ESF/EMRC Exploratory Workshop

Microarray meets diagnostics: Chip technology as an innovative technology to study complex and heterogeneous diseases.

Tübingen, April 15-17, 2005

– Scientific Report –

Olaf Riess and Michael Bonin
Department of Medical Genetics
University of Tübingen, Germany
{olaf.riess │michael.bonin}@med.uni-tuebingen.de

Contents

Introduction

Two recent complementary advances, one in knowledge and one in technology, can be considered as major breakthroughs in molecular biology and medicine. As a result of the Human Genome Project, there has been an explosion in the amount of information available about the DNA sequence of the human genome. The challenge currently facing scientists is to find a way to organize and catalogue this vast amount of information in terms of both, functions of the genes and patterns of sequence variations in the population. The second advance facilitates the identification and classification of this DNA sequence information and assignment of functions to certain genes: the DNA microarray technology. Microarrays are expected to promote a variety of application domains that concern in particular the diagnosis of complex and heterogeneous diseases.
Making this vision a reality, however requires substantial and coordinated research in variety of disciplines, including bioinformatics, biobanking, clinics, quality control and technology. The purpose of the ESF EMRC Exploratory workshop "Microarray meets diagnostics: Chip technology as an innovative technology to study complex and heterogeneous diseases" was to bring together European researchers and industry working in the field of microarray application in diagnostics in order to help form a network and/or community and to initiate the development of a consortium that can apply for major European funding.

Executive Summary

Summary of Presentations

The workshop was co-convened by Olaf Riess and Michael Bonin and took place on April 15-17, 2005 at the University of Tübingen, with 20 people from 11 countries attending the workshop. While the majority of the participants were academic researchers, there were also attendees from academic management and representatives from industry.
In the opening session, Olaf Riess from University of Tübingen described the resources of the convening department of Medical Genetics of the University of Tübingen and shortly presented the motivation for performing the workshop. Besides building a European community/network on microarray application in diagnosis the driving motivation was to form a European consortium that could be able to apply for a major European funding in the seventh Framework Programme of the European Commission. Then the participants were asked to introduce themselves by means of a few numbers of slides. The opening session was accomplished by Carole Moquin-Pattey, the head of unit of ESF EMRC, who gave an overview on the activities of the ESF.
The remainder of the workshop consisted of 5 presentation panels. After each panel time for discussion was scheduled. Each talk was allotted a time slot of 20 minutes, already including time for first questions. Each lecturer was given the opportunity to present his background, his ongoing and future work, and to mark the problems and bottlenecks of the scientific field concerned.
The first panel was aiming at presenting and at comparing various technologies approaches and development in Europe by Academia or SME as complement to the methods linked to the Affymetrix technology. On behalf of Sequenome Michaela Downie presented the *MassARRAY*™ system as a genetic analysis platform that can be applied for genotyping, quantitative gene analysis, for example expression profiling and comparative sequencing. Janna Saarela spoke about "Allele-specific primer extension on microarray" and a specific “Array of arrays” layout that allows several SNPs and samples to be analyzed simultaneously. The first panel was concluded by Snaevar Sigurdsson who presented the tag-array minisequencing or single base primer extension as a method which has proven to be particularly well suited for accurate SNP genotyping at medium throughput.
The following discussion tried to assign the different technologies available to the different requirements of applications of the microarray technology in terms of throughput and costs.

The second panel focused on expression profiling for diagnostics. Charles Decraene presented the Translational and Preclinical Research Department of the Institute Curie Paris to which a Microarray facility belongs and which plays a prominent role in transferring new diagnostic and therapeutic technologies to the hospital and provides a genuine link between research and medical applications. Harriet Wikman then outlined research activities related to the identification of gene expression changes caused by altered DNA copy numbers in lung cancer. Karsten Zieger presented a project on molecular signatures in superficial bladder carcinoma, which may predict clinical outcome. Valter Gattei continued with a presentation on microarrays in molecularly and immunophenotypically well-characterized chronic lymphocytic leukaemia cells. Sakari Knuutila reported on gene amplifications and losses in hematologic malignancies by array CGH. Finally, Miroslav Prucha talked about gene expression signatures in sepsis.
The following discussion round emphasised the need of clinically well-classified and well-defined tumors/tissues. Additionally, it was agreed that long prospective studies are needed. More basically, the question was aired under which conditions research data should be translated into daily medicine. This discussion round ended with the issue of handling and protection of patient data relating to samples studied.
The third panel focused on microarray based resequencing for diagnostic purposes. Neeme Tõnisson presented on research of Asper Biotec related to arrayed primer extension based resequencing of p53. Ionnis Ragoussis' talk focussed on the issue of genome analysis using MALDI-TOF applications. Michael Bonin reported on the design of a microarray for sequencing of the whole CFTR gene.
The fourth panel paid attention to Bioinformatics and database management for microarray based diagnostics. Rani Elkon presented his work on integrative analysis of heterogeneous functional genomics data. The last presentation of the day was done by Magnus K. Magnusson whose talk argued for statistical design of the diagnostic expression array using the principles of randomization, replication and blocking.
The third day of the workshop was devoted to "problems, bottlenecks and politics".
James Hadfield's talk emphasised the sample processing and quality control issue. He outlined how to move array-based diagnostics into the hospital lab. The last presentation of the workshop was given by Holm Graessner who sketched the opportunities to form a European consortium on applications of microarray in diagnostics.
The final discussion on ways to continue was initiated by a brief presentation of Olaf Riess who presented his ideas how microarray analyses could replace and improve cytogenetic prenatal diagnostics. However, the following discussion focused more on horizontal issues such as quality control and standardization. It was agreed that the people present will not apply for a project topic on high throughput molecular diagnostics for hereditary diseases which will be published in the fourth call of the 6. FP of the European Commission. The common understanding was that networking measures are needed before considering such a proposal. Therefore, the attendees agreed on applications making use of two specific funding schemes, which offer the opportunity to receive appropriate funding which are highly appropriate to initiate the forming of a European network/community on microarray applications in diagnostic.
Besides the presentation panels and discussion sessions, the workshop offered many opportunities to get to know and to talk to each other personally and to establish contacts for possible future co-operations during coffee breaks, during two lunches and two workshop dinners in restaurants serving local food.

Meeting Place

The workshop was held in the *Fürstenzimmer* of the Tübingen Castle. This venue allowed meeting in a very pleasant but also historical place in terms of history of genetics. Just the beneath the room were the presentations were given Friedrich Miescher discovered the cellular DNA in the former kitchen of the castle in 1869.

Scientific Content of the event

Presentations and discussions at the workshop mainly covered the following five panels.

PANEL I: TECHNOLOGICAL ADVANCEMENT OF THE MICROARRAY BASED DIAGNOSTICS

Michaela Downie: MassARRAY System: A Versatile Genetic Analysis Platform.
Janna Saarela: Allele-specific primer extension microarrays.
Snaevar Sigurdsson: Four-colour fluorescent minisequencing in an "array of arrays" format.

PANEL II: EXPRESSION PROFILING FOR DIAGNOSTICS

Charles Decraene: Functional analysis of human genome and its application to diagnosis, prognosis and treatment of cancer.
Harriet Wikman: Identification of Gene Expression Changes Caused by Altered DNA Copy Numbers in Lung Cancer.
Karsten Zieger: A molecular signature in superficial bladder carcinoma predicts clinical outcome.
Valter Gattei: Microarrays in molecularly and immunophenotypically well-characterized chronic lymphocytic leukaemia cells.
Sakari Knuutila: Gene amplifications and losses in hematologic malignancies by array CGH.
Miroslav Prucha: Gene expression signatures in sepsis - influences of center-associated factors.

PANEL III: MICROARRAY BASED RESEQUENCING FOR DIAGNOSTICS

Neeme Tõnisson: Arrayed primer extension based resequencing of p53.
Ionnis Ragoussis: Genome analysis using MALDI-TOF applications.
Michael Bonin: Microarray sequencing of the whole CFTR gene.

PANEL IV: BIOINFORMATICS AND DATABASE MANAGEMENT FOR MICROARRAY BASED DIAGNOSTICS

Rani Elkon: Integrative analysis of heterogeneous functional genomics data.
Magnus K. Magnusson: Statistical design of the diagnostic expression array using the principles of randomization, replication and blocking.

PROBLEMS, BOTTLENECKS AND POLITICS

James Hadfield: Moving array based diagnostics into the hospital lab: sample processing and QC.

PANEL I: TECHNOLOGICAL ADVANCEMENT OF THE MICROARRAY BASED DIAGNOSTICS

Michaela Downie: *MassARRAY*™ System: A versatile Genetic Analysis platform.

The *MassARRAY*™ system consists of four key components: Biochemistry (MassExtend assay with typical 24 plex), chip technology (SpectroCHIP), mass spectrometry (MALDI-TOF), and data analysis (software MassARRAY™ QGE). It is an open multi-function platform applicable for genotyping, quantitative gene analysis and comparative sequencing.
Examples of applications:

- Genetic impact of susceptibility genes and gene loci for multifactorial traits: identifying disease associated SNP,
- Real competitive PCR enables differential allele expression analysis.

Today, MassARRAY™ provides a solution for every stage of genetic research.

Janna Saarela: Allele-specific primer extension on microarray.

Description of the method:
The allele-specific primer extension on microarray is based on a microarray technology. "Array of arrays" layout allows several SNPs and samples to be analyzed simultaneously. The method was developed in the National Public Health Institute of Finland in supervision of Ann-Christine Syvänen (Pastinen et al. 2000). Arrays are constructed by printing two 5'amino-modified synthetic oligonucleotides (detection primers) for each SNP in duplicates to a chemically activated microscope slide. The reaction principle includes 1) amplifying the SNP containing sequences in a multiplex PCR reaction with SNP-specific primers containing T7 sequence (4-20 SNPs/reaction), 2) *in vitro* transcribing the PCR products to single stranded RNA molecules, 3) Hybridizing the RNAs to a microarray slide containing the detection oligos, 4) allele-specific primer extension of correctly hybridizing detection primer(s) utilizing reverse transcriptase enzyme and fluorescently labeled nucleotides. After washing and scanning, the genotypes can be read using the SNPSnapper software developed by Juha Saharinen in the National Public Health Institute of Finland.
Examples:
This method has been used in several research projects, such as gene identification (Saarela et al. "Protein Kinase C Alpha shows association to multiple sclerosis in a family based study in two independent populations", submitted), and population screening (Komulainen et al. "Allelic variants of Upstream Stimulatory Factor 1 (USF1)-gene represent a risk factor for cardiovascular events in a prospective, population-based study", submitted). A Finnish biotech company, Jurilab, has developed a commercial product utilizing this technology (DrugMEt™ Pharmacogenetic Test; www.jurilab.com).

Snævar Sigurðsson: SNP genotyping using tag-array minisequencing.

Method:
Tag-array Minisequencing or single base primer extension has proven to be particularly well suited for accurate SNP genotyping. The method is based on SNP-specific detection primer binding to the PCR template with it's 3'-end just downstream of the SNP site, and a DNA polymerase to distinguish between the two alleles of a SNP by extending the primers. Dideoxynucleotides are used so that there will only be a single base extension and the nucleotides can be labeled with one, two or four fluorophores depending on the scanning instrument available.
Each detection primer has specific 5'-Tag sequence that is complementary to one of the Tags that are attached to a microarray in an"array of array" format. From the known locations of the Tags on the microarray, the genotype of the SNP can be deduced.

This system has been used for a wide variety of applications including:
- Quantitative analysis of SNPs in pooled DNA samples for validation and determination of allele frequencies of SNPs in different populations.
- A panel of 50 SNPs distributed on different chromosomes was used for selection of SNP markers for monitoring of chimerism after stem cell transplantation.
- Using MDA, less than 1 µg of genomic DNA is sufficient for genotyping 300.000 SNPs.
- More than 150.000 genotypes were produced when genotyping 40 SNPs in the interferon system in candidate genes for SLE. The success rate was 90-95% and the accuracy was > 99%.
- Imbalanced expression of SNP alleles in cDNA can be measured and quantified.

The microarray-based genotyping system is cost-effective and easy to set up in any laboratory with access to microarray spotter and scanner and is suitable for medium throughput genotyping: 10-500 SNPs, 50-5000 samples
For higher throughput our laboratory has two commercial systems: (i) The SNPstream system from Beckman Coulter. The format is 12-plex PCR and 12-plex primer extension in a 384-well microtiter plate format. (ii) The Illumina bead array system. This system is very high

throughput, 1500-plex in a 96-well format now, and it will be 100.000 plex with an upgrade that will be released in June 2005.

PANEL II: EXPRESSION PROFILING FOR DIAGNOSTICS

Charles Decraene: Functional analysis of human genome and its application to diagnosis, prognosis and treatment of cancer.

Research at the Institute Marie Curie is driven by the urge to improve cancer prevention, diagnosis and treatment. Introducing new hospital practices that have been developed on the basis of research is therefore an essential organizational and conceptual goal.
As a bridge between the hospital and the research centre, the Translational and Preclinical Research Department plays a prominent role in transferring new diagnostic and therapeutic technologies to the hospital and provides a genuine link between research and medical applications. It develops and validates diagnostic technology, biomedical devices and therapeutic approaches based on advances in basic research and up-to-date technical platforms.
In 2005, the first two translational medicine studies that will be performed are about breast cancer in young women and uveal melanoma. Theses studies will be dedicated to identify new biomarkers to help clinicians in diagnosis and/or prognosis of these two cancers and to assist them in choosing the right treatment for each of these diseases.
In the Translational and Preclinical Research Department, the convergence of genomic, proteomic and bioinformatic outlines the possibilities of new and innovative forms of medicine, with cancer management gradually shifting away from statistical probabilities to specific and personalized treatment

Harriet Wikman: Identification of Gene Expression Changes Caused by altered DNA Copy Numbers in Lung Cancer.

Several chromosomal regions are recurrently amplified or deleted in lung tumours, but little is known about the underlying genes, which could be important mediators in tumour formation or progression. By using a combination of different high-throughput array methods we wanted to identify new molecular markers, which could be used in early detection and/or treatment of lung cancer. Genes that are both amplified and over-expressed are especially promising targets for intervention.
By first screening primary lung tumours with comparative genomic hybridisation (CGH) on cDNA microarrays we localised a novel amplicon to a > 0.5 Mbp fragment at 12q13.3-q14 (1). This approach enabled us to identify 10-15 genes with the most consistent amplifications. Semiquantitative RT-PCR analyses of 13 genes at 12q13.3 and gene expression profiles using commercial GeneChips (Affymetrix HU1333A) showed that four genes (*CDK4, CYP27B1, METTL1* and *TSFM*) were also highly up-regulated at the mRNA level. Further functional analysis by immunohistochemical (IHC) staining of a larger number of lung tumour samples on a tissue microarray (TMA) showed that CDK4 was expressed at high level in 22% of lung tumours. 19 percent of the patients with high CDK4 expression were shown to contain the 12q13 amplification by fluorescence in situ hybridisation (FISH).
Amplification of *CDK4* has been described in other tumour types, and now by using combined high-throughput array methods we could for the first time describe this amplification also in lung tumours. Although the *CDK4* amplification seems to be a relatively rare event in lung tumours, it indicates the significance of the RB1-CCND1-CDKN2A pathway in lung tumor genesis. This pathway is the key regulator of the G1/S transition in cell cycle and in lung cancer one of the partners in this pathway seems to be modified in some extent in almost 100% of lung carcinomas.

Karsten Zieger: A molecular signature in superficial bladder carcinoma predicts clinical outcome.

Bladder tumors are common neoplasms in developed countries with a yearly incidence of around 35/100,000. The prediction of the clinical course of superficial bladder cancer is difficult; the clinical challenge is the prevention of both disease progression and overtreatment. Using prospectively collected material from a bladder cancer tissue bank with clinical follow-up data and genome-wide microarray based gene expression profiling of around 100 bladder tumor samples, we were able to develop and validate a gene expression signature to identify the disease stage. Moreover, the results suggested that the molecular profile contained information about the future development of the disease, including recurrence and progression, which were not apparent by histopathological evaluation. Applying the same methods, we constructed a molecular classifier to identify concomitant *carcinoma in situ*, a marker of a distinct molecular pathway of disease development, and a classifier to predict the malignant development, including progression to invasive cancer and metastasis. Both classifiers were transferred to an in-house fabricated gene-expression microarray featuring around 400 individual genes, and validated using an independent sample set of 70 bladder tumors. Our present research aims at stratifying tumors by the molecular signatures, thereby improving classification results. The performance of the classifiers was further validated in an international multicenter study implicating a large sample set and tissue preparation at different laboratories. The perspective is the development of a validated, robust and cost-effective diagnostic tool that will help clinicians in advising their patient and to choose the best individual treatment.

Valter Gattei: Microarrays in molecularly and immunophenotypically well-characterized chronic lymphocytic leukemia (b-cll).

With the aim to identify novel prognosticators to transfer in a clinical setting, our proposal is to investigate the gene expression profiling (GEP) of purified B-CLL cells and to analyse results by comparing B-CLL subsets according to the following findings:
1) IgVH mutational status of B-CLL with different prognosis - B-CLLs with mutated (M) IgV_H genes have a better prognosis than unmutated (UM) cases. We analyzed the IgV_H mutational status of B-CLL according to the biochemical features of canonical somatic hypermutations (SHM), and correlated this data with patient survival. In a series of 141 B-CLLs, 124 cases were examined for IgV_H gene percent mutations and skweing of replacement/silent mutations from FR/CDR as evidence of antigen-driven selection; this allowed the identification of three B-CLL subsets: significantly mutated (sM), with evidence of antigen-driven selection, not significantly mutated (nsM) and UM, without such an evidence and IgV_H gene % mutations above or below the 2% cut-off, respectively. sM B-CLL patients had longer survivals also within the good prognosis subgroup with more than 2% mutations of IgV_H genes.
2) Immunophenotypic signature of B-CLLs with different prognosis - We investigated the expression of 36 surface markers (cell-adhesion molecules, integrins, complement activity regulators, myeloid, T and B antigens) in 123 B-CLLs, all with IgV_H mutations and survivals. Results were analyzed by data mining tools identical to those employed in gene expression profiling. Unsupervised (hierarchical and K-means clustering) algorithms revealed three distinct groups (I, II and III), group I (51/123) with longer survivals, as compared to groups II (36/123) and III (36/123). The immunophenotypic signatures of these groups, as determined by applying the nearest shrunken centroids method as class predictor (supervised analysis), were characterized by the coordinated and differential expression of 12 surface markers, which may represent a set of surface antigens to be employed as additional prognosticators for B-CLL.
GEP is performed by using the Operon (version 2.0) platform that presents 21.329 70*mer* oligonucleotides, designed from Human Unigene clusters, mainly at the 3'-end of genes. Oligonucleotides are deposited in a double replica by the MicroCRIBI, University of Padua, on MICROMAX Glass Slides. Studies are carried out on cDNA from purified B-CLL cells; as reference, RNA from normal B cells purified from a pull of healthy donors is employed.

Sakari Knuutila Gene amplifications and losses in hematologic malignancies by array CGH.

Knuutila presented the use of cDNA- and oligo-based microarrays, consisting of 16,000 and 42,000 probes, respectively, in studies of gene copy number alterations in carcinoma, sarcoma and leukemia cases. The key advantage that oligo and cDNA arrays provide over BAC array is the possibility to directly study amplified genes and their expression on the same array. Some clinical examples were shown of cases recently published by Knuutila's group. Specifically, the group has studied amplicons in gastric cancer, osteosarcoma and leukemias. Their results show the complex nature of amplicons with gene amplifications, losses and fusions. The group has also utilized array CGH in diagnostics to confirm chromosome CGH results. Through this technology they have found small amplicons and deletions in leukemia patients with a normal karyotype. In addition to biological relevance, these findings translate into follow-up signals with direct clinical significance. Moreover, array CGH is a powerful tool for studying marker chromosomes of unknown origin.

Miroslav Prucha: Gene expression signatures in sepsis – influences of center-associated factors.

It is justified that Sepsis is at issue. For the US, 750.000 new cases of sepsis per year have been reported. The mortality is about 30-80% and the costs reach an amount of ca. 17 billions/year.
From scientific point of view two concepts of sepsis are presently predominant: sepsis as immune deficiency and sepsis as hyperinflammatory response to infection. From 2001 on a novel methodology for staging sepsis has been applied that is PIRO (prediction, infection, response, organ dysfunctions). There is hope that testing for differential gene expression using sepsis/inflammation specific microarray technology may soon become part of bedside patient care and may improve clinical research. New diagnostic markers should be able I.) to distinguish between patients with and without infection among those with clinically suspected sepsis, II.) to differentiate between survivors and non-survivors in the early stage of sepsis, and III.) to predict the risk of infectious and non infectious (hemocoagulopathy) complications in elective surgery patients.
A multicentre experimental study was carried out simultaneously in the Czech Republic and Germany. The hypothesis for this study was that the gene expression pattern is highly conserved in patients with sepsis despite strong inter-individual differences and is independent of the admission diagnosis and later treatment (center effects). Such a robust cluster of genes would be able to discriminate between ICU controls and patients with sepsis. The constructed microarray *Sep-Arraytor® human 5.500* (SIRS-Lab GmbH, Jena, Germany) includes 5.226 human genes and additional controls, represented through synthetic polynucleotides with an average length of 70bp. Genes were identified within genome-wide pre-screens relevant to inflammation, immune response and related processes and represent mainly intracellular molecules involved in cell-signalling such as MAPK, CARD, TRAF, and NFκB.
Results.
131 genes were found that express differentially for severe sepsis and do not depend on center effects.
Conclusions:

- Microarrays are advisable as a tool for identifying genes that are up- or down-regulated in septic patients. Despite the center-associated effects we are able to determine genes, which characterize the infectious systemic inflammatory response. These genes can potentially be used as markers for sepsis, or their product, as targets for the therapy.
- INDEED, "Tailored" arrays may be developed in the future which will differentiate between the infectious and non-infectious causes of multiorgan dysfunction.
- These arrays will have reliable platform technology, which allows for POC testing with a limited number of genes.

- Microarrays are currently not advisable as a routine diagnostic tool in septic patients, because of expense and time – consumption.

The following discussion that was initiated by Olaf Riess emphasized the need for clinically well classified and defined tumors/tissues. Tumors should be immuno-histo-pathologically defined. It was stated that for some tumors one might need micro-dissection. DNA analysis for detecting deletions/amplifications should be supplemented optionally. In general, a re-evaluation of the pathological data is needed with view to microarray diagnosis.
Furthermore, the need for long prospective studies was expressed. These require large patient samples, which might only be collected at European level in decent times, which adds to the motivation to form (a) European consortium(s) in this field. Two further issue have been raised. First, the terms of how, when and under what conditions to translate research data into daily medicine were at issue. Here, the number of patients to be studied, the difference of software packages and the influence of treatment on the data were mentioned. Second, the issue of handling patient data was raised. In particular, two scenarios were considered. On the hand it was asked what ethical principals apply if patient data reveal something that was not agreed to investigate in the signed consent form. On the other hand, the need for a European harmonization of sample handling in terms of anonymisation was stressed.

PANEL III: MICROARRAY BASED RESEQUENCING FOR DIAGNOSTICS

Neeme Tõnisson: Resequencing of p53 gene by arrayed primer extension.

P53 is a classical tumour suppressor gene frequently inactivated in cancer by various mutations. Although certain mutation hotspots exist in the p53 gene, the mutations in cancer are scattered allover the region encoding the DNA-binding domain of the protein. Therefore a good and sensitive screening method is required for efficient discovery of the mutations.
Resequencing assay of the full coding region of p53 gene, has been developed by Asper Biotech. The assay is based on four-color primer extension reaction on the oligonucleotide microarray (APEX). In current state, Asper Biotech with its academic collaborators in Norway and France has completed two validation studies of the p53 gene resequencing assay and comparisons of it with other commonly used techniques (DHPLC, TTGE and automated dideoxy sequencing). The results of the validation studies have been promising for various mutation types. The detection limits of APEX were found to be as low as 3 to 6 % of the mutated DNA in the background of wild type sequence. Nevertheless, dideoxy sequencing is still required to complement and partially control the microarray-based resequencing results on somatic mutations.
Based on the APEX technology, a number of mutation screening assays have been developed for various inherited disorders (cystic fibrosis, beta-thalassemia, Stargardt disease, Usher syndrome*, etc*).

Ioannis Ragoussis: Genomic applications of MALDI-TOF MS.

The Wellcome Trust Centre of the Oxford University is unique since it combines the MRC Oxford protein production, the Centre for Cellular & Molecular Physiology, the Facility for the Evaluation for infectious particles, the Magnetic Resonance Imaging Facility and the genomics group. The Centre is focused on three main disease areas i.) neurogenetics, ii.) genetics of inflammation and immunity and iii.) genetics of cardiovascular disease and metabolic syndrome. In addition there are programmes of research in structural biology, bioinformatics, statistical genetics and genomics.
Two MALDI-TOF applications were presented.
1. Pilot study for mutation analysis.

Setting. We screened samples from individuals with Multiple Exostosis for heterozygous mutations in the EXT1 or EXT2 genes. The study investigated in detail EXT1 exon 6 and EXT2 exon 5 with 15 mutations in total. Samples were analysed using DHPLC/WAVE, sequencing and RE digestion were applied where possible. We used different primer designs to produce 200-700bp long fragments. Different RNAse treatment conditions have been tested. 135 assays were evaluated. We applied SNP discovery software and used retrospective simulation analysis to compare spectra to theoretical calculations.
Conclusions. Mutations in 11/15 different samples could be identified correctly. 9/13 different mutations could be characterised. The technology showed high specificity when the fragment length was 200-300 bp. Problems were generated by background/low yield of complete RNAs in vitro. The throughput is 172-960 assays per instrument per run (2 runs per day possible). However, further work towards robust molecular biology and interpretation software is needed.
2. hMC and hME for methylation detection.
Specific aim: Optimize and implement a methylation detection protocol using Sequenome Mass Spec technology, compare the sensitivity of the Mass Cleave (hMC) and the Mass Extend (hME) allelotyping protocols.
Reasons: i.) Sequenom's MALDI-TOF based technology is a high-throughput platform, ii.) the Mass Cleave assay offers the possibility of an extensive methylation profile of a CpG island, at present there is no other technology available that offers the same advantage
Conclusions: As expected both hME and hMC can identify methylated bisulfite treated induced SNPs and the results are consistent using the two methods on the same samples. The hME allelotyping software can be already used to quantify the level of methylation observed (quite challanging assay design in CpG islands). The hMC protocol has more power in producing a more complete profile of the CpG island, but the data are less easy to analyze at this time due to the lack of appropriate software.

Michael Bonin: Microarray-based sequencing of the whole Cystic Fibrosis CFTR gene.

The Microarray facility of the University of Tübingen was founded in December 2001, three scientists and two technical assistants are employed. The facility is integrated in the Medical Genetics Department of the Medical Faculty of the University of Tübingen. It serves as Affymetrix service provider "expression analysis" (since 08/03) and „genotyping" (since 06/04).
Making use of the Affymetrix CustomSeq™ Resequencing Array we developed a whole gene resequencing design for the coding region of the Cystic fibrosis gene. Conventional sequencing of the complete coding region of CFTR needs 4-8 weeks. The array contains 27 exons plus promotor and intron-exon bondaries. The developed CFTR array allows to resequencing the complete coding region of CFTR. Excluded from detection are 1-bp deletions and exon deletions (in the standard design). In comparison to conventional sequencing the costs can reduce upon 20%. The hands on time can reduce upon 15% without any quality reduction. 45 sample (35 patients/10 controls) have been analysed using the designed array. 28 different mutations were detected which were confirmed by traditional sequencing. Additionally 54 SNPs were found. Mean value of N-calls was 0.02%.
In collaboration with our bioinformatics department we plan to put forward the development of a improved software for the resequencing analysis. Our main goal is the implementation of our CFTR Array in the diagnostic routine process. Furthermore, the development of new resequencing arrays for different disease areas and the validation of further resequencing methods (Sequenom, MALDI-TOF) have been taken in view.

PANEL IV: BIOINFORMATICS AND DATABASE MANAGEMENT FOR MICROARRAY BASED DIAGNOSTICS

Rani Elkon: Integrative analysis of heterogeneous functional genomics data.

The DNA damage response network modulates a wide array of signaling pathways, including DNA repair, cell cycle checkpoints, apoptotic pathways and numerous stress signals. The ATM protein kinase, functionally missing in patients with the human genetic disorder ataxia-telangiectasia (A-T), is a master regulator of this network when the inducing DNA lesions are double strand breaks. The *ATM* gene is also frequently mutated in sporadic cancers of lymphoid origin. In this study, we applied a functional genomics approach that combines gene expression profiling and computational promoter analysis to obtain global dissection of the transcriptional response to ionizing radiation (IR) in murine lymphoid tissue. Cluster analysis revealed six major expression patterns in the data. Prominent among them was a gene cluster that contained dozens of genes whose response to irradiation was Atm-dependent. Computational analysis identified significant enrichment of the binding site signatures of the transcription factors NF-κB and p53 among promoters of these genes, pointing to the major role of these two transcription factors in mediating the Atm-dependent transcriptional response in the irradiated lymphoid tissue. Examination of the response showed that pro- and anti-apoptotic signals were simultaneously induced, with the pro-apoptotic pathway mediated by p53 targets, and the pro-survival pathway by NF-κB targets. These findings further elucidate the molecular network induced by IR, and have implications for cancer management. They suggest that restoring the p53-mediated apoptotic arm while blocking the NF-κB-mediated pro-survival arm could effectively increase the radiosensitivity of lymphoid tumors.

Magnus K. Magnusson: Statistical design of microarrays: increasing sensitivity and specificity by introducing novel design features.

Whole genome expression analysis promises to give unique insights into the biology and pathology of cancer. The transcriptional profile of the individual tumor will likely help define i.) new biologically important subtypes, ii.) important regulatory pathways, and iii.) novel therapeutic targets. The crucial question is: How do we translate this research tool into a diagnostically useful tool for clinical practice?
Multiple studies have looked at gene expression profiles in breast cancer for risk prediction. Two major Dutch studies have suggested profiles – *signatures* – that predict clinical outcome in early lymph node negative breast cancer. These studies applied only slight differences in inclusion criteria. But no consensus could be observed. There was a minimal overlap in gene signatures.
Applying expression signature to risk prediction the most striking finding when comparing the signature lists is the virtually complete lack of agreement in the included genes. To some extent this lack may be explained by differences in microarray technology and selection of genes represented on the arrays, as well as experimental issues. Despite these obvious reasons for differing gene signatures, the microarrays used in different studies generally have several thousand genes in common and the underlying principles of the measurement technologies are the same. So we strongly believe the present lack of coherence still warrants further examination.
Hypothesis:
Current gene signatures for risk predictions can serve as platform for development of diagnostic arrays but the gene expression platform needs to be readjusted to allow for testing of smaller gene lists with much better quality. Fluctuation (variability) in individual gene expression levels need to appropriately controlled.
Approach:
We need to step back and improve the quality and power of the statistical analysis of the primary data.
Fundamental design principles of the "Statistical Array":
I.) Block out known sources of variation, ii.) Randomize to remove unknown sources variation and iii.) replicate to dampen uncontrollable variation.
Results:
By applying conventional statistical methods to control for varibility at the microarray design level and the experimental level we can increase significantly the quality of the probe and

gene level data as seen by increase in sensitivity and specificity. We are now applying these principles using clinical breast cancer samples to test whether we can predict oucome
We have access to fresh frozen tissue samples dating 15 years back through the Icelandic Cancer Project linked to extensive clinical and genetic information. The current pilot project will include 80 early stage (lymph node negative) breast cancer with our without metastatic behaviour on follow up.

PROBLEMS, BOTTLENECKS AND POLITICS

James Hadfield: Moving array based diagnostic into the hospital lab: sample processing and QC.

This presentation dealt with the quality control issue covering the full procedure beginning with sample (RNA) quality control and ending with data quality control. Issues touched on were:

- Processing Quality Control,
- Protocols vs SOPs,
- Sample Quality Control,
- Data Quality Control,
- Tracking, and
- Databases.

The emphasis was further on pointing at things at issue and of possible value for the further development of the scientific field such as:

- "Microarrays" as a diagnostic tool or "Microarray" diagnostics for this ESF consortia's favourite disease(s)?
- How to reconcile different group preferences for both area of interest and technology?

More generally:

- pace of technological advancement,
- collaboration: commercial and academic, academic and academic, everyone and clinician/patient,
- quality control of diagnostic test,
- control/analysis of centre to centre variation,
- validation of current tools,
- whole Genome versus focused tests, and
- cost of diagnostic testing.

The discussion on how to continue and/or how to make the best use of this ESF workshop was fed by two presentations First, Holm Graessner chose a more general view on the subject "*Towards a European consortium*". He sketched the general conditions and regulations of the current 6th framework program of the European Commission (FP6) and of the forthcoming 7th framework program (FP7). Referring to the draft of the 4th call of the FP6 he pointed at the included topic "High Throughput molecular diagnostics for hereditary diseases" thought for the formation of an Integrated Project. Furthermore, the general draft structure of FP7 – cooperation, ideas, people, capacities – was presented and related to possible applications coming from a "Microarray/Diagnostics" consortium. Second, Olaf Riess presented his own ideas on how microarray analyses could replace and improve cytogenetic prenatal diagnostics. He outlined that

- strong interaction with gynecology and human genetics is needed,
- ethical aspects are strong,
- the technology is ready but needs to be specifically applied, and that
- data and data bases are lacking.

The concluding discussion concerned the issue how to increase the steadiness of the just launched formation of a Microarray/diagnostics network, what opportunities exist to foster and facilitate this development and what measures can be taken for these purposes. (see *Assessment of the results and contribution to the future direction of the field*)

Assessment of the results, contribution to the future direction of the field

The workshop was a first initiative to bring together people across Europe working in the multidisciplinary area of microarray applications in diagnostics. The workshop enabled the participants to get an overview of existing research activities in Europe and to establish contacts with other participants.
However, the major focus is still to promote microarray as a tool for diagnosis of human diseases. There are already positive aspects of the meeting if one considers horizontal issues and technical levels.
As one important outcome in particular of the last discussion session it was agreed that the workshop functions as the first step towards the forming of an European network on the microarray-diagnostics issue. However, a consensus was reached not to apply for a project topic on high throughput molecular diagnostics for hereditary diseases, which will be published in the fourth call of the 6. FP of the European Commission. The common understanding was that networking measures are needed before considering such a proposal. For this purpose, applications are planned concerning two funding schemes, which could facilitate further steps. First the attending group of academic researchers and industry agreed to apply for an ESF / EMBO Symposium which is high level scientific conference which provides the opportunity for the world's leading scientists and other participants, including young researchers, to meet in an informal setting for discussions at the highest level of the most recent development in their fields of research. Second, the group will try to launch a COST Action, which refers to a research cooperation programme of national funded research at a European level. For both applications the Tübingen group will act as the key player.
Although, the participants proved that excellent work has been done in Europe in this field one could also observe why there is no real European gold standard on microarray applications in diagnosis available yet. This might be due to

- a lack of coordinated collaboration of the microarray facilities and laboratories that are scattered all over Europe,
- missing standards and SOPs for sample handling, microarray analysis, data analysis, data exchange etc.,
- missing technical standardization or missing data on comparability of different technology,
- no established European quality management, for example exchange of samples for interlaboratory tests, and last but not least
- the broad spectrum of diseases studied.

As a first step towards a more coordinated approach the participants agreed to take an inventory of the different technologies and of the sample available in the participating groups. As a second step it has been proposed to set up working groups on quality control, on technology and on cataloging biological material. This structure will be reflected in the COST action to be proposed (see above).

In summary, the workshop highlighted four important subtopic of the microarray/diagnostics field: technological advancement, expression profiling, microarray based resequencing, and bioinformatics and database management. A number of possible research direction with respect to these subtopics were identified:

- development of a validated, robust and cost-effective diagnostic tools using expression profiles,
- need for clinically well classified and defined tumors/tissues,
- translation of research data into daily medicine,
- development of new resequencing arrays for different disease areas,

- validation of additional (to Affymetrix) resequencing methods (Sequenom, MALDI-TOF)
- identification and overcoming of causes for observed minimal overlap in gene signatures of different studies, and
- development and establishment of a comprehensive European quality management system.

Programme

Friday 15 April 2005

18:00 – 18:15	**Carole Moquin-Pattey** Presentation of the European Science Foundation
18:15 – 19:15	**Introduction round**
19:15 – 20:00	**Olaf Riess** Introductory lecture
20:00	*Dinner*

Saturday 16 April 2005

PANEL I: TECHNOLOGICAL ADVANCEMENT OF THE MICROARRAY BASED DIAGNOSTICS

10:00 – 10:20	**Michaela Downie** MassARRAY System: A Versatile Genetic Analysis Platform
10:20 – 10:40	**Janna Saarela** Allele-specific primer extension microarrays
10:40 – 11:00	**Snaevar Sigurdsson** Four-colour fluorescent minisequencing in an "array of arrays" format
11:00 – 11:30	*Coffee break*
11:30 – 11:50	**Discussion**

PANEL II: EXPRESSION PROFILING FOR DIAGNOSTICS

11:50 – 12:10	**Charles Decraene** Functional analysis of human genome and its application to diagnosis, prognosis and treatment of cancer
12:10 – 12:30	**Harriet Wikman** Identification of Gene Expression Changes Caused by Altered DNA Copy Numbers in Lung Cancer
12:30 – 13:30	*Lunch*
13:30 – 13:50	**Karsten Zieger** A molecular signature in superficial bladder carcinoma predicts clinical outcome
13:50 – 14:10	**Valter Gattei**

Microarrays in molecularly and immunophenotypically well-characterized chronic lymphocytic leukaemia cells

14:10 – 14:30 **Sakari Knuutila**
Gene amplifications and losses in hematologic malignancies by array CGH

14:30 – 14:50 **Miroslav Prucha**
Gene expression signatures in sepsis - influences of center-associated factors

14:50 – 15:20 **Discussion**

PANEL III: MICROARRAY BASED RESEQUENCING FOR DIAGNOSTICS

15:20 – 15:40 **Neeme Tõnisson**
Arrayed primer extension based resequencing of p53

15:40 – 16:00 **Ionnis Ragoussis**
Genome analysis using MALDI-TOF applications

16:00 – 16.30 *Coffee break*

16:30 – 16:50 **Michael Bonin**
Microarray sequencing of the whole CFTR gene

16:50 – 17:10 **Discussion**

PANEL IV: BIOINFORMATICS AND DATABASE MANAGEMENT FOR MICROARRAY BASED DIAGNOSTICS

17:10 – 17:30 **Rani Elkon**
Integrative analysis of heterogeneous functional genomics data

17:30 – 17:50 **Magnus K. Magnusson**
Statistical design of the diagnostic expression array using the principles of randomization, replication and blocking

17:50 – 18:00 **Discussion**

19:00 *Dinner*

21:00 *Guided tour through Tübingen*

Sunday 17 April 2005-04-07

PROBLEMS, BOTTLENECKS AND POLITICS

10:00 – 10:20 **James Hadfield**
Moving array based diagnostics into the hospital lab: sample processing and QC

10:20 – 10:40 **Holm Graessner**

Towards a European Consortium

10:40 – 11:00 **Discussion**

11:00 – 11:30 *Coffee break*

11:30 – 13:00 **Discussion**

13:00 *Lunch*

Final list of participants

Co-convenors

Prof. Dr. Olaf Riess

University of Tuebingen
Department of Medical Genetics
Calwerstraße 7
72076 Tuebingen
Germany
phone: +497071-2976458
fax: +497071-295171
olaf.riess@med.uni-tuebingen.de

Dr. Michael Bonin

University of Tuebingen
Department of Medical Genetics
Calwerstraße 7
72076 Tuebingen
Germany
phone: +497071-2972295
fax: +497071-295171
michael.bonin@med.uni-tuebingen.de

Local organiser

Holm Graessner
University of Tuebingen
Department of Medical Genetics
Calwerstraße 7
72076 Tuebingen
Germany
phone: +497071-2972285
fax: +497071-295171
holm.graessner@med.uni-tuebingen.de

ESF representative

Carole Moquin-Pattey
ESF/EMRC
1 quai Lezay Marnésia
F-67080 Strasbourg
France
phone: +33388767118
email: emrc@esf.org
http://www.esf.org/emrc

Participants

Peter Riege
Michaela Downie
Sequenom GmbH
Mendelssohnstr. 15-D
D-22761 Hamburg, Germany
phone: +49 40-899 676-0
fax: +49 40-89967610
mobile: +49 172-9206602
priege@sequenom.de
mdownie@sequenom.de

Janna Saarela, MD, PhD
Senior Scientist, Head of the Microarray Unit
National Public Health Institute, Dept. of Molecular Medicine
Biomedicum Biochip Center
Haartmaninkatu 8, suite 318b
FIN-00290 **Helsinki**
Finland
Email: janna.saarela@ktl.fi

Snaevar Sigurdsson
Dept of Medical Sciences,
Uppsala university
Ing. 70, 3 tr.
SE-751 85 **Uppsala**
Sweden
Snaevar.Sigurdsson@medsci.uu.se

Group:
Ann-Christine Syvänen
Dept of Medical Sciences, Uppsala university
Ing. 70, 3 tr.
SE-751 85 Uppsala

Charles Decraene, PhD
Institut Curie
26 rue d'ULM
75248 **PARIS** cedex 05, France
Email: charles.decraene@curie.fr
tel: 33.1.53.19.74.13
fax: 33.1.53.19.74.18

Group:
Jean Paul Thiery, PhD
CNRS UMR 144
Cell Biology Department
Research Division
and Head Dept Translational Research
Medical Division

Harriet Wikman, Ph.D.
Department of Tumor Biology
Center of Experimental Medicine
University Hospital Hamburg-Eppendorf
Martinistrasse 52
D-20246 Hamburg
Germany
Tel: +49-40-42803 7494
Fax: +49-40-42803 5379
Email: h.wikman@uke.uni-hamburg.de

James Hadfield
Microarray Manager
Genome Laboratory,
John Innes Centre,
Norwich,

NR4 7UH, UK.
Tel: +44(0)1603 450830
Fax:+44(0)1603 450021
Mobile: 07879 844167
www.jicgenomelab.co.uk
Email: james.hadfield@bbsrc.ac.uk

Karsten Zieger, MD
Molecular Diagnostic Laboratory
Dept. Of Clinical Biochemistry
Aarhus University Hospital, Skejby
DK 8200 Århus N
Email: karsten.zieger@ki.au.dk

Group:
Torben F. Ørntoft
MD, DMSci, Chief physician,
Head of Department
Dept. of Clinical Biochemistry
Aarhus University Hospital
Skejby
DK 8200 Aarhus N, Denmark

Valter Gattei, MD
Head Clinical and Experimental Hematology Research Unit
Centro di Riferimento Oncologico, I.R.C.C.S.
Via Pedemontana, 12
I-33081 **Aviano** (PN) – Italy
Phone: +39 0434 659 410
Fax: +39 0434 659 409
E-mail: vgattei@cro.it

Miroslav Prucha MD, PhD
Dept.of Clinical Biochemistry, Haematology and Immunology
Hospital Na Homolce
Roentgenova 2
15130 **Prague** 5
Czech Republic
Phone: +420 257272446, fax: +420 257272981,
Email: miroslav.prucha@homolka.cz

Neeme Tõnisson,
Taavi Einaste
Asper Biotech AS
Oru 3,
51014 Tartu,
Estonia
www.asperbio.com
Phone: +372 7 441 556
GSM: +372 5 270 307
Fax: +372 7 442 343
Email: taavi@asperbio.com

Ioannis Ragoussis
Head of Genomics
Reader in Genomics
Wellcome Trsust Centre for Human Genetics
University of Oxford
Roosevelt Drive
Oxford OX3 7BN
ioannisr@well.ox.ac.uk

Rani Elkon
Department of Human Genetics and
Molecular Medicine
Sackler School of Medicine
Tel Aviv University
Ramat Aviv 69978
Israel
E-mail: ranel@post.tau.ac.il

Group:

Ron Shamir
Head, School of Computer Science
Sackler Professor of Bioinformatics

Magnus K. Magnusson, MD
Department of Hematology &
Department of Genetics and Molecular Medicine
Landspitali - University Hospital,
101 **Reykjavik**
ICELAND
E-mail: magnuskm@landspitali.is

Dirk Koczan
Universität Rostock
Medizinische Fakultät
Institut für Immunologie
Schillingallee 70
18055 Rostock
Germany
Telefon: +49 +381 494 5870
Fax: +49 +381 494 5882
Email: dirk.koczan@med.uni-rostock.de

Sakari Knuutila, Ph.D.
Department of Pathology
Haartman Institute and HUSLAB
POB 21 (Haartmaninkatu 3, 4th flr)
FI-00014 **University of Helsinki**
Helsinki, Finland
Phone: +358-(0)9-191.26527, Mobile +358-50-4272003
Fax: +358-(0)9-191.26788
Email: Sakari.Knuutila@helsinki.fi

Statistical information

Country

Czech Republic: 1
Denmark: 1
Estonia: 2
Finnland: 2
France: 2
Germany: 5
Island: 1
Israel: 1
Italy: 1
Sweden:1
United Kingdom: 3

Sex

Female: 4
Male: 16

Position

Industry:
Researcher: 1
Management: 3
Academic:
Senior Researcher: 9
Junior Researcher: 3
Management: 2
ESF
Management: 1