FLYCAT; Towards a Drosophila Proteome Database
E. Brunner1, S. Mohanty1, F. Potthast2, C. Ahrens2, R. Schlapbach2, R. Aebersold3, and E. Hafen4

1University of Zürich, Zürich, Switzerland; 2Functional Genomic Center, University of Zürich, Zürich, Switzerland; 3Institute for Molecular Systems Biology ETH Zürich, Zürich, Switzerland; 4Zoological Institute, University of Zürich, Zürich, Germany

The field of comparative genomics has revealed a striking conservation of gene sequences from model organisms such as flies or worms to man. It underscores the conservation of basic biological processes and stresses the value of genetic and molecular genetic tools available in model organisms to understand biological processes relevant to human health and disease. Since almost all gene functions are executed by proteins, a thorough understanding of biological processes requires a detailed knowledge of the proteins encoded by a genome, their regulation, and their interactions. As a first step towards this goal, we strive for a comprehensive description of the proteome encoded by the Drosophila genome. To that aim, we have analyzed a large set of protein extracts from flies from different developmental stages or cell cultures using an automated shotgun proteomic approach. To date, a total of approximately 40,000 unique peptides corresponding to about 7,000 proteins have been identified. This large amount of data is stored in a newly developed database denominated FLYCAT, which will be made publicly available. FLYCAT allows one to combine, or compare multiple projects and makes researchers to validate or share protein data in a simple, quick and reproducible manner. Furthermore, it offers sophisticated tools to annotate a complete proteome using pooled experimental data. FLYCAT serves as the prototype proteome data repository, and similar databases are currently built for other model organisms such as Arabidopsis thaliana and Caenorhabditis elegans.

Protein Abundance Versus Identification by Mass Spectrometry; the Yeast Proteome Case
L. De Godoy, G. De Souza1, R. Shi, J. Olsen, P. Mortensen, and M. Mann

Center for Experimental Bioinformatics, University of Southern Denmark, Odense, Denmark

Mass spectrometry is a powerful technology for the analysis of large numbers of proteins in complex samples, like in proteomics. However, due to various analytical challenges, mainly the large dynamic range involved, so far no proteome has been sequenced completely. Yeast remains one of the most popular model organism because of the ease of manipulating it both genetically and biochemically. Recent studies indicate that at least 80% of the yeast genome is expressed in logarithmically growing cells, corresponding to more than 4000 proteins, which range between 1,000,000 to less than 100 copies per cell. To probe the yeast proteome in depth, cells were grown to mid-log phase in YPD medium and proteins were extracted and analyzed by different methods. Mass spectrometric analysis was performed on a linear ion trap-Fourier Transform mass spectrometer with very high mass accuracy and sequencing speed. For unambiguous identification, only those proteins with at least 2 unique peptides were considered in this study. Following this criteria, thousands of proteins with a dynamic range from the most abundant, such as enolases, to those a million-fold less abundant, such as transcription factors, were identified. We have then used knowledge of the copy number of most yeast proteins to estimate the degree of coverage of the proteome using different state of the art proteomic techniques and to estimate how far we are from complete coverage of the yeast proteome by mass spectrometric experiments.

Differential Proteomic Analysis of the Human Pathogen Trichomonas vaginalis
J. De Jesus1, P. Cuervo2, M. Junqueira3, F. Silva-Filho4, O. Fernandes5, E. Cupolillo6, and G. Domont6

1Instituto Oswaldo Cruz, Rio De Janeiro, Brazil; 2Instituto Oswaldo Cruz, DMT, Rio De Janeiro, Brazil; 3Instituto Oswaldo Cruz, DFF, Rio De Janeiro, Brazil; 4Instituto De Biofísica, UFRJ, Rio De Janeiro, Brazil; 5Instituto De Química, UFRJ, Rio De Janeiro, Brazil; 6Instituto de Química, UFRJ, Rio De Janeiro, Brazil

Introduction: Trichomonas vaginalis is the causative agent of human trichomoniasis, the most frequent non-viral sexual transmitted disease. It emerges as an important cofactor in amplifying HIV transmission and is associated to cervical cancer, atypical pelvic inflammatory disease and infertility. This work describes protocols for reproducible extraction and separation by two dimensional gel electrophoresis of T. vaginalis proteins from FMV1 and FF28JT-Rio strains, respectively, a fresh isolate highly cytotoxic and a long-term cultured parasite with low cytotoxicity to MDCK monolayers.

Methods: Protein extraction used 1 x 10⁶ parasite cells lysed by freeze/thawing in PBS buffer pH 7.4, precipitation with TCA and cold acetone wash. Pellets were dissolved in 8 M urea, 2% CHAPS, 20 mg DT and 0.1% ampholytes. IEF gel profiles were obtained in 4–7 pH gradients, 18 cm gel strips, 700 µg sample, 12% polyacrylamide gels, and Coomassie blue staining. Peptides from in gel tryptic hydrolysis were collected, submitted to MS and MS/MS analysis (ABI 4700 MALDI-TOF-TOF) and searched at NCBI using MASCOT software.

Results and Discussion: Mass spectrometry identified cysteine proteinases and cytoskeleton, glycolytic pathway, heat shock, amino acid metabolism and anti-oxidative stress proteins.

Proteomic Analysis of Embryonic Stem Cell-derived Cardiomyocytes and Neonatal Mouse Cardiomyocytes
S. Hosseinkhani1, R. Zonouzi2, S. Kazemi Ashiani2, H. Baharvand2, M. Hajheydari3, G. Hosseini1

1Tarbiat Modarres University, Tehran, Iran; 2RoyaI Institute, Tehran, Iran; 3Agricultural Biotechnology Research Institute, Karaj, Iran

Embryonic stem (ES) cells are totipotent lines derived from the inner cell mass (ICM) of developing blastocystes. ES cells are able to differentiate through three embryonic germ layer (ectoderm, mesoderm and endoderm) to cardiomyocytes (in vitro). The generation of functional cardiomyocytes from embryonic stem cells has several potential applications including myocardial repair through cell transplantation. In this research we compared in vitro differentiation of ES cells with in vivo cardiogenesis, based on their proteins, by combining two dimensional gel electrophoresis with mass spectrometry, ES cell line (Rolyan B1), derived from C57BL/6 mouse strain, were cultured on mouse embryonic feeder cells. For differentiation of embryonic stem cells into cardiomyocytes, cell suspension was prepared and 800 cells/20 µl medium were cultivated in hanging drops on Petri dish for 2 days. Cells were aggregated and formed embryoid bodies, which collected and further cultivated in suspension. Beating areas were mechanically dissected using a pulled-glass micropipette under contrast phase microscopy (Nikon) and washed in PBS (phosphate buffer saline without Ca and Mg) twice. Cardiomyocytes were enriched with discontinuous percoll gradient and selective adhesion by plating on gelatin-coated culture flask and incubating 1.5 h. The non-adhesive myocytes were centrifuged and washed in PBS. Normal and stem cells were studied by 2D gel electrophoresis. Qualitative and quantitative analysis of protein expression will be reported.
25.5 Identification of Protein Markers Separating Chloroquine-resistant and -sensitive Strains of Plasmodium falciparum by SELDI Profiling

S. Koncarevic, M. Meyer, R. Bogumil, and K. Becker
Interdisciplinary Research Center; University Giessen, Giessen, Germany

Tropical malaria, which is caused by the parasitic protozoon Plasmodium falciparum, represents a growing threat to human health. Chloroquine therapy and prophylaxis have become increasingly ineffective, resulting in enhanced malaria morbidity and mortality. Using SELDI (Surface Enhanced Laser Desorption/Ionization-Mass Spectrometry) technology we aimed at gaining further insight into the resistance mechanism and the mode of action of chloroquine. To analyse drug resistant and sensitive strains we established and optimized methods for studying the proteome of P. falciparum cultured in vitro. Protein profiles were measured and compared by SELDI on different chromatographic ProteinChip Arrays. Several hundred distinct protein peaks in Plasmodium lysates were detected and compared to each other to reveal quantitative protein expression differences in the Plasmodium proteome. Hierarchical clustering analysis allowed for the identification of peaks separating resistant from sensitive strains. In addition, we found proteins exhibiting expression differences due to chloroquine treatment. We selected a set of candidate marker peaks for downstream identification. A protein purification and identification strategy was established to enrich markers. Anionic and cationic exchange chromatography in spin-column format was used as a first enrichment step for candidate marker proteins. In addition, purification of proteins enriched in a distinct ionic exchange fraction was achieved by reverse phase chromatography and gel electrophoresis. SELDI was used to guide the enrichment and purification process. Final protein identification was performed by analysis of tryptic fragments using the SELDI PCS4000 and by MS/MS through the ProteinChip Interface (PCI1000) coupled to a Tandem-MS instrument. First identifications have been obtained and will help answering the question how these differentially expressed proteins are connected to chloroquine resistance.

25.6 Functional Proteomics in Zebrafish

V. Link, A. Shevchenko, and C. Heisenberg
Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

Zebrafish (D. rerio) is a widely used model system for the analysis of vertebrate embryogenesis, organ development, and disease. Here, we supplemented the well-established genetic tools with a proteomic approach. The whole workflow of sample preparation, 2D gel electrophoresis, spot quantification and mass-spectrometry-based identification was established, optimized and used to identify proteins involved in germ layer formation during Zebrafish gastrulation. Gastrulation defines the first process in the development of all multi-cellular organisms where a uniform mass of cells transforms into different tissues (germ layers) with specific properties. Developmental biologists have so far mainly focused on the genetic network that triggers this process. In contrast, our approach aims to identify proteins causing changes in migrational activity and morphology during this transformation. The protein expression profiles of the ectodermal and mesendodermal germ layers were compared by 2D gel electrophoresis in combination with the DIGE system. 37 spots with statistically significant changes between ectoderm and mesendoderm were subjected to mass-spectrometry-based identification. We identified several proteins related to cytoskeletal organization, which may play important roles in the transformation of ectoderm to mesendoderm. We are currently testing the function of these proteins for germ layer formation during gastrulation. In this study, we established the core techniques for proteomics in Zebrafish embryos and have successfully used them to identify proteins with a potential role in germ layer formation during gastrulation.

25.7 TIR2 Positional Candidate Genes Analysis

J. Nganga\(^1\) and F. Iraqi\(^2\)
\(^1\)International Livestock Research Institute (ILRI) and Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi, Kenya; \(^2\)International Livestock Research Institute (ILRI), Nairobi, Kenya

Marked differences between inbred strains of mice in their response to T. congolense infection can be exploited in the analysis of the genetic basis of the infection. Using mice, trypanotolerance QTLs have been mapped to chromosome 1, 5 and 17 (Tir1, 2 and 3 respectively). TIR2 and 3 have now been fine mapped to less than 1cM using F\(_2\) and F\(_3\) advanced intercross lines respectively. Tir1 is homologous to a previously mapped bovine QTL. TLR1 and 6 as TIR2 candidate genes were sequenced across resistant and susceptible mice. The sequences were then aligned to each other. Subsequent analyses of SNP haplotype structure showed significant variation between the susceptible A/J and the resistant mouse C57BL/6J inns. A/J strain had a misense substitution on TLR1 at position 1525 and another on TLR6 at position 1703 from the start codon. The two substitutions were predicted to drive a functional variation on each of the genes. Comparative mapping of the SNPs on the bovine genome can now facilitate marker assisted selection. This polymorphisms can also be used to map genes predisposing humans to sleeping sickness hence be used to design preventive and therapeutic strategies for the control of the disease in livestock and man.
25.8 Proteomic Analysis of Dictyostelium discoideum

U. Roth1, S. Müller2, L. Eichinger3, A. Noegel3, and F. Hanisch3

1Medical Faculty, University of Cologne, Cologne, Germany; 2Center for Molecular Medicine, Medical Faculty, University of Cologne, Cologne, Germany; 3Center for Biochemistry, Medical Faculty, University of Cologne, Cologne, Germany

The slime mould Dictyostelium discoideum is a well known model organism for a variety of cellular and molecular studies, like the dynamic reorganisation of the cyto-skeleton resulting in cell movement or signal transduction pathways of chemotaxis. Additionally, during the last years D. discoideum was discovered as a model system for functional genomic studies, because of the completion of the genome sequencing project (Eichinger et al., Nature in press), the availability of gene expression (micro array) data and the applicability of different tools to over-express or to knock-out genes. Interestingly, at the same time only few data describing the D. discoideum proteome have been published so far. Here we present an in-depth proteomic analysis of D. discoideum cells in their vegetative single celled amoeboid phase using high-resolution 2D-PAGE for protein separation and automated protein identification by peptide mass fingerprint (MALDI-TOF). The analysis set covers protein analysis of total cell lysates, as well as sub-proteomes, like putative nuclear or phospho-proteins. In a second step the protein data gained are further analysed bioinformatically by clustering the proteins into different functional groups. In an outlook the methodology is discussed as a tool to study signal transduction events by differential proteomic approaches comparing the protein expression or the occurrence of posttranslational modifications of stimulated vs. non-stimulated cells.

25.9 Proteomic Profiling of the Protozoan Parasite Leishmania major LV 39 Developmental Stages Utilising 2-D Electrophoresis

E. Sage1, R. Almeida2, R. Wait3, A. Ivans4, J. Blackwell2, and M. Dunn5

1Institute of Psychiatry, London, United Kingdom; 2Cambridge Institute for Medical Research, Cambridge, United Kingdom; 3Kennedy Institute of Rheumatology, London, United Kingdom; 4Wellcome Trust Genome Campus, Cambridge, United Kingdom; 5Conway Institute of Biomolecular and Biomedical Research, Dublin, Ireland

Leishmania is a protozoan parasite responsible for a wide range of human diseases. The leishmaniases affect approximately two million people predominantly in the third world each year and there are currently no effective treatments. Therefore the eradication of leishmaniasis is an important target of the World Health Organisation. Here we describe a high resolution proteomics approach utilising 2-D electrophoresis to investigate changes in the Leishmania life-cycle associated with differentiation of the non-infective (promastigote) to the infective (amastigote) form of the parasite. Optimisation of protein solubilisation and recovery was achieved using a lysis buffer consisting of 5 M urea, 2 M thiourea, 2% CHAPS and 2% SB3–10. Initially L. major proteins were separated on broad range pH 4–7 and 6–11 resulting in the detection of 816 and 550 protein spots resolved respectively on large format 2-D gels when samples were applied by in-gel rehydration or anode cup-loading. More detailed analyses were achieved utilising medium range IPG strips for enhanced protein separation and resolution. Proteins exhibiting marked changes in abundance during transformation of L. major LV39 (i.e. qualitative or quantitative alterations in expression) as assessed by bioinformatic software have been identified by electrospray mass spectrometry (ES-MS/MS). Approximately 100 proteins have been identified so far by ES-MS/MS. The resultant data generated will be utilised to investigate their suitability as targets for future drug design or vaccine development.
25.10
Homology-driven Proteomics in Insects and Plants by Automated LC-MS/MS de Novo Sequencing and MS BLAST Searches

P. Waridel1, V. Surendranath1, H. Thomas1, A. Frank2, P. Pevzner2, and A. Shevchenko 1

1Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany; 2University of California, San Diego, California, USA

The characterization of proteomes of many important model organisms, especially within plant and insect kingdoms, is hampered by the paucity of genome sequences and remarkable phylogenetic diversity of proteins in wild-bred species. Here we report a strategy for automated identification of proteins from organisms with unknown genome by a combination of nanoLC-MS/MS, automated de novo sequencing and Mass Spectrometry-driven BLAST (MS BLAST) sequence similarity searching. Tryptic digests of proteins separated by 1D or 2D electrophoresis were analyzed by a linear trap instrument LTQ (ThermoElectron) coupled to a nanoLC system (Dionex). The entire pool of ca 10,000 MS/MS spectra was filtered by a pattern recognition algorithm to remove spectra of trypsin and keratins peptides as well as non-peptide background. The remaining spectra were interpreted de novo by PepNovo software, which takes about 0.1 sec per spectrum. The resulting redundant, degenerate and partially inaccurate sequence candidates were submitted to the web-accessible MS BLAST tool for protein identification. As protein identification relied on the similarity of peptide sequences (rather than on their identity), sequence polymorphism that commonly occurs in wild-bred species was tolerated. The method was validated by automated analysis of proteins from the alga Dunaliella salina, the bug Triatoma infestans and the moth Cerodirphia speciosa, from which a number of proteins were identified by cross-species matching to known protein homologues from other model organisms.

25.11
Proteomic Characterization of Mouse Liver Golgi/Endosome (G/E) Fractions

E. Winstall, C. Lessard, S. Bourassa, I. Kelly, F. Labrie, G. Poirier, and R. Faure

CRCHUL/CHUQ, Université Laval, Quebec, Canada

Endocytic transport is mediated by sub-cellular membranes which encompass early to late endosomal compartments. Highly purified endosomal fractions have been characterized previously by us and others using electron microscopy, protein markers and ligand-mediated endocytosis. The G/E fraction is characterized in this study in proteomic terms. G/E fractions were isolated from liver of fasted adult male mice and were analyzed by electron microscopy and insulin-mediated endocytosis. Solubilized proteins were separated by 2-D or 1-D gel electrophoresis. Gel bands or spots were excised and proteins were digested in-gel using trypsin. In another set of experiments, proteins were directly digested and the resulting peptides were separated by off-line 2D-LC. Tryptic peptides were analyzed using an ESI-ion trap mass spectrometer and proteins were identified using Mascot or X!Tandem software. A G/E membrane permeabilization step resulted in depletion of abundant soluble cargos, thus markedly facilitating analysis of less abundant proteins. We thus identified more than 300 proteins, most of which being associated with vesicular transport, protein folding and processing as well as signal transduction. In addition, a series of proteins of unknown biological function were identified, thus suggesting a possible role of these proteins in G/E function. Each approach gave a complementary view of the G/E proteome. For instance, we found that cytochrome P450 protein family members, were detected only in the peptide 2D-LC workflow. Taken together, our results show that a permeabilization step permits to obtain a representative proteome and thus reveal interesting new features. The present data also emphasize the importance of using complementary protein or peptide separation methods to achieve better coverage of the proteome.