problems due to the repetitive nature of the insert, is the only other option for generating sufficient DNA under these circumstances. Here, we demonstrate the direct sequencing of a 10kb linear DNA fragment starting with as little as 150ng input DNA using PacBio version 2 chemistry. These experiments were performed before PacBio released their megabase station upgrade. The megabase station reduces the amount of input DNA required for polymerase binding and this technical advance will clearly reduce the amount of input DNA required for directly sequencing linear DNA templates over the requirements of our initial experiment.

35 Modified Library Construction Method for 3-5kb Illumina Mate-pair Libraries

Savita Shanker, David Moraga, Neda Panayotova, Xiao H. Zhou, William G. Farmerie
Interdisciplinary Center for Biotechnology Research, University of Florida

NextGen sequencing is a powerful and cost efficient tool for ultra-high-throughput genome and transcriptome analysis. One of the key features of next generation sequencing is de novo whole genome sequencing, but assembly and genome finishing is still a major challenge due to short reads generated by these technologies. The 2kb-5kb mate pair reads combined with Illumina short pair-end reads are used in getting better genomic coverage across the genome. The standard 2kb-5kb Illumina mate-pair library construction protocol does not allow barcoding, and has built-in limitations that prevent getting more than 36bp reads at either end, as increasing read length can lead to elevated error rate. This is due to the fact that the junction reads cannot be identified easily if working with de novo assembly or those reads got discarded, since they would not align to reference sequence. Here, we demonstrate a modified 2kb-5kb mate pair library construction protocol for Illumina technologies that allows long barcoded, mate-paired reads without increasing error rates.
Rapid adoption of new sequencing technologies means that sample QC in this workflow not only needs to have high throughput capability, but flexibility is also critical. The Agilent 2200 TapeStation system meets this demand by providing a simple to use automated electrophoresis platform with truly variable throughput capabilities, from a single sample up to a 96-well plate. The D1K ScreenTape assay on the 2200 TapeStation enables analysis of DNA fragments between 35 and 1000bp, suiting NGS library preparation QC and quantification. This study is an assessment of the performance the D1K ScreenTape assay in high-throughput DNA analysis using the 96-well sample plate capability. Sizing and quantification were assessed for accuracy and precision on a foil-sealed 96-well plate over typical and extended run times. The study concluded that the D1K ScreenTape assay is suited for accurate sizing and quantification of DNA fragments in high-throughput applications.

Increased Sensitivity in Whole-Genome Bisulfite Sequencing (WGBS): A Novel "Post-bisulfite Conversion" Library Construction Method from Sub-nanogram DNA Inputs

Roy Sooknanan1, John Hitchen1, Dixie Gabell1, Nick Caruccio1, Andrew Adesy2
1Epicentre (an illumina® company), 2Department of Genome Sciences, University of Washington, Seattle, WA

Genome-wide analysis of 5-methylcytosines is possible with whole-genome bisulfite sequencing (WGBS), where unmethylated cytosine residues are converted to uracil. However, a major challenge in WGBS is the degradation of DNA that occurs during bisulfite conversion under conditions required for complete conversion. Typically ~90% of input DNA is degraded and thus, is especially problematic with limited starting amounts of DNA. Additionally, regions that are rich in unmethylated cytosines are more sensitive to strand breaks. As a consequence, a majority of DNA fragments contained in di-tagged NGS DNA libraries treated with bisulfite “post library construction” can be rendered inactive due to strand breaks in the DNA sequence flanked by the adapter sequences. These mono-tagged templates are then excluded during library enrichment resulting in incomplete coverage and bias when performing whole genome bisulfite sequencing.

Here, we describe a novel “post-bisulfite conversion” library construction method for preparing NGS libraries from genomic DNA prior to the addition of one or both adapters. This “post-bisulfite conversion” library construction method uses the resulting untagged or mono-tagged single-stranded DNA as template for the subsequent addition of adapter sequences required for NGS. Thus, single-stranded DNA fragments independent of size and position of strand breaks remain as viable templates for library construction, eliminating the loss of fragments and the selection bias associated with a “post-library construction” bisulfite conversion strategy. This novel “post-bisulfite conversion” library construction method exhibits high diversity, increased efficiency and sensitivity (500 picograms human genomic DNA input), and improved coverage required for WGBS.

New and Highly Effective Automated Solutions for Sample Preparation for Next-Generation Sequencing

Ray Ahmed, Salim Essakali, Jörg Dennig, Heike Glowatz, Nan Fang, Peter Hahn, Dominic O’Neil
QIAGEN, Hilden, Germany

Next-generation sequencing requires specialized and often time-consuming methods to select particular nucleic acid fractions and generate libraries suitable for sequencing. The complexity and time requirements of these methods make automation highly desirable, particularly as sequencing becomes more common and higher throughput. We present here new, automatable methods to deplete ribosomal RNA (rRNA) from a total RNA sample for subsequent sequencing and efficient, high-yield library construction. It is desirable to remove rRNA for RNA-seq, since it comprises 85–95% of total RNA, occupies valuable sequencing capacity, and results in a low signal-to-noise ratio that can make detection and analysis of the RNA species of interest difficult. Our method (the GeneRead rRNA Depletion Kit) effectively removes rRNA, while ensuring complete recovery of mRNA and noncoding RNA from various species; including human, mouse, and rat. The method involves specific oligonucleotide probes, designed to hybridize to the large (18s, 28s), small (5s, 5.8s), and mitochondrial (12s, 16s) rRNAs. The rRNA:DNA hybrid is recognized by a hybrid-specific antibody that can be captured on a bead and removed from the sample, depleting the rRNA. This antibody-mediated capture provides a higher level of specificity of rRNA depletion than other methods, works well with fragmented samples, and preserves noncoding RNA. The method can be performed manually or automated on the QIAcube from hybridization, including subsequent RNA cleanup. Kit performance was tested using qRT-PCR and sequencing. Comparison with other rRNA depletion techniques revealed the GeneRead rRNA Depletion Kit effectively eliminates rRNA, while better preserving the natural representation of other RNAs. This method improves the ratio of useful data, decreases bias, and preserves noncoding RNA, providing high-quality RNA highly suited for next-generation sequencing applications. For the Ion Torrent and Illumina platforms, we have developed methods that simplify the library construction process, leading to higher yields and time savings. We have integrated a single-tube protocol for library fragment trimming and adapter-ligation, followed by library purification and adapter-dimer depletion into one straightforward workflow. This enables construction of high-quality libraries from as little as 50 ng of nucleic acid and allows the process to be automated on the QIAcube. Multiple libraries prepared from one sample using the automated procedure on the QIAcube show a very high consistency of the resulting libraries with comparably high yields. The libraries generated also have the full-length library adapters, enabling the preparations obtained to be used directly for sequencing. For the optional library amplification step, a newly developed high-fidelity DNA polymerase can be used that minimizes the amplification-induced sequence biases in AT- and GC-rich regions.
39 High-Throughput Strand-Specific mRNA Library Preparation for Illumina Sequencing from Total RNA Isolated from Normal and Cancersovary Tissue

Shanavaz Nasarabadi, Jacob Berger, Paul Butler, Charles Troup, Tom Olenic, Mary Trounstine, Sayali Salodkar, Stephanie H.I. Yeung and Monika Tomczyk
IntegenX Inc., Pleasanton, CA

The cost benefits of whole genome and transcriptome next generation sequencing (NGS) have revolutionized genetic analysis. Determination of regulatory changes in cancer cells by whole transcriptome analysis has proven useful for tailoring treatment options for patients. An efficient, consistent and reliable whole transcriptome library preparation method is necessary for successful transition of the NGS technologies from research to clinical applications. Due to the labile nature of mRNA and the tedium of processing large numbers of patient samples, automation would greatly improve the reliability and throughput of whole transcriptome library preparation.

We have explored the automation of whole transcriptome library preparation from total RNA of normal and cancersovary tissue. The Apollo 324™ System (IntegenX Inc.) was used to isolate poly A mRNA from 6 to 48 samples of total RNA and prepare strand-specific mRNA libraries for sequencing on the Genome Analyzer Ix (Illumina). The cDNA output was amplified on a bench thermocycler to yield whole transcriptome libraries in a single eight hour day. While most conventional RNA-Seq library preparation methods convert mRNA to cDNA, in our strand specific library preparation, we ligated the adapters directly to fragmented mRNA to preserve strand polarity. Preserving strand polarity of the transcript reduces the bioinformatics bottleneck. A commercially available manual strand-specific library was used as the bench control for comparison and validation of our library preparations.

The gene expression profile of 12 up-regulated and down-regulated genes was equivalent between the high-throughput libraries and the published data. The isolated polyA mRNA had an average of 0.3% rRNA contamination from 500 ng of total RNA. There was a 75% time reduction for automated library preparation from total RNA compared with conventional library preparation methods. There was 90% correlation of gene expression between the automated and bench library preparation methods for up to 48 samples.

40 RNA Sequencing and Quantitation Using Targeted Amplicons

Brian Sanderson, Jeff Schagement, Angie Cheng, Kelli Bramlett
Life Technologies

As Next Generation Sequencing matures, it is quickly moving into translational research applications where it has promise to be a useful tool for diagnosing and treating in a clinical setting. RNA profiling using NGS (RNA-seq) is one of the applications where this potential is currently being realized. RNA-seq experiments have traditionally started with a whole-transcriptome library preparation that produces a sequencing template from all RNA species in a sample. However, in many cases, only a handful of the genes present are necessary to make a clinically relevant diagnosis.

We have demonstrated new technology that allows for RNA-seq from a panel of directed amplicons using an AmpliSeq™ approach with Ion Torrent semiconductor sequencing. This approach offers many advantages over microarray or qPCR such as faster turnaround and data analysis, sample multiplexing, lower RNA inputs, and ability to use degraded or FFPE-derived samples. In addition, the technique simultaneously provides quantitative gene expression information and gene sequence at the single nucleotide level.

We have compiled three gene panels for testing the method including a cancer panel, apoptosis panel, and a panel derived from the Micro Array Quality Control (MAQC) consortium. Starting with 10ng of total RNA, cDNA is made, followed by amplification using primers designed for targeted genes. Resulting amplicons are prepared for sequencing using the AmpliSeq™ technology and sequenced on the Ion Torrent PGM. We demonstrate that the technique produces results that are technically reproducible, quantitative, and have excellent correlation with qPCR using TaqMan® assays. Employing barcodes, we have also tested multiple samples on a single chip thereby increasing the cost-effectiveness of the tool for clinical and research use.

41 Development of a Modified Smart System for Robust Transcriptome Library Preparation from Limited Quantities of Compromised Samples

Magnolia Bostick, C. Chang, A. Farmer
Clontech Laboratories

Next Generation Sequencing has revolutionized biomedical research by providing sequence data on millions of short DNA fragments, in parallel. In particular, NGS has enabled RNA expression analysis over the entire transcriptome with high sensitivity and dynamic range. Currently, the field is seeking methods to utilize challenging samples that are either compromised or are only available in limited amounts. Overcoming these constraints will demand highly sensitive and robust sample preparation methods.

One powerful method for cDNA preparation is SMART™ technology (Switching Mechanism At the 5’ end of the RNA Template), which utilizes the template switching activity of reverse transcriptase to enable the direct addition of a PCR adaptor to the 3’ end of the first-strand cDNA, thus avoiding inefficient ligation steps. One drawback of SMART technology is its current inability to work with compromised samples, owing to its dependence on an oligo dT primer for first strand synthesis.

A modified SMART system has been developed including the use of random primers to work with samples containing compromised or degraded RNA. Data (including gene body coverage, reproducibility, and mappability metrics) will be presented for both chemically degraded RNA samples and FFPE RNA prepared using the modified SMART system.

This modified SMART protocol will be especially useful for small samples of degraded RNA. It is capable of generating cDNA libraries for transcriptome profiling from as little as 1 ng of total RNA.

42 cDNA Library Generation for Transcriptome Analysis From Total RNA Equivalent to a Single Cell

Magnolia Bostick1, C. Chang1, A. Farmer1, G. Schroth2
1Clontech Laboratories, “Illumina Inc.

By providing sequence data on millions of short DNA fragments in parallel, Next Generation Sequencing has revolutionized biomedical research. In particular, the technique has enabled RNA expression analysis over the entire transcriptome with high sensitivity and dynamic range. With this has come a drive to utilize smaller sample inputs—with the goal of analyzing the transcriptome of a single cell.

One powerful method for cDNA preparation is SMART™ technology (Switching Mechanism At the 5’ end of the RNA Template), which utilizes the template switching activity of reverse transcriptase to enable the direct addition of a PCR adaptor to the 3’ end of the first-strand cDNA. The result is a single-tube protocol that enhances library amplification efficiency, while minimizing the chance for contamination, making it ideal for library preparation from small
amounts of starting material. Indeed, the SMARTer™ Ultra Low RNA method allows researchers to readily obtain high quality data from as little as 10 pg of total RNA. Recent publications indicate that this technology is effective in single cell analysis.

Here we present data which demonstrates that SMART faithfully produces full-length cDNA, even as the template library sample preparation. Sequencing results for libraries generated from 0.01 to 10 ng of mouse brain total RNA demonstrate that even with just 10 pg of input RNA, over 90% of the data maps to the genome, and the average transcript coverage is as uniform as that seen with much greater amounts of RNA. Also, under all conditions tested, rRNA reads accounted for only 3–5% of the total reads. Finally, comparing results to quantitative PCR for the MAQC (microarray quality control) gene set shows a high correlation with libraries made with 1 or 0.1 ng RNA. These data indicate that the SMART cDNA preparation method is an ideal choice for single cell transcriptome analysis.

43 Greater than 10 kb Read Lengths Routine when Sequencing with Pacific Biosciences’ XL Release

Primo Baybayan, Cheryl Heiner, Susana Wang, Meredith Ashby, Yan Guo, Jason Underwood Pacific Biosciences

PacBio’s SMRT® Sequencing produces the longest read lengths of any sequencing technology currently available. There have been a number of recent improvements to further extend the length of PacBio® RS reads. With an exponential read length distribution, there are many reads greater than 10 kb, and some reads at or beyond 20 kb. These improvements include library prep methods for generating >10 kb libraries, a new XL polymerase, magnetic bead loading, stage start, new XL sequencing kits, and increasing data collection time to 120 minutes per SMRT Cell. Each of these features will be described, with data illustrating the associated gains in performance. With these developments, we are able to obtain greatly improved and, in some cases, completed assemblies for genomes that have been considered impossible to assemble in the past, because they include repeats or low complexity regions spanning many kilobases. Long read lengths are valuable in other areas as well. In a single read, we can obtain sequence covering an entire viral segment, read through multi-kilobase amplicons with expanded repeats, and identify splice variants in long, full-length cDNA sequences. Examples of these applications will be shown.

44 Rapid and Efficient Methods for Preparing Globin- and rRNA-Depleted Directional RNA-Seq Libraries

Anupama Khanna, John Hitchen, Roy Sooknanan Epicentre

Deep sequencing of cDNA prepared from total RNA (RNA-Seq) or mRNA (mRNA-Seq) has become the method of choice for transcript profiling, discovery of novel transcripts, and identification of alternative splicing events. However, standard whole-transcriptome approaches to RNA-Seq face a significant challenge, as the vast majority of reads map to rRNA. One solution—poly(A) enrichment—does not capture several biologically relevant RNA species, such as microRNA and also requires large amounts of Total RNA. Whole blood RNA samples pose an even bigger challenge due to the presence of reticulocyte globin transcripts that can constitute up to 76% of total mRNA and, unless removed along with the ribosomal RNA transcripts, the detection of low-abundance mRNA transcripts is further compromised. To overcome these challenges, Epicentre developed Globin-Zero™ technology for blood derived RNA. The Globin Zero Gold kit removes rRNA and globin mRNA from intact and fragmented blood total RNA samples (human, mouse and rat), and an improved, more user-friendly version of the ScriptSeq™ RNA-Seq library preparation method. Directional cDNA libraries with ~98% strandedness can be prepared in about 4 hours, in a single-tube workflow, from either intact or fragmented Globin-Zero treated RNA samples, without the need for end-polishing, adapter ligation, cDNA fragmentation or gel size-selection. The combined Globin-Zero and ScriptSeq workflow generates cluster-ready NGS libraries in ~6 hours. Globin-Zero-treated samples contain <2% of rRNA and globin sequences, while maintaining representation of coding and noncoding transcripts independent of polyadenylation. This reduction in rRNA and globin sequence reads improves sequence depth and coverage of mRNA, and increases the percentage of uniquely mapped reads.

45 A Fast Solution to NGS Library Prep with Low Nanogram DNA Input


Next Generation Sequencing (NGS) has significantly impacted human genetics, enabling a comprehensive characterization of the human genome as well as a better understanding of many genomic abnormalities. By delivering massive DNA sequences at unprecedented speed and cost, NGS promises to make personalized medicine a reality in the foreseeable future. To date, library construction with clinical samples has been a challenge, primarily due to the limited quantities of sample DNA available. Our objective here was to overcome this challenge by developing NEBNext™ Ultra DNA Library Prep Kit, a fast library preparation method. Specifically, we streamlined the workflow utilizing novel NEBNext reagents and adaptors, including a new DNA polymerase that has been optimized to minimize GC bias. As a result of this work, we have developed a simple method for library construction from an amount of DNA as low as 5 ng, which can be used for both intact and fragmented DNA. Moreover, the workflow is compatible with multiple NGS platforms.

46 Semiconductor Sequencing of Human Exomes on the Ion Proton System

Gavin Meredith, Srinka Ghosh, Gary Bee, Mitsu Reddy, Loni Pickle, Miroslav Dudas, Guy Del Mistro, Marina Sedova, Brian Reed, Rob Bennett and the Ion Torrent R&D Team Ion Torrent–Life Technologies, Carlsbad, CA

Rapid, accurate, and inexpensive sequencing of exomes is critical to understand DNA variation in human disease. Ion Torrent has developed a benchtop research semiconductor sequencer, the Ion Proton™, that uses a novel CMOS chip with 165 million 1.3mm-diameter microwells, automatically templated sub-micron particles, and integrated hardware and software that enables acquisition of ~5 billion data points per second over a 2-4 hour runtime with on-instrument signal processing.

To illustrate the speed, accuracy, and ease-of-use of the Proton system, analysis of a HapMap familial trio of exomes will be presented. Exome libraries are obtained with high-specificity hybridization probes targeting ~50 Mb of human exons that span 21,700 annotated protein-coding genes, microRNA, key non-coding RNA genes, and 44,000 predicted microRNA binding sites. Exome reads map on-target 75-83% between runs and 10.6 Gb of aligned data, obtained from a single P1 chip, yielded 141X average depth with 30X coverage of >90% of targeted bases. Read mapping, coverage analysis, variant calling and annotation are done with Torrent Suite and Ion Reporter™ software. Each trio dataset yielded ~30,000 SNP calls from single
runs that exceeded 9 Gb of aligned data. The observed Het:Hom ratio of 1.4-1.5 matches the published range of 1.25-1.7 for European ethnicity and the observed Ts:Tv ratio of 2.9 agrees well with the published range of 2.8-3.1 for human exomes. The SNP concordance with dbSNP137 is greater than 98% and Het and Hom concordances with Complete Genomics data are 98% and 96%, respectively. Mendelian inheritance analysis indicates that error for Hets is 0.6% with no errors for homozygotic SNPs. The Proton system delivers high-quality individual exome datasets rapidly and can be used for trio analysis to detect shared germline SNPs with high confidence.

The Ion Proton™ System is for research use only and not for use in diagnostic procedures.

47 Discovering the RNA Transcription Landscape using Directional Approaches

New England Biolabs, Inc.

High-throughput complementary DNA sequencing (RNA-Seq) is a powerful technique that allows for sensitive digital quantification of transcript levels. Moreover, RNA-Seq enables the detection of non-canonical transcription start sites and termination sites, alternative splice isoforms and transcript mutation and edition. Standard “next-generation” RNA-sequencing approaches generally require double-stranded cDNA synthesis, which erases RNA strand information. In this approach, the synthesis of randomly primed double-stranded cDNA followed by addition of adaptors for sequencing leads to the loss of information about which strand was present in the original mRNA template. The polarity of the transcript is important for correct annotation of novel genes, identification of antisense transcripts with potential regulatory roles, and for correct determination of gene expression levels in the presence of antisense transcripts. Our objective was to address this need by developing a novel streamlined, low input method for Directional RNA-Seq that highly retains strand orientation information while maintaining even coverage of transcript expression. This method is based on second strand labeling and excision after adaptor ligation; allowing differential tagging of the first strand cDNA ends. As a result, we have enabled strand specific mRNA sequencing, as well as whole transcriptome sequencing (Total RNA-Seq) from ribosomal-depleted samples. Total RNA-Seq provides a much broader picture of expression dynamics including discovery of antisense transcripts. This work presents a streamlined, fast solution for complete RNA sequencing, with high quality data that illustrates the complexity and diversity of the RNA transcription landscape.

49 Automation of Library Preparation for High-resolution ChIP-seq Profiling

Henriette O’Geen, Isabelle M. Henry, Rebecca Cotterman, Paul Knoepfler, Luca Comai and Ryan W. Kim
UC Davis Genome Center

The dynamic modification of DNA and histones plays a key role in transcriptional regulation through altering the packaging of DNA and modifying the nucleosome surface. These chromatin states, also referred to as the epigenome, are distinctive for different tissues, developmental stages, and disease states and can also be altered by environmental influences.

New technologies allow the genome-wide visualization of the information encoded in the epigenome. For example, the chromatin immunoprecipitation (ChIP) assay allows investigators to characterize DNA–protein interactions in vivo. ChIP followed by high-throughput sequencing (ChIP-seq) is a powerful tool to identify genome-wide profiles of transcription factors, histone modifications, DNA methylation, and nucleosome positioning. The low yield of ChIP assays presents a challenge for reproducible and high quality library preparation for high throughput sequencing. Using the automated library preparation system from IntegenX, we prepared ChIP-seq libraries from as little as 1 ng ChIP DNA material. Sequencing of biological replicates on the Illumina platform confirmed a 95%-98% overlap of identified binding sites. Progress of ChIP experiments using limited tissue amounts from rice seedlings as well as other applications of the library automation system will be discussed.

50 Automating Directional Small RNA Library Preparation for Illumina GA Sequencing

Shanavaz Nasarabadi, Jacob Berger, Paul Butler, Tom Olenic, Stephanie Yeung, Sally Smith
IntegenX Inc., Pleasanton, CA

We have demonstrated an efficient automated method for preparing high quality directional libraries for deep sequencing of small RNA (sRNA) using the Illumina Genome Analyzer. sRNA molecules are approximately 21 nt in length and play an important role in regulating gene expression. They are uniquely challenging to sequence because they degrade faster than other RNA molecules and represent a very small fraction (1% to 2%) of the RNA population depending on sample/tissue type. Although next generation sequencing allows researchers to rapidly sequence entire genomes and process many samples in parallel, the sample preparation can often be very tedious, time consuming, and prone to human errors. We have explored novel methodologies for streamlining and automating directional library preparation processes while enhancing sequencing data quality.

In this study, we demonstrated a process of enrichment for sRNA from raw tissue samples and automated library construction of the
enriched sRNA to increase throughput and reduce human error. While conventional RNA-Seq methods do not permit directional sequencing of RNA, this method of directional library preparation has the advantage of preserving strand polarity of the transcript to provide more valuable sequence data. The raw tissue sample (rabbit brain) was pre-enriched using the mirVana™ kit (Ambion) to purify and concentrate the sRNA. Following the manual enrichment step, the libraries were prepared from as little as 10 ng of enriched sRNA using the PrepXTM RNA-Seq Library Kit for Illumina (IntegenX Inc.) on the Apollo 324™ System to autonomously generate, in parallel, eight RNA-Seq libraries in three hours. The percent mRNA mapped for enriched samples was reduced from 12% to 3% when compared to the non-enriched RNA sample. miRBase analysis of the enriched sRNA samples increased the percentage of sRNA mapped reads from 54% to 74% with a two-fold increase in the number of known and unique miRs.

51 Quality Control and Pre-Qualification of NGS Libraries Made from Clinical Samples

John Langmore
Rubicon Genomics

We have compared sequencing metrics from different types of clinical samples and different methods of making NGS libraries, for purposes of quality control of the samples and sample preps. We have performed the metrics using the HiSeq and MiSeq, however the same QC metrics could be measured on other platforms. By choosing metrics that can be measured from small amounts of data (e.g., 300,000 reads), we can measure the quality of the clinical samples and NGS libraries in a highly multiplexed format, before spending considerable time and money for downstream processes such as sequence enrichment and NGS. More predictive than measurement of insert size and concentration, these metrics predict whether the amount and quality of genomic DNA, as well as the sample preparation method is sufficiently efficient to generate high coverage from deep sequencing. Sequencing results from sonicated human and bacterial DNA, as well as human maternal plasma show that different sample prep methods yield libraries of very different diversity, uniformity of coverage and background. Library quality can also vary considerably in different lots of reagents and also different number of amplification cycles.

By optimizing the efficiency of making genomic DNA into sequencing libraries fewer reads are necessary to achieve reproducible, high quality results, especially from limiting amounts of plasma, FFPE tissue, chromatin immune precipitates, and single cells.

52 Towards Clinical Grade Genomes with Joint Bayesian Variant Identification

Francisco De La Vega, Brian Hildbush Len Trigg, Richard Littin, John G. Cleary
Real Time Genomics, Inc.

The precipitous fall in the cost of sequencing spurred by innovations in high-throughput sequencing (HTS) is bringing the use of genome sequencing closer to the clinic. An important question yet to be answered is whether current HTS protocols provide data that meets clinical standards of quality. While false positives (FP) can be evaluated experimentally, false negatives are more difficult to assess due to the lack of an established gold standard. Sequencing of family pedigrees already enormously simplifies the identification of highly penetrant disease genes. However, joint analysis of family members raw data could also provide a significant boost in variant calling accuracy because related individuals share haplotype blocks. Here we present our Joint Bayesian Calling (JBC) method for pedigrees and show it reduces false positives & negatives and improves accuracy of identified variants in trios and larger pedigrees. Our approach reduces Mendelian variants in trios to 0.1% compared to 2% in singleton calling, and improves specificity of de novo variant identification by reducing FP >50%. We demonstrate JBC scalability to large pedigrees by analyzing sequencing data of a large CEPH pedigree where the genomes of 17 individuals where sequenced to ~50X. As more pedigree members are added accuracy improves and we are also capable of imputing genotypes of missing subjects. Our approach is able to inform trade-offs between depth of coverage and number of family members for research study planning, and coupled with a proprietary fast read mapping algorithm is able to analyze a full depth WGS trio in less than a day (hours for exomes) in commodity hardware. We believe these advances will be crucial for the adoption of genomes & exomes in clinical settings.

53 Which RNA-Seq Processing Algorithm Should I Pick? A Comparison of RNA-Seq Pipelines Based on Experimental Design, Sample Properties and Sequencing Technology

Stuart Levine, Zachary Banks, Vincent Butty
Massachusetts Institute of Technology, Cambridge, MA

In the past few years, the dramatic increase in sequencing throughput has been paralleled by a proliferation of bioinformatic tools designed to analyze and quantitate next generation sequencing data. In the case of RNA-Seq, a number of analysis packages have been developed, but whether some might be better suited to the analysis of samples within different ranges of read coverage or RNA quality has not been extensively studied. Here, we evaluated head-to-head eight different packages on single-end- and paired-end-RNA-Seq data at different read coverages. Gene expression and alternative splicing were quantitated and compared across the different methods using standard accuracy and precision metrics. We demonstrate critical differences in how the different algorithms are able to handle different read lengths and read coverage levels and how these changes impact gene expression and isoform usage values.

FLOW CYTOMETRY

54 Different Sorts for Different Folks: The Importance of Technological Diversity in a Cell Sorting Facility

Monica DeLay1, A. Nicole White1, Edith Janssen2, George Babcock3, Christopher A. Worth1, Sherry Thornton4
1Cincinnati Children’s Hospital, 2Research Flow Cytometry Core, Division of Rheumatology and Division of Cellular and Molecular Immunology, Cincinnati Children’s Hospital, Cincinnati, OH, 3Flow Cytometry Lab, Shriner’s Hospitals for Children-Cincinnati, Cincinnati, OH, 4Stem Cell Institute at James Graham Brown Cancer Center, University of Louisville, Louisville, KY

Flow cytometry core facilities have the challenge of offering appropriate technologies to diverse client needs. Therefore, there is no one-size-fits-all technology that can handle any and all cell types or applications. For cell sorting, this becomes even more of a challenge when a variety of cells need to be purified for an even wider variety of downstream applications. There are two types of technologies currently available for droplet based cell sorting; cuvette and jet-in-air. While there are advantages and disadvantages to both technologies, no cell sorter design can fit all needs, especially when the cell type to be sorted is sensitive to manipulations.
There have been reports (published and anecdotal) of phenotypic and functional changes to dendritic cells after isolation using different techniques. In our core facility, DC populations that were sorted on the FACS Aria II showed an increase in cell death and were found to be nonfunctional in an assay testing their ability to process cell-associated antigen and stimulate proliferation of T cells. In contrast, cells sorted on a FACSVantage, MoFlo Legacy or MoFlo XDP were able to function in the same in vitro assay. We attribute this difference to the differences in fluid dynamics through the sample path that may damage the cells and decrease their ability to function. Here we present data comparing the functionality of DCs with each system. We conclude that it is best to have a diversity of technologies in a cell sorting facility to better meet the needs of all clients.

55 Evaluation of Aerosol Containment of BD FACS Aria Cell Sorter in a Class II Bigneat "LABS" Robotics Enclosure

Lora Barsky1, David Stanley1, George Salazar2, and Stephen Pratt3
1University of Southern California, 2 Biosero LLC, 3 BigNeat Ltd

Steam in air cell sorting of unfixed human cell specimen poses a biosafety hazard to the occupants of the core facility as described by the International Society for Analytical Cytology Biosafety Standard for Sorting of Unfixed Cells. Assessment of aerosol containment with an Aerosol Management System (AMS) is a necessary validation for steam in air cell sorters commonly performed with the Glo Germ protocol as published by Perfetto et al. BD FACS Aria cell sorters may be placed in the Baker BioProtect IV line of Biological Safety Cabinet with integrated AMS, which eliminates the purpose of the Whisper AMS peripheral device and provides both operator and sample sterility. This study evaluates a Bigneat LABS custom robotics enclosure designed to contain a BD FACS Aria cell sorter as an alternative to Baker BioProtect in terms of its ability to contain and evacuate aerosols generated by the instrument. This enclosure provides Class II, type B containment offering operator protection and sample sterility to ISO 5 cleanroom level. Aerosol containment was assessed using the K1 discus test as described in the BS EN 12649:2000 standard to validate the integrity of the air curtain at the time of production and then again after site installation with BD FACS Aria enclosed. We questioned if the Whisper AMS was necessary to contain aerosols within this enclosure and have assessed the AMS utility under stream failure using the GloGerm assay as described by Perfetto et al. The Class II Bigneat "LABS" Robotics Enclosure passed all testing performed to access air curtain integrity and containment of aerosols generated by the FACS Aria cell sorter.

FRAGMENT ANALYSIS


Mark Arsenault1, Natalia Markova2 and Ronan O’Brien2
1GE Healthcare Bio-Sciences AB, Uppsala, Sweden, 2GE Healthcare, Northampton, MA and 3 Sprint Bioscience AB, Stockholm

Background:
Recent developments in ITC instrumentation, namely MicroCal ITC200 and MicroCal Auto-iTC200 have led to an increase in the throughput and decrease in the protein consumption of the technique. In addition, there have been recent methodological Advancements 1 that have extended the affinity range that ITC can measure into the mM range. The combination of all these factors has made the technique ideal for fragment-based drug discovery (FBDD) campaigns.

Method:
This work outlines the role of Microcal Auto-iTC200, in the fragment based drug discovery program of Sprint Bioscience, to identify and optimize potential drug candidates that will inhibit the activity of Vps34. This class phosphatidylinositol3-kinase is central to autophagy and has been shown to play an important role in resistance to cancer drugs2, 3. As such it has been identified as a target for therapeutic intervention. ITC is a generic assay without the need for assay development and as such the affinity of all 50 compounds was measured in less than three days after receiving the purified protein.

Conclusion:
This approach was fast and proved very successful for identifying fragments that co crystallized with the target protein. Of the 14 compounds chosen, based on the ITC data, 12 formed crystals that could be used in the optimization process.

GENETIC VARIATION

57 Uniparental Disomy is Associated with Embryonal Rhabdomyosarcoma in Costello Syndrome and Nonsyndromic Patients: From Core-Side to Bed-Side and Back Again, Impact of Core Laboratories on Biomedical Research

Katie Sol-Church1, Katherine M. Robbins2, Deborah L. Stabler3, Jennifer Holbrook2, Karen W. Gripp2
1 Nemours Center for Pediatric Research, A.I. duPont Hospital for Children Wilmington, DE, 2Division of Medical Genetics A.I. duPont Hospital for Children Wilmington, DE

Embryonal rhabdomyosarcoma (ERMS) is the most prevalent pediatric soft tissue tumor and is characterized at the molecular level by loss of heterozygosity (LOH) at 11p15.5. We aimed to study ERMS in the context of Costello syndrome (CS), a rare syndromic predisposition to ERMS. Because CS arises from a typically paternal germline mutation in the proto-oncogene HRAS, it is considered a rasopathy. CS results in failure to thrive, intellectual disabilities, short stature, coarse facial features, skeletal abnormalities, congenital heart disease, and a predisposition for cancer, most commonly ERMS. Using the Biomolecular Core Laboratory technical expertise and Nemours clinical resources, we molecularly characterized eight ERMS tumors from six unrelated CS patients, carrying paternally derived HRAS clinical deletions. We observed a second primary ERMS tumor showing the CS hallmark LOH at 11p15.5. The six other CS ERMS tumors displayed paternal uniparental disomy (UPD) of the entire chromosome 11. Paternal UPD was also observed in a pure ERMS cell line, established by our core lab using fresh tumor sample from one of the patients with a p.G12A germline mutation. This represents, to our knowledge, the first CS ERMS cell line in existence. Additional cell lines derived from non-CS patients were used in this study and included three ERMS (RD, RH18 and RH36), one Alveolar RMS (RH30). RH36 cells carried an HRAS mutation at position 61 (p.Q61K). Unexpectedly, both ERMS cell lines, RD and RH36, displayed a complete LOH for chromosome 11. The ARMS RH30 cells carried no mutation in HRAS or LOH. RH18 cells
derived from an ERMS had no HRAS mutation and showed LOH on both, the p and q arm of chromosome 11, except at locus D11S1338. Cytogenetic analysis with fluorescence in situ hybridization (FISH) on three CS tumors revealed multiple copies of chromosome 11 leading to trisomies and tetrasoniess in the tumor and cultured cells. Gene expression analysis revealed loss of imprinting in the growth regulating genes H19, IGF2, and CDKN1C on chromosome 11 not only in CS tumors, but also in nonsyndromic ERMS. Elucidating the basis of tumor-specific activation/inactivation pathways in a core lab setting is providing important determinants in patient diagnosis and in novel therapeutic design. This work illustrates one of the critical roles a core laboratory may play when fully integrated in biomedical research endeavors.

**GENOTYPING**

58 Human Cell Line and Tissue Sample Authentication

Kyle Hooper, Margaret M. Ewing, Robert S. McLaren, Kathryn D. Hebble, Kim Ready, and Douglas R. Storts

Promega Corporation

Background: Short Tandem Repeat (STR) genotyping analysis is a proven technology for uniquely identifying virtually all human samples. STR genotyping was adopted as the preferred technology for identification of human tissue culture cell lines by the ATCC Standards Development Organization (ASN-0002: Authentication of Human Cell Lines: Standardization of STR Profiling). We developed new automation-compatible protocols/systems for generating STR profiles from human cell lines or tissue samples.

Methods: We adapted the STR genotyping systems routinely used for forensic and paternity testing to better meet the needs of genomic core facilities. Modifications include balancing for higher amounts of template DNA, configuring the reagents for compatibility with high throughput robotic workstations, and supporting electrophoretic separation and analysis on a wider variety of instrument and software platforms.

Results: The GenePrint® 10 and GenePrint® 21 Systems allow for multiplexed genotyping of 9 or 20 STR loci, plus the amelogenin gender marker. STR genotyping with these loci uniquely identify virtually all human cell lines and tissue samples (body fluids, tissues and extracted DNA) and confirm the absence of cross-contamination or a sample switch. The GenePrint® 10 System is compatible with purified DNA or direct amplification from cells deposited on FTA® Cards (GE Whatman). Both genotyping systems meet the ASN-0002 standard and include the loci represented on the National Center for Biotechnology Information (NCBI) human cell line database: http://www.ncbi.nlm.nih.gov/biosample?term=human%20cell%20line%20STR%20%profile

Conclusions: STR genotyping analysis with the GenePrint® 10 and GenePrint® 21 Systems can establish human cell line or tissue sample identity and confirm the absence of contamination with other human cell lines or tissues. The methods are compatible with DNA concentrations, robotic protocols, instrumentation, and genotyping software typically used in genomics core facilities.

**HPLC OF PROTEINS AND PEPTIDES**

59 Increased Throughput for 2D LC in the Analysis of Human Placental Samples

Martha Stapels¹, Keith Fadgen¹, J. Will Thompson², M. Arthur Moseley¹, James Langridge¹

¹Waters Corporation, Milford, MA; ²Duke University School of Medicine, Durham, NC

Most proteomic samples generate peptides with similar distributions of hydrophobicities and mass (1). The complexity of proteomic samples requires orthogonal methods of separation to identify and quantify all peptides in a sample. Data-independent analysis yields reproducible fragmentation and peak area information for all detectable peptides (2). The use of ion mobility during this analysis inserts an orthogonal separation in the gas phase between chromatographic and mass spectral analyses. In this study, 2D chromatography is combined with ion mobility to resolve peptides in multiple dimensions in a high-throughput manner.

Proteins were extracted from human placenta samples into two solubility fractions using TRizol reagent with sonication. Proteins were then reduced, alkylated, and digested in-solution with trypsin. Samples were injected in triplicate onto a nanoflow liquid chromatography system and analyzed with a data-independent method using alternating low and elevated collision energy on a quadrupole time of flight instrument with ion mobility. Multidimensional chromatographic methods were employed using high-low pH RP-RP (3) with discontinuous step gradients.

A comparison was made between a traditional 2D method and a faster technique that utilized simultaneous gradients in both dimensions. The faster technique took 68% of the time of the traditional 3-fraction method and the percent savings in time will increase as the number of desired fractions increases. Use of the faster method allowed for a 70% increase in the number of ions detected per minute, a 54% increase in the number of peptides identified per minute, and a 46% increase in the number of proteins identified per minute. Incorporation of ion mobility into the analysis yielded an increase in peak capacity of at least another order of magnitude.

References:

**LIGHT MICROSCOPY**

60 Benchmarking the Resolution of 3D-SIM and STORM Using a Well-Characterized Biological System

V. Kaye Thomas, Alison J. North

Bio-Imaging Resource Center, The Rockefeller University, New York, NY

Super-resolution microscopy systems have become more prominent in fluorescence imaging with many researchers interested in implementing these technologies. The three main types of super-resolution systems, structured illumination, serial localization and emission depletions, are expensive, complex and vary in their capabilities and platforms. The resolution gained in these different systems ranges from two- to ten-fold beyond that of the diffraction limit, and in practice is also
dependent upon optimal sample preparation. Limited personal experience with these novel techniques can make it difficult to know which system will be the best purchase option for an institution. As one of the few facilities running both Stochastic Optical Reconstruction Microscopy (Nikon STORM) and 3-D Structured Illumination Microscopy (3D-SIM) (API OMX) systems, we are uniquely positioned to compare and benchmark these different super-resolution techniques in practice using a well-characterized biological sample.

The desmosome is a well-characterized biological structure with parallel cytoplasmic plaques on either side of the cell-cell junction. Each plaque component lines up at a distance from the plasma membrane that has been previously mapped by immunoelectron microscopy. Since resolution is defined as the ability to distinguish adjacent elements as separate components, the labeling of different plaque components is thus ideal for benchmarking the resolution achieved using our super-resolution systems. Two parallel fluorescent lines indicate resolution of the two protein domains, while a single line indicates that the resolution limit has been reached. For example, our results confirm predictions that the two bands of desmoplakin C-termini, ~130 nm separation, can be resolved by both 3D-SIM and STORM, while the desmoplakin N-termini, only ~50nm separation, can be resolved by STORM but not by 3D-SIM. Work is now in progress to maximize the resolution of each system by optimizing sample preparation and labeling, and to compare fixed and permeabilized cells with on-section staining.


61 Improving Dye Brightness and Photostability in Stellaris RNA FISH Assays

Mary Johansson, Sally Coassin, Arturo Orjalo, Matt Lyttle and Ron Cook
Biosearch Technologies

Stellaris RNA FISH (Fluorescence In Situ Hybridization) assays allow visualization of single molecules of RNA in fixed adherent cells and in tissue. A Stellaris FISH probe set contains multiple oligonucleotide sequences each appended with a fluorescent label. These probes collectively bind to a target transcript creating bright spots that can be quantified. Differently colored reporter fluorophores allow multiplexed assays that differentiate multiple RNA transcripts.

As with all new assays, there are opportunities for optimization and improvement. This study focuses on the photochemistry and photophysics and such challenges as dye blinking and bleaching. Dye bleaching can occur within seconds under standard wide field microscopy conditions using a metal halide lamp. We have compared several commercial antifade solutions and tested other formulations. We found that FAM (fluorescein) is incompatible with the widely used antifade glucose oxidase. Multiplexing with FAM can be achieved using other antifade solutions. Buffer composition, pH, antifade concentration, and viscosity can impact antifade and dye performance.

TSQ (triplet state quencher) moieties prevent dye blinking and can improve dye stability and brightness. We prepared and tested a series of Stellaris probes in which the reporter dyes were covalently attached to TSQ molecules. Although initial results were promising, our tentative conclusion is that the TSQ moieties that we tested were too hydrophobic and interfered with target hybridization. This continues to be an active area of research.

62 An Integrated Microscope for Laser Nanosurgery and Optical Trapping

Michelle Duquette, Veronica Gomez, Charlie Chandrawangwana, Marcello Harsono, Nick Hyun, Michael Berns

1University of California, San Diego, 2University of California, Irvine

We describe an advanced RoboLase microscope system that is comprised of laser scissors and two trapping beams; the three can be used alone or in combination in the same cell. For micro or nanoablative laser wavelengths may be in the near UV, green, or near IR. The pulse duration of the laser may be in the nanosecond, picosecond, or femtosecond regimes. The parameters used depend upon such factors as, desired spot size, absorption properties (natural or applied), and ablation mechanisms (single photon or multiphoton). The nanoablation system is used in studies on DNA repair (with or without an absorbing chromophore present) and chromosome microsurgery. The tweezers are used to move cut or whole chromosomes as well to trap and release sperm to study swimming force. The system serves as a “user” research device and can be accessed via the Internet.

63 User Support with Remote Desktop Software in a 24/7 Light Microscopy Facility

Peter Toth, Ke Ma and William Hendrickson
Research Resources Center, University of Illinois at Chicago

Supporting users in a light microscopy facility involves dealing with a large number of specific challenges; one of them is user support outside regular working hours. On one hand, the new incubator fluorescence microscope systems created new demands: multi-day experiments are now common and users want to intermittently check the status of the cells. On the other hand, troubleshooting for technical issues is easier by remotely controlling the mouse and keyboard of an image acquisition computer. This presentation will provide a comparative overview of remote desktop programs, which allow users to remotely and non-interactively check the status of overnight experiments, and more importantly, allow the core managers to remotely check and adjust the imaging parameters. Lastly, we will explain the basic requirements for installing a remote application support program and describe the program’s advantages as well as limitations and provide pointers for core administrators who want to introduce remote user support in their facilities.
FPs, TALEN’s and TALE-FPs) and can be fused (as cDNAs) to different fluorescent proteins and FP biosensors. Many biosensors are available as affinity peptides. I realized that spatial multiplexing of many DNA binding protein-reporters by localizing to different spots in the cell nucleus and distinguished by combinatorial addressing. When using cyan-yellow FRET biosensors, and blue, red and far-red FP’s enable 23 = 8 addresses. Adding orange and very-far-red (mNeptune) enable 25 = 32 addresses. Switching to single FP biosensors (ex. GCaMP6 for Ca++) will enable more addresses (26 = 64), higher signal-to-noise ratio and simplify experiments. Fluorescence nanoscopes will result in better spatial and temporal resolution, enabling use of even more simultaneous biosensors. I have released Tattletales to the public domain1.

4http://home.earthlink.net/~pubspectra/McNamara_20121023Tue_Tattletales_GFP_Public_Domain.jpg

**MACROMOLECULAR INTERACTIONS**

### 65 The CRAPome: a Contaminant Repository for Affinity Purification Mass Spectrometry Data


1Department of Pathology, University of Michigan, Ann Arbor, MI, 2Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, 3Centre for Systems Biology, Samuel Lunenfeld Research Institute at Mount Sinai Hospital, Toronto, ON, Canada, 4Department of Molecular Biology, Princeton University, Princeton, NJ, 5Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland, 6Stowers Institute for Medical Research, Kansas City, MO, 7Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands, 8Netherlands Proteomics Center, Utrecht, The Netherlands, 9Division of Cell Biology, Netherlands Cancer Institute, Amsterdam, The Netherlands, 10Department of Molecular Biology; Faculty of Science; Nijmegen Centre for Molecular Life Sciences; Radboud University; Nijmegen, The Netherlands, 11Institut de recherches cliniques de Montréal (IRCM), Montréal, QC, Canada, 12Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada, 13Department of Biochemistry, Université de Montréal, Montréal, QC, Canada, 14CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria, 15Saw Swee Hock School of Public Health, National University of Singapore, Singapore, 16Department of Pathology & Laboratory Medicine, University of Kansas Medical Center, Kansas City, KS, 17Ontario Cancer Institute, Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada, 18Center for Proteomics and Bioinformatics, and Department of Genetics and Genome Science, Case Western Reserve University School of Medicine, Cleveland, OH

Affinity purification coupled with mass spectrometry (AP-MS) is now a widely used approach for the identification of protein-protein interactions. However, for any given protein of interest, determining which of the identified polyptides represent bona fide interactors versus those that are background contaminants (e.g. proteins that interact with the solid-phase support, affinity reagent or epitope tag) is a challenging task. While the standard approach is to identify nonspecific interactions using one or more negative controls, most small-scale AP-MS studies do not capture a complete, accurate background protein set. Fortunately, since negative controls are largely bait-independent, we reasoned that the negative controls generated by the proteomics research community could be developed as a resource for scoring AP-MS data.

Here we present the Contaminant Repository for Affinity Purification (The CRAPome), currently containing AP-MS data from 343 control purifications conducted by 11 different research groups (www.craponome.org). Users employ an intuitive graphical user interface to explore the database, by either querying one protein at a time, downloading background contaminant lists for selected experimental conditions, or uploading their own data (alongside their own negative controls when available) and performing data analysis. The CRAPome database scores contaminants vs. true interactors based on semi-quantitative mass spectrometry data (normalized spectral counts) embedded in most mass spectrometry experiments. The Significance Analysis of INTeractome (SAINT) scoring scheme, in addition to a simpler Fold Change calculation (FC score) are used to score user-supplied data and return a ranked list of putative interactors. We also describe database structure and composition, provide examples of the use of this resource to filter contaminants with properly chosen controls, and demonstrate the utility of the scoring scheme for identifying bona fide interaction partners. The CRAPome accommodates a variety of purification schemes and, while currently focused on human data, will be expanded to other species.

### Improved Sensitivity in Label Free Technologies Provide New Possibilities for Structure Function Studies

Robert Karlsson1, Olof Karlsson1, Åsa Frostell Karlsson1, Natalia Markova1, Anders Marcussen1, Johan Weigelt2, Andrei Zhukov3, Jörg Bomke4 and Jos Buiks4

1GE Healthcare Bio-Sciences AB, Uppsala, Sweden, 2Karolinska Institutet, Stockholm, Sweden, 3Heptares Therapeutics Ltd, Hertfordshire, UK, 4Merck Serono, Darmstadt, Germany

Background: X-ray crystallography and NMR give detailed information about protein structure, while functional assays provide information on interaction dynamics. Here, three applications and their relevance to structure function studies are discussed.
1) Identification and characterization of fragments binding to protein targets
2) Selectivity of inhibitor binding to Poly ADP ribose polymerases (PARPs)
3) Kinetic analysis of drug like compounds binding to stabilized GPCRs

Method:
Fragment screening on Biacore 4000 system included direct and competitive assays for identification and affinity analysis of specific binders. X-ray crystallography was used for structural analysis. PARP selectivity studies were based on kinetic analysis using Biacore T200 system, while studies on stabilized GPCRs combined kinetic analysis with molecular modeling.

Conclusions:
Data obtained in these applications suggest an iterative workflow, where interaction analysis and structural methods are used in sequence to identify ligands and to improve the resolution in structure function studies.

MASS SPECTROMETRY

Scale-up and Validation of a MALDI-MS-based Inhibitor Screening Technology

Rakesh Ratnare1, Richard J. Kirby1, Sandra L. Nelson2, William L. Seibel1, Ruben Papoian2, and Kenneth D. Greis1

Proteomics & Mass Spectrometry laboratory, Department of Cancer and Cell Biology, 2Drug Discovery Center, College of Medicine, University of Cincinnati, Cincinnati, OH

Recently, mass spectrometry (MS)-based readout has been demonstrated to be a highly effective for high throughput screening (HTS) assays. The major advantages compared to the most common fluorescence readout are the paucity of false readouts, reduced reagent costs, and the ability to multiplex assays such that multiple therapeutic targets can be screened for inhibitor hits with one pass through the compound repository. Previously, we have developed MS-based methods for rapid and accurate compound screening for inhibitors to therapeutic targets. However, the limited use of MS-based methods with small test libraries has been insufficient to validate the overall utility of this readout for large screening campaigns. Thus, in this report, the MS-based readout technology was scaled to include a library of >30,400 compounds to systematically validate the reliability of MALDI-MS readout head-to-head versus a traditional methods of HTS. The target enzyme for these comparative assays is PKC-iota, which plays a role cancer cell survival, tumor growth and potentially invasion. First the MS-based assay was fully integrated into an automated workflow on a PerkinElmer Plate:Explorer HTS system in a 384-well format. Then, the primary screen of 30,400 compounds with both the MS and fluorescence-based readouts yielded a hit rate of 0.3% and 0.9% for the two methods, respectively. Only 29% of the MS-based hits confirmed in triplicate assays; however, 95% of those confirmed hits validated as concentration-dependent inhibitors with IC50 value ranging from low nM to high µM inhibitors. By contrast, 58% of the fluorescence hits were deemed as false positives since they produced fluorescence inhibition even in the absence of PKC-iota. Overall the data validate the utility of the MS-based readout in terms of sensitivity, reproducibility and minimal reagent cost. We are now investigating ways to incorporate screening technologies as an additional service and revenue stream for our core laboratory.

Mass Spectrometry Data Collection in Parallel at Multiple Core Facilities Operating TripleTOF 5600 and Orbitrap Elite/Velos Pro/Q Exactive Mass Spectrometers

Anthony Yeung1, K. Jones1, K. Kim1, B. Patel1, S. Kelsen2, A. Braverman2, D. Swinton2, P. Gafken2, L. Jones2, W. Lane2, J. Neveu2, H. Leung2, S. Shaffer2, J. Leszyk2, B. Stanley3, T. Fox3, A. Stanley4

1Fox Chase Cancer Center, Philadelphia, PA, 2Temple University School of Medicine, Philadelphia, PA,
3Lincoln University, Lincoln, PA, 4Proteomics Shared Resources, Fred Hutchinson Cancer Research Center, Seattle, WA, 5Mass Spectrometry and Proteomics Resource Laboratory, Harvard University, Cambridge, MA,
6Mass Spectrometry and Proteomics Core Facility, Baylor College of Medicine, Houston, TX, 7Proteomics and Mass Spectrometry Facility, UMass Medical School, Shrewsbury, MA, 8Mass Spectrometry Core, Penn State College of Medicine, Hershey, PA

Proteomic research can benefit from simultaneous access to multiple cutting-edge mass spectrometers. 18 core facilities responded to our investigators seeking service through the ABRF Discussion Forum. Five of the facilities selected completed four plasma proteomics experiments as routine fee-for-service. Each biological experiment entailed an iTRAQ 4-plex proteome comparison of immunodepleted plasma provided as 30 labeled-peptide fractions. Identical samples were analyzed by two AB SCIEX TripleTOF 5600 and three Thermo Orbitrap (Elite/Velos Pro/Q Exactive) instruments. 480 LC-MS/MS runs delivered >250 GB of data over two months.

We compare herein routine service analyses of three peptide fractions of different peptide abundance. Data files from each instrument were studied to develop optimal analysis parameters to compare with default parameters in Mascot Distiller 2.4, ProteinPilot 4.5 beta, AB Sciex MS Data Converter 1.3 beta, and Proteome Discover 1.3. Peak-picking for TripleTOFs was best by ProteinPilot 4.5 beta while Mascot Distiller and Proteome Discoverer were comparable for the Orbitraps.

We compared protein identification and quantitation in SwissProt 2012_07 database by Mascot Server 2.4.01 versus ProteinPilot. By all search methods, more proteins, up to two fold, were identified using the Q Exactive than others. Q Exactive excelled also at the number of unique significant peptide ion sequences. However, software-dependent impact on subsequent interpretation, due to peptide modifications, can be critical. These findings may have special implications for iTRAQ plasma proteomics. For the low abundance peptide ions, the slope of the dynamic range drop-off in the plasma proteome is uniquely sharp compared with cell lysates. Our study provides data for testable improvements in the operation of these mass spectrometers. More importantly, we have demonstrated a new affordable expedient workflow for investigators to perform proteomic experiments through the ABRF infrastructure.

(We acknowledge John Cottrell for optimizing the peak-picking parameters for Mascot Distiller)
Alternative Enzymes Lead to Improvements in Sequence Coverage and PTM Analysis

Robert Brazas¹, Kyle Hooper¹, Michael Rosenblatt¹, Marjeta Urh¹, Sergei Saveliev¹, Chris Hosfield¹, Gary Kobs¹, Michael Ford², Richard Jones², Ravi Amunugama², David Allen²
¹Promega Corporation, Madison, WI, ²MS Bioworks, Ann Arbor, MI

The profiling of proteins using biological mass spectrometry (bottom up proteomics) most commonly requires trypsin. Trypsin is advantageous in that it produces peptides of optimal charge and size. However, for applications in which the proteins under investigation are part of a complex mixture or not isolated at high levels (i.e. low ng from an immunoprecipitation), sequence coverage is rarely complete. In addition, we have found that in several cases, like phosphorylation, acetylation, and methylation, alternative proteases are required to prepare peptides suitable for MS detection. This poster will provide specific examples which demonstrate this observation. For example, the application of a combined Trypsin/ Lys-C mixture reduces the number of missed cleavages by more than 3-fold producing samples with lower CV’s (for biological replicates). The mixture is also well-suited for the complete proteolysis of hydrophobic, compact proteins. The addition of chymotrypsin and elastase has been found to be useful for identifying phosphorylation sites on proteins, especially on sequences where the site of phosphorylation inhibits trypsin (i.e. proximal to K or R). Many epigenetic applications have focused on histone modifications, like lysine acetylation and arginine methylation. Alternative proteases like Asp-N, Glu-C, and chymotrypsin have been especially useful given the fact that the modified K and R residues are resistant to c-terminal cleavage by trypsin. Finally, in the case of serum profiling, the addition of the endoglycosidase, PNGase F has been found to improve sequence coverage due to the removal of N-linked glycans.

Visualisation and Identification of Proteins Directly from a Single On-Tissue Tryptically Digested Sections Using MALDI Imaging HDMSE

LeRoy Martin, E Claude, M Towers, K Neeson, J Vissers
Waters Corporation

Mass spectrometry Imaging (MSI) was first reported for the direct analysis of intact proteins from tissue. The need for the identification of proteins and therefore their enzymatic digestion has become an essential step in the discovery of biomarkers from tissue. Structural identification is traditionally carried out after processing untargeted MS image data followed by further manual acquisitions either on the same or consecutive sections.

A data independent MALDI imaging acquisition method is presented, where MS and MS/MS information are acquired within a single experiment, without any precursor selection or selection requirements. The precursors and the fragments are correlated post acquisition on the basis of their common drift time, which is further refined utilizing spatial distribution commonality.

On-tissue digestion step was carried out where a trypsin solution was sprayed on tissue sections using a SunCollect nebulising spray device. After overnight incubation at 37°C, CHCA matrix was evenly applied with the same nebulising spray device. Data were acquired using a MALDI SYNAPT G2 instrument with tri-wave ion guide optics to separate ions according to their gas phase mobility.

The MALDI imaging experiment was designed such that adjacent pixels had low and elevated collision energy applied in the Transfer T-Wave collision cell situated after the drift cell. The low energy function, comprising tryptic peptide precursor ion images, and the elevated energy function, providing fragmentation profiles ion images, were independently visualised within the Analysis tab of HDI software. Using the drift time alignment functionality, tryptic peptide precursors were associated with product ions that share similar drift times. For additional and supplementary refinement, another level of correlation was achieved on the basis of spatial distribution commonality. A .pkl file is generated for MASCOT database search where improvement of the MASCOT score was demonstrated.

MALDI Imaging HDMSE: A Novel Data Independent Technique for the Visualisation and Identification of Lipids Directly from a Single Tissue Section

LeRoy Martin, E Claude, M Towers, K Neeson, J Vissers
Waters Corporation

Lipidomics is an emerging application area that is taken advantage of the mass spectrometry developments that enabled the simultaneous analysis of a wide range of analytes. The spatial localization of lipids within tissue microstructures is however often lost during the lipid extraction process and challenges image analysis. MS imaging developments on the other hand allowed for the mapping of lipid species in entire tissue section. Structural identification is typically performed after processing the untargeted MS imaging data, followed by further manual acquisitions either on the same or consecutive sections.

A data independent acquisition method called MALDI Imaging HDMSE is presented, where MS and MS/MS information are acquired within the same experiment, without any precursor selection. Post-acquisition, precursors and fragments are correlated on the basis of ionic drift in the gas phase, which is further refined utilizing commonality of their spatial distributions.

Proof-of-principal experiments have been carried out using part of a rat whole body section where CHCA matrix was applied evenly to the sample in several coats using a SunCollect nebulising spray device. Data were acquired using a MALDI SYNAPT G2 instrument with tri-wave ion guide optics to separate ions according to their gas phase mobility. Within the same MALDI imaging experiment, the mass spectrometer was set to apply alternate collision energies to the transfer cell between low energy and elevated collision energies, with the latter inducing lipid fragmentation. Since fragmentation occurs post-ion mobility separation, the precursors at low energy have the same drift time as their associated fragments from the elevated energy scan. The dataset was subsequently processed using High Definition Imaging (HDI) MALDI software for enhanced image analysis. Drift time alignment and spatial correlation of the data from both low and elevated energy functions were carried out with the same platform.
72 Novel Approaches to Managing and Contrasting Complex Ion Mobility MALDI Imaging Datasets

LeRoy Martin¹, E Claude¹, M. Towers¹, K Neeson¹, M Djidja², J Erler²
¹Waters Corporation, ²Institute of Cancer Research

Mass spectrometry imaging (MSI) has proved to be a powerful analytical tool for the detection, localization and identification of many analytes, including metabolites, lipids and proteins, originating from complex, biological sample surfaces. MSI experiments can generate vast amount of data, depending on image size and acquisition mass range, which will both directly relate to the number of ions detected, the number of pixels and possibly the addition of ion mobility to improve the specificity of the analysis.

Tissue sections used were xenograph tissue, where the rat animals were administrated an anti-cancer drug called Dasanitib at a concentration of 30 mg/kg, and scanned at different time points (1 and 3-hours). In situ digestion was performed with a trypsin solution being sprayed directly onto the tissue samples and left to incubate overnight. Several layers of matrix, α-cyano-4-hydroxycinnamic acid (CHCA) containing aniline in acentonitrile: water:TFA (1:1:0.1), was also sprayed directly onto the tissue samples.

We are presenting a new approach where HDMS Compare software is used in combination with High Definition Imaging (HDI) software. HDMS Compare is a powerful analytical tool that investigates the data by comparing two datasets based on multi-dimensional differences in the m/z and drift time domains. It measure differences between samples that are believed to be very similar and were the difference cannot by detected by MS only. After Comparing, Inspecting and Detecting the 2-D plot images, a peak list was generated where the tryptic peptides were more abundant in the 3-hours vs. the 1-hour post dose tissue sections. The peak list was used to generate ion images of the contrasted tissue sections in the HDI software to confirm the highest intensities and identity of the tryptic peptides in the 3-hour tissue section.

73 Evaluation of Digested Wheat Gluten as a Proposed Reference Material for ELISA Analysis

Sara C. McGrath, G.M. Sharma, P.V. Rallabandi, J.H. Callahan, K.M. Williams, M.M. Ross
USFD/A/CFSAN, College Park, MD

Introduction: For millions of celiac disease sufferers, foods containing wheat gluten must be totally avoided. Intact gluten proteins can be reasonably quantified using commercially available ELISA kits, but these assays provide a poor measure of gluten when it has been hydrolyzed or fermented. A protocol for digesting gluten using multiple enzymes to mimic hydrolysis and thereby generate a hydrolysate of gluten could be used as a reference standard in competitive ELISA kits. Several methods have been proposed, but the resulting hydrolysate has not been verified using an orthogonal analytical technique.

Methods: Hard Spring wheat was ground into flour and extracted in two stages using 70% ethanol (fraction A) and 2% B-mercaptoethanol/1% acetic acid (fraction B). After extraction, the protein content of the fractions was measured, and 10 mg of each was quantitatively digested using pepsin and trypsin. The resulting peptide mixtures were analyzed via SDS-PAGE and mass spectrometry. Samples were spiked with 1 pmol/L Angiotensin I as an internal standard and quantitative surrogate. MS/MS results were searched against the Pooidae subfamily protein database from NCBI using Mascot and evaluated using Scaffold.

Results: Mascot results indicate highly-confident matches to common gliadins and glutenins from wheat, matching 69 out of 75000 possible proteins. The distribution of peptides had an average mass of 1900 Da, with 50% of detected peptides having masses between 1500-2000 Da. Ten percent of peptides had a mass greater than 3000 Da, with peptides as large as 35 residues detected. Alignment of peptides to intact gluten proteins showed that many peptides contained single AA substitutions and that the low specificity of pepsin produced peptides of variable and inconsistent length. The data were searched for peptides containing epitopes relevant to two commercial ELISA kits: R5 (R-Biopharm) and G12 (Romer Labs). In total, 21 exact peptide matches and 5 known degenerate matches were found.

74 A Global Analysis of the Mycobacterium Tuberculosis Lipidome

Brandy L. Young, Zandile Mlamlal, Julian Peters, Zac McDonald, Jonathan Blackburn
University of Cape Town

Mycobacterium Tuberculosis (Mtb), the causative agent of tuberculosis disease, has the most structurally complex and diverse lipids of any species. However, the flood of information from genomic research and proteomic research has not been matched by the corresponding advancements in lipidomics. The lack of lipidomic assays has been a glaring deficiency in the collection of analytical techniques currently available. We have transferred into our lab a method to characterize the global lipidome of several Mtb strains. This methodology relies on Liquid Chromatography-Mass Spectrometry (LC-MS) coupled with accurate mass determinations from a catalogued database that has 22,606 entries specific to Mtb. We show lipids from all the major lipid classes and we also show changes in the lipid content with variations in the Mtb strain. This methodology can be adapted to any biological sample and represents a huge advancement for characterizing the lipids (molecules that are implicated in promoting virulence and protecting the bacilli) found in Mtb.

75 MALDI Analysis of Modified Oligonucleotides Containing Dark Quenchers

Igor Smirnov, Galina E. Pozmogova
Institute of Physical-Chemical Medicine of Ministry of Public Health, 1a Malaya Pirogovskaya Str., Moscow, Russia

Oligonucleotides containing black hole quenchers are widely used as the probes for quantitative and RT-PCR applications. For reproducible and stable PCR results it is essential to use high grade probes, which makes oligonucleotide purification and QC steps important for success. Typical methods, such as HPLC and capillary electrophoresis yield information on overall purity of the probe, while mass spectrometry provides more chemical information hardly accessible by other analytical techniques, such as nature of synthetic failure products, degree of deprotection, presence of deterioration products due to aminopurination etc. While MALDI-TOF mass spectrometry is a method of choice for QC of synthetic oligonucleotides it’s application toward modified oligonucleotides containing dark quenchers is limited because of complication of spectra due to photo-dissociation reaction and the specific limitation of photo-dissociation step. These in-plume photo-decomposition reactions lead to appearance of spectral artifacts interfering with the actual signals and lead to more propensity for peak broadening at elevated laser powers. Such effects are observable for both nitrogen gas (337nm) and solid Nd-YAG (355nm) lasers typically used in MALDI mass spectrometers. In this work we discussed mechanism of photo-degradation and investigated various ways to minimize it. Short oligonucleotides containing BHQ-1, BHQ-2, and BHQ-3 labels were synthesized for testing purposes. The photo-fragmentation degree was evaluated in presence of various...
MALDI matrices. Different approaches to quench the exited states of quencher groups during ionization were investigated. That includes collisional cooling by using neutral additives (such as sugars) in order to facilitate energy dissipation from BHQ-groups and application of various organic additives bearing red-ox functionality, which could quench in-plume exited triplet states of azido groups. The novel matrix composition has been developed, which inhibits the photo-fragmentation of BHQ-groups during desorption-ionization step, thus eliminating spectrum artifacts and allows for successful analysis of BHQ-labeled oligonucleotides by MALDI MS.

**METABOLOMICS**

**76** Bioenergetics Reprogrammed in a Cellular Model of Friedreich’s Ataxia: Frataxin Deficiency Affects Formation of Respiratory Supercomplexes and Down-regulates Both OXPHOS and Glycolysis

Tao Wang¹, Bennett Van Houten²

¹North Dakota State University, ²University of Pittsburgh

Friedreich’s ataxia (FRDA) is an inherited progressive neuro- and cardiodegenerative disease that is caused by a reduced expression of the mitochondrial iron-binding protein, frataxin. FRDA is associated with decreased ATP level. This has been attributed to mitochondrial dysfunction as a result of frataxin deficiency, specifically decreased activities of respiratory complexes. In a stable human cell model of FRDA created by knocking down frataxin in LNA28 using lentiviral shRNA, we found that glycolysis, as well as OXPHOS, was significantly down-regulated, leading to reduced steady-state ATP levels. In consistence, PCR array analysis revealed an overall down-regulation in genes involved in glucose metabolism. Meanwhile, the fatty acid oxidation was also reduced by frataxin deficiency. And blue native gel electrophoresis analysis indicated that frataxin deficiency affected assembly of respiratory complexes and formation of supercomplexes. We conclude that frataxin deficiency exerts far more complicated effects on energy metabolism than previously thought, and bioenergetics reprogramming plays an important role for the pathological development of Friedreich’s ataxia.

**77** Discovery of Lipid Alterations in Biological Samples using UPLC, Ion Mobility Mass Spectrometry and TransOmics

LeRoy Martin, G Astarita, G Isaac, J Shockcor, J Langridge
Waters Corporation

Lipids play essential roles in health and disease. The discovery of novel alterations in lipid levels related to human diseases could lead to the development of novel biomarkers and shed light on the etiology of many human diseases. The challenge with global lipid analysis is the chemical complexity and the large range of concentrations of thousands of lipid species that are present in biological samples. Here we present a robust workflow for global lipid profiling, which employs UPLC, ion mobility mass spectrometry and TransOmics software for high throughput discovery of lipid alterations in biological samples. Lipid extracts from biological matrices were separated using Charged Surface Hybrid (CSH) C₁₈ UPLC according to acyl chain length and number of double bonds. Major lipid classes were identified using both positive and negative ESI. After chromatographic separation, lipids were analyzed using a hybrid QTOF system upgraded with an Ion Mobility Separation (IMS) cell capable separating lipids according to their size and shape. For example, differences in the acyl chain length or number of double bonds affect the shape and size of lipid molecules, resulting in characteristic migration times across the drift cell. Therefore, IMS provided an additional degree of separation besides chromatography, improving peak capacity and increasing selectivity over conventional UPLC. Data were analyzed using TransOmics software developed for visualization, processing, and interpretation of multi-dimensional MS data.

In conclusion, we present a simple and robust solution for the high-throughput, automated identification of lipid alterations using novel analytical and informatics tools. The use of UPLC coupled to ion mobility provides multiple degrees of orthogonal separation, delivering unprecedented peak capacity required for the confident identification of lipid species in a biological mixture.

**MICROFLUIDICS**

**78** A Novel Assay for Automated Electrophoretic Analysis of Genomic DNA

Marcus Gassmann¹, Delphine Rabiller², Donna McDade-Walker², Arunkumar Padmanaban³, Ruediger Salowsky¹

¹Agilent Technologies, Waldbronn, Germany, ²Agilent Technologies, Loanhead, UK, ³Life Sciences Application Centre, Agilent Technologies, Bangalore, India

The success of any genomic study depends primarily on the quality of the starting material, like genomic DNA (gDNA). The integrity of the extracted gDNA affects the downstream applications like microarray hybridization and next generation sequencing library construction. Since these are expensive and time consuming applications, a quality control (QC) of the genomic DNA has become highly recommended. The integrity of gDNA has traditionally been analyzed using agarose gel electrophoresis, which is manual, cumbersome and can involve exposure to hazardous chemicals like ethidium bromide. The new Agilent gDNA ScreenTape has been developed for the electrophoretic separation, sizing and quantification of large DNA samples (up to >60000 bp). This ready-to-use device, which runs on the Agilent 2200 TapeStation instrument, provides a reproducible QC method for analyzing the integrity and quantity of genomic DNA combined with the convenience of an automated system. With minimal sample preparation, automated loading and a variable throughput system, digital results can be presented as gel image, data table and in an electropherogram view. The ability to overlap and compare electropherograms within the software enables the discrimination of sample quality across different degradation states, sample types and concentrations. We present data that shows the Genomic DNA ScreenTape can easily verify the integrity of gDNA starting material with high sensitivity, precision and accuracy.

**79** High-Throughput Microfluidic Applications

LeRoy Martin, A Doneanu, J Murphy, J Johnson, S Cohen, G Astarita, A Chakraborty
Waters Corporation

A novel platform was developed for the expansion of microfluidic LC-MS beyond typical nanoscale applications into areas traditionally performed at analytical scales. The material used to fabricate the device allows for pressures of 12,000 psi. The integrated metal electrospray emitter supports flow rates from 100s of nl/min up to 8 µl/min and allows for the use of 150 µm in ID separation channels. Peak widths are on par with commercial ultra-high pressure LC instrumentation and cycle times as low as 10 min. System versatility will be demonstrated by several examples including intact proteins analysis, lipidomics and biopharmaceutical applications.
All experiments were performed using a nanoflow system coupled with an oaTOF mass spectrometer fitted with an electrospray ionization source designed to accommodate the microfluidic device. The emitter is connected to a packed channel by zero dead-volume connections and incorporates the use of nebulising gas. The separation channel was 150 µm x 5 cm and packed with sub-2-µm reverse-phase particles of various chemistries.

For intact protein analysis, using Ribonuclease A, Cytochrome C, Holotransferrin and Aponyoglobin, good resolution was achieved using shorter alkyd chain resins. Additionally, light and heavy chains of reduced monoclonal antibodies could be separated. Separations were performed at 80°C, using a 3 µL/min linear gradient from 20% to 55% ACN over 5 min. Appropriability for biopharmaceutical use is shown with minute amounts of Trastuzumab, demonstrating peak widths less than 6 s with 90% sequence coverage. Methionine oxidation and sites of glycosylation were also detected.

The method was also employed for lipid analysis showing improved separation for the major classes, particularly inter and intra forms. The microfluidic system offered high retention time reproducibility with RSD values smaller than 0.2%. This is especially useful for lipidomic analysis, which requires the comparison of a large number of LC-MS chromatograms from multiple sample sets.

80 Genomic DNA Quality Assessment by an Automated Microchip Electrophoresis Platform

Mai Ho, Tobias Wheeler, Ryan Swenerton, Kevin Zhu, Rajendra Singh, Bahram Fathollahi
Caliper (a PerkinElmer company), Hopkinton, MA

Assessment of the quantity and integrity of genomic DNA (gDNA) is an important step in the preparation of Next Generation Sequencing (NGS) libraries. Analysis of these samples prior to downstream NGS applications can prevent wasted time and resources. By implementation of a quality control method degraded gDNA can be excluded from further preparatory steps and from sequencing. A widely used method to assess gDNA quality are agarose gels, but in addition to being labor intensive and not automatable, limited quantitative data is offered. To ease this limit, we have developed a rapid, microfluidic assay for determining the quality of gDNA on the LabChip GX, an analytical platform that is also used for quantifying and sizing DNA and RNA in several other steps of the NGS sample preparation workflow. In this presentation we will describe the use of the assay for quantification and determination of gDNA sample integrity by a quality metric, the gDNA Quality Score (GQS). The GQS ranges from 0 to 5, with 5 representing the highest quality and can be used to establish acceptance criteria for further analysis of a sample.

82 A Comparison of DNA Purification Methods for Sanger Sequencing and Library Size Selection

Michael Zianni, S. Martin, A. McCoy
The Ohio State University

Purification of DNA is a critical process for many aspects of molecular biology including DNA sequencing by automated capillary electrophoresis and library preparation for Next Generation DNA sequencing. Towards this end there are many options including alcohol precipitation, size exclusion chromatography, and solid phase reversible immobilization (SPRI). Two new SPRI reagents were tested for effectiveness and ease of use as compared to these other techniques and a previously used SPRI reagent (SPRI1) by purifying multiple DNA sequencing reaction products by the use of each method and analyzing the quality of the resulting electropherograms. Second, AMPClean was compared to another SPRI reagent for its ability to select certain size fragments by assessing the resulting DNA for its quality with a Bioanalyzer and quantity with quantitative PCR. Third, AMPClean was compared to the other SPRI reagent for its ability to purify an amplicon from residual primer dimers through the use of the Bioanalyzer. DTRClean had advantages as compared to the ethanol precipitation, but performed about the same as the other two methods. AMPClean performed the same as the other SPRI reagent in all three tests. Both new products do show the ability to generate similar quality results to other methods commonly used in the Facility with the possibility of reduced costs.
Optical mapping generates an ordered restriction map from single, long DNA molecules. By overlapping restriction maps from multiple molecules, a physical map of entire chromosomes and genomes is constructed, greatly facilitating genome assembly in next generation sequencing projects, comparative genomics and strain typing. However, optical mapping relies on a method of preparing high purity DNA >250 kb in length, which can be challenging from some organisms and sample types. Here we demonstrate the ability of Boreal Genomics’ Aurora instrument to provide pure, high molecular weight (HMW) DNA 250-1,100 kb in length, ideally suited for optical mapping. The Aurora performs electrophoretic DNA purification within an agarose gel in reusable cartridges, protecting long DNA molecules from shearing forces associated with liquid handling steps common to other purification methods. DNA can be purified directly from intact cells embedded and lysed within an agarose gel, preserving the highest molecular weight DNA possible while achieving exceptional levels of purity. The Aurora delivers DNA in a buffer solution, where DNA can be condensed and protected from shearing during recovery with a pipette. DNA is then returned to its regular coiled state by simple dilution prior to optical mapping. Here we present images showing HMW DNA purification taking place in the Aurora and subsequent images of single DNA molecules on OpGen’s Argus® Optical Mapping System. Future work will focus on further optimizing Aurora HMW DNA purification to bias DNA recovery in favor of only the longest molecules in a sample, maximizing the benefits of optical mapping.

Formalin fixed paraffin embedded (FFPE) samples are commonly used for archiving pathological samples for molecular oncology labs and researchers. Traditional methods for the purification of nucleic acids from FFPE tissue samples are often labor intensive, include the use of hazardous organic reagents, and involve difficult pre-processing steps. Here, we describe an automated method for the purification of DNA and RNA from FFPE tissue sections using the Maxwell® 16 instrument that simplifies pre-processing and minimizes hands-on time. Pre-processing of FFPE tissue sections involved a simplified protocol with no xylene or phenol extraction required. Following pre-processing, samples were placed directly into the Maxwell® 16 cartridges, and purified nucleic acid was ready in approximately 45 minutes. Amplifiability was analyzed by qPCR or RT-qPCR. Both sample processing and amplification were performed in replicate to assess reproducibility. Nucleic acid was successfully purified from a variety of tissue sections including human breast, lung, and colon. DNA or RNA recovery was consistent as determined by real-time amplification. The RNA method also had very low contaminating DNA due to a pre-processing DNase step included prior to addition into the Maxwell® 16. Longer amplicons (200+ bp) were readily amplified using this system. In conclusion, DNA and RNA were successfully purified from FFPE tissue samples using the Maxwell® 16 instrument. Automated purification decreases hands-on time, provides more consistent results from difficult to purify sample types, and reduces the risk of nuclease contamination.

Library construction for whole transcriptome sequencing from low-quantity RNA samples requires additional amplification after reverse transcription into cDNA. Several approaches have been developed to minimize variability and biases.

This study aimed at quantifying and characterizing transcripts amplified using four commercial kits: (1) NuGEN Ovation RNA-Seq System V2, (2) Clontech SMARTer Ultra Low RNA Kit for Illumina Sequencing, (3) Sigma Transplex WT2-SEQ Kit and (4) Miltenyi Biotec µMACS SuperAmp Kit II for NGS.

The amplification reactions for each method were started with input amounts of Universal Mouse Reference RNA (UMRR) equivalent to approximately 10 to 300 cells. Resulting libraries built according to Illumina’s TruSeq procedure were compared to unamplified references prepared from 1-2 mg of UMRR total RNA that was either enriched for poly-A RNA or depleted of rRNA.

Sequencing data were evaluated for read alignment, library complexity, transcript coverage, and gene expression with regard to sensitivity and dynamic range. The Sigma kit-derived samples showed overrepresentation of genes categorized as snoRNAs, snRNAs and pseudogenes. The current Miltenyi Biotec protocol yielded low library complexity and increase of multi-mapped reads. Samples amplified with the Clontech and NuGEN methods performed well across all criteria. These kits will be tested further using cells as starting material.

Ligases are gaining utility in molecular biology applications, such as nucleotide sequence detection, single nucleotide polymorphism (SNP) detection, protein detection and “next generation” sequencing by ligation. With the increased demand for DNA ligases in the field of biotechnology, comes increased demand for ligation fidelity. Described approaches to improved ligation fidelity include ligases from different biological sources, point mutations of key amino acid residues within the ligase, modified reaction conditions and addition of crowding reagents, such as PEG. Although most approaches to improved ligation fidelity have focused on the ligase itself, further improvements are needed and may be attainable by a different approach. Herein a strategy to improve the discrimination between matched and mismatched targets is described which employs chemical modification to the nucleic acid components of the reaction, such as the donor probe, the acceptor probe and the ATP cofactor. The results demonstrate that chemically modified components increase the stringency of DNA ligation-mediated nucleic acid detection, providing a unique approach for SNP genotyping.
87 Chemical Variant of 7-deaza-dGTP for Improved GC-rich PCR and Amplification

Sabrina Shore, Natasha Paul, Gerry Zon, Richard Hogrefe, Alexandre Lebedev
TriLink Biotechnologies Inc.

PCR amplification of nucleic acids is a fundamental technique used in many molecular biology laboratories. Despite its widespread use, GC-rich regions of DNA sequence remain a challenge for amplification. Sequences high in GC content can form strong secondary structure, which prevents strand denaturation and blocks processive DNA polymerase amplification. As a consequence, mis-priming is prominent and can complicate specific product formation. Especially as applied to the molecular diagnosis of inherited diseases, several assay modifications have been developed to improve the specificity of target amplification. These approaches include specialized polymerases, Hot Start assays, addition of organic molecules, and thermal cycling alterations. However, as the GC content increases, the combination of two or three approaches may be required. Here, we show how 7-deaza-dGTP, a commonly used molecule to amplify GC-rich targets, can highly improve results when a thermodlabile protecting group is incorporated at the 3’-hydroxyl. The presence of the protecting group blocks low temperature primer extension and only allows nucleotide incorporation at higher temperatures when the protecting group is removed, improving PCR specificity as a result. Amplification specificity of GC-rich targets is further improved when a Hot Start version of all dNTPs are employed. This 7-deaza-dGTP containing Hot Start mixture allows for challenging targets of more than 85% GC content, such as Fragile X, to be amplified. Another benefit of this technology is in downstream sequencing reactions. PCR amplification of problematic targets with a Hot Start 7-deaza-dGTP mix prior to Sanger dyeexy sequencing can significantly improve the read quality along the entire sequence. In summary, the use of dNTPs modified with thermodlabile chemistry simplifies GC-rich amplification and provides a valuable solution that can improve disease diagnosis.

88 Rapid Approaches Towards DNA Damage Analysis

Manasa Paidipalli, N. Crews
Louisiana Tech University

The severe effects of DNA damage on human health have led to a tremendous amount of research being focused. Owing to the importance of damage detection, different approaches for the detection and quantification of the damaged DNA will be presented. In this work, we have modeled DNA damage using well-known mutagens: UV radiation to create photoproducts and restriction enzymes digestion to create double strand breaks. We will show that quantitative PCR (qPCR), a widely known measure of detecting the presence of the target DNA can be used to quantify photoproducts/intramolecular DNA damage. Our results indicate that a comparison of the initial concentration available in the undamaged and the damaged samples can be used to reveal the effect of damaged DNA in its amplification. By analyzing multiple regions using this technique, their relative susceptibility to damage can be measured. We also show that high-resolution melting analysis (HRMA), a measure of the bond energy between two DNA strands, can be used to quantify double strand breaks. The strand breaks resulted in a change in the overall distribution of the bond energy thus causing variations in the melting profile. HRMA has also been examined to compare DNA damage resulting from UV-A, UV-B and UV-C irradiation. The evaluation techniques demonstrated can potentially be extended to various types of DNA damage.

89 Monitoring Protein Synthesis in Living Cells with Fluorescent Labeled tRNA FRET Pairs

Zeev Smilansky1, Sima Barhoom2, Ian Farrell1, Dvir Dahary3, Marcelo Ehrlich1, Barry S. Cooperman1 and Orna Elroy-Stein1
1Anima Cell Metrology, Inc., 2Department of Cell Research and Immunology, Tel Aviv University, 3Department of Chemistry, University of Pennsylvania, Philadelphia

We introduce Protein Synthesis Monitoring (PSM) – a technique to monitor protein synthesis in live cells. In PSM, we transfect cells with tRNA labeled as FRET donors and acceptors. A FRET signal is generated only when a donor- and an acceptor-labeled tRNA come in close contact (<7nM), as they do on the ribosome during elongation. The intensity of the FRET signal correlates with the number of ribosomes engaged in protein synthesis, providing a real-time, live-cell assay for measuring rates of protein synthesis. PSM can be used to monitor the rate of either total protein synthesis (overall PSM), using bulk tRNAs, or of the synthesis of a specific protein (specific PSM), using specific pairs of tRNA to mark the protein of interest. PSM has sub-micron spatial and sub-second temporal resolutions. Cells continue to live and grow normally, and the synthesized proteins are unchanged since the labeling is on the tRNA itself and not on the amino acid. We have demonstrated specific PSM for monitoring synthesis of a viral protein during viral infection using Isoleucine tRNA, and for monitoring synthesis of collagen during fibrosis in mouse fibroblasts using tRNA-Gly and tRNA-Pro, as collagen is distinguished by many repeats of Gly-Pro dipeptides. We will discuss these results as well as additional applications of PSM in basic research, drug discovery, cell sorting, neurobiology, cancer, biomanufacturing, viral infections, and various protein synthesis-specific diseases.

90 Rapid Protein Analysis: Reagents for Improved GC-rich PCR and Amplification

Hugh White, F. Curtis
Lonza Rockland Inc., Rockland, ME

Protein sample preparation, separation in denaturing gels, detection by staining, and Western blot transfer followed by immunological detection are key elements of basic protein analysis in many lab settings, but can be very time-consuming. Through selection and use of optimized reagents the typical workflow can be dramatically reduced from 8 hours or more, to less than 4 hours in total elapsed time. This process flow uses reagents allowing: lysate preparation in <10 minutes, separations in 20-25 minutes (PAGEr™ EX gels), staining or transfer in 10 minutes (ProSieve™ EX Safe Stain, or ProSieve™ EX Transfer Buffer), and immunological detection in 1-2 hours. This optimized process utilizes reagents that do not require dedicated and costly systems. The reagents used offer further advantages: the PAGEr™ EX gels used for rapid separation are stable for 1 year of refrigerated storage, and the ProSieve™ EX Safe Stain, and ProSieve™ EX Transfer Buffer do not include, or require the use of, any hazardous materials. This optimized workflow has been verified by demonstrating the rapid detection of the widely used control target glyceraldehyde-3-phosphate dehydrogenase (GAPDH), with sensitivities comparable to standard methods.
91 Analysis of Local Dynamics of Human Insulin and a Rapid-acting Insulin Analog by Hydrogen Deuterium Exchange Mass Spectrometry

Joomi Ahn1, Shiori Nakazawa2,3; Noritaka Hashii2; Kenji Hirose6; Nana Kawasaki1,2; Waters Corporation, Milford, MA; 2Hokkaido University, Sapporo, Japan; 3National Institute of Health Science, Tokyo, Japan; 4Nihon Waters, Tokyo, Japan

Human insulin, used by diabetics to regulate blood sugar, was first introduced as a recombinant therapeutic drug nearly 30 years ago. Human insulin and insulin lispro have identical primary structure, except for the transposition of two amino acids. Lispro is one of the rapid-acting insulin analogs, which has higher tendency to dissociate than human insulin. In this study, we present an analytical workflow to allow us to detect the difference in the oligomeric dynamics using Hydrogen Deuteration Exchange Mass Spectrometry (HDX MS). The HDX analysis on insulin and lispro peptides was conducted to identify the location where different deuterium uptakes were observed between human insulin and lispro. The detected areas were illustrated in various formats to help understand their flexibility associated with rapid dissociation of insulin oligomers.

Drug products, human insulin (Humulin R) and lispro (Humalog), were reduced and digested online by pepsin. Deuterium labeling, quenching, and injection to on-line pepsin digestion were prepared using a robotic sample manager. Labeling experiments in 0, 0.5, 5, 10, 60, and 180 min interval were duplicated for both samples. The peptic digests were separated on a UPLC system at 0 °C. Q-TOF MS was used to measure the deuterium incorporation of identified peptides. The amount of deuterium was determined by automated HDX data processing software, DynamX 2.0.

We obtained 98% of sequence coverage for both human insulin and lispro. From peptide HDX determination, two regions were revealed distinctive different values in deuterium uptakes between human insulin and lispro; the N terminus of chain A, and a region adjacent to the C terminus of chain B. We attributed this localized behavior to the relation of hexamerization and dimerization, respectively. Furthermore, characteristic profiles that showed different deuteration margins between two insulins were determined, which was also consistent with their involvement in hexamer and dimer formation.

92 Biochemical Characterization of Human IL-15 Receptor-alpha

Oleg Chertov1, E. Chertova1, R. Sowder1, J. Lifson2, C Bergamaschi3, B. K. Felber4, G. N. Pavlakis5
1Protein Chemistry Laboratory, ATP, SAIC-Frederick, Frederick National Laboratory for Cancer Research, Frederick, MD; 2AIDS and Cancer Virus Program, SAIC-Frederick, Frederick National Laboratory for Cancer Research, Frederick, MD; 3Human Retrovirus Section and 4Human Retrovirus Pathogenesis Section, Vaccine Branch, CCR, NCI, Frederick, MD

The cytokine IL-15 is a growth, mobilization and activation factor for NK cells, CD8+ T cells and intraepithelial L T lymphocytes. IL-15 activates target cells via a heterodimeric receptor consisting of a beta subunit (CD122) and a gamma chain (CD132) common to the IL-2 receptor. Recently, we demonstrated that in vivo in both mice and humans IL-15 circulates as a heterodimeric cytokine in a noncovalently associated 1:1 complex with a soluble processed form of the IL-15 receptor alpha subunit (IL-15Ra). Soluble IL-15Ra is generated by proteolytic cleavage close to the transmembrane region of the membrane spanning molecule non-covalently linked to IL-15. To identify the cleavage site and other potential post-translational modifications we purified soluble IL-15Ra from the supernatant of stable HEK293-derived human cell lines overproducing IL-15/IL-15Ra heterodimers using reverse phase HPLC. MS analysis and Edman sequencing of sIL-15Ra and its Lys-C peptides allowed identification of C-terminus sequence and thereby the cleavage site. Multiple O-glycosylation sites were identified by mass spectrometry of sIL-15Ra and its Lys-C peptides.

Contract No. HHSN26120080001E

93 Use of LC-MS/MS to Identify Matrilysin/Matrix Metalloproteinase-7(MMP7) Cleavage Products of Perlecan (PLN)/Heparan Sulfate Proteoglycan (HSPG2)

Christopher Pennington1, Grindel, B.J.1, Chung, L.W.2, Farach-Carson, M.C.1
1Rice University, Shared Equipment Authority, Rice University, Houston, TX; 2Biochemistry and Cell Biology, Rice University, Houston, TX; 3Uro-Oncology Research Program, Samuel Oschin Comprehensive Cancer Institute at Cedars-Sinai Medical Center, Los Angeles, CA

Introduction
PLN/HSPG2, a 450 KDa heparan sulfate (HS) proteoglycan, is expressed in the basement membrane (BM) underlying epithelial and endothelial cells. Break down of PLN and other BM proteins is thought to play a critical role in prostate cancer (PCA) metastasis. During PCA metastasis, a variety of proteolytic enzymes are expressed. However, proteases associated with PCA, that can cleave PLN, have yet to be identified. An interesting extracellular protease upregulated in PCA invasion, whose ability to cleave PLN has not been investigated, is matrilysin (MMP-7). As part of a broader investigation LC-MS/MS analysis was used to identify cleavage products from MMP-7 PLN domain (DM)-IV3 digests.

Experimental Methods
DM-IV of PLN was recreated recombinantly and purified as three separate pieces (IV1, IV2, IV3). DM-IV3, a 75 KDa PLN fragment, was incubated with MMP-7 and then subjected to LC-MS/MS analysis on a Thermo-Fisher LTQ-Orbitrap MS. The MS was operated in the ESI(+) mode using data dependent MS/MS scanning.

Results
A total of 39 MMP-7 PLN DM-IV3 peptide fragments were identified by LC-MS/MS. They ranged in mass from 722 to 2475 Da with parent mass errors ranging from 0.7 to 3.9 ppm. Peptides were first identified by performing a “no-enzyme” MASCOT search on a small PLN database containing DM-IV3. Mascot de novo sequencing and sequencing via mMass was also employed to confirm the identity of PLN peptide fragments. Seven larger peptides ranging in mass from 3827 Da to 10855 Da were also observed. LC-MS data indicates MMP-7 has a strong preference for cleaving aliphatic amino acids (L, I, V, A) at the P1? site. PLN DM-IV3 sequence coverage was 49% for the 39 identified peptides.

Conclusion
LC-MS/MS data indicates that PLN DM-IV3 is cleaved by MMP-7 with fragments ranging in mass from 722 to 10855 Da being produced.

POSTER ABSTRACTS  ABRF 2013 TOOLS FOR THE ADVANCEMENT OF CONVERGENCE SCIENCE 27
**Assessing mRNA Translation: Deep Sequencing of Ribosome Footprints**

Fraz, Syed; Ramesh Vaidyanathan, Scott Kuersten, Agnes Radek, and Sajani Swami
Epicentre Biotechnologies

A translating ribosome protects a discrete fragment of ~30 nt on its mRNA template from nuclease digestion [Steitz, JA, (1969) Nature 224, 957]. Ribosome profiling determines the positions of active ribosomes on cellular mRNA by deep sequencing the ribosome-protected mRNA fragments [Ingolia, NT et al. (2009) Science 324, 218]. Quantitative measurement of abundance in this complex library is obtained by counting the number of times any given sequence is read by deep sequencing. The rate of protein synthesis is measured by examining the ribosome density of each mRNA.

Current methods to isolate polysomes and monosomes rely on several hours of ultracentrifugation using either a sucrose gradient or a cushion. We have investigated size-exclusion chromatography (SEC) in disposable spin-columns as an alternative to ultracentrifugation to isolate polysomes and monosomes for ribosome profiling. The size-exclusion method is simpler and rapid, and does not require any special equipment. Using this method we have streamlined the steps to convert ribosome-protected RNA fragments into libraries compatible with sequencing on Illumina® instruments.

We present ribosome profiling data generated using these methods in conjunction with bioinformatics tools for analyzing the deep sequencing data in mammalian samples.

**New Automated Systems for Size-fractionation of Protein Samples**

Rosemary O’Brien, Singh, Ezra Abrams, Todd Barbera, Chris Boles
Sage Science, Inc.

Sage Science has developed two new systems for automated preparative electrophoresis of protein samples. In both systems, the user simply loads samples into precast gel cassettes, programs the desired collection ranges into instrument software, and starts the run. At run completion, the user removes the desired protein fraction(s) from membrane-bounded elution modules within the cassette. The fractionated samples are recovered in SDS gel buffer – no gel extraction is required. Instrument software controls of timing of protein fractionation using input from on-board optical detection units, which monitor the progress of fluorescently-labeled molecular weight markers during electrophoresis.

The two system differ significantly in fractionation strategy. The first system is based on Sage’s existing BluePippin electrophoresis instrument for nucleic acid fractionation, but features precast gel cassettes developed specifically for protein fractionation. The protein BluePippin cassettes are useful for collecting one or two targeted protein fractions from a sample.

The second system, the Sage ELF system (for Electrophoretic Lateral Fracti0nator), is a completely new instrument/cassette system designed to fractionate a protein sample into 12 contiguous size fractions.

In both systems, SDS agarose gels are used as the separation matrix, and fractionated proteins recovered in SDS gel buffer. 5% agarose gels are used for proteins 18 to 80 kd in size, and 3% agarose gels are used for sample 50-200 kb in size. The width of the size collection window can be varied by the user, with a minimum window of about 20 kd.

The new Sage systems should provide increased reproducibility and ease-of-use for procedures that require gel isolation of proteins, such as top-down proteomics studies and other protein mass spectrometry methods.

---

**On-line Chip based 2D RP/RP LC-MS/MS Method for Proteomic Analysis**

Xiang Zhu¹, Christie Hunter²; Jenny Albanese²; Remco van Soest³
¹Eksigent, part of AB SCIEX; ²AB SCIEX

Two-dimensional (2D) liquid chromatography is widely used for proteome identification and quantification using the advantage of increasing peak capacity. Herein, we developed an simplified chip based 2D-LC workflow using a high pH/RP first dimensional separation and low pH/RP secondary dimension coupled to mass spectrometer for proteomic analysis.

The 2D LC separation was performed using ekspert™ nanoLC 425(Eksigent, part of AB SCIEX) system. Digested E. coli cell lysates were first loaded onto a 200µm x 15cm C18 column at pH 9.8 with a flowrate of 1 µL/min. Step gradient was used to sequentially elute peptide fractions, which were diluted to pH 2.5 before being captured by a 200µm x 6mm C18 chip trap column. Each fraction was then separated with a 75 µm x 15cm C18 chip column at 300 nL/min and analyzed with TripleTOF® 5600 (AB SCIEX). Data was processed with ProteinPilot™ Software (AB SCIEX).

The preliminary result suggested the on-line 2D RP-RP method as an easy and competitive approach for proteome discovery. The comparison experiments of 1D, 2D-6 fraction and 2D-10 fraction were performed. Using ~1ug of E. coli digested cell lysates, there are 1.8x and 2.1x increase in the identification numbers for 2D-6 and 2D-10 fractions versus the 1D configuration at the peptide level (5% local FDR). The other advantage of 2D workflow is the larger sample loading capacities on the column. When the loading amount was increased by 10x, the number of detected peptides increased by 3.5x and 4x for the 2D 6 and 10 fraction workflows, respectively, over 1D workflow.

The measured retention time of peptides detected in both the 1D 1ug and 2D 10ug experiments showed very good correlation (r2 0.99, slope 1.0). Further optimization of both the first and second dimension is ongoing to further improve the peptide detection rates.

**Determination of the Intra-individual Variability of Protein Expression in Benign/Cancerous Prostate Tissue via LC-MS – Efforts Towards Molecular Classification**

Michael Bereman, Martine Roudier, Larry True, Michael MacCoss
University of Washington

With a 1 in 6 lifetime risk and over 200,000 diagnoses per year in United States, prostate cancer is the most common noncutaneous cancer in men. The discrepancy between the number of diagnoses and the mortality risk (1 in 35) has led to scrutiny in the clinical management of the disease. The incidence of the disease has more than doubled since the FDA approval of prostate specific antigen test. Despite its low specificity, it has decreased the proportion of metastatic cancers at diagnosis. However, it has also increased the total number of cancers diagnosed with a majority of them having either indolent disease or completely benign conditions. This discrepancy between increased treatment and decreased disease aggressiveness has lead to much criticism that prostate cancer is “overdiagnosed” leading to unnecessary treatment. To this end, it is of interest to develop protein markers or panels of markers that are more indicative/specific of disease severity than currently available. Herein, we begin this endeavor by determining the precision of protein expression from different regions of benign tissue and Gleason 6 grade cancer...
tissue within the same individual using laser capture microdissection coupled to LC-MS.

After LCM, tissue were lysed and digested using established laboratory procedures. 500ng of protein was separated by LC and analyzed by a velos orbitrap mass spectrometer operating in a top 10 data dependent mode. Data was searched using Sequest and imported into skyline for differential abundance determination. Statistical design of experiments was used to optimized data dependent settings and afforded a 33% increase in unique peptide identifications from our current default DDA settings. 2000 proteins were identified including prostate specific antigen and prostatic acid phosphatase. Label free methods will be used to develop an intra individual variability index of both benign and cancerous tissue which will aid in statistically identifying differences.

98 Qualitative and Quantitative Characterization of the Metabolome, Lipidome and Proteome of Human Hepatocytes Stably Transfected with Cytochrome P450 2E1 Using Data Independent LC-MS

LeRoy Martin1, S Geenen2, C Coccociuri1, L Gethings1, G Isaac1, L Fernandes1, R Tonge1, J Vissers1, J Langrige1, I Wilson1
1Waters Corporation, 2AstraZeneca

Drug toxicity is a major reason for the failure of candidate pharmaceuticals during their development. It is therefore important to realize the potential for toxicity in a timely fashion. Many xenobiotics are bioactivated into toxic metabolites by cytochromes P450 (CYP). However, the activity of these enzymes typically falls in in-vitro systems. Recently, a transformed human hepatocyte cell line (THLE) became available in which the metabolic activity of specific CYP isoforms is maintained. THLE cells could be an ideal system in which to examine the potential toxicity of candidate pharmaceuticals. The baseline effect of the addition of CYP2E1 into THLE hepatocytes has been characterized to better understand the biochemistry of this model system.

Dedicated and independent sample preparation protocols were applied in order to isolate metabolites lipids and proteins. Three independent replicates of THLE null or THLE +2E1 cells were investigated for all analyte classes. Proteins were recovered and digested with trypsin overnight. The same LC-MS Omics Research Platform was used for all experiments and generic, application dependent LC conditions applied throughout. In all instances, MS data were acquired using a data independent analysis (DIA) approach, whereby the energy applied to the collision cell was switched between a low and elevated energy state during alternate scans. For the proteomics experiments, ion mobility separation (IM) was incorporated into the analytical schema (IM-DIA). Multi-omic data were processed and searched using TransOmics software, allowing for normalized label-free quantitation. Pathway analysis and systems biology experiments were conducted to interrogate the datasets further using various bioinformatics tools.

Comparison of the correlation variance and fold change between the two groups illustrates significant analyte expression. Data interpretation by means of clustering, statistical, and data analysis approaches have shown protein, lipid, and metabolite data to be complimentary and confirmative, which is further supported from the resulting pathway analysis output.

99 Comparison of Proteome Discoverer and PEAKS Studio for Phosphoproteome Analysis

Pawel Olszowy1, Melinda Wojtikiewicz2, Jayme Wiederin2, Pawel Ciborowski2,3
1Department of Environmental Chemistry and Bioanalytics, Nicolaus Copernicus University, Poland, 2Mass Spectrometry and Proteomics Core Facility, University of Nebraska Medical Center, 3Department of Pharmacology and Experimental Neuroscience, University of Nebraska Medical Center

Phosphorylation of proteins is a reversible post-translational modification that is fundamental in the regulation of cellular processes, such as protein function, cellular signaling and protein complex formation. The ability to perform a global analysis of phosphoproteome will give us insight as to how mechanisms of proteins and pathways might become altered due to changing experimental conditions. Because only a fraction of proteins or any given protein is phosphorylated, enrichment of sample for phosphoproteins or phosphopeptides is essential to the success of identification. Briefly, 100 ug of mouse whole cell lysate was digested in-solution with trypsin overnight. Samples were cleaned-up using PepClean columns, and then processed using TiO2 Phosphopeptide Enrichment and Clean-up Kit (Pierce, Inc.). The enriched samples were injected onto nano-LC-LTQ Orbitrap with ETD in configuration of HCD and ETD as two microscans complementing information from MS/MS spectra. The acquisition method was created in data dependent mode with one precursor scan in the Orbitrap, followed by fragmentation of the 4 most abundant peaks in both ETD detected in the LTQ, and in HCD detected in the Orbitrap. Tolerances were set to 10 ppm for the Orbitrap precursors and fragments while 0.8 Da for the ETD fragments. The following dynamic modifications were applied: Phospho / +79.966 Da (S, T, Y), Oxidation / +15.995 Da (M), Carboxymethyl / +58.005 Da (C). A Decoy database was also searched and a target FDR set for 0.01. For protein identification and localization of phosphorylated sites we used two software packages, Proteome Discoverer and PEAKS Studio. We present here a comparison of the search output obtained from these two software packages.

100 Micro-scale Sample Preparation for C-terminal Protein Characterization by Mass Spectrometry Using Combined Liquid- and Solid Phase Derivatization

David Hawke1, Heinz Nika2, Edward Nieves2, Ruth Hogue Angeletti2
1Department of Molecular Pathology, MD Anderson Cancer Center, Houston, TX, 2Laboratory for Macromolecular Analysis and Proteomics, Albert Einstein

The isolation of C-terminal fragments and subsequent structural characterization is of general interest because of the high sequence specificity at the protein’s C-terminal region resulting in an increased success rate of protein identification. In addition, the isolation of C-terminal fragments from peptide mixtures may provide for a10-fold reduction sample complexity. Drawbacks associated with current C-terminal peptide enrichment methods include mainly non-specific peptide-isolation and modest detectability of the isolates. To address these shortcomings, we developed a reaction scheme that involves protein carboxyl groups protection by glycinamide followed by trypsination. The digests are then adsorbed onto C18 reversed-phase supports used as venue for subsequent, sequential peptide α-amine acetylation and EDC-mediated carboxylate condensation
using ethylenediamine as nucleophile. The amino group-functionalized N-terminal and internal peptides are then depleted on N-hydroxysuccinimide (NHS)-activated Sepharose leaving the C-terminal peptides since rendered impervious to aminination in the flow-through fractions. Examples of the successful application of the method to low-level quantities of protein digests will be presented.


101 Integrating de novo Sequencing and Database Search for Monoclonal Antibody Sequencing

Baozhen Shan, Lei Xin
Bioinformatics Solutions Inc

A mass spectrometry-based workflow for sequencing the monoclonal antibodies was presented. It combined multi-enzyme digestion, acquisition of high-resolution on both precursor and fragment steps, and integration of database search and de novo sequencing. All MS/MS spectra were first searched against a database, and the top ranking proteins were selected as protein candidates. Unidentified spectra were performed de novo sequencing. The good de novo sequences were selected for homology matching with candidate proteins. The sequence segments of protein candidates were replaced with the de novo sequences to generate new candidates until no better protein candidates could be found.

A sample of human monoclonal antibody was reduced, alkylated, and digested with six different enzymes (AspN, chymotrypsin, GluC, LysC, pepsin and trypsin). Peptide mixtures were analyzed using nana-LC-MS on LTQ-Orbitrap at high resolution in the Orbitrap. Six raw data (MS/MS spectra) were analyzed with PEAKS. All MS/MS spectra were searched against a public antibody database. The proteins (IGHG1, HV102, IGKC, KV304, etc.) were selected as the candidate for further analysis. At 0.5% of false discovery rate (FDR), the majority of the constant region sequences of heavy chain (Ig gamma) and light chain (Ig kappa) can be found in current database with the coverage of 99% and 100% respectively. However, the variable domains containing complementarity determining regions (CDRs) are not complete in the database, with the coverage of 40% for the heavy chain and 36% for the light chain. By iteratively finding homology matches, full sequences of the antibodies including both constant regions and variable regions were determined with 0.5% of FDR. For example, the sequence of the variable region of the heavy chain was determined by the homology match of KFKSKATLTVDKSASTAYME from de novo sequence to KFKSKATLTVDKSASTAYME from the database.

A workflow enables an alternative approach for sequencing monoclonal antibodies.

102 Effects of Space Flight on the Expression of Liver Proteins in the Mouse

Lawrence Sandberg1, Dalia S. Gridley2, Michael J. Pecaut3,4, Lora M. Green2, E. Clifford Herrmann3, Brandon Bianski2, James M. Slater2, Louis S. Stodieck4, Virginia L. Ferguson5, and Lawrence B. Sandberg1,4

1Loma Linda University, Loma Linda, CA, 2Departments of Radiation Medicine, 3Basic Sciences, 4Rheumatology, Pathology and Human Anatomy, Loma Linda University and Medical Center, Loma Linda, CA, 5Department of Aerospace Engineering, Bioserve Space Technologies, University of Colorado, Boulder, CO

Raw data derived from mass spectrometric (MS) analyses of formalin-fixed paraffin-embedded (FFPE) tissue sections of the essential metabolic organ, liver, allocated by the provider (Amgen) from mice subjected to 13 days of microgravity on NASA Flight STS-118 were analyzed by two different search engines using shotgun proteomics. With the eight statistically significant readouts in hand, Ingenuity Pathway Analysis (IPA) was employed to visualize probable biologic pathway relationships among proteins that might be associated with alterations in liver biochemistry due to space flight. Most noteworthy was the finding of up-regulation of the first urea cycle enzyme carbamoyl-phosphate synthetase, consistent with increased amino acid catabolism resulting from gravitational changes and/or other stress associated with missions in space. Down-regulation of fructose-bisphosphate aldolase B, regucalcin, ribonuclease UK114, alpha enolase, glycine N-methyltransferase and S-adenosyl methionine synthetase isoform type-1 was observed. 60 kDa heat shock protein was elevated.

This protein data set (67 proteins) represents the potential of utilizing FFPE tissue for conducting studies of this type and represents a beginning understanding of what takes place in the mammalian liver with weightlessness followed by stress of landing. Eight of the 67 proteins show statistical differences between FLT and GRD. Six were mapped by IPA. Some of these are linked to detoxification pathways within the liver (carbamoyl-phosphate synthetase, glycine N-methyltransferase, S-adenosylmethionine synthetase) and some to carbohydrate metabolism (fructose-bisphosphate aldolase B, alpha-enolase). 60 kDa heat shock protein was up-regulated, probably because of its relation to stress. Regucalcin was highly down-regulated possibly limiting osteoporosis which is a major problem with space flight. Ribonuclease UK114, also known as heat-responsive protein 12 was down-regulated possibly due to the stress of space flight as well.

103 Addressing Robustness and Reproducibility in Nanospray LC-MS for Peptide Analysis

Gary Valaskovic, Helena Svobodova, Amanda Berg

New Objective, Inc.

Nanospray is an essential tool in high-sensitivity mass spectrometry, but limited robustness and reproducibility have historically challenged the adoption of nanospray in quantitative applications. Recent trends toward MS-based biomarker quantitation have placed strict requirements on the analytical performance of nanobore LC-MS. Nanospray MS and nanobore LC-MS both rely heavily on nanospray source hardware for successful experiments. Nanospray source hardware has matured over the past ten years from simple homemade devices to sophisticated application specific instrumentation featuring stage automation, thermal control and high-resolution imaging. Many of these enhanced features provide robustness (automated tip rinsing, automated emitter change), throughput (multi-channel workflows), ease of use (chip systems) or experimental flexibility. While the market offers nanospray hardware solutions for specific applications, there are limited nanospray source solutions which deliver a broad range of capabilities on a single hardware platform. Here we present a nanospray source solution which delivers enhanced features of stage automation, thermal control and high-resolution imaging, while preserving high-performance and flexibility in experimental design on a single platform. Ease of use has been realized through the incorporation of an integrated nanobore LC-MS consumable onto the existing hardware platform, on which over 400 replicate injections were collected with no loss of chromatographic performance. Data collected on a single nanospray source for flow rates ranging from 20 nl/min. to 10 µl/min will be presented, highlighting wide experimental flexibility. Robustness and reproducibility were demonstrated using automated tip rinsing. Monitoring analyte signal in spiked plasma for flow injection data resulted in reproducible analyte response with a 4.6% RSD for data collected with tip rinsing versus 44% RSD for data collected with no rinsing.
105 Proteomic Methods for Orthopedic Biopharmaceuticals

Argus Sun, Samuel S Murray
UCLA Bioengineering, Research Service-Veteran’s Health Administration

Bone Morphogenetic Proteins are growth factors used in orthopedic procedures to accelerate bone growth at fracture and surgical sites. In particular, BMP-2 has been studied significantly because of its efficacy in mediating these effects. Binding peptides that modulate BMP-2 release when incorporated into a resorbable implanted system is one goal of research efforts. One designed peptide, BMP-binding peptide (BBP), has shown some promise in this application. Structural causes of its bioactivity can be explored with Proteomic methods, such as vinyl-pyridine protein conjugation and Matrix Assisted Laser Desorption Ionization MALDI-MS. Using these methods the existence of an internal disulfide bond is confirmed. The vinylpyridine blocks the formation of disulfide bonds causing a peak shift in the mass spectrum. This cysteine linkage constrains the peptide into a cyclic shape contributing to growth factor binding. Compared to homologues to the cystatin-like protein from which BBP’s sequence is derived, the cyclic conformation is seen frequently and gives an homologous to the cystatin-like protein from which BBP’s sequence is derived, the cyclic shape contributing to growth factor binding.

106 Comparative Proteomic Study of Mouse Liver Exposed to Differing Gravitational Environments

Michelle Salemi1, Brett S. Phinney1, Darren M. Weber1, Charles A. Fuller2
1UC Davis Proteomics Core, UC Davis Genome Center, 2Department of Neurobiology, Physiology and Behavior, UC Davis

It has been shown that long term exposure to altered gravitational environments leads to altered intermediary metabolism with a concomitant reduction in body adiposity. This effect on liver protein expression has been poorly examined. Using gel/c MS/MS on extracted liver proteins we will compare protein profiles from mice flown in space compared to control mice maintained at earth’s gravity on the ground. Livers were obtained from mice that were exposed for 90 days to three different living conditions; zero gravity on the international space station; housing similar to spaceflight in a ground laboratory and mice housed in standard vivarium cages in a ground laboratory. Liver proteins were extracted, run on a 1-D gel, and trypsin digested. Extracts were analyzed by LC MS/MS. Proteins were searched and validated with X! tandem and Scaffold Software. Spectral counts were compared using ANOVA to determine significant differences in expression. Roughly 4,000 proteins with a 7.7% Protein and 0.1% peptide false discovery rate were identified, from nine gel lanes. On average more proteins were identified from the liver samples of zero gravity mice (3439 average) than were identified in the lab (3241 average) and vivarium conditions (3157 average). Whether this difference represents an upick in expression due to altered gravitational conditions, or sample preparation and LC MS/MS variability is yet to be determined. ANOVA showed approximately 500 proteins with a p value of <0.05. Many proteins were only found in the space treated mouse liver including Gly domain-containing linker protein and E3 ubiquitin-protein ligase with p values of (0.00000085) and (0.0000095) respectively. Proteins that show statistically significant differences between conditions will be monitored by Single Ion Monitoring targeted mass spectrometry for verification.

107 An Acetone-based Peptide Labeling and Mass Spectrometry Phosphoproteomics Workflow Enables Identification of Biomolecular Targets Relevant to a Fibroblast Growth Factor Induced Post-ischemic Cardiac Recovery

Aruna Wijeratne1, Janet Manning2, Jo El Schultz2, Kenneth D. Greis1
1Department of Cancer Biology, University of Cincinnati, OH, 2Department of Pharmacology & Cell Biophysics, University of Cincinnati, OH

It has been shown that long term exposure to altered gravitational environments leads to altered intermediary metabolism with a concomitant reduction in body adiposity. This effect on liver protein expression has been poorly examined. Using gel/c MS/MS on extracted liver proteins we will compare protein profiles from mice flown in space compared to control mice maintained at earth’s gravity on the ground. Livers were obtained from mice that were exposed for 90 days to three different living conditions; zero gravity on the international space station; housing similar to spaceflight in a ground laboratory and mice housed in standard vivarium cages in a ground laboratory. Liver proteins were extracted, run on a 1-D gel, and trypsin digested. Extracts were analyzed by LC MS/MS. Proteins were searched and validated with X! tandem and Scaffold Software. Spectral counts were compared using ANOVA to determine significant differences in expression. Roughly 4,000 proteins with a 7.7% Protein and 0.1% peptide false discovery rate were identified, from nine gel lanes. On average more proteins were identified from the liver samples of zero gravity mice (3439 average) than were identified in the lab (3241 average) and vivarium conditions (3157 average). Whether this difference represents an upick in expression due to altered gravitational conditions, or sample preparation and LC MS/MS variability is yet to be determined. ANOVA showed approximately 500 proteins with a p value of <0.05. Many proteins were only found in the space treated mouse liver including Gly domain-containing linker protein and E3 ubiquitin-protein ligase with p values of (0.00000085) and (0.0000095) respectively. Proteins that show statistically significant differences between conditions will be monitored by Single Ion Monitoring targeted mass spectrometry for verification.
Idiopathic nephrotic syndrome (INS) is the most prevalent glomerular disease in children. In spite of progress, its pathogenesis is still unknown and the therapy options are confined to gross immune modulation. A variety of methods for diagnostic and treatment purposes are available for the patients; however, the lack of understanding regarding the pathogenic mechanisms underlying INS can lead to poor therapeutic response and adverse side-effects. Here, we describe quantitative multi-omics approach to reveal new molecular factors involved in pathogenesis of INS with potential diagnostic and therapeutic significance.

Urine samples were collected from 10 children diagnosed with INS receiving no therapy and 10 healthy children. All samples were purified using spin filters followed by affinuty depletion of albumin. The purified proteins were recovered and digested with trypsin overnight. Label-free protein expression data were acquired with Synapt G2 using an ion mobility data independent approach, where the collision energy was switched between low and elevated energy state during alternate scans and associate precursor and product ions by means of retention and drift time alignment. The acquired data were processed and searched against a human database that was amended to account for N-terminal processed peptides. Normalized label-free quantitation results were generated using TransOmsics software. In a similar fashion the diluted neat urine samples were analysed using a small molecule profiling approach. The resulting data was also analyzed using TransOmsics. Interpretation of the data has shown a significant number of proteins to be over-expressed in the urine from INS patients, which includes a high percentage of glycosylated proteins. Metabolites of interest showing statistically significant changes include homocysteine, glutamate and uridine. Pathway analyses of the complimentary datasets strongly suggests correlation with the neuronal system disorders network, specifically acute fatigue.
assess changes in peptide abundance between mouse liver, brain, and embryo to reveal unique signatures of tissue specific signaling. PTMScan Direct is broadly applicable to any experimental system in which quantitative profiling of specific, critical signaling proteins is desired.

111 Quantification of the H7N7 Influenza Hemagglutinin and Neuraminidase by Isotope-Dilution Mass Spectrometry

Wanda Santana, T. Williams, J. Pirkle, J. Barr

Centers for Disease Control and Prevention

Influenza vaccination is the primary method for preventing influenza and its severe complications. Government agencies dedicate much time and resources to surveillance of circulating influenza strains and try to predict which strains would be most likely to be transmitted among humans in order to develop a safe and effective influenza vaccine each year. Influenza viral strains undergo rapid mutations and may experience an abrupt change in the surface proteins hemagglutinin (HA) and neuraminidase (NA), which can result in a virus that is radically different from those previously circulating in human populations. This is the reason that seasonal vaccines must be annually updated. The mutations and their effects on the virus and its subsequent transmissibility to or between humans cannot be adequately foreseen. Thus, the strategy of predicting which strain of influenza may cause the next worldwide pandemic so that vaccines can be manufactured and stocked prior to the emergence of an influenza pandemic may not be successful. It is instead, imperative that the process of producing a vaccine that will protect the public in the face of an unpredicted virus strain be streamlined and as swift as possible. We have developed an isotope dilution mass spectrometry method to quantify HA and NA in H7N7 influenza samples. The IDMS method involves enzymatic digestion of viral proteins and the specific detection of target peptides. Four HA peptides and two NA peptides of the H7N7 influenza strain were used in the analysis to ensure complete digestion of the protein, verify accuracy of the measurement, and provide flexibility in the case of amino acid changes among H7N7 strains. The IDMS method is an accurate, precise, sensitive, and selective method to quantify the amount of HA and NA antigens in crude allantoic fluid, purified virus samples, and final vaccine presentations.

RECOMBINANT PROTEINS

112 New IMAC Media Enabling Purification of Histidine-tagged Proteins Directly From Eukaryotic Cell Culture Supernatants

Helena Hedlund, Marianne Carlsson, Therese Granér, Mattias Algotsson, Lars C Andersson, Maria Björner, Gunnar Glad, Lena Hörnsten, Pierre Le Greves, Helena Lindgren, Staffan Lindqvist and Katarina Öberg

GE Healthcare, Uppsala, Sweden

Immobilized metal affinity chromatography (IMAC) purification of secreted histidine-tagged proteins in eukaryotic cell culture supernatants is often problematic. Incompatibility with the cell culture media appears as stripping of the immobilized metal ions required for binding of target proteins. The stripping effect is aggravated by low target protein concentrations and hence large sample volumes.

In this study, purifications of histidine-tagged proteins were performed using two novel Sepharose™ based IMAC media; Ni Sepharose excel and magnetic His Mag Sepharose excel. Both media have a new type of chelating ligand with exceptionally strong binding of nickel ions. Data showing successful purification of histidine-tagged proteins from CHO cell culture supernatants will be presented. In comparison, conventional IMAC medium showed no recovery of target protein.

Furthermore, the characteristics of new media enabled purification of target protein from insect cell culture supernatants. The purification was easily scaled up from 20 μl His Mag Sepharose excel beads to 1 mL pre-packed columns.

113 Purification of a Miniature Spidroin Expressed in E. coli

John Flensburg1, Gabriella Risberg2, Anton L Lindqvist1, Sara Nystedt1, Ingrid Schenning3 and My Hedhammar2

1 GE Healthcare Bio-Sciences AB, SE-751 84 Uppsala, Sweden, 2 Spiber Technologies AB, Doppingväg12, SE-756 51 Uppsala, Sweden

Spider dragline silk, used by spiders as safety line and web frame, is exceptionally strong, elastic, and biocompatible. These features make it attractive for a large number of applications. Spider silk proteins, spidroins, are highly modular, with long repetitive sequences rich in alanine and glycine flanked by nonrepetitive amino- and carboxy-terminal domains of about 100 amino acids.

The objective of the study was to purify a recombinant miniature spidroin, RepCT, derived from the nursery-web spider Euprosthenops australis. To facilitate chromatographic purification the miniature spidroin was fused to a solubility protein and affinity tagged resulting in HisSolRepCT. After induced protein expression in E. coli, harvesting using centrifugation, cell lysis, and clarification, the clarified lysate was immediately loaded onto an immobilized metal ion affinity chromatography (IMAC) column, captured and eluted using step gradients with increasing concentrations of imidazole. Following proteolytic tag removal and desalting, the miniature spidroin, RepCT, was subjected to a polishing step consisting of anion exchange chromatography.

This purification strategy resulted in a purity of RepCT exceeding 95% and a recovery of 60%. Biological activity of RepCT was confirmed by the induced ability to assemble into macroscopic fibers. However, RepCT in aqueous solutions has a tendency to form amorphous aggregates before fiber formation. Herein, several low molecular weight compounds were investigated for their ability to delay this tendency.

114 Investigation of Erythropoietin Conformation Using HPLC-Chip Based H/DX MS

Gregory O. Staples1, Craig D. Wenger1, Julie Cichelli2, Yunan Miao3, Terry D. Lee3, Debbie Ritchey1, Arpad Horvath1, Hongfeng Yin1, Kevin Killeen1

1 Agilent Laboratories, 2 Agilent Technologies, 3 City of Hope

Erythropoietin (EPO) is a cytokine hormone that regulates blood cell production. EPO is used for clinical treatment of anemia and renal disease, but may be best known as a performance enhancing drug due to its activity as an erythropoiesis stimulating agent. EPO is a small protein that is heavily glycosylated (40% by mass), and displays large N-linked glycans characterized by high sialic acid content and poly-lactosamine repeats. Thus, it is an ideal model for understanding the effect of glycan modifications on protein dynamics and conformation. We have performed H/DX MS experiments on two recombinant EPOs from CHO cells using a custom microfluidic chip MS system. The analytical platform permits low flow (300 nanoliter per minute) H/DX MS on sample quantities in the 200 femtomole per timepoint
range, a sensitivity which is ~100X higher than that provided by a typical H/DX MS system. Analysis of the EPO samples revealed EX1 type kinetics in two of the four alpha helices of the molecule. These data are consistent with the fact that these two alpha helices are not involved in either of the two disulfide bridges which stabilize EPO and are required for biological activity. The data thus indicate that these relatively unthethered secondary structures take part in rare but extensive unfolding events under native conditions. Comparison of the two recombinant EPO samples revealed similar deuterium uptake profiles at the timepoints tested, likely a reflection of their mutual origin from CHO cells. We obtained high sequence coverage from our H/DX MS experiments despite the high glycan content of the molecule. We conclude that H/DX MS is an ideal platform for comparing extensively post-translationally modified recombinant proteins and that H/DX MS experiments can be optimized for situations where sample quantity is limited.

**RESEARCH GROUPS**

**115 The ABRF-Next Generation Sequencing Study: A Five-Platform, Cross-site, Cross-Protocol Examination of RNA Sequencing**

Christopher Mason, Scott Tighe, Don Baldwin, and George Grills

Weill Cornell Medical College, Vermont Cancer Center, Pathonomics LLC, and Cornell University

RNA sequencing is a rich assay for delineating the transcriptome but few RNA-Seq standard data sets exist to help quantification of gene or splice form expression. Moreover, each next-generation sequencing (NGS) platform has unique aspects of library synthesis, sequencing, alignment, and data processing. Little is known about cross-site reproducibility, technical variance and interoperability of NGS platforms for RNA-Seq.

The goals of the ABRF-NGS study are to evaluate the performance of NGS platforms and to identify optimal methods and best practices. The study includes five ABRF Research Groups and over 20 core facility laboratories. To address RNA-Seq issues, we performed sequencing on five NGS platforms at multiple sites using two standardized RNA samples with synthetic RNA spike-ins. Platforms tested included illumina HiSeq 2000/2500, Roche 454 GS FLX, Life Technology Ion PGM and Proton, and PacBio. We evaluated a wide range of variables, including varying input amount (1-1000 ng), alternate library preparation methods, specific size fractionation (1, 2, and 3 kb), and performance on degraded RNA (using heat, sonication, and RNase A). We used a set of 18,250 rt-PCR reactions as an orthogonal tool to gauge the linear and dynamic range of the RNA-Seq results.

Our results show that unique transcripts and isoforms are revealed by each method and NGS platform. We found that the majority of the human transcriptome can be found with each method and platform. We also discovered thousands of transcriptionally active regions (TARs) beyond existing gene annotations, which demonstrate that conservative annotation sets are inappropriate for analysis, versus larger annotation sets. Moreover, while we see high correlation of RNA-Seq within sites, we observed that “site effect” is the largest variance factor outside of biological sources. Additionally, we observed that the “bioinformatics noise” of aligners and annotations contributes substantial variance, underscoring the need for data provenance for long-term studies.

**116 iPRG-2013: Proteome Informatics Research Group Study: Using RNA-Seq Data to Refine Proteomic Data Analysis**


1University of California San Francisco, 2University of California San Diego, La Jolla, CA; 3Vanderbilt University Medical Center, Nashville, TN; 4Matrix Science Ltd., London, UK; 5Institute for Systems Biology, Seattle, WA; 6Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia; 7University of Science and Technology, Hong Kong, China; 8New York University School of Medicine, New York, NY; 9Chinese Academy of Sciences, Beijing, China; 10Purdue University, West Lafayette, IN; 11University of Texas Health Science Center at San Antonio, San Antonio, TX

The Proteome Informatics Research Group (iPRG) this year performed a study to evaluate the benefits of using databases derived from RNA-Seq data for peptide identification. The proteomic dataset provided consisted of high mass accuracy tandem mass spectra acquired when analyzing human peripheral blood mononuclear cells. A variety of different types of sequence databases were supplied. These included a standard protein sequence database; a database containing only sequences of proteins expressed in the sample based on RNA-Seq data; a database that included sequence and splice variants; a database of sequences that could not be reconciled to known expressed gene sequences.

Participants were asked to report spectral identifications in the form of an Excel spreadsheet, highlighting those identifications that were only identified using one of the RNA-Seq derived specialized sequence databases. Participants were also required to complete a web-based questionnaire summarizing the tools and methods they used.

Additional peptide identifications were achieved by the use of each of the different RNA-Seq derived databases, although the number of additional identifications was modest. Nevertheless, these new identifications could have potential biological significance, so this type of analysis may still be worthwhile.

**117 Antibody Technology Research Group (ARG) Presentation of Studies and Initiatives**

Robert Carnahan, Brian Agnew, Frances Weis-Garcia

1Vanderbilt University Medical Center, 2Life Technologies, 3Memorial Sloan-Kettering Cancer Center

The Antibody-technologies Research Group (ARG) was formed to foster a community of individuals focused on developing and using antibodies as well as to educate research scientists about the generation, production, manipulation, and optimization of research grade antibodies. In line with these aims, ARG undertook the following activities, which will be discussed at this session. Firstly, a study was conducted to compare of the impact of various antibody-labeling approaches on the ultimate antibody performance. Three distinct antibody-labeling approaches were compared to determine their impact on the functionality of antibodies previously determined to be “difficult to label”. Results suggest a novel click-chemistry based approach may be superior for difficult labeling situations. Secondly, a multi-year study comparing the effectiveness of various immunization strategies was begun. Several approaches to developing a robust humoral immune response were compared head-to-head across multiple sites. Preliminary results have identified adjuvant formulations...
that drive a stronger antibody response than others. Finally, we will review the progress and challenges associated with the electronic repository for antibody related protocols and unpublished lessons learned from working with antibodies. This session will conclude with a discussion on the research group’s efforts for the coming year.

118 Nucleic Acids Research Group (NRG): The Importance of DNA Extraction in Metagenomics: The Gatekeeper to Accurate Results!

Jennifer Holbrook1, R. Carmall2, V. Nadella1, Z. Herbert1, N. Beckloff1, S. Chitturi1, C. Rosato2, A. Perera3, H. Auer3, M. Robinson1, S. Tighe1

1Nemours/A.I. duPont Hospital for Children, 2University of Texas Medical Branch, 3Ohio University, 4Dana-Farber Cancer Institute, 5Case-Western Reserve University, 6University of Albany-SUNY, 7Oregon State University, 8Stowers Institute for Medical Research, 9Institute for Research in Biomedicine, 10University of Zurich, 11University of Vermont

It is well recognized that the field of metagenomics is becoming a critical tool for studying previously unobtainable population dynamics at both an identification of species level and a functional or transcriptional level. Because the power to resolve microbial information is so important for identifying the components in an mixed sample, metagenomics can be used to study nearly any possible environment or system including clinical, environmental, and industrial, to name a few. Clinically, it may be used to determine sub-populations colonizing regions of the body or determining a rare infection to assist in treatment strategies. Environmentally it may be used to identify microbial populations within a soil, water or air sample, or within a bioreactor to characterize a population-based functional process. The possibilities are endless.

However, the accuracy of a metagenomics dataset relies on three important “gatekeepers” including 1) The ability to effectively extract all DNA or RNA from every cell within a sample, 2) The reliability of the methods used for deep or high-throughput sequencing, and 3) The software used to analyze the data.

Since DNA extraction is the first step in the technical process of metagenomics, the Nucleic Acid Research Group (NRG) conducted a study to evaluate extraction methods using a synthetic microbial sample. The synthetic microbial sample was prepared from 10 known bacteria at specific concentrations and ranging in diversity. Samples were extracted in duplicate using various popular kit based methods as well as several homebrew protocols then analyzed by NextGen sequencing on an Illumina HiSeq. Results of the study include determining the percent recovery of those organisms by comparing to the known quantity in the original synthetic mix.

119 Genomics Research Group (GRG): Elucidating the Effects of the Deepwater Horizon Oil Spill on the Atlantic Oyster Using Global Transcriptome Analysis

Natalia G. Reyero1, Nalini Raghavachari2, Kurt Showmaker1, Poching Liu1, Naderedeh Jafari1, Natalie Barker1, Kristine L. Willett2, Jone Corrales3, Heather K. Patterson4, Ruth H. Carmichael5, Don Baldwin5

1Mississippi State University, 2NIH, 3Northwestern University, 4USACE, 5University of Mississippi, 6DISL, 7University of Pennsylvania

Global transcriptome analysis is of growing importance in understanding how altered expression of genetic variants contributes to complex diseases such as cancer, diabetes, and heart disease as well as the effect of environmental pollutants to living organisms. The Genomics Research Group applied next generation sequencing technologies to study the effects of Deep Water Horizon oil spill on the transcriptome of Atlantic oysters. The Deep Water Horizon oil spill resulted in the release of over 200 million gallons of crude oil into the waters of the Gulf of Mexico. Over two million gallons of chemical were used to emulsify and disperse oil plumes posing further risks to the environment in addition to the direct impacts of crude oil. Biota such as the commercially important Atlantic oyster Crassostrea virginica, were inevitably exposed to spill-related contaminants in the Gulf. The potential effects of oiled water and sediments on oysters range from non-detectable to reduced settlement to impaired immune function, acute intoxication, and death due to bioaccumulation of contaminants. Oil may also affect oxygen diffusion through the water column, and in some cases lead to hypoxic conditions that prompt avoidance migration by mobile species. Sedentary organisms such as oysters are even more susceptible to these negative effects of oil contamination. The mechanisms of toxicity of the oil and spill-related compounds are not well understood. In order to understand these mechanisms, we used RNAsequencing of oyster samples from before and after the spill. As the C. virginica genome is not available, we used two different approaches. First, the sequences were mapped to the Pacific oyster genome, recently released. Secondly, a de novo transcriptome assembly was performed. The de novo transcriptome assembly returned a 66-70% alignment rate. Finally, 9,469 transcripts were identified as homologs between the Atlantic and the Pacific oyster.

120 The 2012/2013 PRG Study: Assessing Longitudinal Variability in Routine Peptide LC-MS/MS Analysis

Maureen Burger1, Tracy M Andacht2, Keiryn L Bennett3, Cory Bystrom4, Larry Dangott5, John Leszyk6, Henrik Molina7, Robert L Moritz8, Brett S Phinney9, J. Will Thompson10, Jason Williams11, Felix Elortza12, Matthew Chambers13, David Tabb13

1Proteovations, LLC, 2Centers for Disease Control and Prevention, 3CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, 4Cleveland Heart Lab, 5Texas A&M University, 6University of Massachusetts Medical School, 7The Rockefeller University, 8Institute for Systems biology, 9Proteomics Core UC Davis Genome Center, 10Duke University, 11NIH-NIAMS, 12CIC bioGUNE, 13Vanderbilt University

The PRG study for 2012-2013 was intended to catalog critical parameters of variability influencing LC-MS/MS data quality within laboratories over a nine month period between March and November, 2012. This study was intended to determine intra-laboratory reproducibility and inform participants of key areas of variability in routine peptide mass spectrometry analyses. Aliquots of a dried, digested protein mixture was sent to all participants with the expectation that once per month a new vial will be reconstituted and analyzed using routine LC-MS and data-dependent MS/MS acquisition settings. Of key importance in the design of this study is the lack of a standard operating protocol. The goal was to measure the degree of reproducibility within a lab as it applies to their established HPLC and MS settings and QC measures. A survey was conducted with each sample submission to catalog individual laboratory practices, instrument configurations, acquisition settings, and routine and non-routine maintenance procedures. Over 80 participants submitted at least one data set, and 36 participants completed the study with 8 or more submissions over the 9 month period. Survey data revealed the vast majority of laboratories (>90%) perform routine QC to determine system suitability, but there was considerable variability
in the type and frequency of QC analysis. Collected raw data was searched using identical parameters by the PRG and analyzed for more than 40 MS and MS/MS metrics using the software QuaMeter. The software tool generates metrics that assess multiple properties of LC-MS/MS, from extracted ion chromatogram peak width to total ion current distribution and MS sampling rates. Both identification-dependent and identification-independent metrics can be generated. The variability within these metrics among study was analyzed for each participant and correlative relationships with survey results will be presented.

121 ABRF-sPRG 2013 Study: Development and Characterization of a Proteomics Normalization Standard Consisting of 1000 Stable Isotope Labeled Peptides and a Qualitative Stability Study of Peptides from the ABRF-sPRG 2012 Study

The Proteomics Standards Research Group (sPRG) is reporting the first year progress in a two-year sPRG 2012-2013 study which focuses on the generation of a standard that can be used for interassay, interspecies, and interlaboratory normalization in both label-free and stable isotope label-based quantitative proteomics analysis. The standard has been formulated as two mixtures: 1000 stable isotope 12C/13C/N-labeled synthetic tryptic peptides alone, and peptides mixed with a tryptic digest from a HEK 293 cell lysate. The sequences of the synthetic peptides were derived from approximately 400 proteins and were conserved across proteomes of the most commonly analyzed species: Homo sapiens, Mus musculus and Rattus norvegicus. The selected peptides represent the full range of hydrophobicities and isoelectric points typical to tryptic peptides from complex proteomic samples. The standard was designed to represent proteins of various concentrations, spanning three orders of magnitude. This year we focused our efforts on selection of appropriate protein and peptide candidates, peptide synthesis, quality assessment and LC-MS evaluation by several sPRG member laboratories. The sPRG study design and initial results of a thorough characterization of the standard using a variety of instrumental configurations and bioinformatics approaches will be presented in this talk.

The sPRG is hopeful that the designed formulation will become a valuable resource in various mass spectrometry-based proteomic applications, including quantitative and differential profiling, as well as general benchmarking (e.g. chromatographic retention time). The sPRG plans to start recruiting participants in April 2013, complete the study by the end of the year 2013, and present the results at the ABRF 2014 meeting. The sPRG encourages proteomics laboratories to participate in the study and sign in at www.abrf.org/sprg.

The second half of the session will discuss the qualitative stability study performed using purified synthetic peptides containing a variety of modifications selected from the 2012 sPRG ABRF sample. The stability of the selected synthetic peptides was evaluated by the sPRG using different storage conditions over a three-month period. After storage at either at room temperature, +4°C or -80°C for one week, one month, or three months. Quantitative LC-MS/MS analysis was used to monitor the stability and degradation of the peptides, and to determine the effect of modifications and storage conditions on peptide degradation rates. The data presented have been built on the quantitative study that was presented at both the 2012 ABRF and ASMS conferences. All forms of degraded peptides were separated and identified using nano-LC Q-Exactive hybrid mass spectrometer on a Thermo Scientific Q-Exactive hybrid mass spectrometer. Integrated extracted ion chromatograms were used to measure relative amounts of degradation to identify which pathways are most prevalent during storage.

122 The New ABRF Flow Cytometry Research Group (FCRG)

The Flow Cytometry Research Group (FCRG) is the latest addition to the ABRF RG family. This RG is currently in its first year and has 9 members several of whom are new to the ABRF but have been very active in and come from the flow cytometry core community. The FCRG has submitted a 3 year research plan that will characterize alterations in both gene expression and ultimately cellular function as a result of the stresses imparted by cell sorting. We will use a variety of cell types, lasers, and sorters to identify optimal conditions and eventually Best Practices for minimal cellular system disruptions. Integration of flow cytometry with other core technologies and ABRF RGs will become even more critical as many new technologies will fully take advantage of the sample processing capability of cell sorting allowing higher resolution targeted downstream molecular applications such as single cell gene expression. The new FCRG will seek to foster collaboration, integration and synergy between experts of diverse technologies the very factors that will become increasingly vital to successful research.

123 DNA Sequencing Research Group (DSRG): Evaluation of RNA Amplification Kits at Subnanogram Input Amounts of Total RNA for RNA-Seq

DNA Sequencing Research Group (DSRG): Evaluation of RNA Amplification Kits at Subnanogram Input Amounts of Total RNA for RNA-Seq

Savita Shanker1; Charles Nicot2; Ariel Paulson3

1University of Florida; 2University of Southern California; 3Stowers Institute

ADDITIONAL AUTHORS


1DNA Analysis Inc., 2Penn State, 3University of Michigan, 4Purdue University, 5University of Illinois, 6Boston University, 7Case Western Reserve University, 8University of Missouri, 9UC Davis, 10UMD in New Jersey, 11Indiana University, 12Oklahoma University Health Science Center, 13Northwestern University, 14Dartmouth, 15National Cancer Institute, 16University of Tennessee Health Science Center
Multiple recent publications on RNA-Seq have demonstrated the power of next generation sequencing technologies in whole transcriptome analysis. The vendor specific protocols used for RNA library construction typically require at least 100ng of total RNA. However, under certain conditions such as single cells, stem cells, difficult to isolate cell types, or fractionated cancer cells, only a small amount of material is available. In these cases, effective transcriptome profiling requires amplification of subnanogram amounts of RNA. Several RNA amplification kits are available for amplification prior to library construction and next generation sequencing but these kits have not been comprehensively field evaluated for accuracy and performance of RNA-Seq for picogram amounts of RNA.

This study conducted by the DNA Sequencing Research Group (DSRG) focuses on the evaluation of amplification kits for RNA-Seq. Four commercial amplification kits were chosen: Ovation v2 (NuGEN Technologies), SMARTer (Clontech), Seqaplex (Sigma Aldrich), and Super-AMP (Miltyen Biotech). Starting material was 5ng, 500pg and 50pg of human total reference RNA (Clontech) spiked with Ambion ERCC control mix (Life Technologies) following the manufacturer’s protocol. Each kit was tested at 3 different sites to assess reproducibility. Total RNA and ERCC RNA spike-in control mixes from the same lots were sent to 12 ABRF lab sites for amplification and cDNA generation. Ideally, this would have resulted in 36 different amplified samples, 3 from each input RNA. Libraries were constructed at one site from the amplified cDNAs using the TruSeq RNA library preparation kit on the Tecan Freedom EVO Liquid Handling Robot. As an unamplified control, ribosomal depletion and PolyA selection were performed separately using 5ng, 100ng and 1ug of total RNA prior to library construction. All libraries were pooled and sequenced using the Illumina HiSeq platform. An overview of the study and the results will be presented.

124 Protein Sequencing Research Group (PSRG): Results of the PSRG 2013 Study Year 2: Terminal Sequencing of Standard Proteins in a Mixture

Robert English1, M. Garfield2, P. Jalili3, V. Katta4, S. McGrath5, H. Remmer6, J. Simpson7, D. Suckau8, J. Walters9
1University of Texas Medical Branch, Galveston, TX,
2Research Technology Branch NIAID/NIH Rockville, MD,
3Sigma-Aldrich, St. Louis, MO, United States,
4Genentech Inc., San Francisco, CA,
5FDA Center for Food Safety and Applied Nutrition, College Park, MD,
6University of Michigan, Ann Arbor, MI,
7United States Pharmacopeia, Rockville, MD,
8Bruker Daltonics, Bremen, Germany

Establishing the N- and C-terminal sequences of intact proteins plays a critical role in biochemistry and drug development. Terminal sequence analysis is necessary for quality control of protein biologics, for determining sites of signal peptide cleavage events, for characterizing monoclonal antibodies, and for elucidating sequences of genes from uncommon species. N-terminal sequencing is in the midst of a technology transition from classical Edman sequencing to mass spectrometry (MS) based terminal sequencing, with C-terminal sequencing largely accessible only through top-down MS. Protein homogeneity (absence of interfering protein, non-protein contaminants and buffer components) is critical to the success of analysis. While it has been straightforward to isolate the protein of interest out of protein mixtures in preparation for Edman sequencing, core facilities now need to adopt novel sample preparation techniques to isolate proteins in high purity and make them amenable for terminal sequencing by MS. There is a lack of easy-to-use, field proven methods for sample clean-up.

In order to address the upcoming change in technology platforms, the PSRG is conducting a two-year study with the goal of sample preparation and terminal sequencing of a protein mixture. The 2013 study is the second phase towards this goal and entails terminal sequencing and identification of fusion proteins, which were provided in mixture. Participants used Edman and/or MS techniques, along with bioinformatics tools, to derive the termini of the sample proteins. Study participants were directed to a website to anonymously upload sequences and supporting data.

Our analysis focused on comparison of results obtained by different technology platforms. The results obtained by Edman sequencing and MS as well as information on instrumentation and protocols will be presented. This data and information will be provided to the community in conjunction with year one of the study, which entailed the same proteins supplied as isolated, homogeneous proteins.

125 Glycoprotein Research Group (gPRG) 2013 Interlaboratory Study: Quantitative N-glycan Profiling of Prostate Specific Antigen

Nancy Leymarie1, Morten Thaysen-Andersen2, Detlev Suckau3, Jonscher Karen4, Daniel Kolarich5, Ron Orlando6, Mark Mccomb7, Joseph Zai8
1Center for Biomedical Mass Spectrometry, Boston University School of Medicine, Boston, MA,
2Department, Dept. of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, Australia,
3Department of MALDI Applications and Proteomics R&D, Bruker Daltonik, Bremen, Germany,
4Department of Anesthesiology University of Colorado, Denver Aurora, CO,
5Department of Biomolecular Systems, Max Plank Institute of Colloids and Interfaces, Berlin, Germany,
6Complex Carbohydrates Research Center, University of Georgia, Athens, GA

The fact that most proteins are glycosylated underlies the key roles played by glycosylation during evolution. The functions of glycoproteins are tremendously broad and include cell attachment recognition, homeostasis, transport of molecules, and enzymatic and immunology recognition domains. One of the principal axes of glycoprotein research is to understand better the correlation between glycan structure and function. Mass spectrometry has emerged as a powerful analytical technique in the field of glycoprotein characterization. Its sensitivity, high dynamic range, and mass accuracy provide both quantitative and sequence/structural information.

The goal of the gPRG 2013 study was to determine the ability of the international glycoproteomics community to compare N-glycosylation between two different sources of prostate specific antigen (PSA) by mass spectrometry. PSA was selected to conduct this study for several reasons: PSA is a relatively small glycoprotein (MW ~30,000Da) characterized by a single site of N-glycosylation and is of significant biological interest as a biomarker for prostate cancer.

A total of 35 samples were sent to laboratories around the world, of which 25 data sets were returned. Here, we present the study results, and a global overview of the approaches and methodologies used for differential glycoprotein characterization. The results highlight the challenges faced by researchers in this glycoprotein characterization. We will present a consensus of the interlaboratory data collection built using statistical treatment so as to compare qualitatively and quantitatively the data submitted. In this presentation we will answer the following questions:

- Which sample preparation, separation and analysis methods produced the most consistent results?
- What is the consensus for the glycosylation patterns of the two sources of PSA?
- For which glycoforms from the two PSA sources are abundances significantly different?
Improvements to RiboMinus™
Eukaryote rRNA Depletion Probe Design and Functionality to Enable a Faster and More Complete Workflow

Charmaine Hnahon, Luming Qu, Natalie Supunpong Hernandez, Laura Chapman, Chris Burnett, Jian Gu, Kelli Bramlett, Jeff Schageman, Joel Brockman
Life Technologies

Cellular RNA is composed mainly of cytoplasmic and mitochondrial ribosomal RNA (rRNA). Since rRNAs are not usually the target of whole transcriptome RNA-Seq studies and can potentially take up a majority of valuable sequencing reads, fractionation of the total RNA to obtain rRNA depleted RNA is a necessary first step. The current RiboMinus™ design has had limited success with partially degraded total RNA, due to an abundance of fragmented rRNA contamination in compromised RNA that the current design of probes do not address. We have made improvements to the design and functionality of the RiboMinus™ Eukaryote Kit for RNA-Seq and have expanded its utility to include rRNA depletion from partially degraded total RNA samples. Improvements include: 1. increased the number of rRNA probes; 2. expanded the probe design to include mitochondrial as well as cytoplasmic rRNA; 3. optimized the streptavidin-biotin hybridization time to bring the overall workflow time to about one hour; 4. developed a bead based concentration step enabling scalability of the protocol to multiple samples. These improvements enable the new RiboMinus™ Eukaryote System v2 to be a competitive option for rRNA depletion upstream of Ion Total RNA-Seq Kit v2 library preparation while preserving the whole transcriptome population including RNA transcripts less than 200 nt in length. This workflow enables discovery of new transcripts and accurate gene expression profiling of biologically relevant RNA species beyond those of polyadenylated mRNA.

SOFTWARE

Copy Number Variation Detection through the Utilization of Unique Algorithms in the CNV Tool in NextGENe Software®

John McGuigan, N. Shouyang, C.S. Liu
SoftGenetics, LLC

One of the most important steps in accurate determination of copy number variants using a case-control comparison is normalizing the coverage level between the different projects. NextGENe’s CNV tool uses unique methods for this global normalization and also for log, ratio calculation. The global normalization method is based on the median allele coverage of selected heterozygous positions. After normalization, log, ratios are calculated by selecting one position in each specified region (based on annotation or amplicon location). The positions are selected if they are adequately representative of the region, and ideally if they are heterozygous, so that allele frequency changes can be noted. In this analysis, the tool is used to analyze data from multiple targeted sequencing technologies. A known variant (deletion in the KCNH2 gene) was detected using HaloPlex™ Target Enrichment System Panels. Large chromosomal abnormalities (such as trisomy 21) were detected using Ion AmpliSeq™ Panels.

A Practical Evaluation of Next Generation Sequencing & Molecular Cloning Software

Christian Olsen, Peter Meintjes, Kashef Qaadri
Biomatters

Laboratories using Next Generation Sequencing (NGS) technologies and/or high-throughput molecular cloning experiments can spend a significant amount of their research budget on data analysis and data management. The decision to develop in-house software, to rely on combinations of free software packages, or to purchase commercial software can significantly affect productivity and ROI. In this talk, we will describe a practical software evaluation process that was developed to assist core facility managers and principal investigators in determining the best tools for DNA/RNA/protein sequence analysis and molecular cloning. Eleven software packages were evaluated using six criteria: interface design, data management, data analysis, feature availability, extensibility, and support. This evaluation recommends software packages that excel within each of the individual criteria and the overall best software package for sequence analysis & molecular cloning.
Differential Interference Contrast (DIC) and phase contrast light microscopy are the two most popular methods of obtaining contrast when acquiring images of unstained cells and tissues. DIC images lack the halo effect found in phase contrast images and can provide optically sectioned images especially for high-resolution (high spatial frequency) cellular features. A DIC image is essentially the derivative of the optical path length, $nt$, defined as the index of refraction times the specimen thickness. This derivative is obtained by splitting polarized light into two closely spaced beams with a Wollaston prism, allowing both beams to interact with the specimen and recombining the beams with a second Wollaston prism and analyzer. The shears are the distance between the beams which are oriented along the shear axis of the DIC optics (usually 45° from horizontal). DIC imaging produces positive and negative peaks at the edges of cell structures while unchanging structure results in a gray background intensity similar to that found outside the cell. Since much of the internal structure of a cell has the same intensity as the image background, standard image segmentation methods are not very effective. Therefore some preprocessing step should be taken to reverse the effects of differentiation. Many methods such as low-pass filtering, line-integration along the shear direction, Hilbert Transformation, Weiner filtering and deconvolution have been suggested to prepare DIC images for further image processing or display. The Hilbert Transform has been shown to reverse the effects of differentiation without introducing high frequency noise into the image. A combination of Hilbert Transformation followed by deconvolution with a bright-field point spread function (psf) produces images that are highly amenable to further processing and display. An ImageJ macro routine has been written to automatically process a series of images using this algorithm so they are ready for further analysis.

**Methodology**

We have demonstrated a novel somatic mutation enrichment methodology demonstrating multiplexed detection of tumor mutations in plasma with sensitivity as low as 0.01% compared to normal DNA.

This highly sensitive detection of low abundance mutations is achieved using electrophoretic separation and enrichment of DNA fragments containing point mutations over their wild-type counterparts. Commercialized as the OnTarget platform by Boreal Genomics, the system enriches nucleic acid samples for specific targets prior to amplification and detection, enabling the use of next-generation sequencing (NGS) or other detection assays for plasma or FFPE-based mutation detection and profiling.

We present data demonstrating highly sensitive and multiplexed detection of panels of up to 100 mutations in plasma samples, improving the sensitivity of NGS assays to below 0.01% mutant content. We also report on concordance studies comparing low tumor content FFPE tissue and matched plasma in human samples demonstrating that OnTarget represents a robust, highly sensitive and multiplexed platform for non-invasive tumor monitoring.