apoptosis and senescence. A DNA damage response was also observed in dysplastic nevi and in human skin xenografts, in which hyperplasia was induced by overexpression of growth factors. Both lung and experimentally-induced skin hyperplasias showed allelic imbalance at loci (common fragile sites) that are prone to formation of DNA DSBs when DNA replication is compromised. Further, in various model systems, oncogene overexpression led to stalling and collapse of DNA replication forks and generation of DNA DSBs.

Conclusion: From its earliest stages, cancer development is associated with DNA replication stress, which leads to DNA double-strand breaks, genomic instability and selective pressure for p53 mutations.

07 July 2008 10:15 - 12:15

SYMPOSIUM

Genetic epidemiology / Whole genome scan

245
Identifying new breast cancer genes through international consortia

B. Ponder[1]
[1]*Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Cambridge, United Kingdom*

Breast cancer, like other common cancers, tends to cluster in families. This clustering is predominantly genetic in origin. Most of the genetic effect is probably polygenic – that is, the result of the combined action of many genetic variants of small effect. We have recently completed a genome-wide scan for common genetic variants that contribute to susceptibility, and have identified 5 new predisposing loci. The possible future applications of this knowledge to breast cancer detection and prevention will be discussed.

246
Genome wide association studies for breast and prostate cancer susceptibility loci in the CGEMS initiative

G. Thomas[1], R. Hoover[1], D. Hunter[2], S. Chanock[1]
[1]*National Cancer Institute, Division of Cancer Epidemiology and Genetics, Rockville, USA;* [2] *Harvard School of Public Health, Department of Epidemiology, Boston, USA*

Three major advances have recently made possible genome-wide association studies (GWAS). First, the realization that locus-specific relative risks are usually small and thus require in order to be detected international consortia able to pool into joint analyses large numbers of patients and controls. Second the establishment of the first repertoire of the human genetic diversity by the HapMap project. Finally, the development of cost-effective techniques enabling the genotyping of a DNA on hundreds of thousand of loci in a single step. The judicious selection of half a million markers provides useful information on 80% of the estimated 7 million SNPs with minor allele frequencies higher than 0.05 present in a population of European origin. With the potential to explore a large fraction of the genome, the initial requirement of functional hypotheses to perform association studies becomes unnecessary. Recognizing the promises of this new approach, the NCI has launched the Cancer Genetic Markers of Susceptibility (CGEMS) initiative which aims at providing to the scientific community the results of GWAS for breast and prostate cancer.

The planned strategy involves for each tumor type three stages. In the first stage, about 1100 cases and 1100 controls nested in a prospective cohort are typed on 500,000 markers. The statistical analysis of genotypic data identifies a set of 25,000 SNPs with p-value for association lower than approximately 0.05 and with low pair-wise correlation among them (r^2< 0.8). In the second stage, these SNPs are typed on about 4,000 cancer cases and 4,000 controls. The subsequent analysis identifies about 150 chromosomal regions, each containing at least one SNP with a p-value smaller than 10^-3. SNPs in these regions are taken to the third stage which involves the genotyping of an additional set of 5000 cases and 5,000 controls. At stages 2 and 3, regions with convincing indication of being truly associated (low p-value in CGEMS and/or reported by others to be associated) are investigated with a dense set of markers. In order to investigate a total of over 40,000 DNAs, collaborations involving multiple European and American groups were established.

The first stage has been completed for both tumor types. The results have been posted on a public web site in October 2006 and April 2007 for the prostate study and in May 2007 for the breast study. They provide genotype counts and and p-values under various statistical models for over 500,000 SNPs. The second stage for the prostate study has been posted in March 2008 and provides follow-up data for 27,000 SNPs The second stage for the breast study will be released in the second half of 2008. Analysis of the second stage of the prostate study revealed 7 loci with p-values less than 2.5 10^-6 including three previously known loci and 4 new ones. In addition, 9 new loci showed suggestive association ($p < 2.5\ 10^{-5}$). The best significance ($p<7.4\ 10^{-13}$) was observed for MSMB, which encodes a primary constituent of semen and is proposed prostate cancer biomarker.

Combined with results from other published GWAS, fifteen loci have been convincingly associated with prostate cancer susceptibility. The risk-allele frequency may reach 0.85, indicating lack of natural counterselection. Most per allele odds ratios are small, typically 1.2. Importantly, to date, none of the functional polymorphisms have yet been unambiguously identified for any of the loci, and for many of them the relevant functional gene(s) remain(s) elusive. Knowledge of these loci provides unique original leads for further investigation in the mechanism of tumorigenesis. We will soon know if similar conclusions may be drawn from the CGEMS breast cancer study. The use of this new information to predict individual cancer risk for improved patient management requires validation by large studies, as little is known on the interaction of multiple risk-alleles co-existing in the same individual.

247
Issues and opportunities in family-based designs for young-onset cancers

C.R. Weinberg[1]
[1]*National Institute of Environmental Health Sciences, Biostatistics Branch, Research Triangle Park, USA*

Conditions with onset early in life, such as childhood cancers, can have complex etiologies, because both genetic and environmental factors can contribute to risk, and also because both the maternal and the fetal genomes can play a role. Classical case-control analysis exploring effects of inherited genetic variants is vulnerable to confounding by the maternal genotype, which can be causally related to the outcome (prenatal effects) and is certainly causally related to the genotype of the offspring. Consider a "triad" design, where one studies offspring with cancer together with their mothers and fathers. Under a simplifying assumption of genetic mating symmetry, a log-linear analysis (i.e. Poisson regression) can efficiently disentangle effects of inherited variants from prenatal effects mediated through the maternal genotype. Taking advantage of the mathematical distortions produced by over-transmissions of risk-related variant alleles to affected offspring, one can estimate effects of autosomal fetal genetic variants, with full robustness against bias due to population stratification or failure of Hardy-Weinberg equilibrium in the source population. One can also take advantage of the asymmetries induced between the maternal and paternal genotypes to identify maternally-mediated effects. Using this design, one can thus efficiently distinguish effects that work through the fetal genes from those that work through expression of maternal genes during gestation. No inheritance model needs to be assumed (e.g. dominant or recessive) and families with a missing parent or a deceased offspring can also be included. Multiplicative models can be assessed for examining joint effects with environmental factors. A limitation of the triad approach involves the fact that although main effects of genetic variants and multiplicative interactions with environmental effects can both be studied, one cannot assess main effects for exposures. A hybrid approach extends the design by including the parents of population-based controls, and provides greatly improved power and flexibility for inference related to joint effects.

248
Detecting new genes for tobacco-related cancers - genomewide association study of lung cancer

R.J. Hung[1], J. McKay[2], V. Gaborieau[2], P. Boffetta[3], C. Chu[4], J.K. Field[5], H.E. Krokan[6], P. Vineis[7], M. Larthrop[8], P. Brennan on behalf of the IARC/CNG lung GWA study[2]
[1]*Samuel Lunenfeld Research Institute, University of Toronto, Toronto, Canada;* [2] *International Agency for Research on Cancer, Genetic Epidemiology Group, Lyon, France;* [3] *International Agency for Research on Cancer, Lifestyle Environment and Cancer Group, Lyon, France;* [4] *Fred Hutchinson Cancer Research Center, Program in Epidemiology, Seattle, USA;* [5] *University of Liverpool Cancer Research Centre, Roy Castle Lung Cancer Research Programme, Liverpool, United Kingdom;* [6] *Norwegian University of Science and Technology, Department of Cancer Research and Molecular Medicine, Trondheim, Norway;* [7] *Imperial College, Division of Epidemiology Public Health and Primary Care, London, United Kingdom;* [8] *Centre National de Genotypage, Institut Genomique Commissariat à l'énergie Atomique, Evry, France*

Lung cancer is the most common cause of cancer death worldwide with over 1 million cases annually. While a heritable component for lung cancer