

($p=5 \times 10^{-20}$ overall), and it was found to account for 14% of lung cancer cases (Hung et al, Nature 2008). The association region contains several genes, including three that encode nicotinic acetylcholine receptor subunits (CHRNA5, CHRNA3, and CHRNB4).

This effect has been identified in two other studies (Thorgeirsson et al, Nature 2008; Amos et al, Nat Gen 2008), one of which concluded that the primary pathway is via addiction and a greater propensity to be exposed to tobacco products. The contracting arguments for the different interpretations will be discussed.

Finally, we have since extended our genome-wide study of lung cancer by including two further studies comprising an additional 750 cases and 800 controls. We subsequently replicated the top 31 independent findings in a further 4 studies comprising 3339 lung cancer cases and 6064 controls (total 5911 cases and 9416 controls). After pooling the genome-wide and replication phase results, two additional variants were strongly associated with lung cancer, suggesting new susceptibility loci.

10-IS

Statistical Approaches for Genome-Wide Association Studies

D. Balding¹

¹Department of Epidemiology and Public Health, Imperial College London, United Kingdom

Genome-wide association studies (GWAS) have in the past one or two years been highly successful at identifying genetic variants associated with complex diseases. Yet in a sense they have also been unsuccessful, because the variants discovered show only modest effect on disease risk, and explain little of the phenotypic variation that has been attributed to genetic effects in heritability studies. Diagnostic prediction of cases remains poor for most complex diseases, even in samples enriched for cases, so that for population samples it remains infeasible in practice. There are many plausible hypotheses concerning where the “missing” genetic information lies, including the possibility that structural variants and epigenetic features will be more highly predictive of disease than are SNPs. However there may be also be more predictive power from SNPs than is currently being realized. I will review the statistical methods that have been used to identify causal variants from GWAS results. These tend to be simple one-SNP-at-time analyses, perhaps with some adjustment for population structure and with cryptic kin removed. I will review some alternatives that could improve power, by analyzing multiple SNPs simultaneously and by more sophisticated adjustments for population structure and cryptic kinship. Gene-gene (epistatic) and gene-environment interactions have not yet been widely reported, in the latter case because the case-control study design does not facilitate this. I will review possibilities for incorporating them in future studies, particularly as attention moves towards GWAS of prospective cohorts that often are rich repositories of phenotype and environmental covariates. Epistatic interactions are beginning to be tackled via network and pathway-based approaches that I will also briefly review.

11-IS

Copy number variants: a common mechanism in complex diseases

X. Estivill¹

¹Genes and Disease Program, Centre for Genomic Regulation (CRG), Barcelona, Spain

Copy number variants (CNVs) may have important implications in complex diseases. Susceptibility to HIV-1 infection, glomerulonephritis associated to Lupus, Crohn's disease and psoriasis have been found to be associated with common CNVs involving the CCL3L1, FCGR3B and DEFB4 genes, respectively (Gonzalez et al., 2005; Aitman et al., 2006; Fellermann et al., 2006; Fanciulli et al., 2007; Hollox et al., 2008). Rare cases of common disorders (pancreatitis, Alzheimer's disease and Parkinson's disease) are also associated with rare CNVs (Le Marechal et al., 2006; Rovelet-Lecrux et al., 2006; Singleton et al., 2003). Since several CNVs are present in each individual and several CNVs contain genes with roles in response to the environment and adaptation, it is likely that regions containing CNVs with a wide range in copy number variation might have important roles in drug-response, susceptibility to infection, inflammation and cancer, among other common traits.

Functional sequences within CNVs might provide new insights on population differences in pharmacogenomics or disease predisposition. Information about the population distribution of CNV frequencies is crucial for association studies. HapMap DNA samples from Caucasian, Asian and African populations have been used to generate CNV data at the genome level (Redon et al., 2006; Wong et al., 2006). Resequencing of DNA samples is also showing that the human DNA contains many differences at the structural level (Korbel et al., 2007), defining what is perhaps an infinite amount of structural variability of the genome organization. The mechanisms by which CNVs could have functional consequences include

a direct gene dosage effect of the gene or genes embedded in the CNV, or a positional effect on genes proximal or distal to the CNV (Estivill and Armengol, 2007). A comprehensive catalogue of the spectrum of alleles at CNV regions of the human genome should provide the appropriate tools to explore the relationship between CNV loci and phenotypes, and will define specific structural changes that modify gene expression and function.

We have studied for CNVs, samples of the HGDP-CEPH Human Genome Diversity Cell Line Panel, a widely used resource for studies of human genetic variation. By comparing samples from twelve population groups we have identified over 200 genomic regions that vary in the genomic structure between groups. Most of these regions coincide with already known CNVs and segmental duplications. These regions contain genes with a role in immune response, adaptation to environment, and metabolic pathways. This data set of CNVs allows the most comprehensive characterization to date of human genetic variation with a potential role in disease susceptibility in different human population groups.

12-IS

Molecular cancer epidemiology and public health

R. Saracci¹

¹ IFC- National Research Council, Pisa, Italy

Molecular epidemiology, i.e. the use in epidemiological studies of techniques of molecular biology, has pervaded in the last two decades all areas of epidemiology, and of cancer epidemiology in particular, ranging from studies of aetiology and pathogenesis to investigations of biomarkers for early detection of minimal disease, potentially applicable to screening programmes. It might be pointed out that enthusiasm for the adoption of molecular epidemiology methods in cancer epidemiology has not been matched by translation into material prevention advances at the population level, as the largest scope for prevention still derives from knowledge on such exposures as tobacco, alcohol, energy balance or occupational and environmental carcinogens acquired by more traditional epidemiological approaches. This view however underrates the time lag often necessarily intervening between novel research approaches and practical applications. These have in fact started to emerge, notably in the area of cancers induced by infectious agents. Cervical cancer is the most relevant example: at the level of exposure biomarkers, the aetiological link with human human papilloma viruses has been established by advanced molecular epidemiology techniques, and at the level of early effect/outcome biomarkers, the detection of the virus in cervical cells has come into discussion as complement to the pap-test for screening purposes. In the area of susceptibility biomarkers expectations are high of major prevention advances, well beyond what is already possible and implemented for the important but comparatively uncommon cancers determined by single genes. Susceptibility biomarkers (in the broad sense) encompass genetic variants affecting not only cancer risk via a multiplicity of direct or mediating paths (e.g. through activation or inactivation of carcinogens) but also conditioning exposure to external agents (e.g. tobacco or alcohol) via behaviour-related effects. The working hypothesis of an “integrative epidemiology” approach is that a combined study of environmental exposures, susceptibility biomarkers with their proteomic and metabolic expressions as well as of biomarkers for molecular sub-typing of cancers may lead to a more and more refined stratification of risk by individual traits and by cancer type, paving the way to individualized prevention. While certainly fertile for research this working hypothesis raises two issues from a public health viewpoint. First it cannot be taken for granted, and needs to be critically examined from different quantitative angles, that the bulk of cases of a common cancer can be prevented by an individualized risk approach, a refined variant of the “high risk” strategy of prevention. Second environmental factors, often more difficult to measure with accuracy than molecular bodily components, may become secondary elements in the constellation of demonstrable determinants in the carcinogenic process: hence prevention may in practice become restricted to molecular determinants as modifiable by pharmacological means.

13-IS

Abstract not received

14-IS

Epidemiology of childhood leukemia: a transdisciplinary approach

J. Ross¹, K.J. Johnson¹, D.A. Largaespada¹, L.G. Spector¹, S. Ognjanovic¹

¹University of Minnesota Cancer Center, Minneapolis, MN, USA

Leukemia, including acute lymphoblastic (ALL) and acute myeloid (AML), is the most common childhood cancer in developed countries; over 3000 new cases are diagnosed each year in the United States. Incidence rates have been increasing significantly since the 1970s. Despite several studies, the etiology remains largely unknown. Steadily growing evidence suggests

many pediatric leukemias are initiated in utero. Epidemiological studies have identified high birth weight, absence of early life infections, pesticides, and low maternal folic acid/vitamin supplementation as risk factors for childhood leukemia. Some relevant genetic polymorphisms, particularly in the folate pathway, have also been implicated. Further, leukemia-associated translocations (e.g., MLL/AF4, MLL/AF9, TEL-AML1, AML1-ETO) in neonatal blood spots have been observed in children who later developed ALL and AML. Unfortunately, primary reliance on the case-control approach and little integration across disciplines is hampering progress in understanding relationships. For example, geneticists have developed MLL/AF4 and MLL/AF9 knock-in mice that develop hematopoietic neoplasms, and TEL-AML1 and AML1-ETO transgenic mice that do not, but none has been used to explore suspected environmental risk factors. We are integrating novel approaches across several disciplines to investigate the causes of childhood leukemia. One initiative includes investigating the potential role of environmental exposures (e.g. maternal folic acid, pesticides) in murine models through tracking disease outcome and gene methylation and expression in appropriate target cells in offspring. In an initial effort to understand the mechanism by which maternal folic acid (FA) impacts the risk of childhood leukemia, we conducted a pilot study using a murine model that examines gene expression in B-lineage cells of the offspring of mothers randomized to low (0.3 mg/kg), control (2.0 mg/kg) and high (8.0 mg/kg) FA diets prior to conception through weaning. Thirty three genes were identified as differentially expressed between offspring of mothers assigned to different dietary groups. Five genes that were up or down regulated by at least 2-fold in the low or high FA groups compared to the control group have been selected for validation by comparative RT-PCR: CDH8, GHR, E2F3, NEU3, and HOXD9. The second initiative involves establishing a pregnancy cohort to investigate the functional relationship between specific exposures and their biomarkers (e.g. cytokines, folate, growth factors) in cord blood and neonatal blood spots in relation to relevant risk factors (e.g., infection, maternal vitamin supplementation, high birth weight) and genetic polymorphisms of interest. Data generated from these initiatives will inform future studies of childhood leukemia and will also be applicable to other cancers.

15-IS

Evaluating cumulative evidence in genomic epidemiology

J. Ioannidis¹

¹ Department of Hygiene and Epidemiology, University of Ioannina, School of Medicine, Ioannina, Greece

The evaluation of cumulative evidence in human genome epidemiology presents several challenges. With the advent of massive-testing high-throughput platforms, the amount and complexity of data is unprecedented. Moreover, the concept of sufficient replication has evolved over time and more stringent criteria are now required to claim that an association with robust credibility has been detected. This may reduce the problem of having too many associations proposed, of which only a trifle stand the test of time. Stringent and systematic criteria are needed to enhance our understanding of the rapidly evolving status of the field of human genome epidemiological associations and to place these associations in context. One interim effort has been the generation of the Venice criteria for grading the credibility of genetic associations (Ioannidis et al., International Journal of Epidemiology 2008). The Venice criteria have three axes on which the evidence is appraised: amount of evidence, replication consistency, and protection from bias. Each of the three axes is rated from A to C. There is also a composite grading across the three axes. Associations that get an A in all three axes are considered to have "strong" epidemiological credibility; those that get at least one B but not C are considered to have "moderate" epidemiological credibility; and those that get at least one C are graded as having "weak" credibility. The advent of large-scale evidence in large scale collaborative studies and in synopses of associations in whole fields has allowed the application of the Venice criteria in various settings. I will present several examples of the application of the criteria in associations on cancer and other fields and will highlight some of the remaining challenges. Further refinement and validation of the performance of these criteria would be useful.

16-IS

Transdisciplinary Science Approaches for Molecular Epidemiology

R. Hiatt¹

¹Department of Epidemiology and Biostatistics and the Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, California

To make further significant advances in cancer molecular epidemiologic research, a transdisciplinary science approach is needed that integrates the study of the biologic nature of cancer and its clinical applications with behavioral and social influences on cancer. Molecular epidemiology could

be more integrated into a broader approach to solving cancer-related problems by seeking collaborative approaches that draw on other disciplines. This transdisciplinary approach may be a more effective way to apply the finding of molecular epidemiology for use in therapeutic, behavioral, and public health interventions to reduce the burden of cancer by taking into account the behavior and social determinants of cancer. Such an approach seeks to discover interactions between social, environmental, behavioral and biologic factors in cancer etiology. In this presentation I present 1) a brief historical perspective on the origins of transdisciplinary thinking in science, 2) a cross-disciplinary, multi-level framework for organizing cancer research along these lines, 3) some examples, and 4) a review some of the challenges in adopting this approach. The goal is to promote a more complete understanding of the causes of cancer that will lead to improved translation and implementation of the results of research.

17-IS

Abstract not received

18-IS

Future perspectives of molecular cancer epidemiology

M.T. Smith¹

¹ School of Public Health, University of California, Berkeley, USA

Our understanding of cancer is evolving rapidly through the application of "omic" technologies and the importance of epigenetic changes is becoming apparent. Genomics is being used to find genetic variants conferring susceptibility to cancer through genome-wide scans of >500K polymorphisms and high-throughput sequencing. High quality DNA is needed for this endeavor. Transcriptomics (gene expression profiling) is providing insights into carcinogenic mechanisms and is being used to subclassify cancers. Methods that preserve and isolate all forms of RNA are critical. Proteomics and metabolomics are newer technologies that are likely to impact our understanding of cancer etiology in the near future. Proteomics holds special promise in understanding the histone code that controls gene expression. Metabolomic profiles are highly dependent upon diet and standardized collection procedures are key, but novel insights could be obtained through this approach. The epigenome has been less explored to date, but new microarrays that measure hundreds of microRNAs or the CpG promoter methylation status of hundreds of genes are now available. Cell heterogeneity poses the same problem for epigenomics as it does for mutation analysis. The mutations or epigenetic changes of importance are usually present in only a small number of clonal cells. Fortunately, lab-on-a-chip technologies have now advanced to the point that single cell genetic analysis is feasible. However, intact cells or nuclei will be needed for these assays, again posing a problem to the epidemiologist. Finally, advances in nanotechnology should allow for multiplexed immunoassays of protein adducts, cytokines and antibodies in very small quantities of human material, such as pieces of dried blood spots and a few microliters of serum. At last, some good news for epidemiologists, as these nano-immunoassays can be combined with microfluidics on lab-on-a-chip microdevices so that they are applicable in large-scale cohort studies and in samples collected in low-resource environments such as developing countries. Lab-on-a-chip microdevices therefore hold promise in revolutionizing molecular cancer epidemiology in the not-too-distant future through the emerging field of exposure biology.

Proffered papers – Abstracts

1-PP

Serologic response to HPV and the risk of head and neck cancer

K.B. Ribeiro¹, J.E. Levi¹, M.P. Curado¹, J. Eluf-Neto¹, S. Koifman¹, V. Wunsch Filho¹, A. Menezes¹, A.W. Daudt¹, E. Matos¹, L. Fernandez¹, P. Boffetta¹, P. Brennan¹

¹IARC, Genetic Epidemiology Group, Lyon, France

Introduction: Head and neck squamous cell carcinomas (HNSCC) are frequent cancers in Latin America. Tobacco and alcohol are the main risk factors, although recent studies support a role for human papillomavirus (HPV) in a subgroup of HNSCC. **Objectives:** To evaluate the association between serologic response to HPV infection and HNSCC in a Latin-American population. **Methods:** A hospital-based case-control study was conducted in three countries from Latin America (Argentina, Brazil, and Cuba) to evaluate risk factors for HNSCC. Using multiplex serology, serum