(p=5x10-20 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. Balding[1]
[1]*Department 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 kins 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. Estivill[1]
[1]*Genes 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. Saracci[1]
[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. Ross[1], K.J. Johnson[1], D.A. Largaespada[1], L.G. Spector[1], S. Ognjanovic[1]
[1]*University 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