1-IS
Successes and failures of molecular cancer epidemiology

F. Perera[1]
[1]*Columbia Center for Children's Environmental Health, Mailman School of Public Health, Columbia University, New York, USA*

Molecular epidemiology was introduced in the study of cancer in the early 1980s, with the expectation that it would help overcome some of the limitations of epidemiology and facilitate cancer prevention. The expectation was that biomarkers would improve exposure assessment, document early changes preceding disease, and identify subgroups in the population with greater susceptibility to cancer, thereby increasing the ability of epidemiological studies to identify causes and elucidate mechanisms in carcinogenesis. An early example of a successful application of molecular epidemiology in cancer research includes the measurement of chemical-specific DNA adducts formed by the interface between environmental exposures, like polycyclic aromatic hydrocarbons (PAHs) and DNA, a toxicological target. Using this type of biomarker as a biologic dosimeter, researchers could identify individuals who, when exposed to specific carcinogens, are likely to be at increased risk for developing health conditions, including cancer. Another success, also involving a genotoxic endpoint, has been the demonstration that increased chromosomal aberrations or specific pre-leukemic changes are associated with risk of cancer. In addition, molecular epidemiology has successfully identified certain subgroups, such as the fetus and young child, or those with specific genetic polymorphisms or nutritional deficits as being more susceptible. However, the failure to measure the full spectrum of pre-clinical alterations resulting from carcinogen exposure has precluded clear gains in terms of cancer prevention. Recently, new epigenetic and "-omic" biomarkers have become available to address these gaps, thanks to the development of high-throughput technologies based on advances in molecular biology. While these new biomarkers hold promise to revolutionize the field of molecular epidemiology, most have not yet been adequately validated and their role in the causal paradigm is not clear. To achieve their potential, there is a need for systematic validation of these newer biomarkers using principles and criteria established over the past several decades in the epidemiology and molecular epidemiology of cancer. They can then be used in combination with the earlier validated biomarkers of exposure, risk and susceptibility to identify "at risk" individuals, increase our understanding of mechanistic carcinogenic pathways, and mount more effective interventions to prevent cancer occurrence.

2-IS
Novel viral markers

M. Pawlita[1], M. Schmitt[1], K. Michael[1], T. Waterboer[1]
[1]*Dept. of Genome Changes and Carcinogenesis, Research Program Infection and Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany*

Infection with distinct types of Human papillomaviruses (HPV), Herpes viruses, Hepatitis viruses, very recently identified Merkel polyoma virus but also bacteria, e.g. Helicobacter pylori, are essential factors in the pathogenesis with specific human tumors. We have developed high-throughput multiplex technologies for the simultaneous, quantitative detection of (i) antibodies to up to 100 different viral or bacterial proteins, or (ii) the genomes of broad varieties of viruses and bacteria.

Multiplex serology uses full-length proteins bacterially expressed as Glutathione S-Transferase (GST) fusion proteins as antigens, affinity-purified in a single step in situ on different sets of glutathione-coated fluorescence-labeled polystyrene beads (Luminex®). Multiplex genotyping uses multiplex PCR to amplify the target sequences followed by hybridization of the PCR products to specific oligonucleotide probes covalently bound to Luminex beads. Antibody bound to antigens or hybridized PCR products on beads are stained with fluorescent secondary reagents and quantified specifically for each antigen or probe in a modified 2-colour FACS machine.

Examples for application of these technologies to large cross-sectional case-control and infection prevalence studies will be described and discussed.

3-IS
Application of epigenetics to cancer epidemiology

Z. Herceg[1]
[1]*Epigenetics Group, International Agency for Research on Cancer (IARC), Lyon, France*

It is now recognized that non-genetic heredity is important for understanding the causes of complex diseases such as cancer. Recent discoveries provided strong evidence arguing that epigenetic mechanisms play key roles in virtually all stages of cancer development and progression. A number of critical processes found in cancer cells, such as silencing of tumour suppressor genes, activation of oncogenes, aberrant cell cycle, and defects in DNA repair, can be a consequence of not only genetic but also epigenetic changes. The term "epigenetic" refers all heritable changes in gene expression and associated phenotypic traits that are not coded in the DNA sequence itself. Epigenetic inheritance include DNA methylation, histone modifications, and microRNAs, all of which are essential mechanisms that allow the stable propagation of gene activity states from one generation of cells to the next. Despite the fact that epigenetic changes induced by environmental, dietary and lifestyle factors are likely to be important mechanisms of cancer development, epigenetic epidemiology is in its infancy, and little is known of the precise contribution of epigenetic changes to cancer burden. A list of genes and gene networks that are targets of epigenetic alterations are likely to grow with the development of powerful screening approaches. Almost spectacular technological advances in epigenetics and epigenomics now allow powerful screening of large series of samples of different cancer types. These approaches are beginning to reveal a number of genes (tumor suppressors and other cancer-associated genes) susceptible to inactivation through epigenetic mechanism. Epigenetic profiling using both genome-wide and candidate-gene approaches in different tumor types will help in elucidation the mechanism underlying tumourigenesis. Technological advances in epigenetics and epigenomics as well as ongoing studies aiming to identify specific epigenetic targets, environmental factors, and the critical windows of vulnerability to environmentally induced epigenetic alterations will be discussed.

4-IS
Expression microarrays and cancer epidemiology; lessons from breast cancer

A.L Børresen-Dale[1], V.N. Kristensen[1]
[1] *Norwegian Radium Hospital, Rikshospitalet University Hospital, Oslo, Norway*

Expression profiling in relation to breast cancer (BC) have been performed on healthy breast tissue and various tumor tissues from benign to malignant in order to identify profiles related to BC risk, aggressiveness, metastatic potential and therapy response. At least five distinct molecular subgroups have been identified associated with different clinical outcome (Sørlie et al PNAS 2001, 2003). There is growing evidence that analyzing blood-derived gene expression profiles may lead to a diagnostic test for BC. A gene signature of 82% accuracy, 87% sensitivity and 76% specificity has been identified (Sharma et al BCR 2005, Aarøe et al 97thAACR and 19thEACR, 2006, Børresen-Dale AL et al 97thAACR 2007). In a current study to investigate the efficacies of a blood based test in another ethnic population, 720 subjects with or without BC from diverse areas of India are recruited. The results of interim analyses of approximately 350 cases indicate that the previously identified 96 gene set efficiently discriminates BC and non-BC samples, providing evidence for a gene expression signature as a potential additional tool in BC diagnostic work-up.

To improve the reliability and accuracy of the various expression profiles, the role an individual's genotype (SNPs, CNVs) and exposure (eg hormones) has on these profiles is of importance to identify. Recently, GWAS analysis of BC has revealed SNPs in 5 novel genes associated to susceptibility. With the notion that the probability of developing a given subclass of breast cancer is genetically determined, we might expect to find that the newly discovered susceptibility genes are differentially expressed in the various tumor subclasses, and that their transcription is regulated in cis by SNPs within them. With this in mind, we retrieved the mRNA expression data of these 5 genes from 112 breast tumors representing all 5 subclasses, and significantly different mRNA levels between the subtypes were found for all the 5 genes by ANOVA analysis (Nordgard et al 2007 BCR, Kristensen and Borresen-Dale 2008 Mol. Onc.). This illustrates the necessity to conduct stratified SNP-disease association studies. Stratification of patients by their molecular subtypes may give much more power to the classical case control studies, and genes of no or borderline overall significance may be highly penetrant for certain subclasses, and therefore identifiable.

5-IS
Integrated gene expression analysis in PBL and in the derived Lymphoblastoid Cell Lines to define individual radiotoxic risk profiles

M.A. Pierotti[1]
[1]*Fondazione IRCCS Istituto Nazionale dei Tumori Milano, Italy*

Microarrays are a powerful tool for the simultaneous analysis of expression of thousands of genes to outline similarities and differences among samples. This technology is extensively used to explore biological causes

and effects and offers great promise for leading to the future development of treatments tailored according to persons' unique genetic characteristics. We focused on identification of expression profiles potentially predictive of response to treatment using an accessible source such as blood. We defined gene expression profiles of PBLs, in particular of non quiescent T-lymphocytes such as those stimulated with Phitohemoagglutinin and interleukin-2 (PHA-PBLs) and of the derived immortalized LCLs, from healthy individuals. These data will be useful to determine the feasibility of developing directly on lymphocytes clinical tests for the prediction of response to treatment, avoiding the need of cell immortalization. The results have been applied to the study of toxicity from ionizing radiation (IR) therapy in cancer, based on the hypothesis that some cases of toxicity may be associated with abnormal transcriptional response to radiation.

Our goal is to establish the basis for a practical clinical test that works directly on peripheral blood and will be used to predict response to treatment, therefore enabling to match the appropriate cure to the appropriate patient. The proposed approach could be extended for the identification of genes signatures associated to molecules that damage DNA in a manner similar to ionizing radiation (genotoxic or radiomimetic agents).

6-IS
Application of proteomics to cancer epidemiology

S. Hanash[1]
[1] *Molecular Diagnostics Program, Fred Hutchinson Cancer Research Center, Seattle, USA*

Although our understanding of the molecular pathogenesis of common types of cancer has improved considerably, the development of effective strategies for cancer prevention, risk assessment and early detection have lagged behind. Current proteomic strategies allow quantitative profiling of cells, tissues and biological fluids and identification of proteins changes resulting from altered levels, post-translational modifications and amino acid substitutions. A major application of proteomics is assessment of health related changes in the plasma proteome. However the vast dynamic range of protein abundance in plasma and the likely occurrence of biomarkers in the lower range of protein abundance represent a major challenge in applying a proteomics to cancer epidemiology. A combination of innovative strategies promises to overcome these challenges. Recent experience in comprehensive profiling of plasma proteins indicates that low abundance proteins may be identified and quantified with high confidence following extensive plasma fractionation and with the use of protein tagging procedures and high-resolution mass spectrometry. From an experimental design point of view, most cancer biomarker studies, including those aimed at identifying markers for early detection, are initiated with analysis of specimens from newly diagnosed subjects. Specimens obtained from large cohorts are becoming available to allow large-scale investigations of risk factors related to common cancers. The current status of the field and emerging findings will be presented.

7-IS
Proteomics; An epidemiological view

R. Vermeulen[1]
[1]*Institute for Risk Assessment Sciences, Utrecht University, The Netherlands*

In general the function of a cell can be described by the proteins that are present in the cell and the abundance of these proteins. Although all proteins are based on mRNA precursors, post translational modifications and environmental interactions make it impossible to predict abundance of specific proteins based on gene expression analysis. As such proteomics holds great theoretical promise. However, major challenges still need to be over won both technologically (ability to reliable detect a wide range of proteins), as epidemiologically (study design, sample collection, information on inter- and intra-individual variability). In light of these challenges it might not be surprising that studies until now have only sporadically identified the same proteins associated with disease state. Furthermore, identified proteins have in general been high abundance (housekeeping) proteins and it could be questioned whether these markers provide any new biological insights in the mechanism of disease. However, new techniques like LC/MSMS and protocols like high abundance protein depletion might improve the ability of proteomic techniques to reliably measure low abundance proteins.

Targeted protein screening based on carefully selected protein markers might (at least for the immediate future) be a more promising way forward. In these techniques ligand-binding reagents, which are usually antibodies but may also be alternative protein scaffolds, peptides or nucleic acid aptamers, coated to beads or solid carriers are used to measure multiple protein markers in a small volume of biological materials. Given that a predetermined panel of protein markers is measured basic information on assay performance (inter- and intra-batch variance), protein stability and inter- and intra-individual variance in the markers of interest can be established a-priori greatly facilitating the correct interpretation of study results. However, these techniques are currently only able to measure up to a few hundred proteins. Furthermore, a number of important technical challenges and bottlenecks in protein array technologies, some of which are unique to proteins while others are common to high throughput methods in general, will need to be solved in order to achieve the maximum capability.

A more fundamental issue might be however the highly variable proteome in combination with the often unknown influence of collection media and storage conditions on the stability of proteins. This would undoubtedly lead to an unfavorable ratio between the inter- and intra-individual variability and as such complicating its application in epidemiological research. This becomes even a bigger issue when one moves from studies targeted to disease-recognition to early (prospective) detection of diseases where one would expect relatively small changes in biological parameters. It is therefore imperative that appropriate, well-powered study designs are employed using highly standardized biological collection protocols. At current full scale proteomics, despite of the theoretical advantages over genomics and transcriptomics might not be ready to be used in large scale(prospective) epidemiological research.

8-IS
Case-control mutation screening – insights and lessons from breast cancer susceptibility

N. Rahman[1]
[1]*Human Genetics, Section of Cancer Genetics, Institute of Cancer Research, Sutton, United Kingdom*

Approximately 1 in 10 women develop breast cancer. Epidemiological studies have demonstrated that first-degree relatives of breast cancer cases are at two-fold risk of developing the disease. Currently, three components of the genetic architecture of breast cancer have been delineated. 1) Rare, high penetrance (>10 fold) autosomal dominant cancer predisposition genes such as BRCA1 and BRCA2. 2) Common, low penetrance (<1.5 fold) susceptibility alleles that have emerged from genome-wide tag-SNP searches in breast cancer. 3) Rare, intermediate penetrance (2-4 fold) susceptibility alleles discovered through large-scale, case-control resequencing analyses. We have identified four DNA repair genes, ATM, CHEK2, BRIP1 and PALB2 which exemplify this final class, through mutation screening of familial breast cancer cases and controls. These susceptibility genes are similar to BRCA1 and BRCA2 as they are characterised by multiple, individually rare, monoallelic, truncating mutations but they are associated with smaller increases in risk, approximately doubling the risk of breast cancer. All four genes encode proteins that function in DNA repair pathways and biallelic mutations in three of them (ATM, BRIP1 and PALB2) cause childhood developmental disorders associated with high risks of cancer, similar to biallelic BRCA2 mutations. Identification of rare, intermediate / low penetrance genes is currently challenging because of the dependence on high-throughput sequencing in large case-control series and correct candidate gene selection. However, this will likely change in the near future with the advent of whole genome sequencing. Moreover, it is highly plausible that this class of susceptibility allele is making a contribution to many diseases.

9-IS
Lessons from genome-scans – the example of lung cancer

P. Brennan[1]
[1]*International Agency for Research on Cancer – Genetic Epidemiology Group, Lyon, France*

Lung cancer is the most common cause of cancer globally, representing 1 in 8 of all cancer cases and 1 in 6 of all cancer deaths occurring in 2002. The predominant risk factor is tobacco smoking, with the risk of developing lung cancer by age 75 in Europe and North America ranging from approximately 15% in lifelong smokers to less than 1% in never-smokers. In populations where the large majority of smokers have quit smoking, such as men in many parts of Europe and North America, an increasing proportion of lung cancer cases now occur among ex-smokers. This trend is likely to continue and emphasizes the importance of elucidating further the aetiology of lung cancer. While a heritable component for lung cancer has long been recognized, progress in identifying susceptibility genes has been slow. To identify genetic factors that modify disease risk, we conducted a genome-wide association study of lung cancer. The initial phase constituted an analysis of 317,139 SNPs in 1,989 lung cancer cases and 2,625 controls from 6 central European countries. We identified a locus in chromosome region 15q25 that was strongly associated with lung cancer (p=9x10-10). This locus was replicated in 5 separate lung cancer studies comprising an additional 2,518 lung cancer cases and 4,752 controls