Journal of Molecular Diagnostics, Vol. 2, No. 2, May 2000

Special Article

Molecular Profiling of Clinical Tissue Specimens

Feasibility and Applications

Michael R. Emmert-Buck,*† Robert L. Strausberg,† David B. Krizman,†‡ M. Fatima Bonaldo,§ Robert F. Bonner,¶ David G. Bostwick,‖ Monica R. Brown,†‡ Kenneth H. Buetow,** Rodrigo F. Chuaqui,* Kristina A. Cole,* Paul H. Duray,‡ Chad R. Englert,* John W. Gillespie,† Susan Greenhut,† Lynette Grouse,† LaDeana W. Hillier,†† Kenneth S. Katz,† Richard D. Klausner,‡‡ Vladimir Kuznetzov,¶ Alex E. Lash,§§ Greg Lennon,¶¶ W. Marston Linehan,‖‖ Lance A. Liotta,‡ Marco A. Marra,*** Peter J. Munson,††† David K. Ornstein,*‖‖ Vinay V. Prabhu,††† Christa Prange,¶¶ Gregory D. Schuler,§§ Marcelo Bento Soares,§ Carolyn M. Tolstoshev,§§ Cathy D. Vocke,‖‖ and Robert H. Waterston††

From the Pathogenetics Unit, Laboratory of Pathology, National Cancer Institute, Bethesda, Maryland; the Cancer Genome Anatomy Project,† Office of the Director, National Cancer Institute, Bethesda, Maryland; the Laboratory of Pathology,‡ National Cancer Institute, Bethesda, Maryland; the Departments of Pediatrics and Physiology and Biophysics,§ University of Iowa, Iowa City, Iowa; the Laboratory of Integrative and Medical Biophysics,¶ National Institute of Child Health and Disease, Bethesda, Maryland; Bostwick Laboratories,‖ Richmond, Virginia; the Laboratory of Population Genetics,** National Cancer Institute, Bethesda, Maryland; the Genome Sequencing Center,†† Washington University, St. Louis, Missouri; the Office of the Director,‡‡ National Cancer Institute, Bethesda, Maryland; the National Center for Biotechnology Information,§§ National Library of Medicine, Bethesda, Maryland; the Integrated Molecular Analysis of Genomes and their Expression Consortium,¶¶ Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, California; the Urologic Oncology Branch,‖‖ National Cancer Institute, Bethesda, Maryland; the Genome Sequence Centre,*** British Columbia Cancer Research Centre, Vancouver, British Columbia, Canada; and the Mathematical and Statistical Computing Laboratory,††† Center for Information Technology, National Institutes of Health, Bethesda, Maryland*

The relationship between gene expression profiles and cellular behavior in humans is largely unknown. Expression patterns of individual cell types have yet to be precisely measured, and, at present, we know or can predict the function of a relatively small percentage of genes. However, biomedical research is in the midst of an informational and technological revolution with the potential to increase dramatically our understanding of how expression modulates cellular phenotype and response to the environment. The entire sequence of the human genome will be known by the year 2003 or earlier.[1,2] In concert, the pace of efforts to complete identification and full-length cDNA sequencing of all genes has accelerated, and these goals will be attained within the next few years.[3–7] Accompanying the expanding base of genetic information are several new technologies capable of global gene expression measurements.[8–16] Taken together, the expanding genetic database and developing expression technologies are leading to an exciting new paradigm in biomedical research known as molecular profiling.

Molecular Profiling

Molecular profiling uses measurement of global expression patterns toward identification of the individual genes and collections of genes that mediate particular aspects of cellular physiology. The method is primarily hypothesis-generating, emphasizing new discoveries and creation of novel postulates based on analysis of expression data sets.[17–21] Much like an astronomer with a new telescope, investigators use molecular profiling to explore and observe, with the goal of producing insights that would not readily be predicted based on the currently available body of knowledge. In humans, molecular profiling efforts hold great promise to advance our understanding and treatment of diseases. Measurement of expression patterns of normal and affected cell populations

Reprinted from Am J Pathol 156:1109–1115

The work of C. P. was supported by the U.S. Department of Energy under contract W-7405-Eng-43.

Primary contributors to this article were M. E.-B., R. S., and D. K.

Accepted for publication February 18, 2000.

Address reprint requests to Dr. Michael R. Emmert-Buck, National Cancer Institute Laboratory of Pathology, National Institutes of Health, Building 10, Room 2A33, 9000 Rockville Pike, Bethesda, MD 20892. E-mail: mbuck@helix.nih.gov.

likely will identify specific sets of genes that are disregulated. Moreover, the availability of full-length mRNA coding sequences will allow prediction of function based on computer modeling algorithms, promoting a more fundamental understanding of the disease process as well as new diagnostic and therapeutic targets for clinical intervention.

There are several experimental systems available for molecular profiling, including human cells *in vitro* and animal models that mimic human pathologies. Each of these approaches has proven to be valuable in past studies and hold excellent potential to produce important discoveries in expression profiling studies. However, in parallel, it is critical that patients be studied directly. Molecular profiles of human cells *in vivo*, as they exist in patients, may lead to unique insights that are not readily evident in laboratory-based investigations, and are the gold standard against which model systems should be compared.[22] Certainly, the ability to peer directly into the molecular anatomy of normal and diseased human cells in their complex tissue milieu is a particularly exciting application of molecular profiling.

However, there are significant technical challenges associated with expression profiling of clinical samples and substantive obstacles that must be addressed. For example, investigators are confronted with the difficulty of procuring specific microscopic cell foci from heterogeneous tissues. Moreover, high-throughput expression studies require recovery of a diverse and complex transcriptome, not a trivial task when using small numbers of cells as template. Although it is exciting in concept, to date there are few experimental data available that support the possibility of this approach. Therefore, a study was designed to answer two key questions. Is molecular profiling of histopathologically defined cell populations from clinical tissue specimens feasible using available technologies and methodologies? If so, what are near-term and long-term applications of global gene expression data sets from patient samples?

Feasibility: Prostate Cancer Study

Molecular profiling studies generate large data sets for analysis, representing a significant challenge for investigators. Moreover, clinical studies ideally include multiple samples, such that molecular findings can be assessed for their frequency among patients and/or correlated with particular features of a disease. Thus, integration of clinical information, histopathology, developing technologies and laboratory methods, and bioinformatics algorithms is essential for profiling efforts. The present study was performed as part of the Cancer Genome Anatomy Project (CGAP) of the National Cancer Institute (NCI).[23–25] CGAP is an interdisciplinary program that aims to establish the information and technological tools needed to decipher the molecular anatomy of cancer cells. All data from the project are immediately made available to the public and can be used without restriction.

The feasibility of molecular profiling of microdissected cell populations was assessed using cDNA library sequencing as an initial gene expression platform and prostate cancer progression as a disease for study. Sample collection, microdissection, and library production were performed at the NCI (for additional information on the technical features of the study, see "Molecular Profiling of Prostate Cancer" below). The libraries were subsequently arrayed at Lawrence Livermore National Laboratories, and selected clones were sent to the Genome Sequencing Center at Washington University. The sequence data were returned to the National Center for Biotechnology Information where they were filtered and entered into the database of expressed sequence tags (dbEST). The flow of reagents and information essentially followed that initially designed by the Integrated Molecular Analysis of Genomes and their Expression consortium.[5]

Twelve microdissection-based libraries were produced from epithelial components of radical prostatectomy or biopsy specimens, including normal epithelium, premalignant foci, locally invasive cancer, and metastatic cancer (see Table 1). A total of 29,183 successful sequences was performed. Analysis of the number and frequency of genes expressed showed that all of the libraries exhibited a high level of complexity. The majority of genes were observed only once or twice in each library, and the overall gene diversity (number of genes identified/number of sequences analyzed) averaged 39.1%, which compares favorably with standard libraries derived from whole tissue specimens or cultured cells. Moreover, a wide range of expression was seen, from genes observed at high levels (prostate-specific antigen, β-microseminoprotein) that are known to be abundant in prostate epithelium, to a large number of low-abundance genes that were observed infrequently. Thus, the data clearly demonstrate the feasibility of recovering complex transcriptomes from microdissected cell populations, encouraging news for investigators interested in molecular profiling studies of clinical samples.

Applications of Molecular Profiling Data Sets

The first goal of the data analysis was to determine a prostate epithelial unigene set, ie, a catalogue of genes expressed in normal and malignant prostate epithelium. Clustering analysis of sequences derived from the libraries revealed expression of more than 6000 different epithelial genes, representing 35 to 50% of the estimated total, presumably including all of the genes that are expressed at high levels. The epithelial unigene set serves as a foundation for multiple analyses of gene expression. Five separate examples are briefly described below.

Prostate-Unique Gene Expression

Comparison of the expression patterns in the prostate libraries with all of the library sequence information in dbEST permits identification of genes that are unique to prostate epithelium as well as those that are expressed at significantly higher levels in prostate than in other cell types. These genes are of biological interest, due to their presumed specialized function in the gland, as well as

Table 1. Summary of Microdissection-Based Libraries

Library no.	Library Name	Patient	Sample type	Sequences	No. of new genes discovered	% Diversity
281	NCI_CGAP_Pr1	1	Normal epithelium	5689	152	35.2
515	NCI_CGAP_Pr5	2	Normal epithelium	805	8	40
526	NCI_CGAP_Pr9	3	Normal epithelium	1104	10	46.1
529	NCI_CGAP_Pr11	4	Normal epithelium	1376	15	45.2
282	NCI_CGAP_Pr2	1	Premalignant lesion	5688	119	34.9
511	NCI_CGAP_Pr6	2	Premalignant lesion	1462	24	42.6
538	NCI_CGAP_Pr7	2	Premalignant lesion	468	5	39.1
544	NCI_CGAP_Pr4/4.1	1	Premalignant lesion	1928	24	37.8
283	NCI_CGAP_Pr3	1	Adenocarcinoma	5209	135	29.6
513	NCI_CGAP_Pr8	2	Adenocarcinoma	1100	14	42.4
527	NCI_CGAP_Pr10	3	Adenocarcinoma	1139	15	42.6
523	NCI_CGAP_Pr12	5	Metastatic adenocarcinoma	3215	38	33.6
				29,183	559	39.1

potentially useful as diagnostic or therapeutic targets. For example, prostate-specific proteins localized to the cell surface may serve as targets for antibody-mediated delivery of therapeutic compounds.[26] Alternatively, knowledge of the promoter regions of prostate-unique genes could have value for virally mediated gene therapy by restricting transcription to prostate epithelial cells. For new serum protein markers of cancer, transcripts that are both highly expressed in tumors and unique to prostate epithelium have the most potential, because their gene products will be the easiest to detect and monitor based on levels of abundance. As an example, prostate-specific antigen is the current standard as a serum marker for prostate cancer, and its transcript was consistently observed at high levels in the libraries.

Integration: Genome, Expression, and Disease

Expression profiles of the prostate epithelial libraries can be integrated with GeneMap'99 to examine specific areas of the genome implicated in cancer. For example, chromosomal arms 1q, 8q, 8p, 13q, 16q, and Xq have been identified as important in prostate tumorigenesis based on linkage studies or chromosomal abnormalities observed in tumors.[27–33] The responsible gene at each of these regions has yet to be identified. The standard approach to finding such genes involves narrowing the physical size of the candidate interval using techniques such as meiotic recombination or marker disequilibrium in affected families, or tumor deletion/amplification in sporadic cases.[34–36] An adjunct approach is to use expression patterns to narrow the region, ie, to prioritize the subset of genes for analysis that map to the minimal search interval and are expressed in the involved tissue. The *MEN1* and *PTEN* genes are examples of recently identified tumor suppressor genes that are found in appropriate libraries (*MEN1*, NCI CGAP Lu5; *PTEN*, NCI CGAP Pr3/Pr22).[37–39] Integration of cell type-specific gene expression and transcript map location is likely to become an increasingly valuable approach for disease gene hunting as molecular profiling databases grow and sequencing and mapping of all human genes are completed.

cDNA Microarray-Based Profiling

Investigators using expression arrays to study prostate tumorigenesis can prioritize the prostate epithelial unigene set for study. This has both short-term and long-term advantages. In the near term, a practical strategy is to use the prostate unigene set on an expression array and focus on measuring the genes of moderate or high abundance whose expression levels change substantially during tumorigenesis. To facilitate these efforts the prostate expression data were used to create a commercially available prostate cDNA expression microarray, which includes a majority of the epithelial unigenes, including those uniquely expressed in prostate.[40] The major long-term challenge of array-based studies will be quantitative measurement of small expression changes, particularly for those genes present at low levels. Refinement of experimental strategies will likely be required, such as gene-specific primers to prepare cDNA for analysis and careful selection of sequences used on the array to avoid cross-hybridization. Efforts to design such custom arrays will be facilitated by a successive reduction in the number of genes required for analysis, beginning with prioritization of the relevant unigene set and eventually reducing to the specific set of genes that mediate the pathways and processes under study.

Single Nucleotide Polymorphisms (SNPs)

The genetic variation in genes that are found to be important in prostate cancer can be determined through the Genetic Annotation Initiative (GAI) section of the CGAP website. The GAI focuses on identifying SNPs in genes expressed in cancers.[25,41] Analysis of the frequency and transmission of SNPs can be used for many genetic studies, including traditional linkage mapping and dissection of complex pathways. Gene-specific SNPs are also valuable polymorphic markers for finely mapping regions of allelic loss in tumor loss of heterozygosity studies. The GAI identifies candidate SNPs through an analytical software package called SNPpipeline and then verifies the variation by sequencing DNA from several individuals. To date, more than 10,000 candidate SNPs

have been identified. To make the information easy to access, all SNP data are placed on an integrated genetic/physical SNP map available through the GAI website.

Differential Gene Expression

An important use of molecular profiling data sets is to compare and contrast the expression profiles that occur during evolution of a disease process. Thus, we analyzed the sequence data from the normal epithelial, premalignant, and invasive tumor libraries using a variety of statistical methods and identified the genes that were differentially expressed during tumor progression. The transcripts that showed the largest change between normal and tumor cells were a subset of mRNAs that encode for ribosomal proteins. This finding is expected in cancer cells due to their requirement for increased protein synthesis for cell division.[12] Interestingly, though, these ribosomal protein mRNAs were not increased in libraries from premalignant cells that showed expression levels similar to normal epithelium. This finding is at odds with most current thinking, which presumes that premalignant foci develop due to a marked increase in growth rate, with subsequent transition to cancer primarily involving acquisition of an invasive phenotype. Based on the present gene expression data set, one can propose two alternative hypotheses for testing. First, premalignant cells do not proliferate at a rate near that of invasive tumor cells, and fundamental alterations in oncogene and/or tumor suppressor gene pathways that substantially increase the rate of cell division are still required for their progression to cancer. Second, a decreased rate of apoptosis is an important early event in prostate tumor progression; ie, it is a decreased rate of cell death, as opposed to an increase in cell division, that mediates the development of premalignant foci.

In addition to expected findings such as increased ribosomal protein transcripts in cancer, several unanticipated discoveries were made, including both quantitative and qualitative alterations in gene expression. For example, the transcript for T cell receptor γ was found in normal and cancerous prostate epithelium, and observed at statistically elevated levels in cancer libraries. The presence of T cell receptor γ mRNA in prostate epithelium and the high level of expression in tumor cells is both surprising and puzzling. A second example was detection of a novel splice variant of PB39 transcript in a library derived from premalignant cells. PB39 mRNA was previously reported to be overexpressed in prostate cancer, but was not known to exist in an alternative splice form.[42] Interestingly, based on a search of all cDNA libraries and sequences in dbEST the novel splice variant is primarily expressed in fetal tissues and tumors and thus may be associated with the loss of cellular differentiation that occurs during prostate tumor progression.[43] Additionally, PHDhtm and SignalP computer-based analysis of the predicted amino acid sequence of PB39 indicates the N-terminus contains a secretory signal peptide sequence for a secreted protein. Thus, the protein product of the alternative splice form could potentially serve as a serum marker of early prostate cancer development.

Certainly, the significance of ribosomal protein mRNAs, T cell receptor γ mRNA, and PB39 splice variant mRNA in prostate tumors and premalignant lesions remains to be determined in follow-up studies. However, the larger implication of these findings is immediately clear. There is much yet to be learned with respect to gene expression profiles in complex human tissues. Thus, exploratory studies using developing expression technologies and the information provided by the Human Genome Project are likely to have a unique and important role in the study of normal cell physiology and the development of diseases.[17] In this regard, the present study is encouraging and indicates molecular profiling of clinical tissue specimens is a feasible and promising experimental approach.

Molecular Profiling of Prostate Cancer

Case Selection

Samples from five different patients were included in the study to determine whether molecular profiling could be routinely performed on clinical specimens. The five cases were randomly selected from the NCI frozen tissue bank, and 12 libraries were produced. Each specimen was snap-frozen within 15 minutes of surgical resection, but no other special procedures were used for handling the tissues.

Tissue Acquisition

The goal of molecular profiling of human tissue specimens is to measure global gene expression levels as they exist in cells in patients. In the present study the libraries were created from tissues that had been surgically removed; thus, it is possible that alterations in gene expression profiles occurred during or after the resection, eg, transcription of new genes due to environmental stress or loss of transcripts during tissue handling. This is an important issue that needs to be addressed experimentally in the future by comparing molecular profiles of needle biopsy samples (immediate removal and freezing) with surgically resected samples of the same tissue type. If molecular alterations are shown to occur in surgical specimens, then two potential scenarios arise that will affect how samples should be acquired for future molecular profiling studies. In the first scenario, the induced changes are minimal and occur reproducibly, and thus can be predicted and factored into subsequent data analyses. In this case surgically resected samples will be useful templates for study as long as they are appropriately acquired and processed. In the second, the induced changes are substantial and cannot reliably be predicted. In this case, future molecular profiling efforts will need to use biopsy or cytology samples as templates, and/or will need to be performed like intraoperative diagnostic frozen section analysis; ie, at the outset of the operation the surgeon will need to procure and immediately freeze several small tissue samples for molecular profiling studies.

Microdissection

Cells were procured by either manual microdissection or the initial prototype laser capture microdissection instrument.[44,45] Based on careful histopathological review of the tissue sections, it is estimated each sample contained >90% of desired cells. Newer laser-based dissection systems and associated methodologies currently allow for dissections approaching 100% purity.[46–48] Following are some technical observations made during the course of the study. Rapid dehydration of cryostat sections is important to inactivate endogenous RNases. Staining with hematoxylin and eosin allows microscopic visualization during microdissection and does not significantly diminish mRNA recovery. Approximately 5000 microdissected cells are required to produce a library with acceptable numbers of recombinants (>100,000) and gene diversity (>20%).

Library Preparation and Characteristics

Detailed protocols for all of the 156 CGAP libraries are indicated on the web page. Each of the 12 prostate libraries in the present study was made using microdissection library protocol no. 1.[49,50] Evaluation of the library sequence data showed two important characteristics that impact on the overall utility of a microdissection-based approach. First, the clone insert size averaged only 500 to 600 bp in length due to the fragmented mRNA recovered from tissue samples. Technical attempts to increase the insert size were not considered a high priority, because the libraries were intended solely for gene profiling and not as templates for full-length gene cloning. Second, the number of recombinants ranged from approximately 100,000 to 200,000 per library, substantially less than in traditional libraries. Additional PCR cycles of cDNA could increase the number of recombinants significantly; however, because the libraries were prepared for expressed sequence tag (EST) analysis as opposed to traditional screening, the number of PCR cycles was limited to 10 to minimize amplification bias.

Assessment of Library Quality

Measurement of one or a few genes from small numbers of cells using RT-PCR is relatively straightforward to perform. However, global expression profiling studies are significantly more challenging, because the recovered mRNA and subsequent cDNA must contain a complex set of genes reflective of the native abundance of the transcript population. Gene diversity (number of genes observed/number of sequences) was used as the measure of cDNA library quality and was determined by sequencing a minimum of 500 randomly selected clones per library. This was sufficient to provide a statistically reliable indicator of complexity and was a useful tool that provided a rigorous measure of library diversity. Additionally, the expression frequency of all individual genes observed was calculated to determine relative levels of abundance.

Informatics Analysis

Several analysis tools and all of the present prostate data are provided on the CGAP website (www.ncbi.nlm.nih.gov/ncicgap/) to allow statistical comparison of gene expression profiles in the libraries. For additional information, relevant website links include:

NCBI, www.ncbi.nlm.nih.gov/
Unigene, www.ncbi.nlm.nih.gov/UniGene/index.html
LibraryBrowser,www.ncbi.nlm.nih.gov/UniGene/lbrowse.cgi?ORG=Hs
GeneMap'99, www.ncbi.nlm.nih.gov/genemap/
CGAP GAI www.lpg.nci.nih.gov/GAI/
dbEST, www.ncbi.nlm.nih.gov/dbEST/index.html
Genes and Diseases, www.ncbi.nlm.nih.gov/disease/
CGAP Update, www.nih.gov/news/pr/aug99/nci-10a.htm

The dbEST and Unigene sites are continually updated. Investigators should query the data sets using the latest Unigene build for the most up-to-date information. As with all projects using EST data, one must use caution in interpreting results, and candidate genes of interest should be subjected to rigorous follow-up analysis.

Prostate-Unique and Prostate-Specific Genes

CGAP website tools were created to be capable of generating two different classifications. Prostate-unique genes are those that have been observed only in libraries derived from prostate and are precalculated on the website (query "prostate" under the "Summary Tables of Libraries, Genes and Sequences" section). Prostate-specific genes include those expressed at statistically elevated levels in prostate epithelial libraries compared to libraries from other cell and tissue types. These can be determined using the Digital Differential Display tool.

(Prostate-unique genes are included in this category based solely on detection in cDNA libraries used as part of EST projects. A subset of these genes may have been observed in non-prostate tissue in other studies. It is anticipated that with additional EST sequencing some of these genes will shift to the "prostate-specific" category or will drop out of both classifications.)

cDNA Microarray-Based Studies

We have observed two noteworthy features of expression array studies. First, intense artifactual hybridization signals can be problematic for 3′ cDNA clone-based arrays due to hybridization to polyA sequences and repetitive DNA elements, ie, samples can appear to hybridize successfully to a large number of genes on an array when in fact the majority of signal is artifactual. Thus, one must be careful in evaluating the apparent success of methods to prepare array samples from small numbers of microdissected cells, and must also be cautious in using array results to construct a unigene set of expressed genes from a given cell type. Secondly, individual transcripts often hybridize strongly to at least a few additional DNAs on arrays besides the intended DNA and hybridize less strongly to many DNAs. This cross-hybridization re-

duces the overall sensitivity of arrays to detect expression changes, and importantly, must be carefully considered when using gene cluster algorithms to analyze array results.

Differential Gene Expression

This effort was considered a lesser priority goal of the project, since it was thought that relatively few statistically valid differences in gene expression could be determined based on the amount of sequencing planned for the study. In fact, analysis of the library data proved this to be the case. Even with completion of nearly 30,000 total sequences, the majority of epithelial genes were not expressed at sufficient levels or in enough libraries to permit a reliable statistical assessment of differential expression. However, the gene distribution profile in the libraries indicates that comparison of expression levels of a significant fraction of the prostate epithelial unigene set could be achieved by using substantially greater sequencing depth.

Initially, it was presumed that contamination of lymphocytes during the dissection step was responsible for the presence of T cell receptor γ transcript in the prostate libraries.[51] However, epithelial localization was confirmed by *in situ* hybridization studies of tissue sections and appears selective for the γ chain, as other components of the T cell receptor were not observed in the prostate libraries.[52]

References

1. Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, Walters L, and the members of the DOE and NIH planning groups: New goals for the U.S. human genome project, 1998–2003. Science 1998, 23: 682–689
2. Venter JC, Adams MD, Sutton GG, Kerlavage AR, Smith HO, Hunkapiller M: Shotgun sequencing of the human genome. Science 1998, 280:1540–1542
3. Adams MD, Krelavage AR, Fields C, Venter JC: 3,400 new expressed sequence tags identify diversity of transcripts in human brain. Nat Genet 1993, 4:256–267
4. Hillier LW, Lennon G, Becker M, Bonaldo MF, Chiapelli B, Chissoe S, Dietrich N, DuBuque T, Favello A, Gish W, Hawkins M, Hultman M, Kucaba T, Lacy M, Le M, Le N, Mardis E, Moore B, Morris M, Parsons J, Prange C, Rifkin L, Rohlfing T, Schellenberg K, Marra M: Generation and analysis of 280,000 human expressed sequence tags. Genome Res 1996, 6:807–828
5. Lennon G, Auffray C, Polymeropolous M, Soares MB: The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. Genomics 1996, 33:151–152
6. Deloukas P, Schuler GD, Gyapay G, Beasley EM, Soderlund C, Rodriguez-Tome P, Hui L, Matise TC, McKusick KB, Beckmann JS, Bentolila S, Bihoreau M, Birren BB, Browne J, Butler A, Castle AB, Chiannilkulchai N, Clee C, Day PJ, Dehejia A, Dibling T, Drouot N, Duprat S, Fizames C, Bentley DR, et al: A physical map of 30,000 human genes. Science 1998, 23:744–746
7. Strausberg RL, Feingold EA, Klausner RD, Collins FS: The mammalian gene collection. Science 1999, 286:455–457
8. Schena M, Shalon D, Davis RW, Brown P: Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 1995, 270:467–469
9. DeRisi JL, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM: Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat Genet 1996, 14:457–460
10. Nowak R: Entering the postgenome era. Science 1995, 270:368–371
11. Velculescu V, Zhang L, Vogelstein B, Kinzler KW: Serial analysis of gene expression. Science 1995, 270:484–487
12. Zhang L, Zhou W, Velculescu VE, Kern SE, Hruban RH, Hamilton SR, Vogelstein B, Kinzler KW: Gene expression profiles in normal and cancer cells. Science 1997, 276:1268–1272
13. Fodor SP, Rava RP, Huang XC, Pease AC, Holmes CP, Adams CL: Multiplexed biochemical assays with biological chips. Nature 1993, 364:555–556
14. Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ: High density synthetic oligonucleotide arrays. Nat Genet 1999, 21:20–24
15. Luo L, Salunga RC, Guo H, Bittner A, Joy KC, Galindo JE, Xiao H, Rogers KE, Wan JS, Jackson MR, Erlander MG: Gene expression profiles of laser-captured adjacent neuronal subtypes. Nat Med 1999, 1:117–122
16. Marton MJ, DeRisi JL, Bennett HA, Iyer VR, Meyer MR, Roberts CJ, Stoughton R, Burchard J, Slade D, Dai H, Bassett DE Jr, Hartwell LH, Brown PO, Friend SH: Drug target validation and identification of secondary drug target effects using DNA microarrays. Nat Med 1998, 4:1293–1301
17. Brown PO, Botstein D: Exploring the new world of the genome with DNA microarrays. Nat Genet 1999, 21:33–37
18. Collins FS: Microarrays and macroconsequences. Nat Genet 1999, 21:2
19. Phimister B: Going global. Nat Genet 1999, 21:1
20. Lander ES: Array of hope. Nat Genet 1999, 21:3–4
21. Shibata D: Pattern recognition and arrays: the times are a-changing. Am J Pathol 1999, 154:979–980
22. Cole KA, Krizman DB, Emmert-Buck MR: The genetics of cancer: a 3D model. Nat Genet 1999, 21:38–41
23. Strausberg RL, Dahl CA, Klausner RD: New opportunities for uncovering the molecular basis of cancer. Nat Genet 1997, 15:415–416
24. Pennisi E: A catalog of cancer genes at the click of a mouse. Science 1997, 276:1023–1024
25. Strausberg RL, Buetow KH, Emmert-Buck MR, Klausner RD: The Cancer Genome Anatomy Project: building an annotated gene index. Trends Genet 2000, 16:106–116
26. Lorimer IA, Wikstrand CJ, Batra SK, Bigner DD, Pastan I: Immunotoxins that target an oncogenic mutant epidermal growth factor receptor expressed in human tumors. Clin Cancer Res 1995, 1:859–864
27. Bova G, Isaacs WB: Review of allelic loss and gain in prostate cancer: World J Urol 1996, 14:338–346
28. Dong JT, Isaacs W, Isaacs JT: Molecular advances in prostate cancer. Curr Opin Oncol 1997, 9:101–107
29. Emmert-Buck M, Vocke CD, Pozzatti RO, Duray PH, Jennings SB, Florence CD, Zhuang Z, Bostwick DG, Liotta LA, Linehan WM: Allelic loss on chromosome 8p12–21 in microdissected prostatic intraepithelial neoplasia (PIN). Cancer Res 1995, 55:2959–2962
30. MacGrogan D, Levy A, Bostwick DG, Wagner M, Wells D, Bookstein R: Loss of chromosome arm 8p loci in prostate cancer: mapping by quantitative allelic imbalance. Genes Chromosomes Cancer 1994, 10:151–159
31. Smith J, Freije D, Carpten JD, Gronberg H, Xu J, Isaacs SD, Brownstein MJ, Bova GS, Guo H, Bujnovszky P, Nusskern DR, Damber JE, Bergh A, Emanuelsson M, Kallioniemi OP, Walker-Daniels J, Bailey-Wilson JE, Beaty TH, Meyers DA, Walsh PC, Collins FS, Trent JM, Isaacs WB: Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. Science 1996, 274:1371–1374
32. Trapman J, Sleddens HF, van der Weiden MM, Dinjens WN, Konig JJ, Schroder FH, Faber PW, Bosman FT: Loss of heterozygosity of chromosome 8 microsatellite loci implicates a candidate tumor suppressor gene between the loci D8S87 and D8S133 in human prostate cancer. Cancer Res 1994, 54:6061–6064
33. Xu J, Meyers D, Freije D, Isaacs S, Wiley K, Nusskern D, Ewing C, Wilkens E, Bujnovszky P, Bova GS, Walsh P, Isaacs W, Schleutker J, Matikainen M, Tammela T, Visakorpi T, Kallioniemi OP, Berry R, Schaid D, McDonnell S, Schroeder J, Blute M, Thibodeau S, Trent J: Evidence for a prostate cancer susceptibility locus on the X chromosome. Nat Genet 1998, 20:175–179
34. Emmert-Buck M, Lubensky IA, Dong Q, Chandrasekharappa C, Guru SC, Manickam P, Keseter M, Olufemi S-E, Agarwal S, Burns AL, Spiegel AM, Collins FS, Marx SJ, Zhuang Z, Liotta LA, Debelenko LV: Localization of the multiple endocrine neoplasia type I (MEN1) gene based on tumor deletion mapping. Cancer Res 1997, 57:1855–1858

35. Emmert-Buck M, Debelenko LV, Kester M, Agarwal S, Manickam P, Zhuang Z, Guru SC, Olufemi S-E, Burns AL, Chandrasekharappa S, Lubensky IA, Liotta LA, Skarulis MC, Spiegel AM, Marx SJ, Collins FS: 11q13 Allelotype analysis in 27 northern american MEN1 kindreds identifies two distinct founder chromosomes. Mol Genet Metab 1998, 63:151–155
36. Debelenko L, Emmert-Buck MR, Manickam P, Guru SC, Kester M, DiFranco EM, Olufemi S-E, Agarwal S, Lubensky IA, Zhuang Z, Burns AL, Spiegel AM Liotta LA, Collins FS, Marx SJ, Chandrasekharappa S: Haplotype analysis defines a minimal interval for the multiple endocrine neoplasia type 1 (MEN1) gene. Cancer Res 1997, 57: 1039–1042
37. Chandrasekharappa S, Guru SC, Manickam P, Olufemi S, Collins FS, Emmert-Buck MR, Debelenko LV, Zhuang Z, Lubensky IA, Liotta LA, Crabtree JS, Wang Y, Roe BA, Weisemann J, Boguski MS, Agarwal SK, Kester MB, Kim YS, Heppner C, Dong Q, Spiegel AM, Burns AL, Marx SJ: Positional cloning of the gene for multiple endocrine neoplasia type 1. Science 1997, 276:404–407
38. Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, Puc J, Miliaresis C, Rodgers L, McCombie R, Bigner SH, Giovanella BC, IM, Tycko B, Hibshoosh H, Wigler MH, Parsons R: PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science 1997, 275:1943–1947
39. Debelenko LV, Brambilla E, Agarwal SK, Swalwell JI, Lubensky IA, Zhuang Z, Guru SC, Manickam P, Olufemi S-E, Chandrasekharappa SC, Crabtree JS, Kester MB, Kim YS, Heppner C, Burns AL, Spiegel AM, Marx SJ, Liotta LA, Collins FS, Travis WD, Emmert-Buck MR: Identification of *MEN1* gene mutations in sporadic carcinoid tumors of the lung. Hum Mol Genet 1997, 6:2285–2290
40. Carlisle AJ, Prabhu VV, Elkahloun A, Hudson J, Trent JM, Linehan WM, Williams ED, Emmert-Buck MR, Liotta LA, Munson PJ, Krizman DB: Development of a prostate cDNA microarray and statistical gene expression analysis package. Mol Carcinogen 2000, 27:1–11
41. Buetow KH, Edmonson MN, Cassidy AB: Reliable identification of large numbers of candidate SNPs from public EST data. Nat Genet 1999, 21:323–325
42. Chuaqui R, Englert CR, Strup S, Vocke CD, Zhuang Z, Duray PH, Bostwick DG, Linehan WM, Liotta LA, Emmert-Buck MR: PB39: Identification of a novel gene up-regulated in clinically aggressive human prostate cancer. Urology 1997, 50:302–307
43. Cole KA, Chuaqui RF, Katz K, Pack S, Zhuang Z, Cole CE, Lyne JC, Linehan WM, Liotta LA, Emmert-Buck MR: cDNA sequencing and analysis of PB39: A novel gene up-regulated in prostate cancer. Genomics 1998, 51:282–287
44. Emmert-Buck M, Roth MJ, Zhuang Z, Campo E, Rozhin J, Sloane BF, Liotta LA, Stetler-Stevenson WG: Increased gelatinase A and cathepsin B activity in invasive tumor regions of human colon cancer samples. Am J Pathol 1994, 145:1285–1290
45. Emmert-Buck M, Bonner RF, Smith PD, Chuaqui RF, Goldstein SR, Zhuang Z, Weiss RA, Liotta LA: Laser capture microdissection. Science 1996, 274:998–1001
46. Bonner R, Emmert-Buck MR, Cole KA, Pohida T, Chuaqui RF, Goldstein SR, Liotta LA: Laser capture microdissection: molecular analysis of tissue. Science 1997, 278:1481–1483
47. Fend F, Emmert-Buck MR, Chuaqui R, Cole KA, Lee J, Liotta LA, Raffeld M: Immuno-LCM: laser capture microdissection of immunostained frozen sections for mRNA analysis. Am J Pathol 1999, 154: 61–66
48. Schutze K, Lahr G: Identification of expressed genes by laser-mediated manipulation of single cells. Nat Biotech 1998, 16:737–740
49. Krizman D, Chuaqui RF, Meltzer PS, Trent JM, Duray PH, Linehan WM, Liotta LA, Emmert-Buck MR: Construction of a representative cDNA library from prostatic intraepithelial neoplasia (PIN). Cancer Res 1996, 56:5380–5383
50. Peterson LA, Brown MR, Carlisle AJ, Kohn EC, Liotta LA, Emmert-Buck MR, Krizman DB: An improved method for construction of directionally cloned cDNA libraries from microdissected cells. Cancer Res 1998, 58:5326–5328
51. Vasmatzis G, Essand M, Brinkmann U, Lee B, Pastan I: Discovery of three genes specifically expressed in human prostate by expressed sequence tag database analysis. Proc Natl Acad Sci USA 1998, 95:300–304
52. Essand M, Vasmatzis G, Brinkmann U, Duray P, Lee BK, Pastan I: High expression of a specific T-cell receptor gamma transcript in epithelial cells of the prostate. Proc Natl Acad Sci USA 1999, 96: 9287–9292