tion than has been identified by other groups. The proportion of germline TP53 splicing mutations in LFS families may be markedly under-reported. We have examined the biological consequences of these mutations to determine the functional properties of the mutant p53 and to confirm that they are causative mutations. Previous studies have indicated that five main component tumors are associated with LFS: sarcomas, brain and breast tumors, leukemia and adrenocortical carcinoma. We have analyzed the observed incidence of cancers in first- and second-degree relatives of probands in comparison with that expected, on the basis of United Kingdom national cancer rates. The observed pattern of cancers in germline TP53 carriers was highly significantly different from that expected across all ages. Carcinoma of the breast, brain and spinal cord tumors, bone and soft-tissue sarcomas, adrenocortical carcinoma, Wilms tumor and malignant phylloides tumors were all strongly associated with a germline TP53 mutation. Common carcinomas (such as lung, colorectal, bladder, prostate, cervix and ovarian) did not occur to excess. Germline TP53 mutations are associated with a subset of specific, mainly rare cancers.

Vidal, Marc

[53]

The gene number dilemma: direct evidence for at least 19,000 protein-encoding genes in *Caenorhabditis elegans* and implications for the human genome

David E. Hill, J. Reboul¹, P. Vaglio¹, N. Tzellas¹, C. Jackson², T. Moore², Y. Kohara³, J. Thierry-Mieg⁴, D. Thierry-Mieg⁴, J. Hitti⁵, L. Doucette-Stamm⁵, J. Hartley⁶, G. Temple⁶, M. Brasch⁶, D.E. Hill¹ & M. Vidal¹

¹Dana-Farber Cancer Institute and Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA

²Research Genetics, Huntsville, Alabama, USA

³Center for Genetic Resource Information, National Institute of Genetics, Shizuoka, Japan

⁴NCBI, NIH, Bethesda, Maryland, USA⁵Genome Therapeutics Corp., Waltham, Massachusetts, USA

⁶Life Technologies Inc., Rockville, Maryland, USA

Estimates of the number of genes for organisms with a wide range of biological complexity will emerge from genome sequencing projects. This information will be crucial to correlate gene number with biological complexity. Computer-based annotations of the genome sequences of *Saccharomyces cerevisiae*, *Caenorhabditis elegans* and *Drosophila melanogaster* predict approximately 6,000, 19,000 and 13,600 genes, respectively. It is seemingly paradoxical that *C. elegans* contains 50% more genes than *D. melanogaster* and that the latter only contains twice as many genes as *S. cerevisiae*. However these numbers should be considered with caution until gene predictions are verified using experimental approaches. To experimentally verify gene predictions, we PCR amplified predicted open reading frames from a *C. elegans* cDNA library, cloned them and generated an ORF sequence tag (OST) for each of them. In a random sample ($n = 1,200$) of the ~10,000 genes predicted from the entire *C. elegans* genome but experimentally unverified so far, 90% of these predicted genes could be unambiguously verified by OST sequencing. When added to the list of ~9,500 genes previously verified by individual laboratories and/or by EST sequencing, our data confirm the existence of at least 19,000 protein-encoding genes in *C. elegans*. We suggest that ORFeome cloning coupled with OST analysis could be used to verify gene number predictions for other organisms. Furthermore, the demonstration and identification of 10,000 novel genes provides a resource for comparative genomics and gene expression studies. We also discuss our findings in the context of the current debate on the human gene number.

Villuendas, Raquel

[54]

Expression microarrays in the prediction of treatment resistance: The CNIO program

Raquel Villuendas, Jose Palacios, Manolo Morente, Orlando Dominguez, Javier Benitez & M.A. Piris

Centro Nacional de Investigaciones Oncologicas, 28220 Madrid, Spain

The Molecular Pathology Program of the Spanish National Cancer Research Center (CNIO) performs molecular analyses of human tumors, allowing prediction of the treatment response to commonly used drugs and at the same time permitting the identification of relevant cancer genes. Biological therapy has so far been based on the study of unique genes or proteins, such as the administration of tamoxifen to endoplasmic reticulum positive tumors, Herceptin for cancer arising from overexpression of *c-erb-B2*, Cetuximab for cancers associated with overexpressed epidermal growth factor, Rituximab in CD20-positive lymphomas or RAR in progressive multifocal leukoencephalopathy. Although such biologically oriented treatment is making significant progress against different tumor types, it ignores the fact that cancer is a social and multigenic disease. The more common forms of cancer are the result of complex genetic alterations that allow the tumor cells to develop multiple capacities, such as tissue invasion and metastasis, angiogenesis promotion, unlimited replication, apoptosis escape, production of autonomous growth signals, insensitivity to cell cycle control signals and generation of genomic instability. One of the goals of the Molecular Pathology Program is to contribute to an individualized therapeutic strategy, according to the multiple and varied characteristics of the disease. The CNIO is analyzing resistance to commonly used drugs, through molecular study of targets and messenger RNA expression analysis using the microarray technology. The CNIO is also sponsoring the creation of a network of tumor banks in leading hospitals, which will allow massive application of these molecular techniques to large series of patients, diagnosed and treated using standardized protocols. Current projects include the following: (1) A study of the molecular mechanisms of therapeutic failure in cutaneous T-cell lymphoma and malignant melanoma treated with interferon- α , PUVA (a combination of psoralen and long-wave ultraviolet radiation) or both. The objective of this project is to identify the genes involved in interferon- α resistance in myelofibrosis and malignant melanoma. (2) Development of a complementary DNA biochip for the analysis of molecular changes associated with treatment resistance in breast cancer. The CNIO is open to establishing collaborations with other oncology groups to perform integrated analyses of the molecular factors predictive of treatment resistance.

Vinals, Carla

[55]

Discovery of new potential antigens for cancer immunotherapy using "custom tissue" screening of a database of expressed sequence tags

Carla Vinals, Swann Gaulis, Jean-Pol Cassart & Thierry Coche

SmithKline Beecham Biologicals, 89 rue de l'Institut, 1330 Rixensart, Belgium

Immunotherapeutic approaches to fighting cancer are based on the principle of mounting an immune response (either an antibody response or a cytotoxic or T-helper response) against a self-antigen expressed by the tumor cells. To reduce potential autoimmunity side effects, the antigens used should be as tumor-specific as possible. A complementary approach to experimental tumor antigen discovery is to screen the human genome *in silico*, particularly the databases of expressed sequence tags, in search of tumor-specific and tumor-associated antigens. We describe a new method of 'customized' screening of a database of expressed

Abstracts: Session III

sequence tags that allows efficient selection of new potential tumor-associated genes. The resulting ranked list of candidates is analyzed in terms of its content in known genes, used as controls, and the new candidates are further tested for selective expression in normal and tumor tissues by real-time polymerase chain reaction with reverse transcription. The most highly tumor-specific candidates are then pursued to full-length cloning. We present the discovery of a series of new genes that are upregulated in colon cancer.

Visvanathan, Jaya

[56]

Genomic aberrations in gastric adenocarcinoma

Jaya Visvanathan¹, Pulivarthi H. Rao², Christina Rudduck³, Wai Keong Wong³, Chow Yin Wong³, Jean Ho³, Choon Wei Wee¹, Oi Lian Kon¹ & Dennis T.H. Lim¹

¹National Cancer Centre, Singapore

²Baylor College of Medicine, Houston, Texas, USA

³Singapore General Hospital, Singapore

Gastric adenocarcinoma (GA) presents a challenge in understanding the complex interactions of environmental factors with genotype in carcinogenesis. The National Cancer Centre has taken a multifaceted approach to this problem with comparative genomic hybridization of freshly resected GAs. Data from nine cases of intestinal and five cases of diffuse histological types showed the common aberrations to be loss of 4pq (four cases) and gains of 17pq (five cases) and 20pq (four cases). We also documented amplifications in 22q (six cases), 20q (five cases) and 1p (two cases). High-level gains were present in 20q11.2q12 (three cases), 10q24.3q25.1 (one case) and 12q14 (one case). Two cases of diffuse GA showed no change. High-level amplifications in 10q24.3q25.1 and 12q14 have not been reported previously. High-level amplification of 12q14 could be associated with overexpression of *MDM2*, *CDK4* and *RAP1B*. Three cytochrome P450 hydroxylases map to 10q24. Loss of 4pq is a common abnormality in cases of GA; its biological significance may be related to loss of function of caspase 3, caspase 6 and PTPN13, which is frequently deleted in liver and ovarian cancers. High-level gain of 20q or 20pq was another striking finding of this preliminary study. Amplified genes of probable importance from this genomic region are those for DNA methyltransferase 3B, E2F transcription factor 1, EHT and matrix metalloproteinase 5. Gain of 17pq or 17q has been frequently reported in cases of GA, and candidate genes of relevance to gastric carcinogenesis are *HER-2/neu*, and those encoding survivin and gastrin.

Volik, Stanislav

[57]

A 20q13.2 cancer amplicon: novel sequence analysis tools

Stanislav Volik, Joe Gray, David Kowbel & Colin Collins

University of California San Francisco Cancer Center, San Francisco, California

Progress in mammalian genome sequencing demands development of efficient, high-throughput methods for functional annotation of megabase-scale regions. We report the first complete sequence of a tumor amplicon that is associated with aggressive tumor behavior and occurs in a wide range of tumors. The overwhelming amount of data generated for this region by traditional analysis tools (such as GenScan, BLAST and RepeatMasker) required development of efficient annotation methods. For a comprehensive biological annotation of the amplicon we developed Genome Cryptographer. This tool integrates information on distribution and type of repetitive elements, results of database searches, comparative sequence data, expression data, array comparative genomic hybridization copy

number data and information on CpG dinucleotide distribution. Genome Cryptographer revealed an uneven distribution of genes and repetitive elements in the 1-megabase region. The proximal amplicon peak as defined by array comparative genomic hybridization is composed of 60% repetitive sequence and encodes one protein-coding gene, *ZNF217*. A 14-kb duplication was identified close to the 5' end of *ZNF217*; it is 97% homologous to a 14-kb element on chromosome 22q13; contains one of three identified CpG islands, a conserved putative RNA gene and a single repetitive element; and is present at the murine *ZNF217* locus. The distal amplicon peak contains two genes, *CYP24* and *PFDN4*. Breakpoints for both amplicon peaks are located in regions of local repeat maxima. We have also developed software to create digital expression profiles of multi-megabase regions. This allowed us to confirm that *ZNF217* has the highest relative expression in tumors and to identify three additional putative cancer-related genes.

Watson, Mark

[58]

Expression profiling of ductal carcinoma *in situ* by laser capture microdissection and high-density oligonucleotide arrays

Veronica Luzzi, Victoria Holtschlag & Mark Watson

Department of Pathology and Immunology and Alvin J. Siteman Cancer Center, Washington University School of Medicine, St. Louis, Missouri 63110, USA

Gene expression profiling with nucleic acid arrays is a powerful method for the molecular classification of human neoplasms. Laser capture microdissection is an equally useful technique to isolate defined cell populations selectively from heterogeneous histological tissue sections. We demonstrate how a modest use of laser capture microdissection is sufficient to isolate nanogram quantities of high-quality RNA. Together with several internal standards and microcapillary electrophoresis of input RNA, we used two rounds of linear molecular amplification to generate sufficient quantities of labeled target for hybridization to high-density oligonucleotide expression arrays. Our results demonstrate that the technique is reproducible, generates only modest biasing of the original transcript population and offers sensitivity comparable to that achieved using standard methodology. Using this approach, we have compared the expression profiles of nonmalignant human breast epithelium and adjacent ductal carcinoma *in situ* lesions from breast cancer patients. Several genes, including those previously implicated in human breast cancer progression, demonstrate differential expression among the microdissected cell populations. This study demonstrates the feasibility of applying laser capture microdissection and nucleic acid expression array technology to histologically complex clinical specimens to classify human breast cancer molecularly and gain a better understanding of its molecular pathobiology.

Watts, George

[59]

Microarray analysis of a human multiple myeloma cell line selected with melphalan or doxorubicin identifies a set of commonly affected genes

George Watts, Denise Tye, Mary Gleason-Guzman & Bernard Futscher

Arizona Cancer Center, Tucson, Arizona 85724, USA

Multiple myeloma is a monoclonal plasma cell neoplasm that is frequently treated with melphalan and doxorubicin. The development of resistance to drugs such as melphalan and doxorubicin poses a hurdle to their successful clinical use.