LECTURE 6

Jeremy Squire

Future directions: the human genome project and automation in molecular diagnostics

In this lecture we will focus on the importance of "cancer genomics" or the application of the techniques and resources of the human genome project (HGP) in the future molecular pathology of cancer. We will review the goals of the human genome project, the issues that impact on the biomedical and lay public community because of the work that has come out of the HGP. We will then focus on why the technologies developed in the HGP are so important to cancer diagnostics in the future. We will examine microarray analysis of tumors and laser capture microdissection techniques, as applied to heterogeneous tumors such as prostate cancer. All the approaches of doing gene chip analysis will be reviewed together with how Bioinformatics techniques such as 'cluster analysis' of gene expression can help identify new tumor subtypes. The future direction of the genome project such as proteomics and pharmacogenomics will be discussed. Students are encouraged to submit questions for discussion at the end of this lecture.

Slide 1 TITLE: Structure of DNA
TEXT and IMAGE

- human DNA structure elaborated by James Watson (shown) and Francis Crick
- Dr. Watson and others were inspired to start the Human Genome Project as a large scale international cooperation so that all of life science could share data and have a common resource for the future
- Many doubters that this would (a) work and (b) be useful
- In now largely completed…doubters remain but most researchers cannot understand how it would have been possible to do their current work without it!

Slide 2 TITLE: Goals of the Human Genome Project (HGP)
TEXT

- identify all the approximately 30,000 genes in human DNA
- determine the sequences of the 3 billion chemical bases that make up human DNA
- Make electronic databases and develop tools for data analysis
- address the ethical, legal, and social issues that may arise from the project

Slide 3 TITLE: Dec 1, 1999 Chromosome 22 Completely Sequenced
IMAGE
- progress was surprisingly fast.
- Helped perhaps by good economic times and improvements in computers and the internet

Slide 4 TITLE: June 25, 2000 First Survey of Entire Human Genome Completed
IMAGE
•HGP represents a cooperation between the government funded laboratories represented by Dr Francis Collins (right) and a private company called Celera (this word also means "Speed")
•The relationship between these public and private sector initiatives has often been one of suspicion

Slide 5 Science/Nature 2001
IMAGE
The publishing of the human genome sequence was simultaneously performed in Science and Nature by each of the two groups

Slide 6 TITLE: *Celera* Private Database of Human Genome Sequence
IMAGE OF CELERA WEB SITE

- Some feel that competition helped to speed up progress
- Others feel Celera use the public information but do not share the benefits they derive with the international community
- There is clearly financial rewards in addition to scientific rewards to be derived from the fruits of the HGP

•

Slide 7 TITLE: "Pharmacogenomics" appeared
IMAGE OF BRISTOL MYERS SQUIB WEB SITE

- Text from site: " The future of medicine has never held as much promise as it does today. We are standing at the brink of a new golden age of medical research. The sequencing and subsequent analysis of the human genome will revolutionize the practice of medicine. They will change the way we think about health and disease, and they will alter our expectations of the whole heath care system… ."

Slide 8 Cloning Dolly the Sheep
IMAGE

- advances in gene handling techniques helped to clone Dolly the Sheep
- useful for understanding development al biology, early growth and cancer
- but the public and media remain suspicious

Slide 9 Cloning Noah's Ark
IMAGE
-cloning could help preserve endangered species and ensure we do not lose our vital genetic bio-resources from the past

Slide 10 Human Cloning Now a Possibility
SCHEMATIC
Using the same methods applied to Dolly it is theoretically possible to clone humans

Slide 11 TITLE: Telomere erosion and aging
IMAGE/SCHEMATIC
Media quote -"Turning back the strands of time-scientists have found a factor that controls whether a cell dies or thrives"
-Part of telomere research in the private and public laboratories is aimed at understanding the aging process.
-Telomeres (chromosome ends) are very important parts of chromosomes and in cancer cells it is well established that when the stabilizing sequences at the end of a chromosome are lost a cell becomes genomically unstable and can develop some of the chromosomal aberrations associated with cancer.

Slide 12 Title: Benefits of the human genome project in laboratory medicine?
TEXT

- Automation for advanced molecular diagnostics
- Increases in computer speed
- Internet and international networks
- Development of neural networks and Artificial Intelligence
- Specialized digital microscopy

Slide 13 Title: 3-D reconstruction of gene expression in different regions of the prostate
IMAGE
Specialized in situ hybridization techniques used to perform a three dimensional reconstruction of gene expression in a whole prostate.

Today I wanted to return to the question of prostate cancer (CaP) and illustrate how we are using genomic technologies to address the clinical problem of determining prognosis.

Slide 14 Title: What is the pathologist looking for?
IMAGE OF HISTOLOGY
-we can see morphologic features the pathologist subjectively associates with cancer e.g. shape of cells , prominent nucleoli, general disorganization, mitotic figures.
DNA, RNA and protein are key molecular for future dignostic and array methods allow some objectivity to be used in assessing tumors

Slide 15 Title: Identification of differentially expressed genes
SCHEMATIC
-in this slide the letters represent different genes
-in this model we imagine that gene “D” is exp0ressed in normal prostate cells and gene”E” is expressed in CaP
-how can we apply genomic methods to identify genes D and E and use them to diagnose cancer and determine how aggressive it is?

Slide 16 Tile : Microarray analysis on glass slide
IMAGE SPOT ARRAY A GLASS SLIDE MICROARRAY

DNA microarray technology, in which thousands of different DNA sequences are arrayed in a defined matrix on a glass or silicon support, is rapidly becoming the method of choice for the detection of genomic gene expression* (Schena, M.,1996).
The applications of microarrays extend beyond the boundaries of basic biology into diagnostics, environmental monitoring and pharmacology. For example, microarrays are now being exploited by pharmaceutical companies to expedite the drug discovery process, by using gene expression as an indicator of efficacy and toxicity. Environmental biologists will soon be able to use microarrays to monitor the fate and transport of biological material in the environment and to monitor the toxicity of environmental pollutants on human cells. The diagnostic industry will use microarrays that contain genetic material from all known human pathogens for rapid identification of disease causing organism(s) in patients

- PNAS vol. 93, Issue 20, 10614-10619, October 1, (1996) Parallel human genome analysis: Microarray-based expression monitoring of 1000 genes Mark Schena, Dari Shalon, Renu Heller, Andrew Chai, Patrick O. Brown, and Ronald W. Davis

-this glass slide contains 1700 expressed-tagged sequences (ESTs) or genes
it is made at the microarray centre at our institute. I will first review how this slide was made and then return to how it is analyzed.

Slide 17 TITLE: http://www.uhnres.utoronto.ca/services/microarray/index.html
IMAGE

Large-scale genomic analyses has been termed "Functional Genomics". Functional genomics involves the integration of molecular biology with robotics and high throughput approaches to molecular analysis. The aims of functional genomics are to derive as much information about as many genes as possible, as fast as possible. The analysis of global gene expression patterns is a key new area of functional genomics because the development and differentiation of a cell or organism as well as its progression to the disease state, is determined in large part, by its profile of gene expression.

Among the important questions in this area are: in which tissue are given genes expressed? How is the expression of each gene affected by changes in the surrounding neighbouring cells ? What set of genes is

expressed during the development of an organism? How does gene expression change during development and differentiation? How does the misregulated expression of one gene affect others? Are there differences in gene expression profile that can be correlated with, or even contribute to, disease progression.

Slide 18 Tile : Printing arrays 1
CLOSE UP IMAGE OF PRINTING
-DNA is spotted onto the slide in duplicate on glass slides at high density.
Human 19K microarrays

* At 170 micron centre-to-centre spot resolution, ie. spot distance (5 spots/0.85millimetre)
* 2 slides each with 32 grids of 600 spots
* Each grid or sub array contains 24 columns and 25 rows

Slide 17 Tile : Printing arrays 2
OVERVIEW OF IMAGE OF PRINTING
Slides cost about $150 (US) for cost recovery for the 19K slide
The center is 'not for profit '(but always has a backlog for supplies of arrays!)

Slide 19 Title : Results of CaP
IMAGE OF ARRAY EXPERIMENT
Each pair of spots = a gene or EST
Cy5 =red and Cy3 =green signal.
Slide shows gene D is overexpressed in normal prostate
gene E is overexpressed in prostate cancer
yellow signal means a gene is expressed in both normal and cancerous cells
No signal means gene is not expressed in either cell type

Slide 20 Title : none
IMAGE OF AFFYMETRIX ARRAY EXPERIMENT
Probe arrays are manufactured by Affymetrix's, light-directed chemical synthesis process, which combines solid-phase chemical synthesis with photolithographic fabrication techniques employed in the semiconductor industry. Using a series of photolithographic masks to define chip exposure sites, followed by specific chemical synthesis steps, the process constructs high-density arrays of oligonucleotides, with each probe in a predefined position in the array.

Slide 21 title : Affymetrix oligonucleotide probe arrays
SCHEMATIC IMAGE OF AFFYMETRIX OLIGONUCLEOTIDE PROBE ARRAYS

Multiple probe arrays are synthesized simultaneously on a large glass wafer. This parallel process enhances reproducibility and helps achieve economies of scale. The wafers are then diced, and individual probe arrays are packaged in injection-molded plastic cartridges, which protect them from the environment and serve as chambers for hybridization. The color read out is scanned as are the 'spot arrays' described earlier

Slide 22 Scanners
IMAGE OF A SCANNER

ScanArray 4000	GSI Lumonics	Packard BioScience
DNAscope	Gene Focus	Gene Focus
GenePix 4000A	Axon Instruments, Inc.	Axon Instruments, Inc.
ChipReader	Virteck Vision Inc.	Virteck Vision Inc.

Usually laser light is used to measure and determine the ratio of Cy3: Cy5 and these data are stored in a database in which gene identity is also sourced.

Slide 23 Title: cDNA interrogates Microarray
IMAGE OF SPOT ARRAY
-in this experiment the red colour is from RNA extracted from metastatic melanoma and the green colour from the same primary tumor. Thus red spots are genes whose expression is elevated in metastasis relative to the primary tumor.

Slide 24 Title: Alphagene Microarray
IMAGE OF SPOT ARRAYFROM COMPANY ALPHAGENE
-in this experiment cDNA from Alzheimers brain (green) is compared to age matched control (red)
-an increasing concern of those doing microarrays is that the tissue heterogeneity in experiments may make considerable 'noise' and irrelevant signals making the interpretation difficult if not meaningless

Slide 25 Title: Progression of early lesions of prostate cancer
IMAGE HISTOLOGY OF CaP and PROSTATIC INTRAEPIITHELIAL NEOPLASIA (PIN) LESIONS
-we are searching for patterns of gene expression associated with the Gleason score (histological measure of aggressiveness) in CaP biopsies.
-gene expression may be diagnostic or in the case of genes that continue to be expressed in advanced cancer prognostic
-one concern is the marked cellular heterogeneity and we will next examine the way to address this using microdissection.
Under the microscope, tissues are heterogeneous complicated structures with hundreds of different cell types locked in morphologic units exhibiting strong adhesive interactions with adjacent cells, connective stroma, blood vessels, glandular and muscle components, adipose cells, and inflammatory or immune cells. In normal or developing organs, specific cells express different genes and undergo complex molecular changes both in response to internal control signals, signals from adjacent cells, and humoral stimuli. In disease pathologies, the diseased cells of interest, such as precancerous cells or invading groups of cancer cells, are surrounded by these heterogeneous tissue elements. Cell types undergoing similar molecular changes, such as those thought to be most definitive of the disease progression, may constitute less than 5% of the volume of the tissue biopsy sample.

Slide 26 Title : Laser capture Microdissection (LCM)
IMAGE
Laser Capture Microdissection is a method for procuring pure cells from specific microscopic regions of tissue sections. Microdissection is essential to apply molecular analysis methods to study evolving disease lesions in actual tissue. The microdissected cDNA will approximate the true pattern of gene expression of the pure cell subpopulations in their actual tissue. Because limited amounts of RNA and DNA are obtained it is necessary to amplify the genetic material.

Slide 27 Title : none
IMAGE OF SYSTEM AT OCI
-this is Dr Ming Tsao (one of our Patholgists) performing LCM

Slide 28 Title : Laser capture Microdissection
SCHEMATIC
-a laser source not that different to this pointer pulses through the microscope and melts a transparent polymer sitting on top of the tissue section
- the plastic melts and migrates into the tissue so that the cells can be lifted off and placed directly into an extraction buffer

Slide 29 Title : Principles of LCM
SCHEMTIC
Slide on we can how the LCM system works
See how cleanly the LCM removes a PIN lesion for molecular genetic analysis

Slide 30 Title : none
IMAGE OF PIN DISSECTION BY LCM
-similarly we can see how even a small area of normal tissue is left behind.
-of course for tumors of the prostate tumor – stroma interaction is very important so in this way it is possible to examine pure populations of each to study this dynamic association at the molecular level

Slide 31 Title: how each individual patient sample is processed
SCHEMATIC
-at our Institute the samples go for routine pathology but the program for molecular genetic analysis of CaP that I direct will have dissected material as will another group looking at differential radiosensitivity in CaP.
-Using networked research databases the data from each branch of this study can be correlated.

Slide 32 Title: none
SCHEMATIC OF DOP PCR

In this slide the way we amplify the dissected DNA is depicted. Using 'global amplification methods' based on short random oligonucleotides and PCR the amount of DNA obtained from a biopsy can be increased by 10 to 100 X without distorting the represententation (ie biasing the data). Similar methods calledT7 mRNA amplification are being used to amplify RNA for analysis

Slide 33 Title : CGH Analysis after Laser capture Microdissection (LCM)
TABLE
Comparison between PIN (blue) and carcinoma (red) from the same specimen shows some similarities in chromosomal imbalance. Notice how 8q is gained in both and parts of 16 and 17 are lost.

Slide 34 TITLE: - Ovarian cancer SKY and CGG analysis results combined
This slide was shown yesterday. The regions 3p13-p21, 3q11-q12, 3q22-q23, and 8q23-q24 and parts of 6, 11 and 17 had elevated frequencies of structural and numerical aberrations by SKY and CGH. We can now look at the same tumors by expression microarray analysis. Notice how 6p has some changes but not especially many alterations.

Slide 35 TITLE Analysis of microarray data
SCHEMATIC AND TEXT

-jargon (terms used to describe) in cluster analysis: hierarchical clustering, K-means algorithm, tree-structured vector quantization, principal component analysis, self-organizing map, multi-dimensional scaling, block clustering.

Intensity for each feature of the array is captured and a single raw expression level for each gene is derived from the ratios representing each gene using an analytical mean algorithm. Intensity values are then re-scaled such that overall intensity for both Cy3 and Cy 5 are equivalent..

Normalization
Normalization prior to clustering allows to compare genes with different absolute intensities but having the same shape or signature across a set of experiments.

Slide 36 Title none
EISEN HIERARCHICAL CLUSTERING USING OVARIAN CANCER DATASET
-identifies groups of tumors
-shows that different samples from different sites of the same tumor cluster together
-shows that clustering stratifies by degree of differentiation rather than by subtype based on histology
Dendrogram derived from Cluster Analysis. This algorithm successively joins gene expression profiles to form a dendrogram based on their pairwise similarities. The poorly differentiated ovarian tumors (yellow/green highlights) in general cluster to the right

Slide 37 Title Expected genes emerge from low-density screen
EISEN HIERARCHICAL CLUSTERING CLOSE UP
-identifies expected cluster of metallothionien and retinoid signaling pathway
-both pathways implicated in ovarian cancer. Metallothionien associated with cis platinum resistance

Slide 38 Title Utility of Genetic Analysis
TEXT
ADVANTAGES
-identification of novel genes
-increased understanding of tumorigenesis
-Diagnostic information
-Prognostic information
-Potential for novel therapeutic approaches
DISADVATAGES
-expensive
-potential for error

Slide 38 Title: J.C. Boileau Grant
IMAGES
- my Position at OCI is due to a bequest by a very famous Canadian Pathologist called J.C. Boileau Grant. His book 'Grant's Anatomy' is well known to medical students. The intention in having a Chair in his name is to foster the idea that the next phase of anatomy will be based on the Genome Project –the anatomy of the genome to understand disease processes from fundamental principles.
- I would like to thank….list many people here!